2014

Glycosyl rotation and distortion by key residues in Endocellulase Cel6A from *Theromobifida fusca*

Tao Lu
University of Nebraska-Lincoln

Zuoming Zhang
Jilin University, Changchun, China, zmzhang@jlu.edu.cn

Chi Zhang
University of Nebraska - Lincoln, zhang.chi@unl.edu

Published in *Glycobiology* 24:3 (2014), pp. 247–251; doi: 10.1093/glycob/cwt105

Submitted July 11, 2013; revised November 19, 2013; accepted November 22, 2013; published online November 28, 2013.

Glycosyl rotation and distortion by key residues in Endocellulase Cel6A from *Theromobifida fusca*

Tao Lu,[1] Zuoming Zhang,[2] and Chi Zhang [1]

1 School of Biological Sciences, University of Nebraska-Lincoln, Lincoln, NE 68588, USA

2 Key Laboratory for Molecular Enzymology and Engineering of Ministry of Education, Jilin University, Changchun 130023, China

Corresponding authors — Z. Zhang, zmzhang@jlu.edu.cn ; C. Zhang, czhang5@unl.edu

Abstract

Endocellulases are one kind of the important biodegrading cellulose enzymes. Experimental results show that a rotated and distorted preactivated structure exists before the substrate entering the transition state. The molecular dynamic simulation of endocellulase Cel6A models revealed a correlation between the rotation and distortion of pyranoside ring in −1 glycosyl unit of the substrate. The two key residues, Tyr73 and Ser189, in Cal6A cooperate to rotate and distort the pyranoside ring in the cellulose hydrolysis.

Keywords: cellobiohydrolase, cellulose, endocellulase, hydrolysis mechanism, molecular dynamic simulations

Abbreviations: CPCM, conductor-like polarizable continuum model; GU, glycosyl unit; Glu2, cellobiose; MD, molecular dynamic; MM, molecular mechanical; PES, potential energy surface; QM, quantum mechanical; TS, transition states; 3D, three-dimensional.

Introduction

Cellulose, a polymer of β-1,4 linked glucosyl units, very similar to the α-1,4 linkages in starch, is an abundant renewable biomass. Native cellulose is difficult to degrade because of their strong β-1,4 linkages with the *k* values on the order of 10^{-15}/s at room temperature. On the other hand, some microbes produce cellulases that can degrade cellulose. Cellulase is not a solo enzyme, which contains two main sorts of β-1,4 glycosidic hydrolyze enzymes: Cellobiohydrolases (EC.3.2.1.91) and endoglucanases (EC.3.2.1.4). Cellobiohydrolases cut a cellulose chain into cellobiose at the chain end, while endoglucanases cut cellulose chains into short chains in the middle of cellulose. Cel6A, one of the most widely studied endocellulase, is produced by the thermophilic soil bacterium, *Theromobifida fusca*. Cel6A has two domains, a catalytic one at the amino terminus and a cellulose-binding one at the carboxy terminal end. The catalytic domain is the main focus of this paper, especially the residues related to the substrate structural transformations.

The Cel6A hydrolyzes the β-1,4 linkage by a single-displacement mechanism (Knowles et al. 1988). Crystallography studies of the Cel6A ligand complex structures showed that the pyranoside ring distortion in the −1 glycosyl unit (GU) is a key component of the hydrolysis mechanism (Knowles et al. 1988; Damude et al. 1996; Varrot et al. 1999; Larsson et al. 2005). Conformational analyses of the reaction coordinate showed that the best substrate preactivation in β-D-glucopyranose is a skewboat, $^{2}S_{0}$ or $^{1}S_{3}$ (JCBN 1980), which were identified by both experimental and simulation studies (Knowles et al. 1988; Damude et al. 1996; Zou et al. 1999). The hydrolysis transition states (TS) conformation is a preactivated $B_{2,5}$, which appears in the $^{2}S_{0} \leftrightharpoons B_{2,5}$ transformation of −1 GU(Davies et al. 2012). The pyranoside ring of −1 GU in TS is distorted away from its lowest-energy conformation to one that favors orbital overlap (Davies et al. 2012). The preactivated substrate formation is a key process in the catalytic pathway of cellobiohydrolase, because, with the plane rotation and distortion, the −1 GU is pushed to a high energy level and easy to trap in the catalytic process (Davies et al. 2012). The conformational free-energy change in the preactivated conformation transformation is ~5–6 kcal/mol (Davies et al. 2012). Barnet et al. (2011) found that the free-energy barrier is 17 kcal/mol to break the cellulose chain of Cel6A.

A crystallography study of the Cel6A showed that to form the distortion structure of the pyranoside ring, −1 GU first rotates to a nearly perpendicular position to +1 GU (Zou et al. 1999). The conserved tyrosine that binds to the −1 GU determines the pyranoside ring distortion (Koivula et al. 2002b; Varrot et al. 2003). Moreover, the mutation of the Tyr73 causes the ring distortion to disappear, and this mutation even reduces the rotation angle of −1 GU (Koivula et al. 2002b; Larsson et al. 2005). However, when all the key tyrosines are mutated, crystal structures still present more or less some pyranoside ring distortion at the −1 GU, though the enzyme activity is reduced (Barr et al. 1998, Zou et al. 1999; Larsson et al. 2005). Therefore, there may exist other residues that affect the −1 GU distortion.

