# Probing the Active Site Chemistry of β-Glucosidases along the Hydrolysis Reaction Pathway

**ARTICLE** *in* BIOCHEMISTRY · OCTOBER 2012

Impact Factor: 3.02 · DOI: 10.1021/bi300675x · Source: PubMed



**3 AUTHORS**, INCLUDING:

Somayesadat Badieyan
Virginia Polytechnic Institute and State Univer…
**10** PUBLICATIONS **108** CITATIONS
SEE PROFILE

Chenming Zhang
Virginia Polytechnic Institute and State Univer…
**59** PUBLICATIONS **1,050** CITATIONS
SEE PROFILE

BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

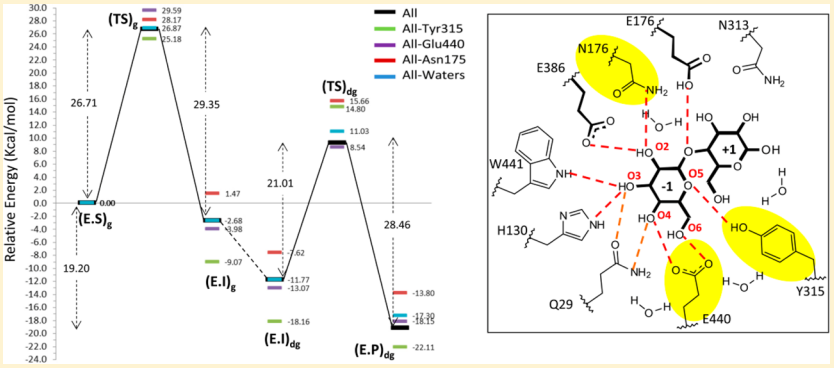
pubs.acs.org/biochemistry

# Probing the Active Site Chemistry of β-Glucosidases along the Hydrolysis Reaction Pathway

Somayesadat Badieyan,[†,‡] David R. Bevan,*[,‡] and Chenming Zhang*[,†]

[†]Department of Biological Systems Engineering and [‡]Department of Biochemistry, Virginia Tech, Blacksburg, Virginia 24061, United States

Supporting Information




**ABSTRACT:** β-Glucosidases (EC 3.2.1.21) can be found in all domains of living organisms, where they play essential roles in the removal of terminal glucosyl residues from nonreducing ends of saccharides and glycosides. Two active site amino acid residues, a nucleophile and a proton donor, play key roles in the hydrolytic mechanism. Besides these two highly conserved catalytic residues, there are other conserved amino acids in the active site of β-glucosidases that make direct hydrogen bonds to the glycosyl moiety at the −1 subsite. In this study, the catalytic mechanism of a GH1 β-glucosidase (BGlu1) is systematically studied. On the basis of the quantum mechanical studies, the side chain of Tyr315 in an interaction with both O5 of the glucose ring and the nucleophilic glutamate contributes significantly to the energy profile. Glu440 and the conserved Asn175 are the other residues in the polar interaction with −1 glucose with considerable influence on the free energy of the reaction. Gln29, His130, and Trp441, which also form hydrogen bonds to the glycosyl moiety, are found to have relatively a minor effect on the reaction. Different arrangements of active site residues in the high-level [quantum mechanics (QM)] and low-level [molecular mechanics (MM)] regions during the hybrid QM/MM calculations indicate that Tyr315 lowers the energy barrier in the deglycosylation step (by 11.95 kcal/mol) while Glu440 mainly reduces the energy barrier of the glycosylation step. Exclusion of either of these two residues from the QM region results in deviation of the geometric parameters of the enzyme−substrate complex from those expected for the preactivated distorted structure of the substrate.

β-Glucosidases (β-D-glucopyranoside glucohydrolases) are enzymes that hydrolyze β,1→4,O-linked (or S-linked in thioglucosidase) glycosidic bonds linking two glucose or glucose-substituted molecules to release nonreducing terminal glucosyl residues from glycosides and oligosaccharides.[1] These enzymes are found in abundance in all domains of living organisms. Their functions include biomass conversion in microorganisms, breakdown of glycolipids and glucosides, and catabolism of cell wall oligosaccharides in fungi and plants. Mammalian β-glucosidases such as cytoplasmic β-glucosidase and human acid β-glucosidase are thought to play roles in the metabolism of glycolipids and dietary glucosides or in signaling pathways.[2] β-Glucosidases in plants are found to display a large range of biological functions, including roles in defense and plant−herbivore and plant−insect interactions, symbiosis, and cell wall degradation.[1] Family 1 of glycoside hydrolases (GH1)[3,4] includes the largest number of well-characterized and structurally defined β-glucosidases (EC 3.2.1.21). Generally, the substrates of GH1 enzymes are termed type G-O/S-X, where G indicates the glycosyl residue at the nonreducing ends of the substrates and X can be either a glycosyl unit or a non-glycosyl aglycone bound to the glycosyl moiety by an S or O link. As a result, many β-glucosidases are able to hydrolyze glucosides with a range of glycosyl or non-glycosyl aglycone groups and nonphysiological substrates. The residues of β-glucosidase at subsite ≥+1 (toward the reducing terminus of the substrate)[5] are mainly involved in the substrate recognition and specificity of the enzyme, particularly toward the aglycone part of the substrate, and play a far less significant role in the hydrolysis reaction.[6−9] Thus, the combinations of residues in this substrate recognition site of β-glucosidases are highly

**Received:** May 23, 2012
**Revised:** October 5, 2012
**Published:** October 8, 2012

   dx.doi.org/10.1021/bi300675x | *Biochemistry* 2012, 51, 8907−8918

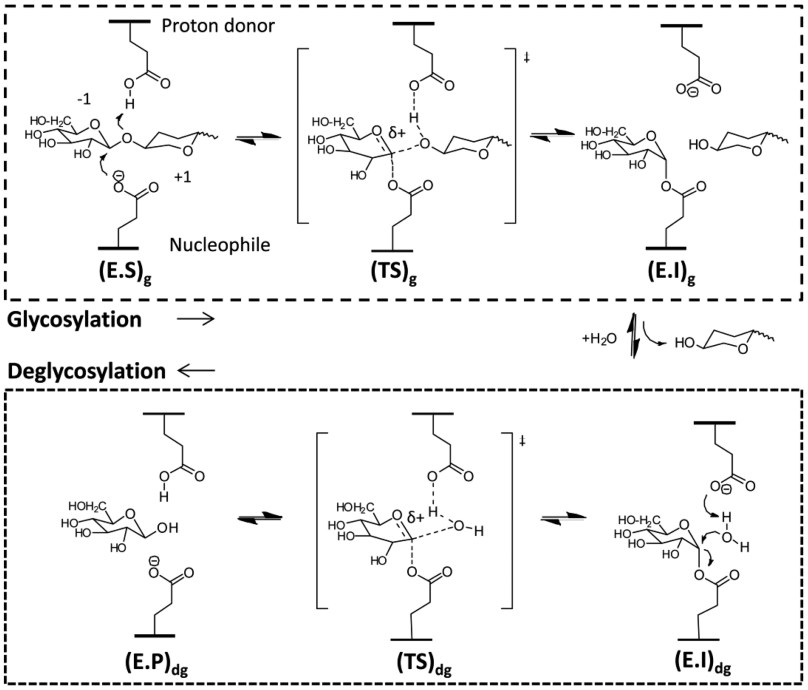



**Figure 1.** Double-displacement reaction mechanism of retaining glycosidases, including members of family 1. The first step of the hydrolysis reaction, known as the glycosylation step, starts with the nucleophilic attack of one glutamate on the anomeric carbon (C1) of the glycosyl moiety followed by formation of an $\alpha$-linked covalent intermediate. The second step is deglycosylation, in which a water molecule activated by the general base (the same glutamate that functioned as the general acid) serves as the second nucleophile to release the glucose from the intermediate state.

diverse and generally defined based on the nature of the aglycone moieties. The interactions with the aglycone at subsite >+1 are primarily hydrophobic,[10] and the hydrogen bonds to the substrate in this region are mediated by water molecules.[11] On the other hand, the residues interacting with the glucose moiety at the −1 subsite are highly conserved and expected to be involved in the hydrolysis activity of $\beta$-glucosidases. GH1 enzymes are retaining enzymes and follow a classical Koshland $S_N2$ mechanism.[12] The mechanism of the hydrolysis reaction is illustrated in Figure 1. Two active site amino acid residues, usually glutamates, play key roles in the hydrolytic mechanism, with one functioning as a nucleophile and the other as a proton donor.[13] The reaction proceeds via the formation of a covalent glycosyl−enzyme intermediate[14] through a transition state (TS) that features extensive $sp^2$ hybridization and partial positive charge on the oxygen of the glucose ring and involves pyranoside distortion to half-chair ($^3H_4$) or envelope conformations. The two catalytic carboxylates directly involved in the hydrolysis reaction are highly conserved in GH1 and many other GH families, generally part of semiconserved (I/V)-(T/X)-**E**-N-G and T-(F/I/L)-N-**E** motifs for the nucleophile and acid/base, respectively. Besides these carboxylate groups in the active site of $\beta$-glucosidases, there are other noncatalytic conserved polar residues that are involved in direct hydrogen bond interactions with the hydroxyl groups of the glycosyl moiety at the −1 subsite. Although many crystal structures of $\beta$-glucosidases in the presence of substrates, inhibitors, or substrate analogues provide insight into the importance of these noncatalytic highly conserved residues, more details about their role in transition state stabilization can be attained by site-specific mutagenesis of active site residues. However, such a comprehensive study of the active site of $\beta$-glucosidases is lacking, and only the data from a few different mutagenesis studies of some individual $\beta$-glucosidases are available.[15−17] Besides mutagenesis, the other approach used to study the general effect of polar interactions at the −1 glycan subsite is replacing glycone hydroxyls with hydrogen and/or fluorine followed by an evaluation of these substitutions using enzyme kinetics.[18,19] Although these studies have provided data about the relative importance of the glycone hydroxyls to the $\beta$-glucosidase activity, they did not contribute to the determination of the relative importance of each subsite −1 residue to the substrate specificity of $\beta$-glucosidases.

