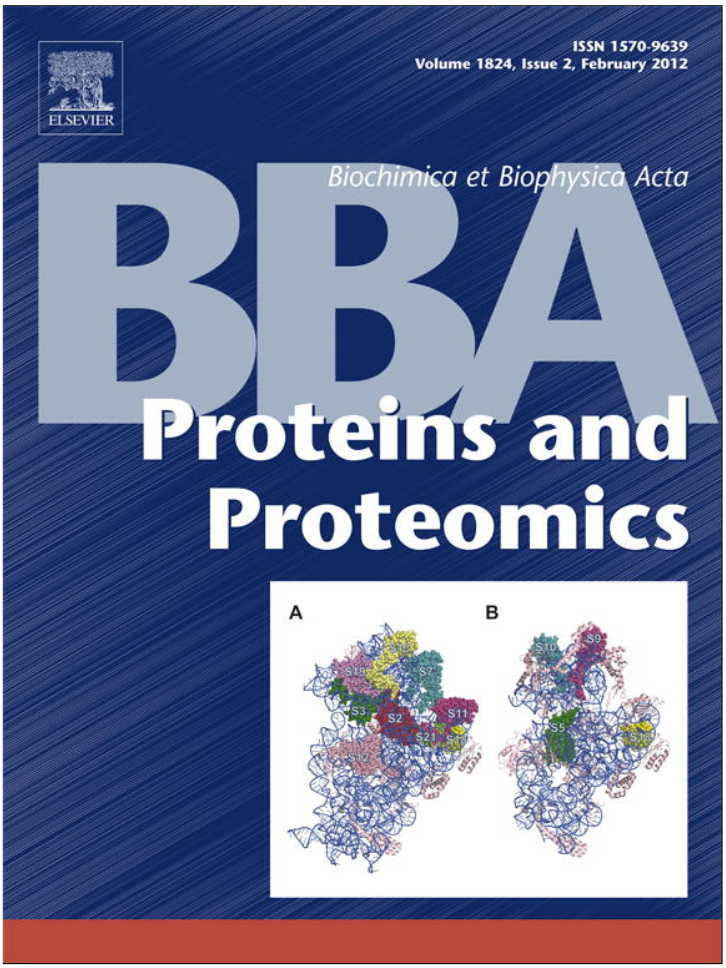

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Biochimica et Biophysica Acta 1824 (2012) 750–758

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

QM/MM investigation on the catalytic mechanism of *Bacteroides thetaiotaomicron* α-glucosidase BtGH97a

Jinhu Wang [a], Xiang Sheng [a], Yi Zhao [b], Yongjun Liu [a,b,*], Chengbu Liu [a]

[a] Key Lab of Theoretical and Computational Chemistry in University of Shandong, School of Chemistry and Chemical Engineering, Shandong University, Jinan, Shandong 250100, China
[b] Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining, Qinghai 810001, China

ARTICLE INFO

Article history:
Received 10 January 2012
Received in revised form 28 February 2012
Accepted 12 March 2012
Available online 19 March 2012

Keywords:
BtGH97a
pNP-Glc
QM/MM
Reaction mechanism

ABSTRACT

Bacteroides thetaiotaomicron α-glucosidase BtGH97a is an inverting enzyme. In this paper, the hydrolysis mechanism of p-nitro-phenyl α-D-glucopyranoside (pNP-Glc) catalyzed by BtGH97a was firstly studied by using quantum mechanical/molecular mechanical (QM/MM) approach. Two possible reaction pathways were considered. In the first pathway, a water molecule deprotonated by a nucleophilic base (here E439 or E508) attacks firstly on the anomeric carbon of pNP-Glc, then a proton from an acid residue (E532) attacks on the glycosidic oxygen to finish the hydrolysis reaction (named as nucleophilic attack-first pathway). In the second pathway, the proton from E532 attacks firstly on the glycosidic oxygen, then the water deprotonated by the nucleophilic base attacks on the anomeric carbon of pNP-Glc (named as proton attack-first pathway). Our calculation results indicate that the nucleophilic attack-first pathway is favorable in energy, in which the nucleophilic attack process is the rate-determining step with an energy barrier of 15.4 kcal/mol in the case of residue E508 as nucleophilic base. In this rate-determining step, the deprotonation of water and the attack on the anomeric carbon are concerted. In the proton attack-first pathway, the proton attack on the glycosidic oxygen is the rate-determining step, and the energy barrier is 24.1 kcal/mol. We conclude that the hydrolysis mechanism would follow nucleophilic attack-first pathway.

1. Introduction

There is a community of 10–100 trillion microbes colonizing in our own gut, which provides us with important physiological attributes, including the ability to break down indigestible nutrients that deliver to the distal gut. *Bacteroides thetaiotaomicron* (*B. thetaiotaomicron*), a Gram-negative anaerobe, is a bacterial symbiont. It is a dominant member of the intestinal microbiota of human gut [1]. It provides us with a range of beneficial metabolic tasks, which has not been encoded in our human genome [2,3], such as the ability to carry out a series of biochemical reactions. This bacterium is able to salvage energy from nutrients, particularly carbohydrates [4]. Furthermore, this flora plays a major role in the breakdown of polysaccharides ingested in the diet into a form that could be absorbed and utilized by the host [5,6].

Family GH97 is a kind of glycoside hydrolases (GHs) families, comprising about 80 members at present, and 10 of which are from *B. thetaiotaomicron* [7,8]. One member, BtGH97a (GenBank accession number: AAO78808; denoted as SusB) was reported recently [9,10] at 1.90 Å resolution. A calcium ion (Ca^{2+}) is present in the active-site region of this enzyme, coordinated by four glutamate residues (E194, E508, E526 and E532). It displays high apparent affinity and partial removal of the calcium ion causes a substantial reduction of activity. BtGH97a was shown to hydrolyze a range of substrates from maltose to maltoheptaose, and both α-1,4 and α-1,6 linkages [11]. Although the catalytic mechanism of BtGH97a was not clear, it had been predicted to be inverting in an insightful bioinformatics analysis (Fig. 1) [10], in which the protonation of glycosidic oxygen by the acid residue and the departure of leaving group were presumably concomitant. E532 was predicted to be the catalytic acid, while the residues E439 and E508 were thought to be the nucleophilic base residues in the inverting mechanism. In this experimental study, the important glutamate residues in the active-site (E439, E508 and E532) were mutated in turn to alanine [10]. The substrate p-nitro-phenyl α-D-glucopyranoside (pNP-Glc) was used to determine the full Michaelis–Menten kinetics constant (K_m) and dinitrophenyl α-D-glucopyranoside (dNP-Glc) to determine the k_{cat}/K_m. For all mutations, the K_m was higher and the k_{cat} was lower than that of the wild-type enzyme. The k_{cat}/K_m was reduced around 250,000-fold for E439A, and 4-fold for E532A. As to the mutation of E508A, no activity could be measured.

To better understand the reaction mechanism, two kinds of pathways are designed, as shown in Fig. 2. In the first pathway

Abbreviations: BtGH97a, *Bacteroides thetaiotaomicron* α-glucosidase; pNP-Glc, p-nitro-phenyl α-D-glucopyranoside; dNP-Glc, dinitrophenyl α-D-glucopyranoside; QM/MM method, quantum mechanical/molecular mechanical method; Nu, nucleophilic; PT, proton transfer; R, reactant; IM, intermediate; TS, transition state; P, product; RMSD, root-mean-square deviation; MD, molecular dynamic

* Corresponding author. Fax: +86 531 885 644 64.
E-mail address: yongjunliu_1@sdu.edu.cn (Y. Liu).