Results and discussion

In order to study the −1 GU rotation and distortion in the catalytic pathway, we conducted molecular dynamic (MD) simulations to a wild-type (WT) Cel6A-ligand complex structure and a series of its mutants. In the simulation trajectory, the relationship between the rotation and distortion of pyranoside ring in −1 GU was studied. The −1 GU plane rotation of each conformation in the trajectory was measured by the angle between the normal vectors of two pyranoside ring planes on −1 and +1 GUs. The planes of two pyranoside rings were obtained by a least-square-fitting to six atoms, and the rotation angle from 0 to 90° means

that the pyranoside ring in −1 GU rotates from the parallel position with the pyranoside ring in +1 GU to the perpendicular position. The distortion to $^2S_0$ is measured by the angle between the normal vectors of two planes: The least-square fitted plane of C_{B1}-C_{B2}-C_{B3}-C_{B4}-O and the triangle plane of C_{B4}-C_{B5}-O in −1 GU ring (Figure 1). The distortion angle changes from 20 to 70°, which corresponds to a change from no distortion to full distortion, and Supplementary data, Figure S4 shows that different structures with the same distortion angle have different standard Cremer–Pople ring pucker parameters (Cremer and Pople 1975). The distribution of pyranoside ring conformations with different rotation and distortion angles is shown in Figure 2. A correlation between the rotation and distortion of pyranoside ring in −1 GU is revealed. The more perpendicular a conformation of −1 GU pyranoside ring is to the +1 GU pyranoside ring, the more its distortion. This correlation indicates that the rotation of pyranoside ring in −1 GU is a necessary step for its distortion. From quantum mechanical/molecular mechanical (QM/MM) trajectories, the change in potential energy from the normal $^2S_0$ conformation to $B_{2,5}$ distorted conformation is ~8.66 kcal/mol, and the TS structure, the $B_{2,5}$ conformation, is shown in Supplementary data, Figure S3. The potential energy surface (PES) was scanned along the C_{B2}-C_{B1}-C_{A4}-C_{A3} dihedral angle in cellobiose (Glu2), and it showed that the pyranoside ring distortion in −1 GU needs less energy than the ring plane rotation (Supplementary data, Figure S2). This discovery also supports the same conclusion that rotation of pyranoside ring in −1 GU is a necessary step for its distortion.

According to the complex structures, there are only five residues within close proximity of −1 GU, defined as having closest atoms within 4 Å distance from each other, and they are Tyr73, Ser189, Asp117, Lys259 and Asp265. The distances between these residues and −1 GU were collected, and different behaviors were uncovered. The residues Aps117, Lys259 and Asp265 have relatively constant distances from −1 GU. For example, there is an H-bond between oxygen on carboxyl of Asp and hydroxyl on C_3 of −1 GU, and the distance between these two atoms does not have significant change (Figure 3). Therefore, these residues do not directly contribute to the rotation of the pyranoside ring.

The distances between Tyr73/Ser189 and −1 GU vary with time (Figure 3). The oxygen atom of the hydroxyl in Tyr73 and the oxygen atom of the hydroxyl in Ser189 were used to represent the whole residues of Tyr73 and Ser189, respectively, because they have the shortest distances to −1 GU and have the potential to form hydrogen bonds with hydrogen atoms on −1 GU. The distance between the oxygen of the hydroxyl in Tyr73 and the hydrogen of the hydroxyl on C6 in −1 GU fluctuates in a range from 2.0 to 7.5 Å, while the distance between the hydrogen on the hydroxyl of C_6 in −1 GU and the oxygen of the hydroxyl in Ser189 fluctuates from 1.5 to 6.7 Å. Interestingly, the fluctuation directions of the distances of Tyr73 and Ser189 are

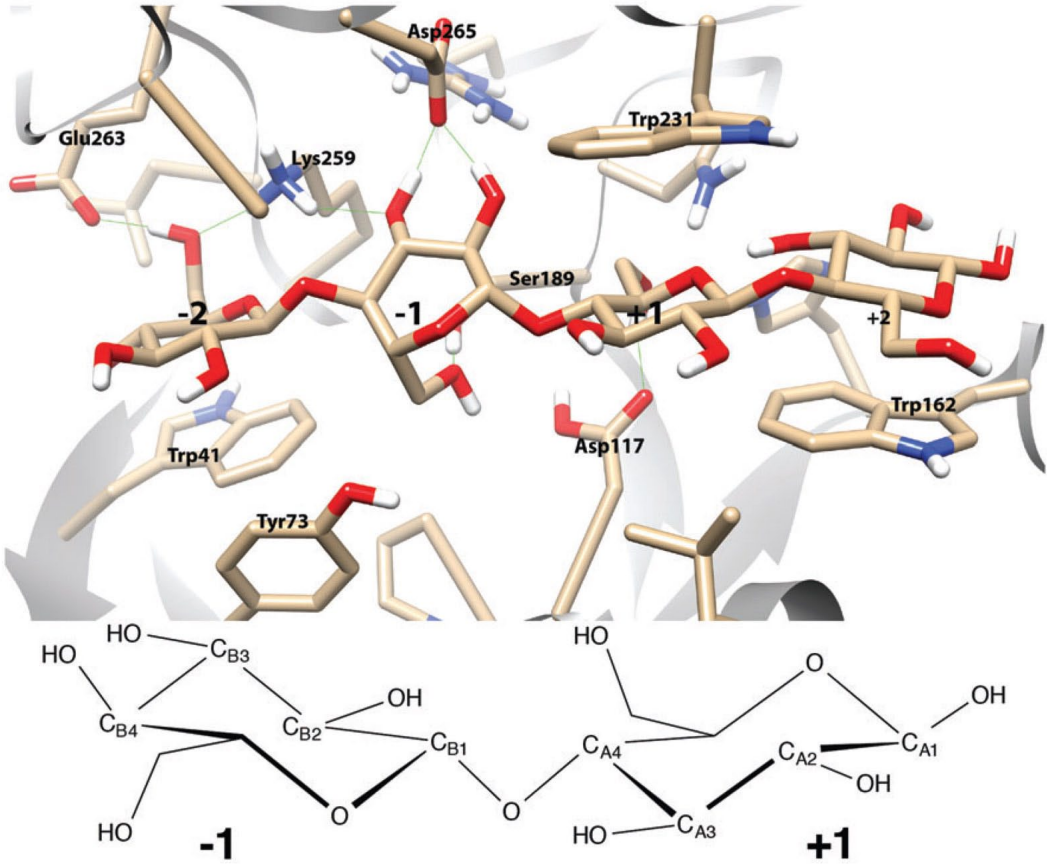


Figure 1. The catalytic cleft of Cel6A and two glycosyl units of cellulose.