To investigate the role of residues that can form hydrogen bonds to the glycosyl moiety at subsite −1 in the hydrolysis reaction of the glycosidic bond by GH1 $\beta$-glucosidases, we performed extensive density functional theory (DFT)[20] and hybrid quantum mechanics/molecular mechanics (QM/MM) (ONIOM)[21] calculations to address the stabilizing role of these conserved residues along the reaction coordinates in both glycosylation and deglycosylation steps. The methods of quantum mechanics have been proven to be useful in the mechanistic characterization of glycoside hydrolases.[22−24] One of the well-studied and highly expressed GH1 $\beta$-glucosidases, BGlu1, from rice (*Oryza sativa* L.) was selected for this study. Many high-quality structures of BGlu1 in complex with different natural substrates[11] or covalently bound to an intermediate analogue[25] make it a good candidate for computational studies. Glu176 and Glu386 of BGlu1 function as the general acid/base and nucleophile, respectively. The other residues that interact with the glucose moiety at the −1 subsite are Q29, H130, N175, Y315, E440, and W441. The details of these interactions are summarized in Table 1. Figure S1 of the Supporting Information displays the schematic representation of polar interactions between the BGlu1 active site residues and glucose at the −1 subsite. Upon formation of the glucosyl−enzyme intermediate of BGlu1, the positional

**Table 1. Distances between Atoms of Hydrogen Bonding Residues of the BGlu1 Active Site and the Corresponding Oxygen Atom on the Glucose Moiety at Subsite −1 in both the Enzyme−Substrate Complex (PDB entry 3F5J) and the Enzyme−Intermediate Complex (PDB entry 3AHV)**

| | | distance (Å) | |
|---|---|---|---|
| substrate atom (glucose) | protein atom | in the E·S complex | in the intermediate |
| O1 | Glu176 $O^{\varepsilon 1}$ | 2.66[a] | not available |
| | Glu176 $O^{\varepsilon 2}$ | 3.51 | not available |
| | water691 | 2.84 | not available |
| O2 | Glu176 $O^{\varepsilon 1}$ | 3.82[b] | 3.48[b] |
| | Glu386 $O^{\varepsilon 1}$ | 2.71 | 4.12[c] |
| | Asn175 $N^{\delta 2}$ | 2.92 | 2.96[c] |
| | – | – | |
| O3 | His130 $N^{\varepsilon 2}$ | 2.84[c] | 2.91 |
| | Gln29 $O^{\varepsilon 1}$ | 2.73 | 2.65 |
| | Trp441 $N^{\varepsilon 1}$ | 3.07 | 2.89 |
| O4 | Gln29 $N^{\varepsilon 2}$ | 3.08 | 2.90 |
| | Glu440 $O^{\varepsilon 1}$ | 2.39 | 2.63 |
| O5 | Tyr315 $O^{\eta}$ | 2.92 | 2.90 |
| O6 | Glu440 $O^{\varepsilon 2}$ | 2.40 | 2.61 |

[a]Adjusted value after Q176 had been mutated back to glutamate in the crystal structure of the enzyme−substrate complex. [b]O2 makes a hydrogen bond to Glu386 $O^{\varepsilon 1}$ in the ground state enzyme−substrate complex and to Glu176 $O^{\varepsilon 1}$ in the intermediate complex. [c]Adjusted value after substitution of OH by F2 in the intermediate complex.

changes of some active site residues are illustrated in Figure S2B of the Supporting Information.

Gln29 and Glu440 of BGlu1 are conserved in all GH1 family members (except for 6-phospho-$\beta$-glucosidases, in which Glu440 is replaced with a Ser). These residues make bidentate hydrogen bonds to −1 glucose (Table 1). Bidentate bonds have been reported to be formed by adjacent hydroxyls of the glycone moiety either when both are equatorial or when one is equatorial and the other axial. Therefore, GH1 $\beta$-glucosidases are able to accommodate substrates derived from the glucose and galactose rings and may exhibit $\beta$-galactosidase activity as well. O2 of the glucose moiety forms a hydrogen bond to Glu386 $O^{\varepsilon 1}$ (nucleophile) and Asn175 $N^{\delta 2}$ in the enzyme−substrate complex. In the intermediate complex, the hydrogen bond between O2 and Glu386 $O^{\varepsilon 1}$ is disrupted, and instead, O2 moves toward Glu176 $O^{\varepsilon 1}$ (proton donor) while keeping the hydrogen bond to Asn175 $N^{\delta 2}$. Asn175 is also a highly conserved residue in clan A of glycoside hydrolases (excluding GH26), and this asparagine is always followed by the proton donor residue. O3 of glucose is interacting with highly conserved His130 ($N^{\varepsilon 2}$) and semiconserved Trp441. The absolutely conserved Tyr315 forms a hydrogen bond to both O5 of the pyranoside ring and the nucleophile (Tyr315 $O^{\eta}$−Glu386 $O^{\varepsilon 1}$, 2.57 Å) and is critical for keeping the substrate and nucleophile in suitable positions for nucleophilic attack. Asn313 does not make direct polar interaction with the glucose moiety at the −1 subsite but is in direct hydrogen bonding distance with both the nucleophile and proton donor residues (Asn313 $N^{\delta 2}$ is 3.11 and 2.91 Å from Glu386 $O^{\varepsilon 2}$ and Glu176 $O^{\varepsilon 1}$, respectively). To compare the effect of this semiconserved residue, which is in the hydrogen bond network of the catalytic residues, with the other residues in direct polar interaction with the glucose moiety, we performed DFT calculations, including Asn313 along with the other conserved active site residues. Figure S2A of the Supporting Information illustrates the position of the named residues around the glucose moiety at subsite −1. Except for one water molecule that is within hydrogen bonding distance of the oxygen of the glycosidic bond, no other water molecule interacts directly with glucose in this subsite, though some water molecules within 4 Å of this glucose have conserved positions in most of the GH1 $\beta$-glucosidases with structures that have been determined. These water molecules are also considered in our computational calculations. The purpose of this study was to analyze the relative energy contribution of each hydrogen bonding residue to glucose in the active site of $\beta$-glucosidases during the glycosylation and deglycosylation steps of the hydrolysis reaction. The results of this computational study will help us to understand and compare the role of key active site residues along the reaction coordinate, an understanding of which is necessary for the successful rational design of glucosidase activity and the design of inhibitors for GH enzymes involved in different diseases.[26]

## MATERIALS AND METHODS

**Initial Protein Structure and Molecular Dynamics (MD) Simulations.** The model system was based on the crystal structure of the *O. sativa* $\beta$-glucosidase (BGlu1)[11] in complex with cellotetraose (CTT) [Protein Data Bank (PDB) entry 3F5J]. This structure is of a mutant (E176Q) enzyme−substrate complex, and Q176 was changed back to glutamate to generate the actual enzyme−substrate complex for the glycosylation step of the hydrolysis reaction, which is denoted as $(E{\cdot}S)_g$. The reaction of 2-deoxy-2-fluoroglucoside with BGlu1 produces a covalent complex (PDB entry 3AHV) that was used to represent the covalent intermediate in the reaction sequence. Because this complex represents the reactant for the deglycosylation step, it is denoted as $(E{\cdot}I)_{dg}$. The protonation states of ionizable residues of structures were determined using H++[27] at pH 5.5, the activity optimal pH of wild-type BGlu1.[11] Structures were added to a box of TIP3P waters[28] with no less than 12 Å from the edge of the water box to the nearest protein atom, and the net charge of the systems was neutralized by adding $Na^+$ and $Cl^-$ ions to a final concentration of 100 mM. Energy minimization using the steepest descent method was followed by equilibration at 300 K in the *NPT* ensemble for 500 ps. In both steps, a large harmonic force constant of $1.0 \times 10^6$ kJ $mol^{-1}$ $nm^{-2}$ was applied to restrain the positions of heavy atoms of the entire protein, substrate, and crystal water molecules within 6 Å of the substrate. This strategy allowed equilibration of the hydrogen atoms and bulk water molecules without any significant change in the position of active site atoms from the original crystal structures. The MD simulations were conducted with Gromacs 4.5.[29] The Amber FF03 force field[30] was applied to the protein, and GLYCAM_06[31] was used to simulate polysaccharides. Short-range nonbonded interactions were truncated at 1.4 nm, and long-range electrostatics were treated using the particle mesh Ewald (PME) method.[32] The bond distances and angles of the solvent were constrained using the LINCS algorithm[33,34] with a coupling time of 2 fs. General preparation steps were common to all systems. The temperature was held constant by coupling the system to an external bath at the chosen temperature, using a Nosé-Hoover thermostat[35,36] during equilibrium. In addition, another set of unrestricted simulations were performed on BGlu1 in its liganded and covalent intermediate states to prove the stability of protein and enzyme−substrate complex during


