

Article

Systematic Functional and Computational Analysis of Glucose-Binding Residues in Glycoside Hydrolase Family GH116

Meng Huang^{1,2}, Salila Pengthaisong^{1,2}, Ratana Charoenwattanasatien^{2,3}, Natechanok Thinkumrob⁴, Jitrayut Jitnom^{4,*} and James R. Ketudat Cairns^{1,2,*}

- ¹ School of Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; baggiohm@gmail.com (M.H.); mai.salila@yahoo.com (S.P.)
² Center for Biomolecular Structure, Function and Application, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; ratana@slri.or.th
³ Research Facility Department, Synchrotron Light Research Institute (Public Organization), Nakhon Ratchasima 30000, Thailand
⁴ Unit of Excellence in Computational Molecular Science and Catalysis, and Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand; byunbaw13@gmail.com
* Correspondence: jitrayut.ji@up.ac.th or jitrayut.018@gmail.com (J.J.); cairns@sut.ac.th (J.R.K.C.)

Abstract: Glycoside hydrolases (GH) bind tightly to the sugar moiety at the glycosidic bond being hydrolyzed to stabilize its transition state conformation. We endeavored to assess the importance of glucose-binding residues in GH family 116 (GH116) β -glucosidases, which include human β -glucosylceramidase 2 (GBA2), by mutagenesis followed by kinetic characterization, X-ray crystallography, and ONIOM calculations on *Thermoanaerobacterium xylanolyticum* TxGH116, the structural model for GH116 enzymes. Mutations of residues that bind at the glucose C3OH and C4OH caused 27–196-fold increases in K_M for *p*-nitrophenyl- β -D-glucoside, and significant decreases in the k_{cat} up to 5000-fold. At the C6OH binding residues, mutations of E777 decreased the k_{cat}/K_M by over 60,000-fold, while R786 mutants increased both the K_M (40-fold) and k_{cat} (2–4-fold). The crystal structures of R786A and R786K suggested a larger entrance to the active site could facilitate their faster rates. ONIOM binding energy calculations identified D452, H507, E777, and R786, along with the catalytic residues E441 and D593, as strong electrostatic contributors to glucose binding with predicted interaction energies $> 15 \text{ kcal mol}^{-1}$, consistent with the effects of the D452, H507, E777 and R786 mutations on enzyme kinetics. The relative importance of GH116 active site residues in substrate binding and catalysis identified in this work improves the prospects for the design of inhibitors for GBA2 and the engineering of GH116 enzymes for hydrolytic and synthetic applications.

Keywords: β -glucosidase; glucose-binding residues; mutagenesis; enzyme kinetics; X-ray crystallography; ONIOM



Citation: Huang, M.; Pengthaisong, S.; Charoenwattanasatien, R.; Thinkumrob, N.; Jitnom, J.; Ketudat Cairns, J.R. Systematic Functional and Computational Analysis of Glucose-Binding Residues in Glycoside Hydrolase Family GH116. *Catalysts* **2022**, *12*, 343. <https://doi.org/10.3390/catal12030343>

Academic Editors: Yane-Shih Wang, Rung-Yi Lai and Yan-Jiun Lee

Received: 15 February 2022

Accepted: 14 March 2022

Published: 17 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Enzymes are particularly malleable catalysts due to their ability to change specific functional groups and the catalytic environment by the genetic engineering of the amino acid sequence. An enzyme's catalytic reaction often occurs in a buried groove or pocket-like active site, where the functional groups of the surrounding amino acid residues build a unique environment that promotes the highly specific and efficient catalysis of reactions. In the active site and the surrounding region outside this site, a relatively small number of specific amino acid residues are involved in substrate binding, and few of these act directly in the catalytic reaction [1]. The roles of amino acid residues during catalysis have been delegated into seven basic classes—stabilization, steric roles, activation, proton shuttling, hydrogen shuttling, electron shuttling residues, and covalent catalysis [2].

β -Glucosidases (β -D-glucopyranoside glucohydrolases, E.C. 3.2.1.21) are enzymes that release the nonreducing terminal β -D-glucosyl residue from glucoconjugates, including glucosides, 1-O-glucosyl esters, and oligosaccharides, by glycosidic bond hydrolysis [3]. β -Glucosidases have been categorized into the protein sequence-based glycoside hydrolase families GH1, GH2, GH3, GH5, GH16, GH30, GH39, and GH116 (<http://www.cazy.org>, accessed on 14 February 2022) [4], and the structure and catalytic mechanism are generally conserved within each family. Their specificity toward different substrates varies, depending on the enzyme and its biological function.