1570-9639/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2012.03.005

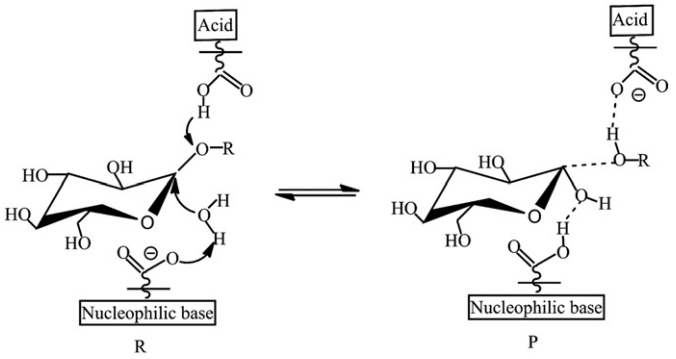

Fig. 1. Proposed inverting mechanism for BtGH97a.

(Fig. 2a), the nucleophilic attack occurs at the anomeric carbon by one water molecule that is deprotonated by a base residue, generating an intermediate (IM^{Nu}) via a transition state (TS^{Nu}). We call this process as "nucleophilic attack." Then, the proton attack takes place on the glycosidic oxygen via another transition state (TS^{PT}) to the final product P1, in which the acid/base residue acts as an acid. We call this process as "proton attack." In the second pathway (Fig. 2b), the "proton attack" on the glycosidic oxygen occurs firstly, experiencing a transition state (TS^{PT}) to the intermediate (IM^{PT}). Then, the "nucleophilic attack" process follows with a transition state (TS^{Nu}) to the product P2. The two pathways focus on the order of the two processes, i.e., the "nucleophilic attack" occurs firstly or the "proton attack" occurs firstly.

There were many theoretical studies on the hydrolysis mechanisms of remaining GHs enzymes [12–16]. However, few studies were focused on the hydrolysis of inverting ones. The goal of this work was to investigate the inverting hydrolysis mechanism of BtGH97 using the quantum mechanical/molecular mechanical (QM/MM) method. In the QM/MM calculations, a small subsystem that refers to bonds formation and cleavage is considered as the QM region and treated by quantum mechanics, where the remainder of the system is treated by molecular mechanics [17–21]. Our work has answered some meaningful questions, such as which residue is the most favorable nucleophilic base to assist the hydrolysis reaction, and which pathway is most possible.

2. Computational details

2.1. Automated docking setup

The initial structure for this calculation was taken from a recently published crystal of one member of family GH97 (BtGH97a, PDB ID: 2JKE) [10]. Since the crystal in complex with polysaccharide had not been determined by experiment, and some key kinetic parameters for BtGH97a were determined mainly with substrate pNP-Glc, the ligand pNP-Glc was selected as the disaccharide molecule and was docked into the binding pocket using Autodock 4.0 program [22]. Before docking, the substrate of pNP-Glc was optimized at the B3LYP/6-31G(d) level with Gaussian 03 package [23]. When docking, the grid scale was set as 60 Å × 60 Å × 60 Å based on Grid module, with a spacing of 0.375 Å between the grid points. Gasteiger charges [24] were set for both the ligand and protein. 50 Independent docking runs were performed. In the whole calculation, the protein was kept rigid, and all the torsional bonds of the ligand were kept free. Based on a root-mean-square deviation (RMSD) criterion of 10 Å, the docking results were clustered. Finally, the conformation with the more cluster member and the lowest protein–ligand interaction energy was chosen as the bioactive structure. Since BtGH97a superposed well with the closest GH27 structure (PDB ID: 1UAS) [25] over 277 residues with a root-mean-square deviation (RMSD) between Cα atoms of 3.3 Å, and there was a ligand complex (galactose) presenting in the GH27 structure, the obtained docking conformation was overlapped with this enzyme in order to check the reasonability of our method. The whole superposition was shown in Fig. 3.

2.2. Computational model

The obtained docking structure was used as the initial structure of molecular dynamic (MD) simulation. According to the experimental condition, the protonation states of all residues were selected carefully with the help of the VMD program [26]. As the crystal was obtained at the optimum pH value of 6.6, our calculations were carried out at this pH value. The catalytic acid/base, E532, was modeled in its protonated state, while other glutamate residues were modeled in their deprotonated, charged state. The protonation states and hydrogen atom positions of all other ionizable amino acid residues were selected based on their physiological protonation/deprotonation states. As there was no histidine residue in the QM region, the histidine residue was kept neutral and the proton was set on the ND atom. Hydrogen atoms were added via the HBUILD facility in the CHARMM package [27]. Besides, the crystallographic water molecules found in the protein were retained at their original positions and the extra 7122 water molecules were added to form a 47 Å water sphere centered on the residue E532. Finally, a neutral system of 32,627 atoms was then generated after the adding of 16 Na^+ ions. To equilibrate the prepared system, several minimizations followed by a 1000 ps MD simulation were performed with the CHARMM2227 force field [28].

During the subsequent QM/MM calculations, a total of eighty-nine atoms in the active site were selected as the QM region (Fig. 4), including the substrate pNP-Glc, two crystal waters Wat1 and Wat2, a calcium ion (Ca^{2+}), and parts of six important glutamate residues (E194, E391, E439, E508, E526 and E532). The atomic coordinates of these atoms (QM part) are listed in Supporting information. The remaining 32,538 atoms of the enzyme and waters were set as MM region. Any residue with at least one atom within 10 Å of E532 (including the QM region and part of MM region) was kept loose, while the remaining part was kept frozen. The non-frozen atoms are total 775 (the coordinates were also provided in Supporting information). In the geometry optimizations, the QM region was treated with quantum mechanics by Turbomole module [29] and the MM part with molecular mechanics under the CHARMM2227 force field by DL-POLY program [30]. The whole calculations were carried out at B3LYP/6-31G(d)/CHARMM22 level. For the QM/MM optimization, the convergence criterion of was set to 3×10^{-6} a.u. Frequency calculations were performed only for the QM part and the finite-difference method was used, in which the displacement and SCF convergence criterions were set as 0.001 a.u and 1.0×10^{-7} a.u, respectively. To avoid hyperpolarization of the QM wave function, the electronic embedding scheme [31] was used to incorporate the MM atomic partial charge into the one-electron Hamiltonian of the QM calculation. The charge shift model with hydrogen link atoms was adopted to simulate bonds and satisfy the valencies of covalently bonded atoms across the QM/MM boundary [32]. The ChemShell package [33] incorporating Turbomole and DL-POLY programs were used to perform the QM/MM calculations. Geometry optimization was carried out by hybrid delocalized internal coordinates (HDLC) optimizer [34]. During the optimizations, stationary points were searched by the quasi-Newton limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [35,36], which is suitable for optimization problems with a large number of variables. Transition states were searched by the algorithm of partitioned rational function optimization (P-RFO) [37,38]. This method could locate transition states by following eigenmodes of the Henssian, and the transition states are characterized by single negative eigenvalues.