dx.doi.org/10.1021/bi300675x | *Biochemistry* 2012, 51, 8907−8918

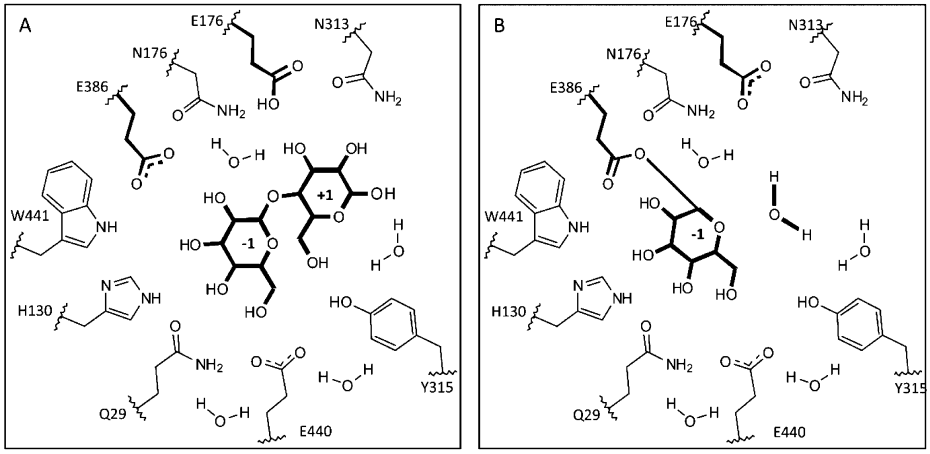



**Figure 2.** Arrangement of hydrogen bond-forming residues around the glucose moiety at subsite −1 in the active site of BGlu1 in (A) the enzyme−substrate complex, $(E{\cdot}S)_g$ (PDB entry 3F5J), and (B) the enzyme−intermediate complex, $(E{\cdot}I)_{dg}$ (PDB entry 3AHV). $(E{\cdot}S)_g$ and $(E{\cdot}I)_{dg}$ were considered as references for glycosylation and deglycosylation, respectively. In the deglycosylation step, the glucose at subsite +1 was replaced with a water molecule. More details about the hydrogen bond made by these residues to the substrate are illustrated in Figure S1 of the Supporting Information.

MD simulation (Figures S3−S5 of the Supporting Information).

**Minimal Model System and DFT Analysis.** The first model system for identifying the role and energy contribution of each hydrogen bonding residue in the active site of BGlu1 consisted of side chains of residues that form hydrogen bonds to the CTT in the −1 subsite as described in the introductory section. Thus, the system included the two catalytic glutamates, the proton donor (E176) and the nucleophile (E386). In addition, the side chains of Gln29, His130, Asn175, Asn313, Tyr315, Glu440, and Trp441 and the four crystal water molecules interacting with active site conserved residues (Figure 2A) were part of the systems. To reduce the computational cost, Trp441 was represented as a pyrrole ring, and C$\alpha$ atoms of the active site amino acid main chain were constrained during the optimization, except for Glu386 because its main chain is expected to move slightly upon moving to the transition state. Furthermore, because BGlu1 is capable of hydrolyzing cellobiose as well as other soluble cellulose chains with a higher degree of polymerization, only the glucose moieties in subsites −1 and +1 were considered for calculation in the E·S complex for the glycosylation step. For the deglycosylation step, the glucose at subsite +1 in the intermediate complex, $(E{\cdot}I)_{dg}$, was replaced with a water molecule that acts as a nucleophile in the hydrolysis reaction (Figure 2B). This water molecule is water625 in the intermediate−complex structure (PDB entry 3AHV), which is 3.58 Å from C1 of glucose. The total numbers of atoms in the glycosylation and deglycosylation steps were 186 and 165, respectively, during the QM analysis.

The B3LYP (default spin) method at the DFT level of theory[37] with the 6-31G(d) basis set was used for geometry optimization of reactants, $(E{\cdot}S)_g$ and $(E{\cdot}I)_{dg}$, transition states, $(TS)_g$ and $(TS)_{dg}$, and the final products of the glycosylation and deglycosylation steps, $(E{\cdot}I)_g$ and $(E{\cdot}P)_{dg}$, respectively. The calculation of vibration frequencies was also conducted at the same level of theory. To find a good guess for geometries at the transition state, we started with a very simple model system including just substrates and catalytic residues, and the Transit-Guided Quasi-Newton method,[38] QST3, was applied to find the actual transition states in the absence of all the other active site residues. The stabilization of saddle points in the TS was checked with vibration analysis and intrinsic reaction coordinate (IRC) methods to verify the transition states are well connected to the reactant and product. The geometric parameters of transition states taken from this simplified system were used for the redundant optimization of restrained structures of the optimized reactant structures of the glycosylation and deglycosylation steps in the presence of all the other noncatalytic residues (minimal system). Then full transition state geometry optimizations were conducted once these redundant optimized structures and formation of accurate transition states were validated by vibration frequency calculation.

A larger basis set, 6-311+G(2df,2pd), was used for single-point energy calculations on the optimized geometries. It has been demonstrated by other investigators that the 6-311++G(3df,3pd) basis set for energy calculation changes the final result by <0.5 kcal/mol relative to the value with the 6-311+G(2d,2p) basis set.[22] Considering the time required for computation, the basis set selected for energy calculation in this study seems to be reasonably accurate. To evaluate the relative energy contribution of each hydrogen bonding residue to the reaction, we followed the method suggested by Tian et al.[39] In the optimized structures of reactants, transition states, and products, the residues one by one were replaced with dummy atoms, and single-point energy calculations [at basis set 6-311+G(2df,2pd)] were conducted on each system without further optimization. The "so-called" relative energy contribution (REC) of each residue, e.g., residue *i*, to the reaction energy profile was evaluated based on the difference between the relative energies of the whole system (optimized minimal system with all residues) and that system without residue *i*. For example, the REC value for residue *i* at $(TS)_g$ was evaluated by $|E'_{(TS)_g} - E_{(TS)_g}|$, where $E_{(TS)_g}$ is the relative energy of the whole system (TS of the glycosylation step) and $E'_{(TS)_g}$ is the relative energy of the system when residue *i* was replaced with a dummy atom. The same calculation was conducted for each residue for the deglycosylation step to calculate the REC value of each residue at that step as well. Because replacement of the leaving group (glucose) with water molecules could not feasibly be modeled by QM calculations, the two steps of glycosylation


dx.doi.org/10.1021/bi300675x | *Biochemistry* 2012, 51, 8907−8918

and deglycosylation were studied separately. The relative energies for the glycosylation and deglycosylation steps were calculated on the basis of the reactant of each step, $(E{\cdot}S)_g$ and $(E{\cdot}I)_{dg}$, respectively. All subsequent DFT and hybrid QM/MM calculations (next step) were conducted using Gaussian 09.[40]

**Full Enzyme System and Hybrid QM/MM (ONIOM).** The equilibrated structures from MD simulations were used for the hybrid QM/MM (ONIOM) study.[21,41] The residues with the highest REC from the DFT study were selected to be part of the high-level (QM) region of the ONIOM system. A two-layer ONIOM (QM/MM) with mechanical embedding was used for geometry optimization. Hydrogen link atoms were used to treat QM and MM interface regions. In geometry optimizations, the higher-level layer included the glucose moieties at subsites −1 and +1 and the glycosidic bond between subsites +1 and +2, side chains of the two catalytic residues (E176 and E386), side chains of N175, Y315, and E441, and water molecules, for a total of 107 atoms for the glycosylation step. The high-level region for the deglycosylation step included all the residues and water molecules listed above, with the glucose residue at subsite +1 being replaced with a water molecule, for a total of 87 atoms. This region was treated with DFT at the B3LYP/6-31G(d) level for geometry optimization. The rest of the system (including glucose moieties of CTT at subsites +2 and +3) was treated at the MM level with the parm99[42] and Glycam_06[31] force fields. The geometry-optimized transition states were located with the quadratically coupled QM/MM geometry optimizer[43] as described by Tao et al.[44] The atoms in the outer shell of the protein were kept frozen, and atoms (including bulk water and protein atoms) within 15 Å of the protein core (the protein core consists of catalytic residues and substrate) were allowed to move.

Residues constituting the core of the reaction (E176, E386, glucose at subsite −1, and glucose at subsite +1 or a water molecule) were maintained at a high level, while each of other high-level region residues, including N175, Y315, and E441 or water molecules, was transferred one by one to the low level. Optimization was again conducted with this new arrangement of high- and low-level regions. The atom number in each ONIOM system that was studied is summarized in Table 2.

**Table 2. Numbers of Atoms in the Minimal System and High-Level Region of Different ONIOM Partitioning Schemes**

| | no. of atoms | | |
|---|---|---|---|
| | high-level glycosylation | high-level deglycosylation | residues[a] |
| minimal system (QM) | 185 | 165 | Q29, H130, N175, N313, Y315, E440, W441, waters |
| ONIOM-All | 107 | 87 | N175, Y315, E440, waters |
| ONIOM-N175 | 99 | 79 | Y315, E440, waters |
| ONIOM -Y315 | 92 | 72 | N175, E440, waters |
| ONIOM-E440 | 98 | 78 | N175, Y315, waters |
| ONIOM-waters | 95 | 75 | N175, Y315, E440 |

[a]E176, E386, and glucose at subsite −1 for the glycosylation step and E176, E386, and a nucleophilic water for the deglycosylation step were kept in all DFT or ONIOM systems. Rearrangement or deletion of residues was applied to only noncatalytic conserved active site residues as indicated.