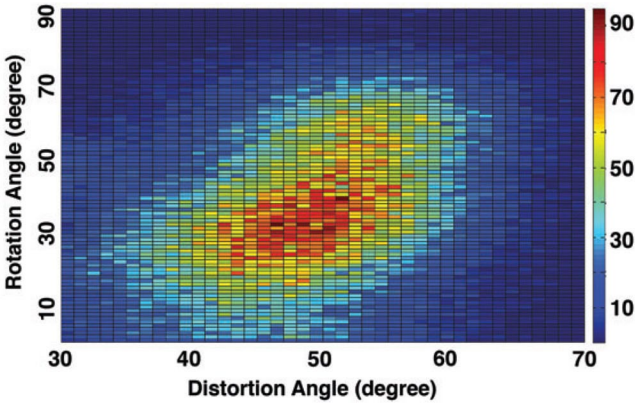


Figure 2. The distribution of numbers of conformations with certain pairs of the rotation and distortion of pyranoside ring in −1 GU in the simulation trajectory.

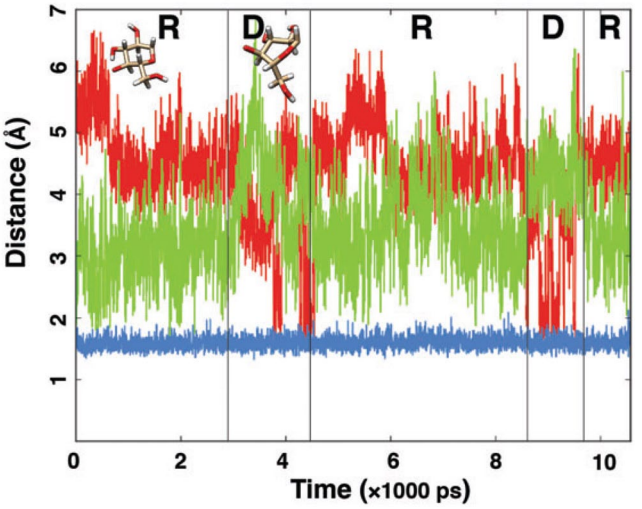


Figure 3. The distances between oxygen of hydroxyl in Tyr73 and hydrogen of hydroxyl on C_6 in -1 GU (green), oxygen of hydroxyl in Ser189 and hydrogen of hydroxyl on C_6 in -1 GU (red), oxygen of hydroxyl in Asp265 and hydrogen of hydroxyl on C_6 in -2 GU (blue). "R" region is for the relaxed structure of -1 GU, and "D" region is for completely distorted structure of -1 GU.

opposite. When two rings on −1 and +1 GU have parallel positions, Tyr73 has the shortest distance to −1 GU, whereas Ser189 has the shortest distance to −1 GU when it rotates to the perpendicular position to +1 GU. Since pyranoside ring in −1 GU is distorted at the perpendicular position with +1 GU, one can conclude that Tyr73 contributes to the rotation of pyranoside ring in −1 GU more than to its distortion, while Ser189 plays an important role for the distortion of pyranoside ring in −1 GU. An MD simulation on a mutant structure of Cel6A (Y73S) was conducted to study the rotation of the pyranoside ring in −1 GU (details of the simulation setup are in the Supplementary data). In the simulation trajectory, there are only a few conformations that have large rotation angles of pyranoside ring (Supplementary data, Figure S5). This study agrees with previous works (Koivula et al. 2002a; Larsson et al. 2005) and further supports the statement that Tyr73 is a key residue for the rotation of the pyranoside ring in −1 GU.

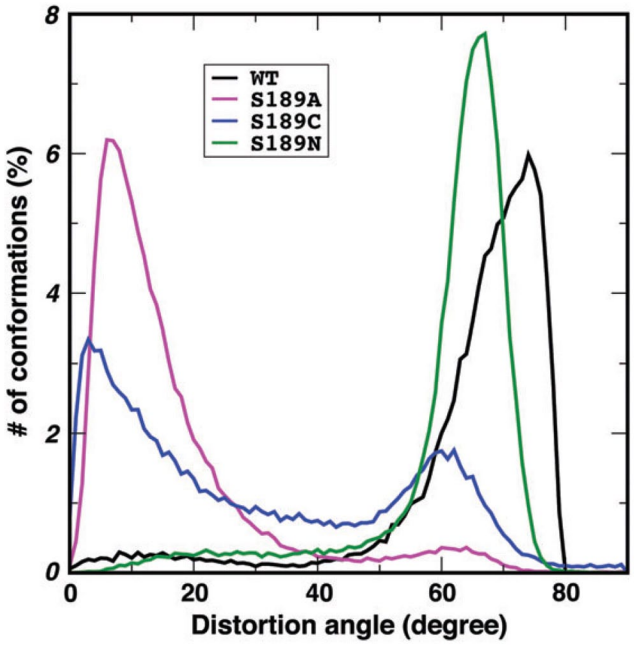


Figure 4. The distributions of conformations with rotation angles >45° for WT and three mutants were extracted from the simulation trajectories. The *y*-axis is the percentage of all conformations with rotation angles >45° in a trajectory