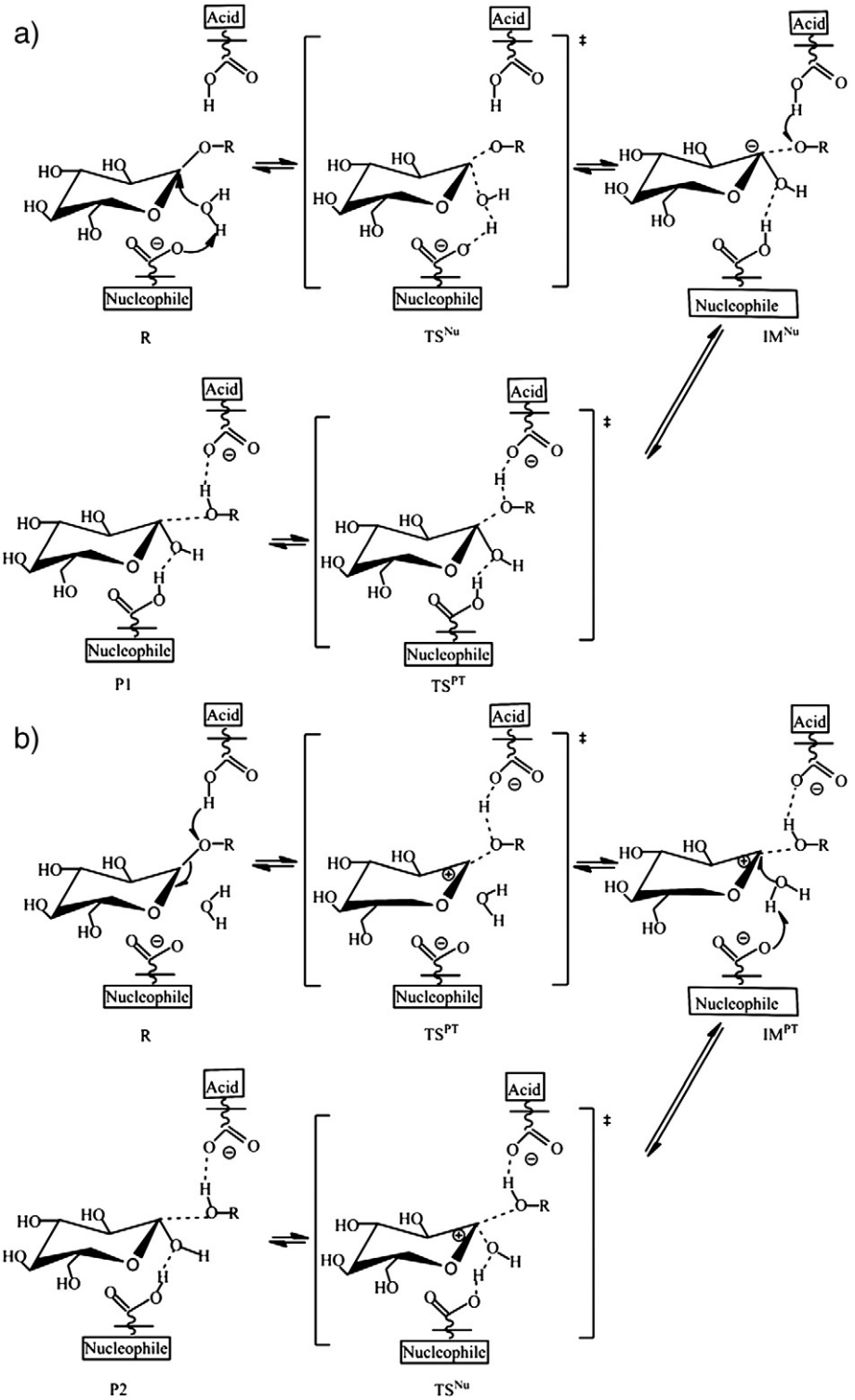

Fig. 2. The possible reaction mechanisms of inverting BtGH97a: (a) the nucleophilic attack-first process; (b) the proton attack-first process.

3. Results and discussion

3.1. Docking structure

A global overlap of the docking structure of BtGH97a (gray) with the GH27 α-galactosidase (blue) was established to validate the reasonability of our docking structure (as shown in Fig. 3a). In fact, totally about 277 residues could be superimposed together. The pocket residues are given in Fig. 3b, which shows that the active-site residues are partial conservation. Residues E391, K467 and E532 (acid residue) in BtGH97a superpose well with D51, K128 and D185 (acid/base residue) in GH27, respectively. The docked ligand pNP-Glc of BtGH97a overlaps well with that of the ligand galactose ring in the crystal of GH27. All these residues of E391, E439 and E532 can form hydrogen bonds (HBs) with pNP-Glc. The possible catalytic base residues E439 and E508 fix the crystal water Wat1 with distances of 2.54 and 2.61 Å, respectively. The hydroxyl group at C1 position of the substrate has an electrostatic interaction with calcium ion (Ca^{2+}) with

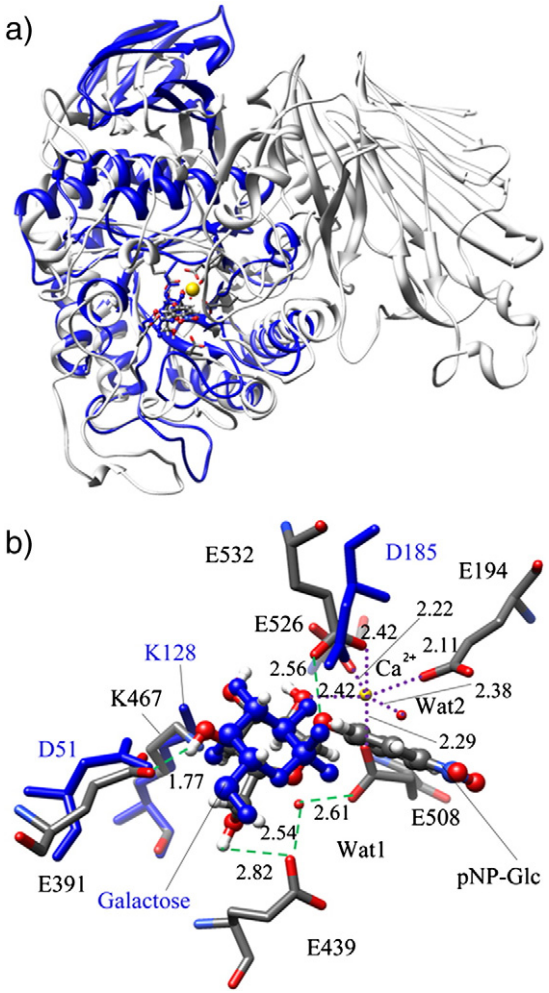

Fig. 3. (a) Superposition of the docking structure of BtGH97a (gray, PDB ID: 2JKE) with the closest GH27 α-galactosidase (blue, PDB ID: 1UAS); (b) pocket residues of the superposition. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a distance of 2.42 Å. Furthermore, four glutamate residues (E194, E508, E526 and E532) and a crystal water Wat2 have electrostatic interactions with Ca^{2+} ion with distances of 2.11, 2.29, 2.22, 2.42 and 2.38 Å, respectively. It is shown that the Ca^{2+} ion adopts an octahedral coordination, as was observed in the native structure [10]. All these descriptions indicate that the docking structure is reliable for the following QM/MM calculations.