Single-point energy calculations of each optimized geometry were conducted at the B3LYP/6-311+G(2df,2pd) level with electronic embedding incorporation of the partial charges of the MM region into the quantum mechanics Hamiltonian.[45] The ONIOM-EE energies were calculated using the equation $E^{ONIOM} = E^{QM,high} + E^{MM,low} - E^{QM,low}$ as described by Morokuma and co-workers.[21,46] Geometries were not optimized with the electronic embedding scheme because of its high computational cost. The atomic point charges were taken from a Mulliken population analysis of the electronic density of the higher-level region.[47] File preparation for ONIOM calculations was conducted as described previously.[48]

## RESULTS AND DISCUSSION

**DFT Calculations on Active Site Residues in a Minimal System.** To identify the residues with a significant impact on the energy profile along the reaction coordinate of hydrolysis by GH1 β-glucosidases, a system containing side chains of the catalytic residues (Glu176, proton donor, and Glu386, nucleophile), side chains of all residues hydrogen bonding to the glycosyl moiety at subsite −1, and a cellobiose as a natural substrate was treated by DFT/B3LYP calculations. DFT methods have been shown to produce the most reliable data relative to those of the other theoretical methods.[22,49] The main chains of the residues listed above are at least 8 Å from the reaction center and may not affect the reaction significantly, so they were excluded from calculations (Figure 2 and Figure S1 of the Supporting Information). The results from the relative energy contribution (REC) calculations are presented in Figure 3. The enzyme−substrate complex, $(E{\cdot}S)_g$, and the glycosyl−intermediate complex, $(E{\cdot}I)_{dg}$, were considered as the references for glycosylation and deglycosylation, respectively, for calculation of REC values for each step. The average values of the REC for each residue in each step of hydrolysis are summarized in Table 3. On the basis of Figure 3 and Table 3, it is clear that Tyr315 plays a major role in the energy profile for both glycosylation and deglycosylation. This high-energy

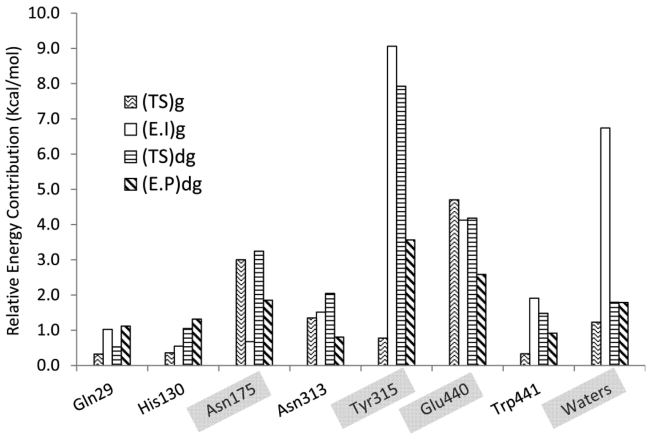



**Figure 3.** Relative energy contribution (REC) of the active site residues to each stage of hydrolysis. The reference for the glycosylation step is the enzyme−substrate complex, and for deglycosylation, it is the enzyme−intermediate complex. The energies are calculated using the B3LYP/6-311+G(2df,2pd) level of theory. $(TS)_g$ represents the transition state of glycosylation, $(E{\cdot}I)_g$ the enzyme−intermediate complex as the product of glycosylation, $(TS)_{dg}$ the transition state of deglycosylation, and $(E{\cdot}P)_{dg}$ the enzyme−glucose complex as the final product of deglycosylation.


dx.doi.org/10.1021/bi300675x | Biochemistry 2012, 51, 8907−8918

**Table 3. Average Values of the Relative Energy Contribution (REC) in Kilocalories per Mole for Each Residue and the Water Molecules in Glycosylation and Deglycosylation Steps Using the Minimal Model System[a]**

| residue | glycosylation | deglycosylation |
|---|---|---|
| Gln29 | 0.67 | 0.82 |
| His130 | 0.45 | 1.18 |
| Asn175 | 1.84 | 2.55 |
| Asn313 | 1.43 | 1.42 |
| Tyr315 | 4.92 | 5.75 |
| Glu440 | 4.41 | 3.38 |
| Trp441 | 1.12 | 1.20 |
| waters | 3.98 | 1.79 |

[a]Geometry optimizations were performed with the B3LYP method at the DFT level of theory with the 6-31G(d) basis set. Large basis set 6-311+G(2df,2pd) was used for single-point energy calculations on the optimized geometries.

contribution for Tyr315 looks reasonable because the orientation of this residue makes it possible to form a hydrogen bond to the nucleophilic Glu386 $O^{\varepsilon 1}$ and/or to O5 of the glucose moiety. The analyses of geometry-optimized structures illustrate that in the enzyme−substrate complex, $(E{\cdot}S)_g$, as well as the enzyme−final product complex, $(E{\cdot}P)_{dg}$, the Tyr-OH group forms a hydrogen bond to Glu386 $O^{\varepsilon 1}$ (Figure S3A of the Supporting Information). For the first transition state, $(TS)_g$, Tyr315 still mainly keeps the hydrogen bond to Glu386, although it moves closer to O5. A similar orientation for Tyr69 in BCX xylanase (in a position equivalent to that of Tyr351 in BGlu1) has been reported in the transition state of the glycosylation step.[23] In the intermediate state and upon formation of the covalent bond between Glu386 $O^{\varepsilon 1}$ and C1, and also because of the reorientation of the carboxylate group of the nucleophile, Tyr315-OH makes its hydrogen bond to O5 of glucose (Figure S3B of the Supporting Information) and keeps this hydrogen bond in the second transition state, $(TS)_{dg}$. The maximal REC values for Tyr315 are for $(E{\cdot}I)_g$ and $(TS)_{dg}$, indicating that this residue contributes mainly to the energy profile of hydrolysis by making a hydrogen bond to the O5 glucose moiety. The role of this Tyr in activation energies will be discussed in the next section.

Glu440 is the next residue with considerably high REC values in both steps of hydrolysis. This residue makes bidentate hydrogen bonds to O6 and O4 of glucose at subsite −1 and almost equally and significantly contributes to the energy profile at different steps of hydrolysis (Figure 3). Interestingly, Gln29 is the other residue with bidentate interactions to glucose and makes hydrogen bonds to O3 and O4 of the glucose moiety, but with a relatively small contribution of energy to the profile. The results are in very good agreement with a site-specific mutagenesis study of Gln39 and Glu451 of a GH1 β-glucosidase from Lepidoptera, equivalent to Q29 and E440, respectively, in BGlu1, where the energy of the noncovalent interaction between the glycone equatorial O4 and Glu451 in the transition state was found to be ∼60% higher than its energy of interaction with Gln39.[15] The catalytic specificity of an E451A mutant of this enzyme was 4−10 times less than the catalytic specificity of the Q39A mutant on several different substrates, comparable to the value of ∼5.2 for the average $REC_{E440}/REC_{Q29}$ ratio calculated in this study. As Marana et al.[15] noted, the higher energies of the hydrogen bonds involving Glu451 (in comparison with Gln39) may result from the fact that hydrogen bonds with charged groups (like Glu451) are stronger than those with uncharged ones (like Gln39), even if they are similar in length.

Another residue with considerable REC values was Asn175 (Figure 3 and Table 3). Asn175 displayed its role primarily in both transition states, $(TS)_g$ and $(TS)_{dg}$, relative to the other enzyme complexes (Figure 3). An explanation for this observation can be made by considering the distance between Asn175 $N^{\delta 2}$ and O2 in reactants, transition states, and intermediates. Because of the distortion of the glucose ring at subsite −1 in TS, the Asn175 $N^{\delta 2}\cdots$O2 distance in $(TS)_g$ and $(TS)_{dg}$ decreases by ∼0.07 Å relative to those in E·S and E·I complexes. Although these differences in the hydrogen bond distance are not large, they may provide greater stabilization in transition states than in the other stages. Because Asn175 is involved in the interaction with O2 of the glucose moiety, as expected it plays a significant role in the determination of the specificity of GH1 glycosidases toward mannose or glucose. Mutation of N206 (equivalent to N175 in BGlu1), in GH1 β-glucosidase CelB, increased the specificity for mannosides and reduced the rates of hydrolysis toward glucose by 10-fold.[17]

His130 is a residue with a relatively low REC (Figure 3). His130 $N^{\varepsilon 2}$ makes a hydrogen bond to O3 of glucose in the E·S and E·I complexes. By examining all the other structures of GH1 enzymes deposited in the PDB, we noted that the hydrogen bond made by this histidine to O3 is not highly conserved in the GH1 family, while the histidine residue itself is highly conserved at that position in this family. When the hydrogen bond is absent, the imidazole ring is rotated such that the new orientations of $N^{\delta 1}$ or $N^{\varepsilon 2}$ of His are not at an appropriate distance to make the hydrogen bond to O3. Examples of this orientation of His in GH1 families are β-glucosidase B from *Bacillus polymyxa*,[50] a β-glucosidase from Triticeae,[51,52] and Os3BGlu6, a β-glucosidase from *O. sativa Japonica*.[53] In these enzymes, this histidine has no direct hydrogen bond to O3 of glucose at subsite −1 or one mediated by water. In this orientation, histidine with the protonated $N^{\varepsilon 2}$ atom may interact with $O^{\varepsilon 2}$ of the nucleophilic glutamate. In some other members of this family, this histidine is oriented in a way that it interacts with both O3 and O2 simultaneously.[54] Figure 3 shows even though the hydrogen bond made by His130 of BGlu1 to O3 of glucose is maintained during the glycosylation and deglycosylation steps, its effect on the potential energy profile of hydrolysis was small compared to those of other active site residues. However, it seems that depending on the orientation of this histidine in the active site of GH1 glucosidases, they may play different roles in the energy profile.[16] Moreover, in addition to its interaction with the glucose moiety in the active site of glucosidases and stabilization of the transition state, this histidine may be involved in pH tolerance in different glycosidases and in keeping the appropriate protonation state of the nucleophile and proton donor in the hydrolysis reaction by switching between charged and neutral states.[55]