The MD simulation on the WT structure shows that, when the −1 GU pyranoside ring is distorted, some interactions are formed between Ser189 and the −1 GU, including an H-bond interaction and a Coulomb interaction. A theoretical study already showed that the positive charge accumulates on the pyranoside ring of TS substrate (Barnett et al. 2011). In the simulation model, Ser189 attacks the hydroxyl of C_6 on −1 GU, which is important for forming the ring distortion conformation. To further test the importance of Ser189 for ring distortion, in our study MD simulations were conducted for three mutants of Ser189: S189A, S189C and S189N, with the same setup as the simulation for the WT (see the Supplementary data for more details about simulation). The distribution of distortion angles in each simulation was studied. The same method was used to calculate the rotation and distortion angles. For three mutated models of Ser189, the distributions of the distorted conformations of pyranoside ring with a rotation angle of >45° were calculated, and the histogram is shown in Figure 4. When Ser189 is mutated to alanine, the distortion angles have a narrow peak of ~10°, while the WT has a peak ~75°. This indicates that the pyranoside ring is not distorted in the S189A mutated model. The reason is that alanine cannot form an H-bond with the hydrogen or oxygen of the hydroxyl on C_6 in −1 GU, and it also does not have negatively charged atoms to have Coulomb interaction with the pyranoside ring in −1 GU. Therefore, the pyranoside ring could not be distorted. When Ser189 is mutated to cysteine, the conformations have a smaller peak around distortion angle = 60° compared with the WT. At the same time, there is another small peak ~10°, which is similar to the mutation of S189A. The reason is that, although oxygen atom (serine) and sulfur atom (cysteine) have similar chemical properties, oxygen atom has a much higher electronegativity than sulfur atom, which causes cysteine

to have less chance to form H-bonds and weaker Coulomb interactions with the pyranoside ring to distort it. However, compared with serine, asparagine has more electronegative atoms, nitrogen and oxygen. Therefore, the mutant, Ser189Asn, could distort the pyranoside ring in −1 GU more easily than the WT because it has higher chance to form H-bonds and stronger Coulomb interactions with the pyranoside ring in −1 GU. The MD simulation supported this hypothesis; the distribution of rotation angles has a higher peak ~65° than that for the WT at 75°. The distributions of rotation angles of the pyranoside ring in −1 GU for three mutants were collected as well. All mutated models have similar distributions as the WT; there is only one peak ~40° for rotation angles (Supplementary data, Figure S6). Therefore, one can conclude that Ser189 is a key residue for pyranoside ring distortion.

In summary, the rotation pyranoside ring in −1 GU is important for its distortion, the $B_{2,5}$ preactivated conformation. The key residues Tyr73 and Ser189 in Cal6A cooperate to rotate the pyranoside ring and distort it. The residue Tyr73 pushes the pyranoside ring to rotate it to a position close to Ser189, and Ser189 affects the pyranoside ring distortion of −1 GU with H-bond interactions and coulomb interactions.

Materials and methods

The three-dimensional structure of Cel6A was obtained from PDB (ID: 2BOD) (Larsson et al. 2005), and the initial complex structure was built by replacing the sulfur atom in the inhibitor with a carbon atom. The complex structures were put into a cubic water box (15 × 15 × 15 $Å^3$) with TIP3P water (Jorgensen et al. 1983), and sodium ions were added to neutralize charges, after adding hydrogen atoms on both substrate and protein. The GROMACS MD package (Hess et al. 2008) with GLYCAM06 force field (Kirschner et al. 2008) and AMBER force field (Lindorff-Larsen et al. 2010) were used for cellulose and for protein, respectively, in the isothermal–isobaric ensemble. GROMACS, a widely used open source MD simulation package for biological molecules, supports various types of force fields, and can naturally conduct simulation with a combination of several different force fields. The GLYCAM06 force field is a force field specifically designed for carbohydrate molecule simulation and it is dependent on and compatible with AMBER force field to simulate the carbohydrate interaction with proteins. The trajectories of 10 ns were built with steps of 1 fs for each structure. In order to improve the simulation precision and to identify the TS structure of hydrolysis to compare our results with experimental data and previous theoretical studies, amino acid residues, Tyr73, Asp117 (protonated), Ser189, Lys259, +1 GU and −1 GU are treated with the QM/MM method. The density functional tight binding method (Elstner et al. 1998; Han et al. 2000), implemented in Gaussian 09 (Frisch et al. 2009), is used as the QM layer. A Glu2 structure model with lowest energy was established and the relaxed PES scan was performed using the B3LYP method, 6 − 31 + G (d,p) basis, and conductor-like polarizable continuum model implicit water solvation model in Gaussian 09 package (Frisch et al. 2009), to assess the structural parameters and potential energy changes in the QM/MM model.

Supplementary data are presented following the *References*.

Funding — This project was supported by funding under Dr. Chi Zhang's startup funds from University of Nebraska, Lincoln, NE, and Dr. Zuoming Zhang's grant from the National Basic Research Program of China (973 Program, 2012CB721003).

Acknowledgments — This work was completed utilizing the Holland Computing Center of the University of Nebraska. Dr. Xiao Cheng Zeng, Department of Chemistry, University of Nebraska–Lincoln, supplied Gaussian 09 package.

Conflict of interest — None declared.

References

Barnett CB, Wilkinson KA, Naidoo KJ. 2011. Molecular details from computational reaction dynamics for the cellobiohydrolase I glycosylation reaction. *J Am Chem Soc*. 133:19474–19482.

Barr BK, Wolfgang DE, Piens K, Claeyssens M, Wilson DB. 1998. Active-site binding of glycosides by *Thermomonospora fusca* endocellulase E2. *Biochemistry*. 37:9220–9229.

Cremer D, Pople JA. 1975. General definition of ring puckering coordinates. *J Am Chem Soc*. 97:1354–1358.