3.2. The optimized structure of reactant

The final solvated model of prepared BtGH97a after QM/MM optimization is displayed in Fig. 5a, and the active site of the optimized structure of is shown in Fig. 5b. For clarity, only the important residues interacting with the substrate pNP-Glc are listed. Fig. 5b reveals that residues E439 and E508 are all hydrogen bonded with the water Wat1 with distances of 1.93 and 1.90 Å, respectively. In addition, another strong hydrogen bridge forms between E391 and the C4-hydroxyl group of pNP-Glc with a distance of 1.56 Å, which is 1.77 Å in the docking structure (Fig. 3b). The side chain of the catalytic acid residue E532 is hydrogen bonded to the glycosidic oxygen (O7) atom of ligand with a distance of 2.04 Å.

3.3. Analyses of different reaction pathways

Fig. 2 gives two possible reaction pathways. Each pathway contains two elementary steps: the nucleophilic attack step and proton attack step. In Fig. 2a the nucleophilic attack occurs firstly followed by the proton attack, while in Fig. 2b the proton attack occurs firstly then the nucleophilic attack takes place. The putative nucleophilic base residues are E439 and E508, and proton donor residue is E532. Different order may result in different results. So, in the following sections, two different pathways are discussed individually.

The geometries of the reactants, transition states, intermediates, and products were firstly optimized at 6–31G(d) level and the corresponding structures are shown in Figs. 5–8. The calculations of frequency were performed at the same level. Calculation results indicate that all the optimized geometries correspond to a local minimum that has no imaginary frequency mode or a saddle point that has only one imaginary frequency mode. The unique imaginary frequency for the corresponding transition state obtained from QM/MM optimizations is listed in Supporting information (Table S1). Furthermore, the energies were then recalculated with single point (SP) calculations with the higher 6-31++G(d,p) basis set. The obtained energy profiles for the two pathways are shown in Fig. 9.

3.3.1. Nucleophilic attack-first pathway

E439 and E508 are the most possible nucleophilic base residues. When the nucleophilic base residue is E439, the optimized structures of transition states ($TS^{Nu}a1$ and $TS^{PT}a1$), intermediate ($IM^{Nu}a1$) and product (P1) are displayed in Fig. 6.

In the nucleophilic attack-first pathway, E439 firstly gets closer to Wat1 proton followed by the approach of Wat1 to the anomeric carbon (C6) atom (distance shorts to 2.26 Å), forming a transition state $TS^{Nu}a1$ (Fig. 6a). The distance between COO^- of E439 and one proton of Wat1 shortens from 1.93 Å in reactant R to 1.81 Å. Moreover, the glycosidic bond (C6—O7) elongates from 1.48 Å in R to 2.63 Å, meaning this bond is almost broken at $TS^{Nu}a1$. The glycosidic oxygen establishes a hydrogen bridge (distance of 1.59 Å) with the catalytic acid residue E532, and the bond length of O—H in E532 increases to 1.02 Å. However, the configuration and relative position of Ca^{2+} ion are similar to the reactant, still coordinating with residues E194, E508, E526 and E532, C1-hydroxyl group of pNP-Glc, and water Wat1. Therefore, in the following discussion, the distance changes of Ca^{2+} with other ligands are omitted unless they are necessary.

Fig. 6b gives the intermediate structure of $IM^{Nu}a1$. In this state, the Wat1 proton has been abstracted by E439 (valence bond distance of 0.98 Å) with its O—H group attaching to the anomeric C6 of pNP-Glc (1.42 Å). The attachment of OH group increases the length of glycosidic bond by 0.54 Å compared with $TS^{Nu}a1$. The long distance of this bond (3.17 Å) implies that the bond cleavage has finished.

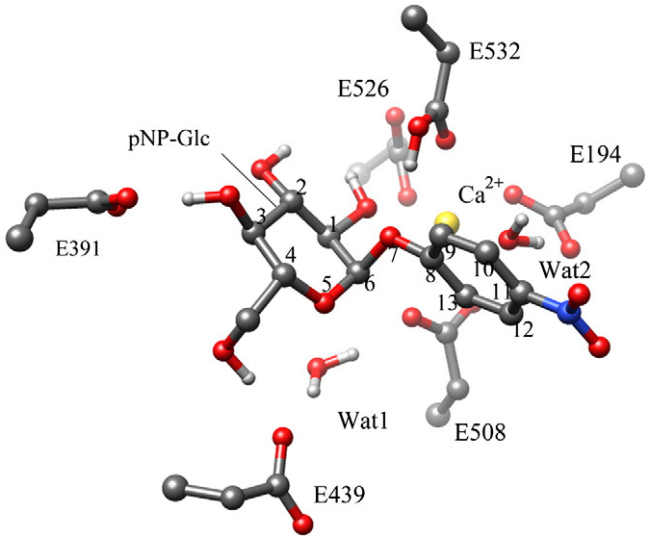

Fig. 4. The selected quantum mechanics (QM) region in QM/MM calculations.

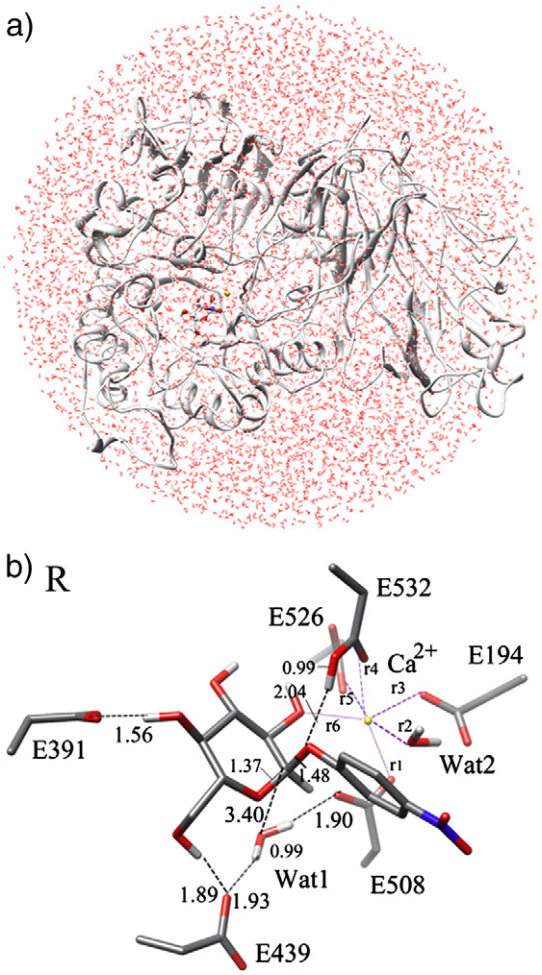

Fig. 5. (a) The solvated model of BtGH97a after QM/MM optimization (the reactant R); (b) the corresponding residues in the active site. The labels *r*1, *r*2, *r*3, *r*4, *r*5 and *r*6 represent the distances between the Ca^{2+} ion and oxygen atoms of E508, water Wat2, E194, E532, E526 and C1-hydroxyl group of pNP-Glc, respectively.

Furthermore, the interaction between E532 and glycosidic oxygen has been strengthened. The covalent H—O of E532 increases to 1.07 Å and the hydrogen bridge between E532 and O7 decreases to 1.42 Å after the cleavage of glycosidic bond C6—O7.