Trp441 is the other residue interacting with O3 of glucose (Figure S1 of the Supporting Information) that has a small REC value in the energy profile (Figure 3). This residue in some members of the GH1 family is replaced by Phe/Ala or is not hydrogen bonded to O3,[56] further demonstrating the minor effect of the hydrogen bond made by this tryptophan on the reduction of the energy barrier in GH1 enzymes. Asn313, which is hydrogen bonded to both catalytic carboxylates (Glu176 and Glu386), was also considered in the QM

calculations. Figure 3 and Table 3 show even though Asn313 is not in direct interaction with the glucose moiety at subsite −1, its average contribution to the energy profile may be higher than or at least at the same level as that of residues making direct hydrogen bonds to hydroxyl groups of glucose such as His130 and especially Gln29, which makes a bidentate hydrogen bond to glucose. This observation suggested that some other active site residues such as Asn313 may have their influence on the energy profile by polarizing catalytic residues along the reaction coordinates and facilitating hydrolysis. The water molecules that were considered in the calculations demonstrated a large REC value in the $(E{\cdot}I)_g$ complex (Figure 3). The energy-optimized structure of $(E{\cdot}I)_g$ illustrates that three of these four water molecules link the leaving glucose (the product of glycosylation step) to the rest of system by mediating hydrogen bonds, which is the reason for that large REC value in $(E{\cdot}I)_{dg}$. Figure 4 displays the REC value from

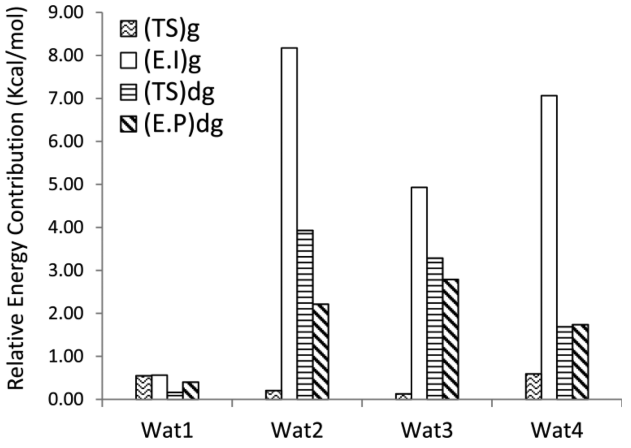



**Figure 4.** Relative energy contributions (RECs) of the four water molecules that are within 4 Å of the glucose in subsite −1 and were considered in the QM/DFT calculation. Labels are described in the legend of Figure 3.

water molecule decomposition. The three water molecules mentioned above had large values for the REC, while one water molecule that interacts with Glu440 made an insignificant energy contribution in all steps. To determine the role of these waters in the potential energy profile, they were evaluated in the protein environment where all other residue and water molecules were considered in the calculation.

**Hybrid QM/MM Analyses of Selected Active Site Residues in the Full System.** To assess the role of each active site residue in the energy profile along the reaction coordinate by GH1 $\beta$-glucosidases, we conducted hybrid QM/MM (ONIOM) calculations. Side chains of Asn175, Tyr315, and Glu440 [residues with significant relative energy contributions (Table 3)], three water molecules, glucose moieties at subsites −1 and +1 of cellotetraose, and side chains of the two catalytic residues were selected to be treated in a high-level (QM) region, and the rest of the protein and crystal or added water molecules were treated by the AMBER force field (MM). Geometry optimizations were performed on this ONIOM system (ONIOM-All) and its subsequent partitions [through transferring one by one each noncatalytic residue to the low-level region (Table 2)] for reactants, transition states, and products of both glycosylation and deglycosylation steps to explore if the optimized structures and energy profiles are affected by alternative QM/MM partitioning schemes. The final geometry-optimized structures from the ONIOM-All scheme are illustrated in Figure 5 for glycosylation and deglycosylation, and only the atoms in the QM region are presented. Figure 6 shows the relative energy profiles plotted on the basis of single-point energy calculations at the Amber/6-311+G(2df,2pf) level of theory for each of these ONIOM model systems (Table 2). This graph illustrates that the energy barriers are very sensitive to different QM/MM partitioning schemes and shows how the presence of each residue in the high-level region during the catalytic reaction affected the energy profile. The two steps of glycosylation and deglycosylation were connected by including the $\Delta G$ of desolvation of glucose from the protein environment to aqueous media.[57] The most dramatic change in the energy profile in Figure 6 happens in All-Y315, when Tyr315 was removed from the high-level (QM) region and moved to the low-level (MM) region. In particular, the activation energy of the deglycosylation step increased by 11.95 kcal/mol (Table 4). The same trend was also observed in the QM study (see above), where in the absence of Tyr315, the activation energy of the deglycosylation step increased by ∼8.0 kcal/mol (Figure S4 of the Supporting Information), in agreement with the large REC value of Tyr315 in $(TS)_{dg}$ (Figure 3). The other considerable effect of this tyrosine along the reaction coordinate is on the polarization of the carboxylate group of the nucleophilic glutamate when this residue was transferred to the MM region in ONIOM calculations. As highlighted in Table S6 of the Supporting Information, the CD-(Glu386)···$O^{\varepsilon 2}$(Glu386) bond was shorter by ∼0.015 Å on average when Tyr315 was removed from the high-level region relative to the equivalent bond length in other ONIOM partitioning. Although the change in bond length is small, it still highlights the role of tyrosine in the polarization of the nucleophilic bond by increasing its length.

Table 4 shows that exclusion of two other noncatalytic active site residues (Asn175 and Glu440) and waters from the QM region caused an increase in activation energies in both the glycosylation and deglycosylation steps. The effect of water molecules on the energy profile of glycosylation was near zero, indicating that the value of REC for waters of $(E{\cdot}I)_g$ shown in Figure 3 cannot be accurate. As explained before, those four water molecules attached the exiting glucose to the rest of system, while in the full systems, many other water molecules interacted with the leaving glucose, which reduced the final energy contribution of those four selected waters. Elimination of Asn175 from the high-level region induced destabilization at all stages of the reaction; however, because the energy of these stages increased relatively by the same amount, Asn175 did not significantly affect the energy barriers. The absence of this residue in the high-level region made the intermediate less stabilized than the reactant in the glycosylation step by 4.12 kcal/mol relative to those of the other studied residues (Figure 6). The exclusion of Glu440 from the QM region had the maximal effect on the energy barrier of the glycosylation step and increased it by 2.88 kcal/mol, with a small effect on the deglycosylation energy barrier (Table 4).

The extent of positive charge distribution between the C1 (anomeric carbon) and O5 (oxygen of glucose ring) atoms among different ONIOM partitioning schemes was also calculated using the 6-31+G(2df,2pd) basis set based on Mulliken atomic charges to evaluate how inclusion of each noncatalytic active site residue influenced the charge variation of C1 and O5 at both transition states. Table 4 lists the