Damude HG, Ferro V, Withers SG, Warren RA. 1996. Substrate specificity of endoglucanase A from *Cellulomonas fimi*: Fundamental differences between endoglucanases and exoglucanases from family 6. *Biochem J*. 315(Pt 2):467–472.

Davies GJ, Planas A, Rovira C. 2012. Conformational analyses of the reaction coordinate of glycosidases. *Acc Chem Res*. 45:308–316.

Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G. 1998. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys Rev B*. 58:7260–7268.

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, et al. 2009. *Gaussian 09, Revision B.01*. Wallingford CT: Gaussian, Inc.

Han WG, Elstner M, Jalkanen KJ, Frauenheim T, Suhai S. 2000. Hybrid SCC-DFTB/molecular mechanical studies of H-bonded systems and of N-acetyl-(L-Ala)(n) N′-methylamide helices in water solution. *Int J Quantum Chem*. 78:459–479.

Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput*. 4:435–447.

JCBN I-I. 1980. IUPAC-IUB Joint-Commission on Biochemical Nomenclature (JCBN) conformational nomenclature for 5 and 6-membered ring forms of monosaccharides and their derivatives—Recommendations 1980. *Eur J Biochem*. 111:295–298.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 79:926–935.

Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J Comput Chem*. 29:622–655.

Knowles JKC, Lentovaara P, Murray M, Sinnott ML. 1988. Stereochemical course of the action of the cellobioside hydrolase I and hydrolase II of *Trichoderma reesei*. *J Chem Soc Chem Commun*. 21:1401–1402.

Koivula A, Ruohonen L, Wohlfahrt G, Reinikainen T, Teeri TT, Piens K, Claeyssens M, Weber M, Vasella A, Becker D, et al. 2002a. The active site of cellobiohydrolase Cel6A from *Trichoderma reesei*: The roles of aspartic acids D221 and D175. *J Am Chem Soc*. 124:10015–10024.

Koivula A, Ruohonen L, Wohlfahrt G, Reinikainen T, Teeri TT, Piens K, Claeyssens M, Weber M, Vasella A, Becker D, et al. 2002b. The active site of cellobiohydrolase Cel6A from *Trichoderma reesei*: The roles of aspartic acids D221 and D175. *J Am Chem Soc*. 124:10015–10024.

Larsson AM, Bergfors T, Dultz E, Irwin DC, Roos A, Driguez H, Wilson DB, Jones TA. 2005. Crystal structure of *Thermobifida fusca* endoglucanase Cel6A in complex with substrate and inhibitor: The role of tyrosine Y73 in substrate ring distortion. *Biochemistry*. 44:12915–12922.

Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct Funct Bioinformat*. 78: 1950–1958.

Varrot A, Macdonald J, Stick RV, Pell G, Gilbert HJ, Davies GJ. 2003. Distortion of a cellobio-derived isofagomine highlights the potential conformational itinerary of inverting beta-glucosidases. *Chem Commun*. 8:946–947.

Varrot A, Schülein M, Davies GJ. 1999. Structural changes of the active site tunnel of Humicola insolens cellobiohydrolase, Cel6A, upon oligosaccharide binding. *Biochemistry*. 38:8884–8891.

Zou J, Kleywegt GJ, Ståhlberg J, Driguez H, Nerinckx W, Claeyssens M, Koivula A, Teeri TT, Jones TA. 1999. Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Ce16A from *Trichoderma reesei*. *Structure*. 7:1035–1045.

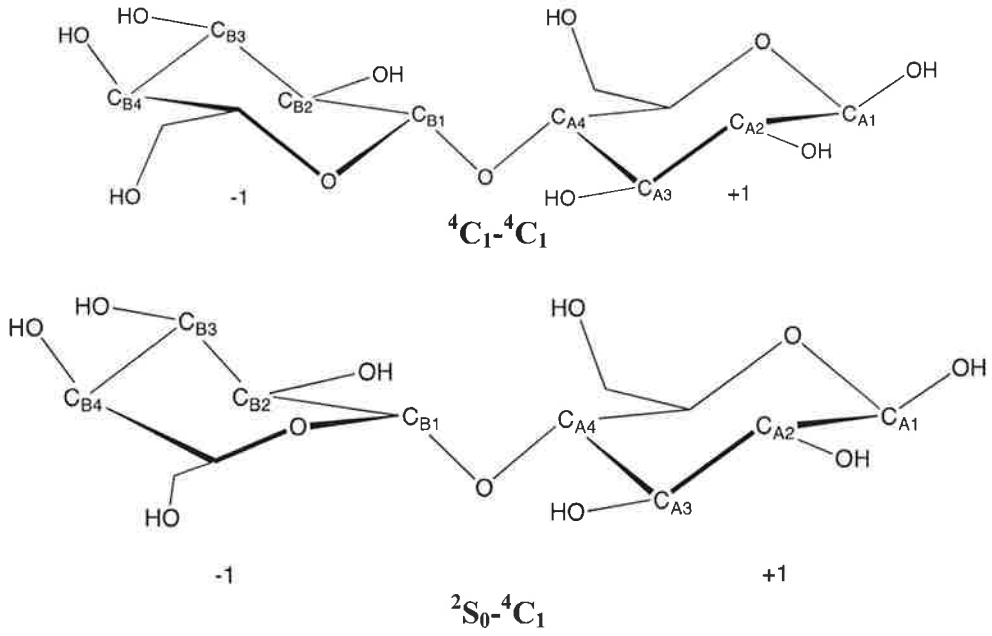
Glycosyl rotation and distortion by key residues in Endocellulase Cel6A from *Theromobifida fusca*

Tao Lu[+], Zuoming Zhang[□]* and Chi Zhang[+]*

Supplementary materials

1. QM study

In order to assess the substrate structures in MD simulation, cellobiose (Glu2) structures were optimized using the popular hybrid density functional theory B3LYP method, which does not have dispersion corrections, 6-31+G (d,p) basis, and CPCM implicit water solvation model in GAUSSIAN 09 package (Frisch, M.J., Trucks, G.W., et al. 2009). Another hybrid density functional theory B2PLYPD, which includes dispersion correction, also used to optimize the models with the same basis and solvation model. All parameters used Gaussion 09 default setting (converged threshold is SCF=tight, water eps=78.3553 and water molecule radius =1.4 Å). Both rigid and relaxed potential energy surfaces were scanned for Glu2. There are two structure models: one is $^4C_1$-$^4C_1$ and the other is $^2S_0$-$^4C_1$ (Supplementary Figure 1).