Then, the proton attack (E532 as the donor group) occurs to generate the final product (P1) via a transition state ($TS^{PT}a1$). In $TS^{PT}a1$ (Fig. 6c), the H—O bond length in E532 changes from 1.03 Å in R to 1.36 Å, meaning this bond is much weakened. Obviously, the proton has come into contact with O7 atom, because the distance between the two atoms is approximately a covalent bond (1.10 Å). Another clear change is the glycosidic bond. Its length increases further to 3.24 Å. In the structure of P1 (Fig. 6d), the H—O bond in E532 increases to 1.43 Å, indicating the cleavage of this bond is completed. As to other bond distances, only slight changes take place compared with $TS^{PT}a1$.

The energy profile is shown in Fig. 9 with black line. When base E439 abstracts a proton from Wat1 (the nucleophilic attack step), the energy barrier is calculated to be 18.2 kcal/mol. The relative energy of intermediate state $IM^{Nu}a1$ is 12.2 kcal/mol, indicating that the step of nucleophilic attack is endothermic. In the second step (proton attack process), the energy barrier is only 1.0 kcal/mol, implying the proton attack process is much easier and the nucleophilic attack is the rate-determining step.

When the nucleophilic base residue is E508, the structures of transition states ($TS^{Nu}a2$ and $TS^{PT}a2$), intermediate ($IM^{Nu}a2$), and product (P2) are also optimized, as shown in Fig. 7.

In $TS^{Nu}a2$ (Fig. 7a), E508 forms a strong hydrogen bridge with Wat1 with a distance of 1.64 Å. At the same time, Wat1 comes closer to C6 (1.95 Å) of the substrate, and the glycosidic bond elongates to 2.85 Å. The acid residue E532 elongates its O—H to 1.03 Å, establishing a strong HB interaction with glycosidic oxygen (distance of 1.53 Å).

In intermediate state $IM^{Nu}a2$ (Fig. 7b), the abstraction of the proton from Wat1 by E508 has finished, and the O—H group of Wat1 has attached to the substrate with a bond distance of 1.47 Å. Besides, the cleavage of glycosidic bond has also finished (distance of 3.15 Å). The side chain of catalytic acid E532 forms a strong hydrogen bridge with O7 atom with a length of 1.40 Å, and the previous covalent H—O bond distance of E532 increases to 1.08 Å. Clearly, the HB interaction between E532 and O7 atom has strengthened, which is feasible for the following proton attack.

A transition state ($TS^{PT}a1$) is obtained when the proton attack of E532 takes place, as shown in Fig. 7c. The original H—O bond in E532 weakens with a length of 1.24 Å and the hydrogen bond between the side chain of E532 and glycosidic oxygen strengthens (distance of 1.18 Å). In the final structure of P2 (Fig. 7d), the catalytic acid E532 has transferred its proton to O7 atom and formed a new hydrogen bond of 1.58 Å and a new covalent H(E532)—O7(pNP-Glc) bond of 1.03 Å. Though the glycosidic bond length decreases slightly compared with that of $TS^{PT}a1$ (3.07 vs. 3.18 Å), it still remains at the cleavage state.

The energy profile is shown in Fig. 9 with red line, which shows that the energy barrier of nucleophilic attack is 15.4 kcal/mol, and that of proton attack process is 0.2 kcal/mol. Therefore, nucleophilic attack is still the rate-determining step.

Compared with E439, E508 as nucleophilic residue corresponds to a lower energy than that of E439 (15.4 vs. 18.2 kcal/mol), so E508 is the most possible nucleophilic base residue for the nucleophilic attack-first pathway.

3.3.2. Proton attack first pathway

When the proton donor is residue E532, the proton attack corresponds one transition state ($TS^{PT}b$) and an intermediate structure ($IM^{Nu}b$). Since there are two possible nucleophilic base residues (E439 and E508), two different nucleophilic attack steps are considered. All the optimized structures are shown in Fig. 8.

In the structure of $TS^{PT}b$ (Fig. 8a), the E532 partially donates its proton to O7 atom with a distance of 1.14 Å, which is approximately a covalent bond. At the same time, the H—O bond length of E532 increases to 1.31 Å, meaning this bond weakens greatly. With the approach of E532 proton to O7 atom, the glycosidic bond is also weakened with a bond length increasing to 2.28 Å. Obviously, the cleavage of glycosidic bond and protonation of glycosidic oxygen proceed in a concerted manner in the proton attack step. Besides, Wat1 approaches to anomeric C6 atom when glycosidic oxygen moves away, but the covalent bond has not yet formed in this structure (distance shorts to 2.70 Å). The water still forms two hydrogen bridges with side chain of E439 and E508 with distances of 1.85 and 1.81 Å, respectively.

In the structure of $IM^{PT}b$ (Fig. 8b), the acid E532 donates its proton to O7 atom. The old O—H bond of E532 and the newly formed H(E532)—O7(pNPGlc) bond distances are calculated to be 1.42 and 1.08 Å, respectively, showing the process of PT is completed. With the protonation of glycosidic oxygen, the glycosidic bond distance elongates further to 2.38 Å, a little longer than that in $TS^{PT}b$. Obviously, the proton transfer process promotes the cleavage of glycosidic bond. In addition, the water Wat1 gets closer to anomeric C6 (length of 2.58 Å) and the hydrogen bond interactions between Wat1 and base residues E439 and E508 strengthen. In this stage, the proton transfer of Wat1 is ready to occur (to E439 or E508).

When E439 acts as the proton acceptor, a transition state $TS^{Nu}b1$ (Fig. 8c) is obtained. In $TS^{Nu}b1$, the COO^- group of E439 comes closer

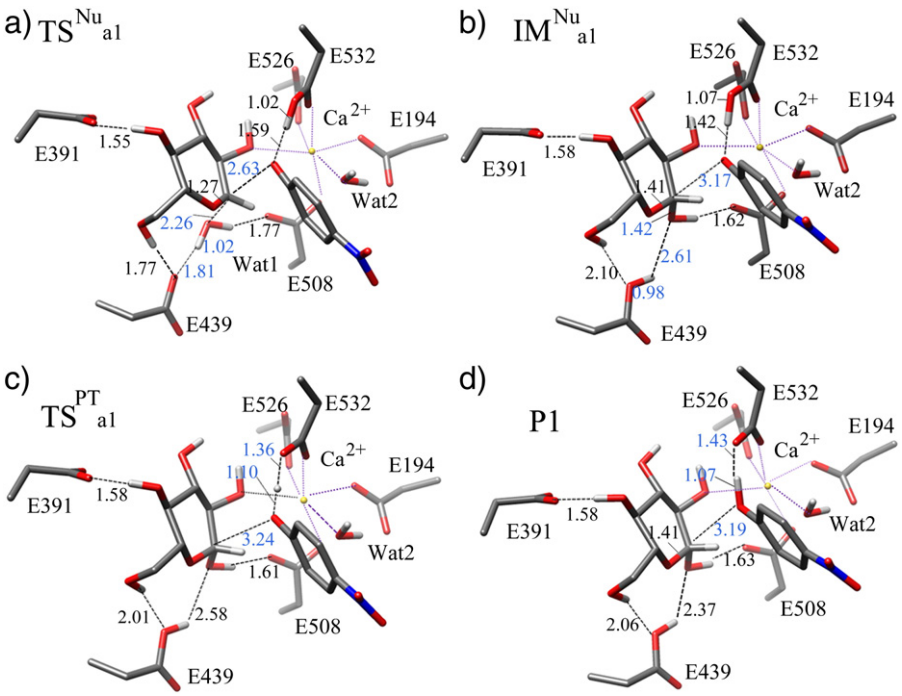

Fig. 6. Optimized structures of the nucleophilic attack-first (E439 as the nucleophile base) pathway. The bond distances which are formed or broken in each state are colored blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to Wat1 (distance of 1.74 Å) and Wat1 approaches to anomeric C6 (distance of 2.02 Å) accompanying with the weakness of the glycosidic bond (length of 2.72 Å). In addition, the old O—H bond in acid E532 elongates to 1.49 Å accompanying with the formation of a new O—H bond (distance of 1.05 Å).