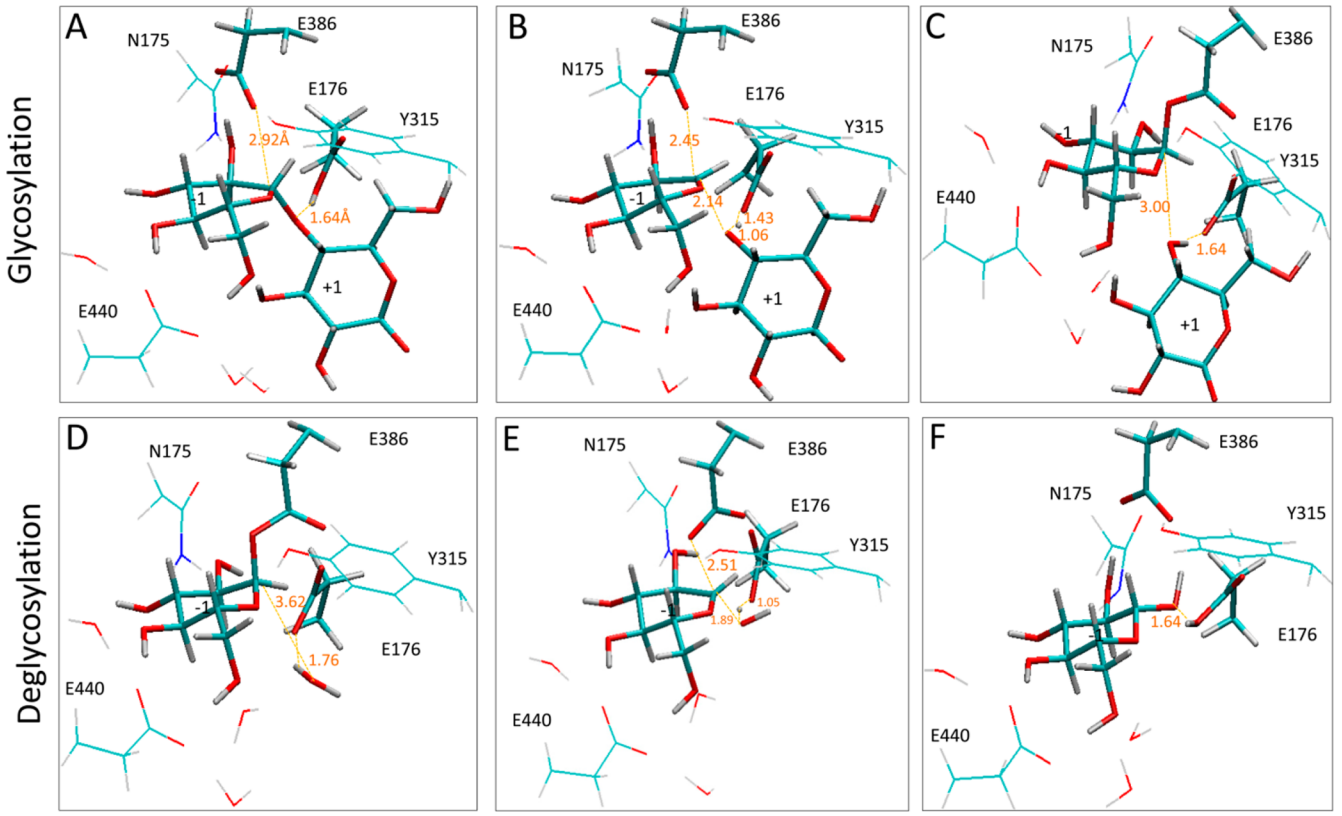



**Figure 5.** Optimized geometry for (A) reactant $(E{\cdot}S)_g$, (B) transition state $(TS)_g$, and (C) product $(E{\cdot}I)_g$ of the glycosylation step and (D) reactant $(E{\cdot}I)_{dg}$, (E) transition state $(TS)_{dg}$, and (F) product $(E{\cdot}P)_{dg}$ of the deglycosylation step calculated using the B3LYP/6-31G(d) theoretical method in the ONIOM-All partitioning scheme. Thick lines indicate the catalytic residues and substrate, and thin lines are used for noncatalytic residues and water molecules. The distances are in angstroms.

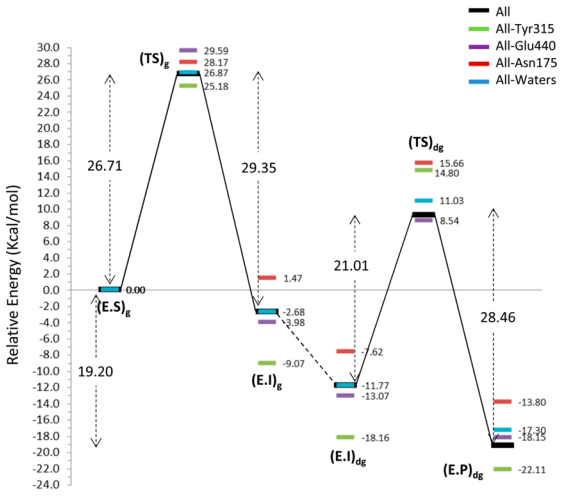



**Figure 6.** Relative energy profiles for the hydrolysis reaction by BGlu1 based on the ONIOM partitioning indicated in Table 2. "All" refers to "ONIOM-All", and all of the other residues named in the legend refer to the exclusion of that residue from the high-level region in the ONIOM calculations.

differences in charges between the transition state and the reactant for both glycosylation, $(TS)_g - (E{\cdot}S)_g$, and deglycosylation, $(TS)_{dg} - (E{\cdot}I)_{dg}$, steps in each ONIOM partitioning system (more details on the absolute values of atomic charges can be found in Tables S4−S8 of the Supporting Information). Exclusion of any residue or waters did not influence significantly the Δcharges for C1 in either of the two steps (Table 4). Δcharges for C1 were on average 0.075 and 0.124 au for glycosylation and deglycosylation, respectively. This difference in Δcharges for the two steps of hydrolysis was due to the more positive charge on C1 in the $(E{\cdot}S)_g$ complex that generated a smaller value for $(TS)_g - (E{\cdot}S)_g$, which is plausibly a direct consequence of preactivation of the glycan substrate in the active site of glucosidases.[58] O5 of the glucose ring also showed a considerable decrease in its negative charge for the two transition states [which is apparent from the positive values for differences in charges between transition states and reactants for both steps (Table 4)]. In contrast to a more positive charge on C1 in $(TS)_{dg}$ than in $(TS)_g$, the decrease in the negative charge of O5 in both transition states was almost the same. Interestingly, Δcharge for O5 when Tyr315 was excluded from the QM region in the ONIOM system was less than when Tyr315 was treated quantum mechanically, particularly in the deglycosylation step, again signifying the considerable role of tyrosine in efficient catalysis by interacting with O5 and decreasing its negative charge. This result is in good agreement with the results of QM/MM molecular dynamics simulations of BCX xylanase when Tyr69 (in the same position as Tyr315) is mutated to Phe in the high-level region.[23] The role of tyrosine was also reflected in the C1−O5 bond length when the difference in the length of this bond between the reactant and transition states was considered. This Δlength for the C1−O5 bond was always a negative value for any combination of QM and MM regions in the ONIOM systems (Table 4), indicating the C1−O5 bond was shorter in transition states than in reactants for each hydrolysis step, demonstrating the partial double bond


dx.doi.org/10.1021/bi300675x | *Biochemistry* 2012, 51, 8907−8918

**Table 4. Relative Energies and Structural Properties of Each ONIOM Partitioning after Geometry Optimization[a]**

| ONIOM partitioning | | Δenergy (kcal/mol) (relative to ONIOM-All) | | O5−C1 Δlength (TS reactant) (Å) | C5−O5−C1−C2 Δdihedral (TS reactant) (deg) | Δcharge (TS reactant) | |
|---|---|---|---|---|---|---|---|
| | | activation energy | reaction energy | | | C1 | O5 |
| ONIOM-All | Glyc | 0.00 | 0.00 | −0.092 | −15.967 | 0.079 | 0.134 |
| | Deglyc | 0.00 | 0.00 | −0.104 | 36.769 | 0.108 | 0.111 |
| ONIOM-N175 | Glyc | 1.46 | 4.12 | −0.091 | −12.952 | 0.072 | 0.105 |
| | Deglyc | 2.28 | 1.28 | −0.100 | 37.569 | 0.129 | 0.121 |
| ONIOM-Y315 | Glyc | −1.53 | −6.42 | −0.084 | −9.789 | 0.079 | 0.094 |
| | Deglyc | 11.95 | 3.51 | −0.079 | 41.738 | 0.122 | 0.050 |
| ONIOM-E440 | Glyc | 2.88 | −1.34 | −0.092 | −9.432 | 0.077 | 0.108 |
| | Deglyc | 0.61 | 2.38 | −0.104 | 36.625 | 0.136 | 0.112 |
| ONIOM-waters | Glyc | 0.16 | −0.04 | −0.092 | −13.029 | 0.075 | 0.098 |
| | Deglyc | 1.80 | 1.94 | −0.099 | 34.783 | 0.132 | 0.111 |

[a]Designation of a residue indicates that residue was excluded from the high-level region and transferred to the MM region in that partitioning. The absolute structural properties, and charges along with relative energies of each ONIOM partitioning, can be found in Tables S2−S8 of the Supporting Information. The geometry optimizations were performed at the B3LYP/6-31G(d)/parm99-GLYCAM_06 level of theory. Single-point energy calculations of each optimized geometry were conducted at the B3LYP/6-311+G(2df,2pd) level with electronic embedding.

character of the C1−O5 bond in $(TS)_g$ and $(TS)_{dg}$. However, the value of Δlength for the C1−O5 bond for glycosylation was smaller by ∼0.01 Å than this length for deglycosylation, which can be explained by the shorter length of this bond in the distorted conformation ($^1S_3$) of the glucose ring in the $(E{\cdot}S)_g$ complex.[58,59] In contrast to other ONIOM partitioning, when Tyr315 was excluded from the high-level region, the length of the C1−O5 bond in $(TS)_g$ or $(TS)_{dg}$ did not decrease as observed in the other systems when Tyr was included (Table 4 and Table S6 of the Supporting Information; the values can be compared to the equivalent values in Tables S4, S5, S7, and S8 of the Supporting Information), signifying the role of Tyr in the distortion of the glucose ring and the formation of the oxocarbenium-like transition state. Surprisingly, such effects on the C1−O5 bond length were also observed for the ONIOM system excluding Glu440. The system excluding Glu440 in the high-level region showed a longer C1−O5 bond in both $(E{\cdot}S)_g$ and $(TS)_g$ relative to the bonds in the other systems (Table S7 of the Supporting Information). The other impact of excluding Glu440 on optimized geometries can be observed in the large deviation of the glucopyranose ring conformation in the optimized $(E{\cdot}S)_g$ complex from its skew boat conformation ($^1S_3$) in the crystal structure (Table S7 of the Supporting Information, C5−O5−C1−C2 dihedral). The distorted conformation of the glucose ring in the E·S complex relative to the relaxed $^4C_1$ structure is mainly due to the geometry of the active site in subsites −1 and +1 (Figure S5 of the Supporting Information) and partially a result of the electronic interactions in subsite −1 that stabilize the distorted structure.[58,59] Deviation from the original $^1S_3$ conformation was observed in other ONIOM partitions as well as minimal systems (QM), most likely because of the exclusion of aromatic residues from QM calculations such as Trp433 and especially Trp358, which are in a $\pi$ interaction with glucose moieties at subsites −1 and +1, respectively. These aromatic residues have been shown to play a substantial role in ring distortion at the −1 site in the enzyme−substrate complex.[60] However, the maximal distortion of the ring from the initial $^1S_3$ conformation was observed when Glu440 was moved to the MM region in the $(E{\cdot}S)_g$ optimized structure. This structure also showed a shorter C1−O4 (glycosidic bond) length relative to other optimized structures with Glu440 in the high-level region. Elongation of the C1−O4 bond in the preactivated conformation of the glucopyranose ring in the $(E{\cdot}S)_g$ complex was observed in other studies with a distorted glucose ring.[58] All these results together indicate the interactions of this glutamate with O6 and particularly O4 of the glucose ring may significantly affect the geometry of glucose, including the length of the O5−C1 or O4−C1 bond and the C5−O5−C1−C2 dihedral angle. These deviations from the crystal structure in the geometry-optimized $(E{\cdot}S)_g$ conformation in the absence of Glu440 in the high-level region possibly can be explained by a contribution of the negative charge of Glu440 to stabilize the positive charge on the C1/O5 group in the glucose ring in both $(E{\cdot}S)_g$ and $(TS)_g$. However, a better explanation for the role of this glutamate in preactivation of $(E{\cdot}S)_g$ and stabilization of $(TS)_g$ may need further investigation such as in silico mutagenesis.