Supplementary Figure 1: Two different conformations of Glu2.

1.1 Rigid Scan

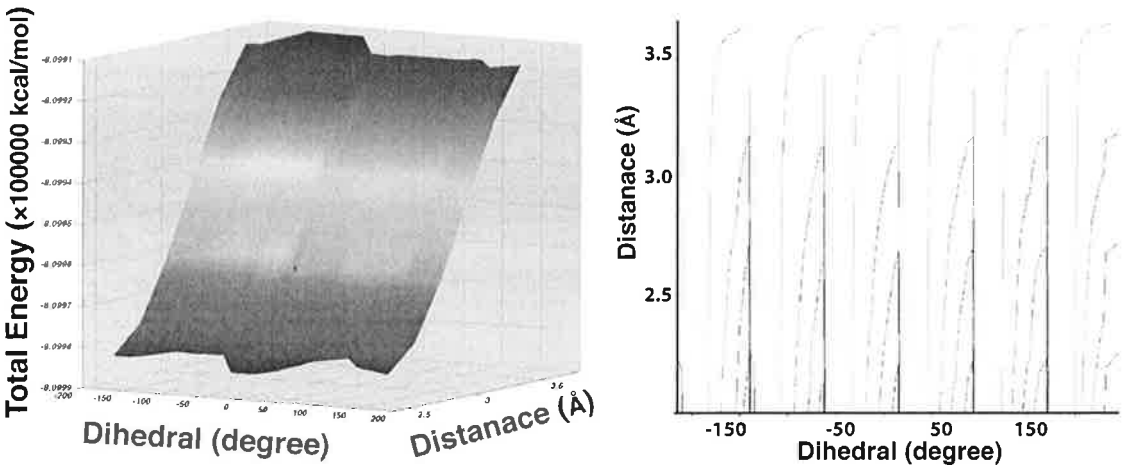
While keeping the -1 and +1 GU conformation, the dihedral angle of C_{B2}-C_{B1}-C_{A4}-C_{A3} was rotated along the axis of C_{B1}-C_{A4} and the point potential energy was calculated with b3lyp/6-31+g(d,p) in water. The dihedral angle of $^4C_1$-$^4C_1$ is -131° while $^2S_0$-$^4C_1$ is -118° (Supplementary Table 1). The angle of the two pyranoside ring planes was calculated with the normal vectors of the two least-square-fitted planes, (C_{A1},C_{A2},C_{A3},C_{A4},C_{A5},O) and (C_{B1},C_{B2},C_{B3},C_{B4},O). Supplementary Table 1 shows that the rotation along the C_{B1}-C_{A4} axis needs less energy in $^2S_0$-$^4C_1$ than that in $^4C_1$-$^4C_1$.

Supplementary Table 1. Energy and the planes rotation.

Dihedral Angle (degree)	Angles between two ring planes (degree)	HF Energy (a.u.)	Δ E (a.u.) Between two steps	Δ E(kcal/mol) Between two steps
$^2S_0$				
(opt)-118	27.24	-1298.043129		
-150	33.60	-1298.029331	0.013798	8.658358
-160	37.52	-1298.019238	0.010093	6.333521
-170	43.41	-1298.007036	0.012202	7.656814
-180	50.26	-1297.992509	0.014527	9.115901
$^4C_1$				
(opt)-131	2.21	-1298.049775		
-150	32.81	-1298.040785	0.008991	5.641780
-160	40.26	-1298.029006	0.011779	7.391252
-170	43.17	-1298.013056	0.015950	10.00897
-180	50.86	-1297.993426	0.019630	12.31783