When E508 acts as the proton acceptor ($TS^{Nu}b2$ in Fig. 8d), its COO^- group forms a short hydrogen bond with Wat1 (length of 1.67 Å). At the same time, Wat1 also gets closer to the sugar ring (distance of 2.22 Å). The position of Wat1 is similar to that in $TS^{Nu}a2$.

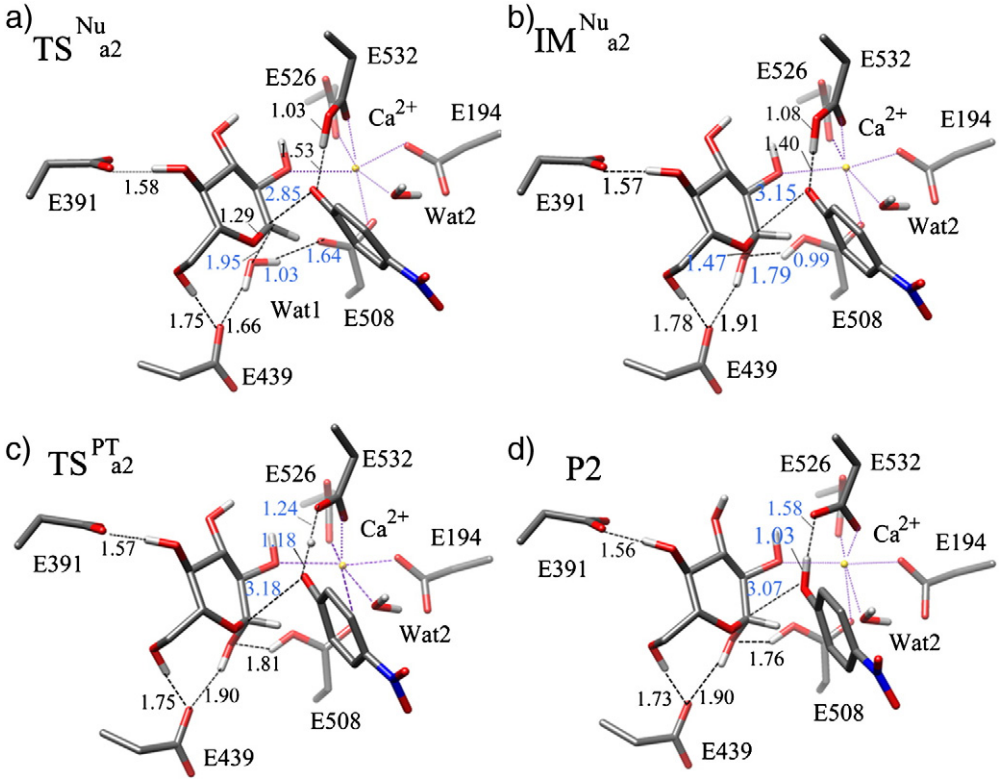

Fig. 7. Optimized structures of the nucleophilic attack-first (E508 as the nucleophile base) pathway. The bond distances which are formed or broken in each state are colored blue.

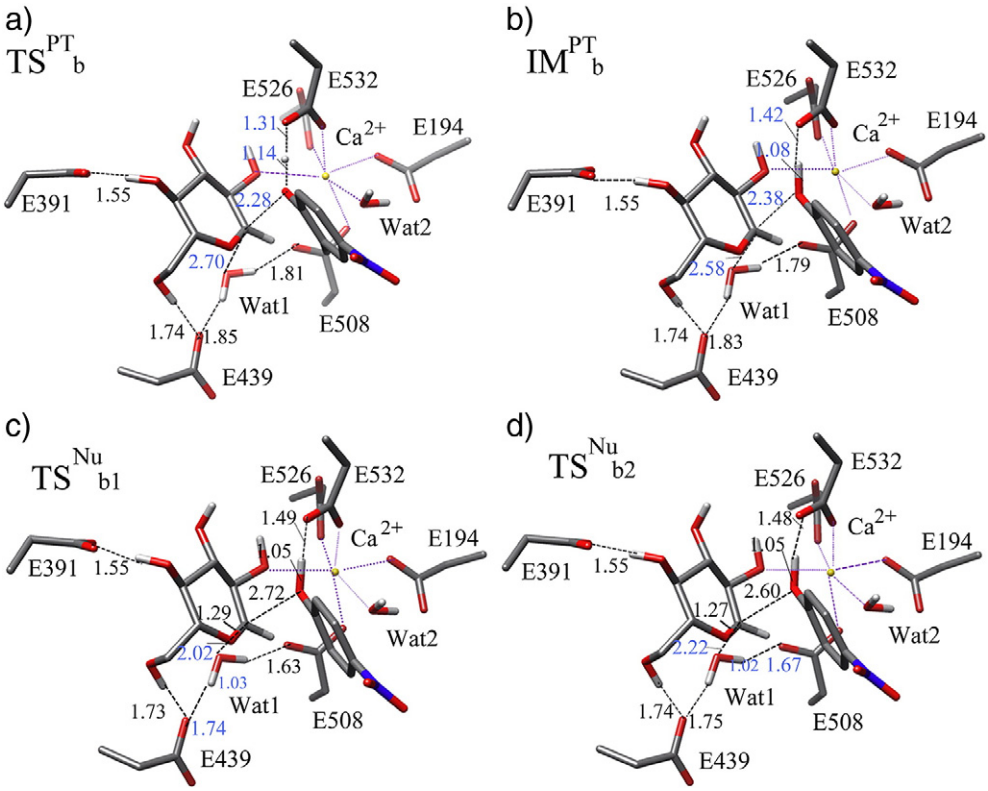

Fig. 8. Optimized structures of the proton attack-first pathway (E532 as the donor). The bond distances which are formed or broken in each state are colored blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The corresponding energy profile (lines colored in blue and green) is also shown in Fig. 9. In the proton attack step, the energy barrier is calculated to be 24.1 kcal/mol. In the next nucleophilic attack steps, both of the energy barriers are negligible (no more than 1.0 kcal/mol), indicating that the proton attack process is the rate-determining step.

From the energy barrier point of view, the nucleophilic attack-first pathways are favorable, and E508 is the most possible nucleophilic base.