In Figure 7, the energy profile along the reaction coordinates for the ONIOM-All system is compared with the energy profile

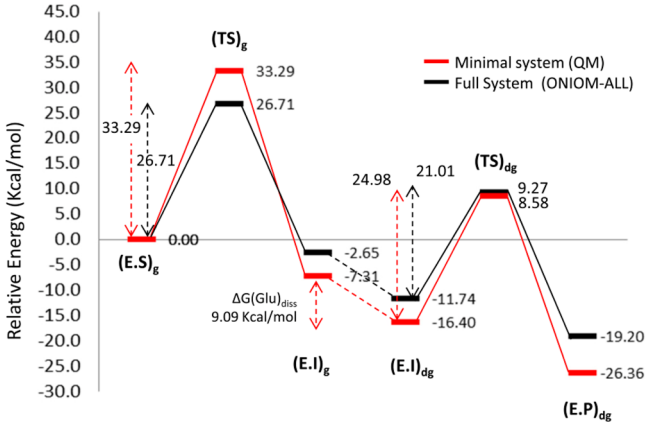



**Figure 7.** Relative energy profiles for the hydrolysis reaction by BGlu1. The energy barriers of the minimal system (QM/DFT) are compared with those of the full protein system (ONIOM-All).

calculated from a minimal (QM/DFT) system at the same level of theory. Although the number of residues hydrogen bonded to the glucose moiety that are treated by QM/DFT in the minimal system was greater than (twice) the number of residues treated in the high-level region of full system calculations, the energy barriers in the full system (whole


dx.doi.org/10.1021/bi300675x | Biochemistry 2012, 51, 8907−8918

protein) calculated in a two-layer ONIOM (QM/MM) are lower than that in the minimal system by 6.58 and 3.97 kcal/mol for glycosylation and deglycosylation steps, respectively. The values of energy barriers in glycoside hydrolysis in other studies cover a vast range of values and are highly dependent on the substrates, the residues included in the high-level region, the theoretical method, and the basis set used for the calculation.[22,49] Additionally, in our study, the dynamics of the enzyme were not considered, and it is possible that the orientation of the active site residues and crystal water molecules will change in the TS. It may be necessary to use alternative methods such as QM/MM molecular dynamics. No experimental data for the activation energy of BGlu1 are available; however, the energy barriers calculated for glycosylation and deglycosylation in this study using the ONIOM method of hybrid QM/MM fall into reasonable ranges when compared to calculated activation energies of glycosidic bond hydrolysis of other retaining glycoside hydrolases (Table S9 of the Supporting Information).

## CONCLUSION

In the active site of $\beta$-glycosidases, in addition to two catalytic residues (two glutamates for GH1 glucosidases), there are some highly conserved, noncatalytic polar residues that make direct hydrogen bonds to the glucose moiety at subsite −1 or to the oxygen of the glycosidic bond. Even though numerous studies have been performed with a number of glucosidases through in silico or in vitro mutagenesis to determine the role of these conserved active site residues in the energy profile of the hydrolysis reaction, many questions remain unanswered. The few available in vitro studies[15−17] showed the mutation of active site polar residues affects the $k_{cat}$ of glycosidase activity far more dramatically than the affinity of the enzyme for the substrate ($K_m$). In this study, using this pure QM calculation, we found Tyr315, which is hydrogen bonded to both the nucleophilic glutamate and O5 of the glucopyranose ring, made the maximal contribution to the energy profile of hydrolysis. In agreement with the QM/DFT study, excluding Tyr315 in the ONIOM study was found to influence substantially the energy barrier of hydrolysis, particularly in the deglycosylation step. The other residues with a considerable relative energy contribution are Glu440 and Asn175. Interestingly, residues such as Gln29 with bidentate bonds to O3 and O4 of the glucose moiety and His130 and Tyr441, which hydrogen bond to O3, contributed relatively little to the energy profile. On the basis of these results, the interactions involving subsite −1 and O3 of glucose are less significant to the stabilization of transition states and $\beta$-glucosidase activity, consistent with some other studies based on the method of hydroxyl group substitution.[18,61] Analysis of geometry-optimized structures at each step in the ONIOM system showed that the absence of Tyr315 in the high-level region influenced the preactivation of the substrate in the enzyme−substrate complex and also affected the accumulation of positive charge on the C1/O5 group.

## ASSOCIATED CONTENT

### Supporting Information

Details of relative energies of glycosylation and deglycosylation steps calculated in the minimal system (QM) or full system (QM/MM) as well as structural properties and values of charge variations among geometry-optimized structures of the full system at different stages of reaction at alternative QM/MM partitioning schemes (Tables S1−S8) and activation energies of glycosylation and deglycosylation steps for BGlu1 in this study compared with the recently published QM/MM-based calculation of activation energies for enzymes from different families of GH (Table S9). This material is available free of charge via the Internet at http://pubs.acs.org.

## AUTHOR INFORMATION

### Corresponding Author

*C.Z.: Department of Biological Systems Engineering, Virginia Tech, Blacksburg, VA 24061; e-mail, chzhang2@vt.edu; phone, (540) 231-7601. D.R.B.: Department of Biochemistry, Virginia Tech, Blacksburg, VA 24061; e-mail, drbevan@vt.edu; phone, (540) 231-5040.

### Funding

This work was supported by the Biodesign and Bioprocessing Center and an integrated, internal competitive grant from the College of Agriculture and Life Sciences at Virginia Tech.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We are grateful to the Virginia Tech Terascale Computing Facility for computing time on the SystemX supercomputer and SGI-Altix systems. We greatly acknowledge Reza Sohrabi and Theresah N. Korbieh for their help and technical assistance.

## REFERENCES

(1) Cairns, J. R. K., and Esen, A. (2010) $\beta$-Glucosidases. *Cell. Mol. Life Sci. 67*, 3389−3405.

(2) Elstein, D., and Zimran, A. (2011) Recent Advances in Treatment Approaches to Gaucher Disease. *Curr. Pharm. Biotechnol. 12*, 854−860.

(3) Henrissat, B., and Davies, G. (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol. 7*, 637−644.

(4) Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for Glycogenomics. *Nucleic Acids Res. 37*, D233−D238.

(5) Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J. 321*, 557−559.

(6) Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B., and Esen, A. (2000) The mechanism of substrate (aglycone) specificity in $\beta$-glucosidases is revealed by crystal structures of mutant maize $\beta$-glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. *Proc. Natl. Acad. Sci. U.S.A. 97*, 13555−13560.

(7) Tribolo, S., Berrin, J. G., Kroon, P. A., Czjzek, M., and Juge, N. (2007) The crystal structure of human cytosolic $\beta$-glucosidase unravels the substrate aglycone specificity of a family 1 glycoside hydrolase. *J. Mol. Biol. 370*, 964−975.

(8) Khan, S., Pozzo, T., Megyeri, M., Lindahl, S., Sundin, A., Turner, C., and Karlsson, E. N. (2011) Aglycone specificity of *Thermotoga neapolitana* $\beta$-glucosidase 1A modified by mutagenesis, leading to increased catalytic efficiency in quercetin-3-glucoside hydrolysis. *BMC Biochem. 12*, 11.

(9) Tsukada, T., Igarashi, K., Fushinobu, S., and Samejima, M. (2008) Role of subsite +1 residues in pH dependence and catalytic activity of the glycoside hydrolase family 1 $\beta$-glucosidase BGL1A from the basidiomycete *Phanerochaete chrysosporium*. *Biotechnol. Bioeng. 99*, 1295−1302.

(10) Verdoucq, L., Moriniere, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B., and Czjze, M. (2004) Structural determinants of substrate specificity in family 1 $\beta$-glucosidases: Novel insights from the