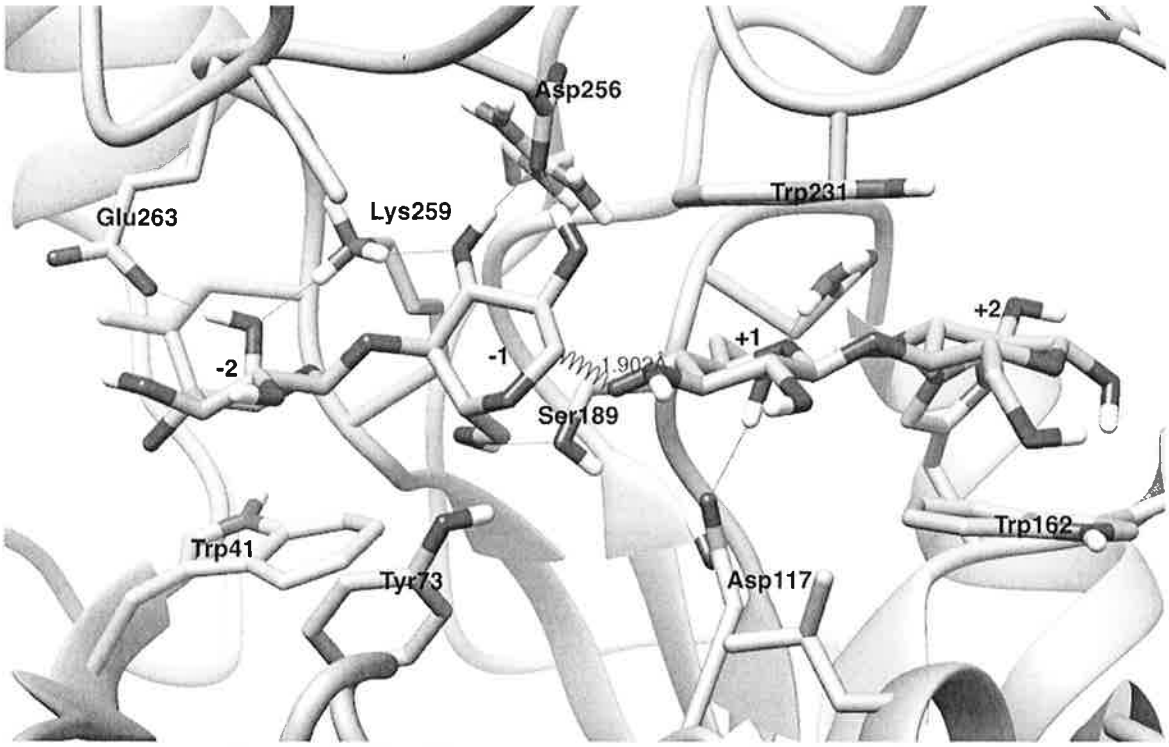
1.2 Relaxed Potential Energy Scan

The relaxed potential energy surface (PES) scan was conducted by changing the C_B1-C_A4 distance and rotating the C_B2-C_B1-C_A4-C_A3 dihedral angle with the C_B1-C_A4 axis in Glu2. When the distance between C_{B1}-C_{A4} was increased from 2.48 Å to 3.48 Å and the angle between two glucose planes (dihedral C_{B2}-C_{B1}-C_{A4}-C_{A3}) changed from -150° to 170°.



Supplementary Figure 2: The potential energy surface with changing the C_B1-C_A4 distance and rotating the C_B2-C_B1-C_A4-C_A3 dihedral angle by the relaxed potential energy scan is shown in colored plane plot (left) and contour plot (right).

2. The TS structure of Glu4



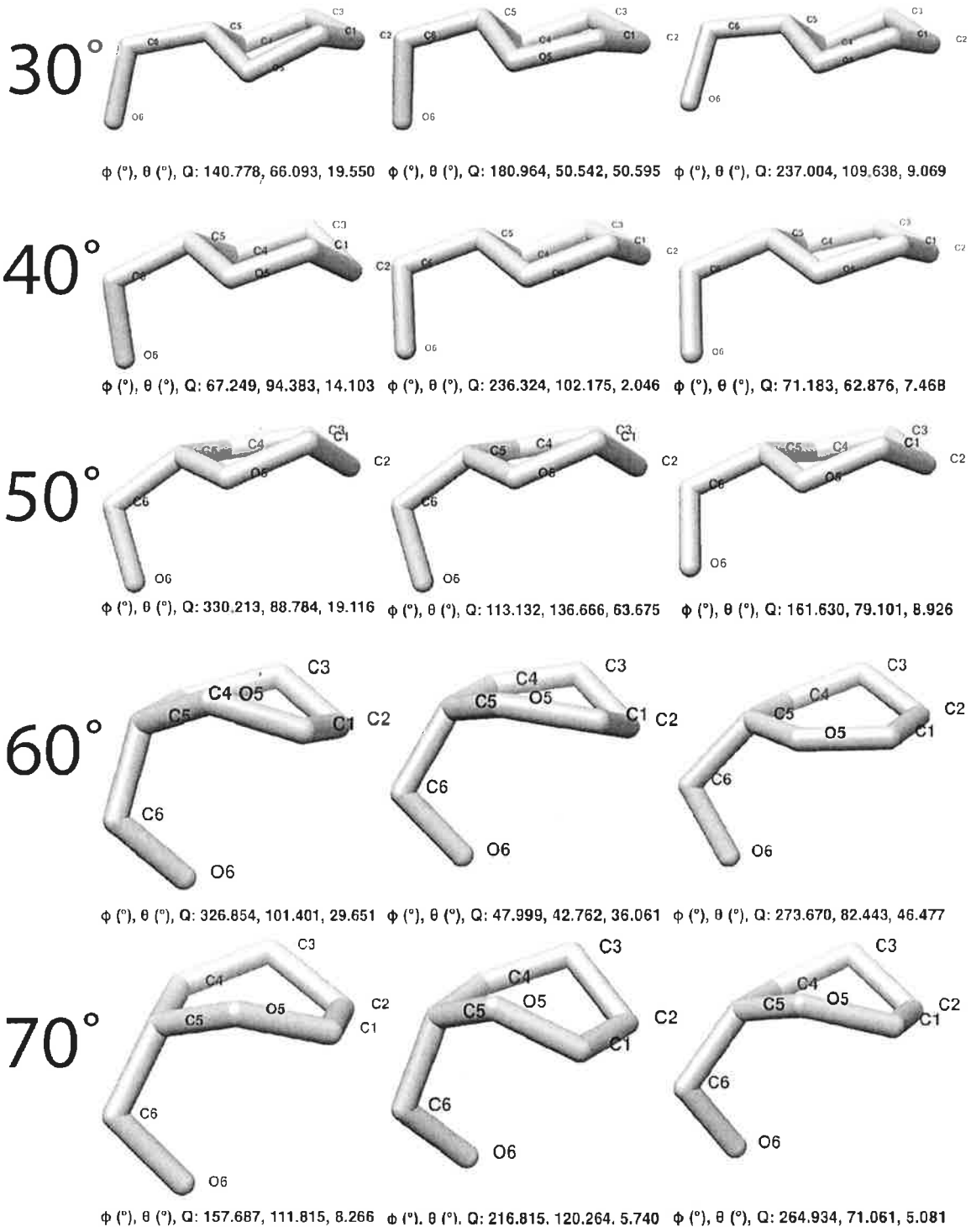
Supplementary Figure 3: The TS structure (above) was identified by searching for a structure with a single imaginary frequency in QM/MM simulation. The TS structure shows that the -1GU ring is rotated and distorted.