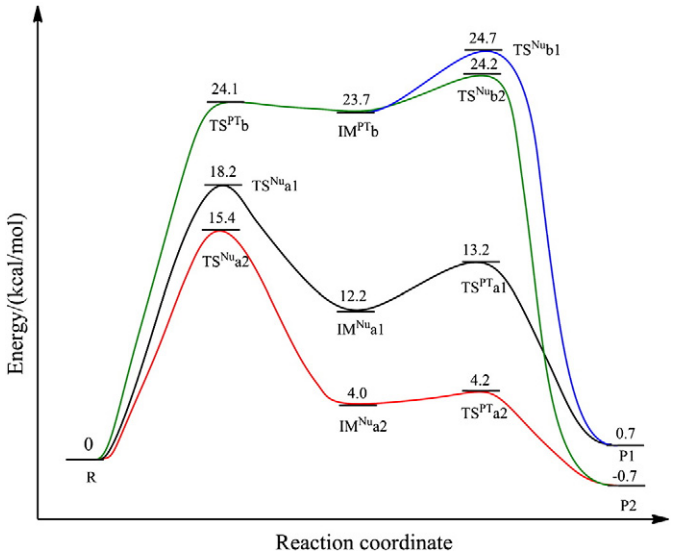

Fig. 9. Energy profiles for the nucleophilic attack-first pathway (black line for E439 and red for E508) and the proton attack-first pathway (blue line for E439 and green for E508). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Obviously, in the step-wise mechanism mentioned above, the residue E508 firstly extracts a proton from the water Wat1 and at the same time the deprotonated water attacks on the C6 followed by a proton attack step. So, residue E508 plays an important role in enzyme catalytic reaction. This is consistent with the mutation experiment [10]. No activity could be measured when the glutamate residue of E508 was mutated to alanine. As to another possible nucleophilic base E439, the mutation of E439A made the k_{cat}/K_m value reduce about 250,000-fold.

Furthermore, a concerted hydrolysis mechanism was also investigated, in which the protonation of glycosidic oxygen, the cleavage of glycosidic bond and the attack of Wat1 to C6 proceed in a concerted manner. The corresponding transition state ($TS^{PT}Nu$) and energy profile are shown in Supporting information (Fig. S1 and Fig. S2). In $TS^{PT}Nu$, the E532 has partially donated its proton to O7 atom, and

Table 1
QM/MM optimized Ca^{2+} ion–oxygen bond lengths (Å) in the substrate-binding site.

	r1[a]	r2	r3	r4	r5	r6
R	2.43	2.43	2.11	2.47	2.38	2.53
$TS^{Nu}a1$	2.42	2.40	2.11	2.43	2.40	2.55
$IM^{Nu}a1$	2.49	2.42	2.13	2.44	2.42	2.58
$TS^{PT}a1$	2.51	2.44	2.12	2.40	2.42	2.58
P1	2.50	2.44	2.12	2.39	2.42	2.59
$TS^{Nu}a2$	2.41	2.40	2.11	2.42	2.40	2.55
$IM^{Nu}a2$	2.55	2.40	2.10	2.40	2.36	2.50
$TS^{PT}a2$	2.58	2.42	2.10	2.33	2.40	2.50
P2	2.61	2.46	2.10	2.38	2.40	2.48
$TS^{PT}b$	2.44	2.45	2.12	2.39	2.41	2.63
$IM^{PT}b$	2.45	2.45	2.12	2.37	2.41	2.63
$TS^{Nu}b1$	2.43	2.45	2.11	2.35	2.42	2.61
$TS^{Nu}b2$	2.44	2.45	2.11	2.36	2.42	2.57

[a] r1, r2—denote the distances between Ca^{2+} ion and oxygen atoms as shown in Fig. 5b.

the glycosidic bond becomes much weakened (distance of 2.40 Å). Besides, Wat1 approaches to anomeric C6 atom (distance shorts to 2.36 Å) and its two O—H bonds elongate a little. The energy barrier of the concerted process is calculated to be 26.7 kcal/mol, much higher than that of the step-wise mechanism. So, the nucleophilic attack-first pathway is favorable, and E508 is the most possible nucleophilic base.

3.4. Changes of calcium ion–oxygen bond lengths in the substrate-binding site

Table 1 lists the calculated Ca^{2+}-oxygen bond lengths in the QM/MM optimized structures. We have mentioned that Ca^{2+} ion coordinates with four residues (E194, E508, E526 and E532), C1-hydroxyl group of pNP-Glc, and water Wat2. For clarity, we use *r*1, *r*2, *r*3, *r*4, *r*5 and *r*6 (which are labeled in Fig. 5) to represent the distances between the Ca^{2+} ion and oxygen atoms of E508, water Wat2, E194, E532, E526 and C1-hydroxyl group of pNP-Glc, respectively. As shown in Table 1, those bond lengths change slightly compared with R. Take the nucleophilic attack-first pathway as an example. The distance between the Ca^{2+} ion and COO^- group of E508 (labeled as *r*1 in Table 1) is 2.43 Å in R, and changes to 2.42, 2.49, 2.51 and 2.50 Å in $TS^{Nu}a1$, $IM^{Nu}a1$, $TS^{PT}a1$ and P1, respectively, indicating that calcium ion keeps very similar coordination with the oxygen atoms as the reaction going.

3.5. Insight into the catalytic effect

BtGH97a could hydrolyze a wide range types of α-glucosidic linkage not only α-1,4-glucosidic linkages but also α-1,6-, α-1,3-, and α-1,2-glucosidic linkages [9]. pNP-Glc is the best model substrate, and the highest k_{cat}/K_m value can be obtained. The hydrolytic activity of enzyme depends on the pH and temperature. As observed from the experiment [10], the optimum pH was about 6.6 and the optimum temperature was 45 °C, indicating that the hydrolysis reaction of pNP-Glc by BtGH97a is typical enzymatic catalysis. We have mentioned that the hydrolysis of pNP-Glc follows an inverting mechanism with energy barriers of 15.4 kcal/mol for the nucleophilic step and 0.2 kcal/mol for the proton attack step. Although both E439 and E508 can be general catalytic bases to assist the nucleophilic attack of one water on C6 atom of pNP-Glc, E508 acts as the most possible nucleophilic base, because the energy barrier of the nucleophilic step for E508 is a little lower than that of E439 (15.4 vs. 18.2 kcal/mol). Also, the mutation experiment supported our results [10]. The mutation of E439A exhibited little reactivity (reducing about 250,000-fold) but no activity could be measured when the glutamate residue of E508 was mutated to alanine. We hope our results could give some insight on the further study of β-glucosidase in experiment.

4. Conclusions

The enzymatic hydrolysis of pNP-Glc by BtGH97a has been studied by using QM/MM approach. Since the whole reaction contains two processes (the nucleophilic attack and proton attack steps), two possible reaction pathways were considered. In the nucleophilic attack-first pathway, when E439 acts as the nucleophilic base, the nucleophilic attack process is the rate-determining step with an energy barrier of 18.2 kcal/mol, while when E508 is the nucleophilic base, the energy barrier is calculated to be 15.4 kcal/mol. In the proton attack-first pathway, the proton attack process is the rate-determining step and the corresponding energy barrier is 24.1 kcal/mol. Since the energy barrier of nucleophilic attack-first pathway is much lower than that of proton attack-first pathway, the hydrolysis mechanism would follow the nucleophilic attack-first pathway.

Acknowledgments

This work was supported by the Natural Science Foundation of China (21173129).

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.bbapap.2012.03.005.