crystal structure of sorghum dhurrinase-1, a plant β-glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem. 279*, 31796−31803.
(11) Chuenchor, W., Pengthaisong, S., Robinson, R. C., Yuvaniyama, J., Svasti, J., and Cairns, J. R. K. (2011) The structural basis of oligosaccharide binding by rice BGlu1 β-glucosidase. *J. Struct. Biol. 173*, 169−179.
(12) Koshland, D. E. (1953) Strereochemistry and the Mechnaism of Enzymetic Reactions. *Biol. Rev. 28*, 416−436.
(13) Vuong, T. V., and Wilson, D. B. (2010) Glycoside Hydrolases: Catalytic Base/Nucleophile Diversity. *Biotechnol. Bioeng. 107*, 195−205.
(14) Vocadlo, D. J., and Davies, G. J. (2008) Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol. 12*, 539−555.
(15) Marana, S. R., Terra, W. R., and Ferreira, C. (2002) The role of amino-acid residues Q39 and E451 in the determination of substrate specificity of the *Spodoptera frugiperda* β-glycosidase. *Eur. J. Biochem. 269*, 3705−3714.
(16) Huber, R. E., Hlede, I. Y., Roth, N. J., McKenzie, K. C., and Ghumman, K. K. (2001) His-391 of β-galactosidase (*Escherichia coli*) promotes catalyses by strong interactions with the transition state. *Biochem. Cell Biol. 79*, 183−193.
(17) Kaper, T., van Heusden, H. H., van Loo, B., Vasella, A., van der Oost, J., and de Vos, W. M. (2002) Substrate specificity engineering of β-mannosidase and β-glucosidase from *Pyrococcus* by exchange of unique active site residues. *Biochemistry 41*, 4147−4155.
(18) Namchuk, M. N., and Withers, S. G. (1995) Mechanism of *Agrobacterium* β-glucosidase: Kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry 34*, 16194−16202.
(19) Fernandez, P., Canada, F. J., Jimenezbarbero, J., and Martinlomas, M. (1995) Substrate-Specificity of Small-Intestinal Lactase: Study of the Steric Effects and Hydrogen-Bonds Involved in Enzyme-Substrate Interaction. *Carbohydr. Res. 271*, 31−42.
(20) Kohn, W., Becke, A. D., and Parr, R. G. (1996) Density functional theory of electronic structure. *J. Phys. Chem. 100*, 12974−12980.
(21) Vreven, T., Morokuma, K., Farkas, O., Schlegel, H. B., and Frisch, M. J. (2003) Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem. 24*, 760−769.
(22) Bras, N. F., Moura-Tamames, S. A., Fernandes, P. A., and Ramos, M. J. (2008) Mechanistic Studies on the Formation of Glycosidase-Substrate and Glycosidase-Inhibitor Covalent Intermediates. *J. Comput. Chem. 29*, 2565−2574.
(23) Soliman, M. E. S., Pernia, J. J. R., Greig, I. R., and Williams, I. H. (2009) Mechanism of glycoside hydrolysis: A comparative QM/MM molecular dynamics analysis for wild type and Y69F mutant retaining xylanases. *Org. Biomol. Chem. 7*, 5236−5244.
(24) Biarnes, X., Ardevol, A., Iglesias-Fernandez, J., Planas, A., and Rovira, C. (2011) Catalytic itinerary in 1,3−1,4-β-glucanase unraveled by QM/MM metadynamics. Charge is not yet fully developed at the oxocarbenium ion-like transition state. *J. Am. Chem. Soc. 133*, 20301−20309.
(25) Chuenchor, W., Pengthaisong, S., Robinson, R. C., Yuvaniyama, J., Oonanant, W., Bevan, D. R., Esen, A., Chen, C. J., Opassiri, R., Svasti, J., and Cairns, J. R. K. (2008) Structural insights into rice BGlu1 β-glucosidase oligosaccharide hydrolysis and transglycosylation. *J. Mol. Biol. 377*, 1200−1215.
(26) Gloster, T. M., and Davies, G. J. (2010) Glycosidase inhibition: Assessing mimicry of the transition state. *Org. Biomol. Chem. 8*, 305−320.
(27) Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005) H++: A server for estimating $pK_a$s and adding missing hydrogens to macromolecules. *Nucleic Acids Res. 33*, W368−W371.
(28) Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulation liquid water. *J. Chem. Phys. 79*, 926−935.
(29) Hess, B., Kutzner, C., van der Spoel, D., and Lindahl, E. (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput. 4*, 435−447.
(30) Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., and Kollman, P. (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J. Comput. Chem. 24*, 1999−2012.
(31) Kirschner, K. N., Yongye, A. B., Tschampel, S. M., Gonzalez-Outeirino, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem. 29*, 622−655.
(32) 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*, 10089−10092.
(33) Hess, B. (2008) P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput. 4*, 116−122.
(34) Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.
(35) Nosé, S. (1984) A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys. 81*, 511−519.
(36) Hoover, W. G. (1985) Canonical Dynamics: Equilibrium Phase-Space Distribution. *Phys. Rev. A 31*, 1695−1697.
(37) Lee, C., Yang, W., and Parr, R. G. (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys. 37*, 785−789.
(38) Peng, C. Y., Ayala, P. Y., Schlegel, H. B., and Frisch, M. J. (1996) Using redundant internal coordinates to optimize equilibrium geometries and transition states. *J. Comput. Chem. 17*, 49−56.
(39) Tian, B., Strid, A., and Eriksson, L. A. (2011) Catalytic roles of active-site residues in 2-methyl-3-hydroxypyridine-5-carboxylic acid oxygenase: An ONIOM/DFT study. *J. Phys. Chem. B 115*, 1918−1926.
(40) Frisch, M. J., et al. (2009) *Gaussian 09*, revision A.02, Gaussian, Inc., Wallingford, CT.
(41) Dannenberg, J. J., Salvador, P., Asensio, A., and Wieczorek, R. (2006) Solvation energies of amides and peptides obtained from hybrid DFT/ONIOM/AM1 calculations and the Cramer/Truhlar (SMx) solvation methods. *Abstracts of Papers*, Vol. *232*, Abstract 169, American Chemical Society, Washington, DC.
(42) Cheatham, T. E., Cieplak, P., and Kollman, P. A. (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. *J. Biomol. Struct. Dyn. 16*, 845−862.
(43) Vreven, T., Frisch, M. J., Kudin, K. N., Schlegel, H. B., and Morokuma, K. (2006) Geometry optimization with QM/MM methods II: Explicit quadratic coupling. *Mol. Phys. 104*, 701−714.
(44) Tao, P., Fisher, J. F., Shi, Q. C., Vreven, T., Mobashery, S., and Schlegel, H. B. (2009) Matrix Metalloproteinase 2 Inhibition: Combined Quantum Mechanics and Molecular Mechanics Studies of the Inhibition Mechanism of (4-Phenoxyphenylsulfonyl)-methylthiirane and Its Oxirane Analogue. *Biochemistry 48*, 9839−9847.
(45) Vreven, T., Byun, K. S., Komaromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K., and Frisch, M. J. (2006) Combining quantum mechanics methods with molecular mechanics methods in ONIO. *J. Chem. Theory Comput. 2*, 815−826.
(46) Vreven, T., and Morokuma, K. (2006) Chapter 3 Hybrid Methods: ONIOM(QM:MM) and QM/MM. *Annu. Rep. Comput. Chem.*,, 35−51.
(47) Cioslowski, J. (1989) A New Population Analysis Based on Atomic Polar Tensors. *J. Am. Chem. Soc. 111*, 8333−8336.
(48) Tao, P., and Schlegel, H. B. (2010) A toolkit to assist ONIOM calculations. *J. Comput. Chem. 31*, 2363−2369.
(49) Bras, N. F., Ramos, M. J., and Fernandes, P. A. (2010) DFT studies on the β-glycosidase catalytic mechanism: The deglycosylation step. *THEOCHEM 946*, 125−133.
(50) Isorna, P., Polaina, J., Latorre-Garcia, L., Canada, F. J., Gonzalez, B., and Sanz-Aparicio, J. (2007) Crystal structures of *Paenibacillus*

*polymyxa* β-glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I Glycosidases. *J. Mol. Biol. 371*, 1204−1218.
(51) Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H., and Miyamoto, T. (2006) Molecular and structural characterization of hexameric β-D-glucosidases in wheat and rye. *Plant Physiol. 141*, 1237−1247.
(52) Sue, M., Nakamura, C., Miyamoto, T., and Yajima, S. (2011) Active-site architecture of benzoxazinone-glucoside β-D-glucosidases in Triticeae. *Plant Sci. (Amsterdam, Neth.) 180*, 268−275.
(53) Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., and Cairns, J. K. (2009) Structural and Enzymatic Characterization of Os3BGlu6, a Rice β-Glucosidase Hydrolyzing Hydrophobic Glycosides and (1→3)- and (1→2)-Linked Disaccharides. *Plant Physiol. 151*, 47−58.
(54) Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B., and Esen, A. (2001) Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with p-nitrophenyl β-D-thioglucoside. *Biochem. J. 354*, 37−46.
(55) Schubert, M., Poon, D. K. Y., Wicki, J., Tarling, C. A., Kwan, E. M., Nielsen, J. E., Withers, S. G., and McIntosh, L. P. (2007) Probing electrostatic interactions along the reaction pathway of a glycoside hydrolase: Histidine characterization by NMR spectroscopy. *Biochemistry 46*, 7383−7395.
(56) Zouhar, J., Vevodova, J., Marek, J., Damborsky, J., Su, X. D., and Brzobohaty, B. (2001) Insights into the functional architecture of the catalytic center of a maize β-glucosidase Zm-p60.1. *Plant Physiol. 127*, 973−985.
(57) Bras, N. F., Fernandes, P. A., and Ramos, M. J. (2010) QM/MM Studies on the β-Galactosidase Catalytic Mechanism: Hydrolysis and Transglycosylation Reactions. *J. Chem. Theory Comput. 6*, 421−433.
(58) Biarnes, X., Ardevol, A., Planas, A., Rovira, C., Laio, A., and Parrinello, M. (2007) The conformational free energy landscape of β-D-glucopyranose. Implications for substrate preactivation in β-glucoside hydrolases. *J. Am. Chem. Soc. 129*, 10686−10693.
(59) Biarnes, X., Nieto, J., Planas, A., and Rovira, C. (2006) Substrate distortion in the Michaelis complex of *Bacillus* 1,3−1,4-β-glucanase. Insight from first principles molecular dynamics simulations. *J. Biol. Chem. 281*, 1432−1441.
(60) Payne, C. M., Bomble, Y. J., Taylor, C. B., McCabe, C., Himmel, M. E., Crowley, M. F., and Beckham, G. T. (2011) Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation. *J. Biol. Chem. 286*, 41028−41035.
(61) Marana, S. R. (2006) Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life 58*, 63−73.