3. MD detail

3.1. MD (QM/MM) simulation method

The wild type (WT) structure used the PDB structure 2BOD. The structure with a Tyr73 to Serine mutation is the structure of PDB 2BOF. Other 3 mutation models were built based on the structure 2BOF using GROMACS (Hess, B., Kutzner, C., et al. 2008) tools to simply replace Ser189 to other types of amino acids. In all models, the sulfur atoms in crystal structures of inhibitors were replaced with carbons to rebuild the substrate structures (Glu4). Asp117 is a key catalytic residue in Cel6A (Koivula, A., Ruohonen, L., et al. 2002) and the Asp117 is protonated. After adding hydrogen atoms on substrates and proteins and rebuilding disulfide bonds, the complex structures were put into a cubic TIP3P water box (15 ×15 ×15 $Å^3$) (Jorgensen, W.L., Chandrasekhar, J., et al. 1983), and Na+ ions added to neutralize charges. Before conducting simulation, these systems underwent 500 ps balance simulation at 330K and 1 atmosphere (NPT ensemble) with leapfrog MD method in GROMACS. Five 10 ns long trajectories were built finally. All the MD simulation used GROMACS MD package with AMBER force field (Lindorff-Larsen, K., Piana, S., et al. 2010) for protein and GLYCAM06 force field (Kirschner, K.N., Yongye, A.B., et al. 2008) for cellulose in NPT ensemble. The interaction cutoffs are 8.0 Å, and the simulation step is 1.0 femtosecond. The QM layer, including Y73, D117, S189, D258, +1 GU and -1 GU, treated using DFTB implemented in Gaussion 09 with skf parameters downloaded from www.dftb.org.

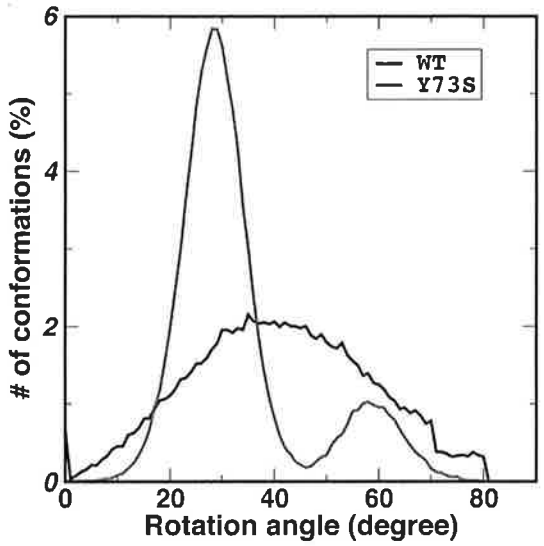
3.2 Distortion and Cremer-Pople parameters



Supplementary Figure 4. Three different structures in each distortion angle (30,40,50,60 and 70) are shown in the above figure. For each structure, its Cremer-Pople Parameters is shown below the structure. One can see, while the structures with the same distortion angle shows similar distorted conformations, their Cremer-Pople Parameters are very different. Therefore, the distortion angle defined in the manuscript is used for a straightforward illustration

4. MD simulation of mutation T73S

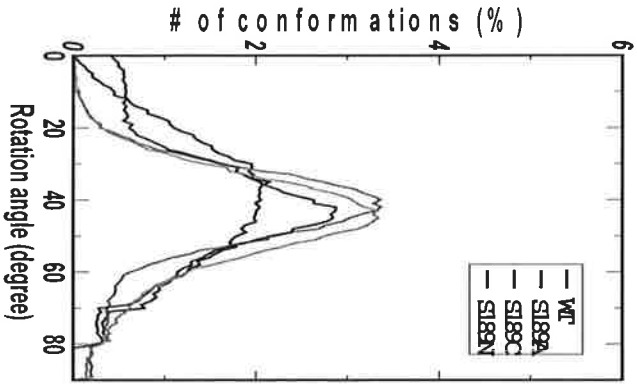
An MD simulation on a mutant structure of Cel6A (Y73S) was conducted to study the rotation of the pyranoside ring in -1 GU. In the simulation trajectory, there are only a few conformations that have large rotation angles of pyranoside ring.



Supplementary Figure 5: The distribution of numbers of conformations with different rotation angles of pyranoside ring in -1 GU.

5. Distribution of rotation angles for Ser189 mutations

When the Ser189 mutated to Cysteine, Alanine, and Asparagine, these mutations do not effect the substrate rotation.



Supplementary Figure 5: The distribution of numbers of conformations with different rotation angles of pyranoside ring in -1 GU.

References:

Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, *et al.* 2009. Gaussian 09, Revision B.01. Wallingford CT.

Hess B, Kutzner C, van der Spoel D, Lindahl E. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, 4:435-447.

Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. 1983. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.*, 79:926-935.

Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. 2008. GLYCAM06: A generalizable Biomolecular force field. Carbohydrates. *J. Comput. Chem.*, 29:622-655.

Koivula A, Ruohonen L, Wohlfahrt G, Reinikainen T, Teeri TT, Piens K, Claeyssens M, Weber M, Vasella A, Becker D, *et al.* 2002. The active site of cellobiohydrolase Cel6A from Trichoderma reesei: the roles of aspartic acids D221 and D175. *J. Am. Chem. Soc.*, 124:10015-10024.

Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinformat.*, 78:1950-1958.