References

[1] J.I. Gordon, M.K. Bjursell, E.C. Martens, Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period, J. Biol. Chem. 281 (2006) 36269–36279.
[2] J.I. Gordon, J.L. Sonnenburg, J. Xu, D.D. Leip, C.H. Chen, B.P. Westover, J. Weatherford, J.D. Buhler, Glycan foraging in vivo by an intestine-adapted bacterial symbiont, Science 307 (2005) 1955–1959.
[3] S.R. Gill, M. Pop, R.T. DeBoy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, D.A. Relman, C.M. Fraser-Liggett, K.E. Nelson, Metagenomic analysis of the human distal gut microbiome, Science 312 (2006) 1355–1359.
[4] B. Henrissat, A. Bairoch, Updating the sequence-based classification of glycosyl hydrolases, Biochem. J. 316 (1996) 695–696.
[5] L.E. Comstock, M.J. Coyne, Bacteroiades thetaiotaomicron: a dynamic, niche-adapted human symbiont, Bioessays 25 (2003) 926–929.
[6] J.I. Gordon, J. Xu, M.K. Bjursell, J. Himrod, S. Deng, L.K. Carmichael, H.C. Chiang, L.V. Hooper, A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis, Science 299 (2003) 2074–2076.
[7] D.G. Naumoff, GH97 is a new family of glycoside hydrolases, which is related to the alpha-galactosidase superfamily, BMC Genomics 6 (2005) 112–123.
[8] D.G. Naumoff, GH97 is a new family of alpha-glucosidases, FEBS J. 272 (2005) 94–95.
[9] M. Yao, M. Kitamura, M. Okuyama, F. Tanzawa, H. Mori, Y. Kitago, N. Watanabe, A. Kimura, I. Tanaka, Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*, J. Biol. Chem. 283 (2008) 36328–36337.
[10] T.M. Gloster, J.P. Turkenburg, J.R. Potts, B. Henrissat, G.J. Davies, Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora, Chem. Biol. 15 (2008) 1058–1067.
[11] K.A. Smith, A.A. Salyers, Characterization of a neopullulanase and an alpha-glucosidase from *Bacteroides thetaiotaomicron* 95–1, J. Bacteriol. 173 (1991) 2962–2968.
[12] J.H. Li, L.K. Du, L.S. Wang, Glycosidic-bond hydrolysis mechanism catalyzed by cellulase Cel7A from *Trichoderma reesei*: a comprehensive theoretical study by performing MD, QM, and QM/MM calculations, J. Phys. Chem. B 114 (2010) 15261–15268.
[13] N.F. Bras, P.A. Fernandes, M.J. Ramos, QM/MM studies on the beta-galactosidase catalytic mechanism: hydrolysis and transglycosylation reactions, J. Chem. Theory Comput. 6 (2010) 421–433.
[14] J.L. Liu, X.M. Wang, D.G. Xu, QM/MM study on the catalytic mechanism of cellulose hydrolysis catalyzed by cellulase Cel5A from *Acidothermus cellulolyticus*, J. Phys. Chem. B 114 (2010) 1462–1470.
[15] W.J. Huang, J. Llano, J.W. Gauld, Redox mechanism of glycosidic bond hydrolysis catalyzed by 6-phospho-alpha-glucosidase: a DFT study, J. Phys. Chem. B 114 (2010) 11196–11206.
[16] J.H. Wang, Q.Q. Hou, L.H. Dong, Y.J. Liu, C.B. Liu, QM/MM studies on the glycosylation mechanism of rice BGlu1 beta-glucosidase, J. Mol. Graph. Model. 30 (2011) 148–152.
[17] A. Warshel, M. Karplus, Calculation of ground and excited state potential surfaces of conjugated molecules. I. Formulation and parametrization, J. Am. Chem. Soc. 94 (1972) 5612–5625.
[18] A. Warshel, M. Levitt, Theoretical studies of enzymatic reactions: dielectric electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme, J. Mol. Biol. 103 (1976) 227–249.
[19] M.J. Field, P.A. Bash, M. Karplus, A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations, J. Comput. Chem. 11 (1990) 700–733.
[20] J.L. Gao, D.G. Truhlar, Quantum mechanical methods for enzyme kinetics, Annu. Rev. Phys. Chem. 53 (2002) 467–505.
[21] H.M. Senn, W. Thiel, QM/MM methods for biomolecular systems, Angew. Chem. Int. Ed. 48 (2009) 1198–1229.
[22] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.
[23] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari,

J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, Gaussian 03, Revision C. 02, Gaussian, Inc., Wallingford CT, 2004.

[24] J. Gasteiger, M. Marsili, Iterative partial equalization of orbital electronegativity—a rapid access to atomic charges, Tetrahedron 36 (1980) 3219–3228.

[25] Z. Fujimoto, S. Kaneko, M. Momma, H. Kobayashi, H. Mizuno, Crystal structure of rice alpha-galactosidase complexed with D-galactose, J. Biol. Chem. 278 (2003) 20313–20318.

[26] W. Humphrey, A. Dalke, K. Schulten, VMD—visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38.

[27] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.

[28] A.D.J. MacKerell, D. Bashford, M. Bellot, R.L. Dunbrack Jr., J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F.T.K. Lau, C. Mattos, S. Michnick, T. Ngo, D.T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J.C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, All-atom empirical potential for molecular modeling and dynamics studies of proteins, J. Phys. Chem. B 102 (1998) 3586–3616.

[29] R. Ahlrichs, M. Bar, M. Haser, H. Horn, C. Kolmel, Electronic structure calculations on workstation computers: the program system TURBOMOLE, Chem. Phys. Lett. 162 (1989) 165–169.

[30] W. Smith, T.R. Forester, DL_POLY_2.0: a general-purpose parallel molecular dynamics simulation package, J. Mol. Graph. Model. 14 (1996) 136–141.

[31] D. Bakowies, W. Thiel, Hybrid models for combined quantum mechanical and molecular mechanical approaches, J. Phys. Chem. 100 (1996) 10580–10594.

[32] A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto, G.J. Kramer, Zeolite structure and reactivity by combined quantum-chemical–classical calculations, J. Phys. Chem. B 103 (1999) 6133–6141.

[33] P. Sherwood, A.H. de Vries, M.F. Guest, G. Schreckenbach, C.R.A. Catlow, S.A. French, A.A. Sokol, S.T. Bromley, W. Thiel, A.J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S.C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, C. Lennartz, QUASI: a general purpose implementation of the QM/MM approach and its application to problems in catalysis, J. Mol. Struct. (THEOCHEM) 632 (2003) 1–28.

[34] S.R. Billeter, A.J. Turner, W. Thiel, Linear scaling geometry optimisation and transition state search in hybrid delocalised internal coordinates, Phys. Chem. Chem. Phys. 2 (2000) 2177–2186.

[35] J. Nocedal, Updating quasi-Newton matrices with limited storage, Math. Comput. 35 (1980) 773–782.

[36] D.C. Liu, J. Nocedal, On the limited memory BFGS method for large scale optimization, Math. Program. 45 (1989) 503–528.

[37] A. Banerjee, N. Adams, J. Simons, R. Shepard, Search for stationary points on surfaces, J. Phys. Chem. 89 (1985) 52–57.

[38] J. Baker, An algorithm for the location of transition states, J. Comput. Chem. 7 (1986) 385–395.