Hydrolysis of the glycosidic bond can occur with either inversion or retention of the anomeric configuration, and this property is generally conserved within one enzyme family [5,6]. Most characterized β -glucosidases catalyze hydrolysis via a two-step retaining mechanism, in which the first step entails the departure of the aglycone and the formation of a covalent intermediate with the enzymatic nucleophile residue, while the second step is a reversal of this process, with water serving as the enzyme-displacing nucleophile [7–9]. A catalytic acid/base residue serves as an acid to protonate the glycosidic oxygen as the aglycone departs in the enzyme glycosylation step and as a base to extract a proton from water in the enzyme deglycosylation step. In each step, the glucose is thought to pass through an oxocarbenium cation at or near the transition state, which requires distortion to a conformation that allows sp^2 hybridization at C1 and O5, typically a 4H_3 half-chair or closely related 4E envelope in β -glucosidases [10]. In most β -glucosidases, the formation of the transition state shape appears to be primed by the distortion of the pyranose ring to a 1S_3 conformation. This induced fit of the substrate implies that the surrounding residues, which generally form tight hydrogen bonds with the hydroxyl groups and aromatic-sugar stacking interactions with the ring hydrogens of the β -D-glucopyranose, are essential to the transition state formation in catalysis, rather than simply serving to bind the substrate. It has been noted that the electrostatic stabilization of the oxocarbenium ion-like transition state by nearby residues also contributes to catalysis [11,12]. Indeed, some previous studies have shown that the contributions of enzyme active-site residues not directly involved in bond making and breaking could be as essential to the enzyme's function as catalytic residues [13–16]. Nonetheless, systematic kinetic analysis of the contributions of the active-site residues of β -glucosidases is rarely carried out, due to the low activities of such mutants [17,18].

The enzyme used in this work, *Thermoanaerobacterium xylanolyticum* GH116 (*TxGH116*), is a thermostable β -glucosidase which consists of an N-terminal domain formed by a two-sheet β -sandwich and a C-terminal (α/α)₆ solenoid domain containing the slot-like active site [17]. As the first family GH116 enzyme to have its X-ray crystal structure determined, *TxGH116* has helped to elucidate the molecular basis of pathogenic mutations in the human GH116 member, GBA2 glucosylceramidase, that lead to hereditary ataxias and paraplegias [19,20]. Aside from its mutations resulting in these rare genetic disorders, the over- and under-production of GBA2 have been shown to affect cellular sphingolipid balance [21–23] and the depletion of GBA2 has also been shown to ameliorate some of the effects of loss of lysosomal glucosylcerebrosidase (GBA) activity, suggesting the inhibitors of GBA2 may have therapeutic applications [24–26]. High-resolution crystal structures of *TxGH116* in complex with glucose and inhibitors allow for the identification of the active-site residues involved in substrate binding and catalysis, which are completely conserved with human GBA2 [17,27–29]. Aside from its role as a model structure for human GBA2, mutants of the *TxGH116* enzyme have been used to efficiently generate glucosylazides for the production of α - and β -glucosidase inhibitors, which have potential applications in treating human diseases [30,31].

Our previous work verified that *TxGH116* utilizes a retaining mechanism by identifying the initial hydrolysis product as β -D-glucose by monitoring the reaction time course by NMR spectroscopy [17]. The identities of E441 as the catalytic nucleophile and D593 as the catalytic acid/base were proven by the mutation of these residues to alanine, followed by chemical rescue using small nucleophiles. *TxGH116* is thought to catalyze hydrolysis

via the usual β -glucosidase conformational itinerary from 4C_1 to 1S_3 to 4H_3 to 4C_1 in the glycosylation step, but no distortion of the pyranose ring has been observed for glucose and glucose-like inhibitors in the active site complexes. Glucoimidazole binds in the transition state-like 4H_3 conformation, but this is the low energy conformation for this inhibitor, so it does not provide strong evidence for distortion toward this transition state conformation during catalysis [17,32]. The yet-to-be-verified transition state and the unusual orientation of the catalytic acid/base compared to β -glucosidases from other families make the roles of GH116 active-site residues particularly worthy of investigation.

Although the $TxGH116$ catalytic acid/base and nucleophile residues were analyzed [17], the importance and roles of other residues within the $TxGH116$ and GH116 family active-site pocket remain to be clarified. In this work, we systematically mutated amino acids that form the remainder of the active-site residues interacting with the glucose ligand characterized the kinetics of the mutants to quantify the contribution of each side chain toward the catalytic efficiency of $TxGH116$. We solved the structures of mutations that increased the turnover number of the reaction (k_{cat}) to show the influence of active-site remodeling. Moreover, the theoretical energy contributions of the individual residues to glucose binding and the energetic effects of their mutations were calculated to further elucidate their roles. The insight into the relative importance of active site residues in substrate and catalysis facilitates the modification of $TxGH116$ β -glucosidase for improved application, the design of human GBA2 inhibitors, and a better understanding of human disorders caused by mutations in GBA2.

2. Results

2.1. Analysis of Glucose-Binding Residues

Glycone-binding residues were identified in the complex of $TxGH116$ wild type with glucose (PDB: 5BX5) [17], and show high conservation in GH116 family enzymes from different species in Figure 1a,b. The polar interactions between these residues and glucose include H507 with C3OH, D452 with C3OH and C4OH, T591 with C4OH, and E777 and R786 with C6OH. In addition, E730 interacts with R786 and W732, which makes a hydrophobic interaction with C6 in Figures 1a and S1. The catalytic nucleophile, E441, also makes a hydrogen bond with the glucose C2OH, but this is difficult to separate from its catalytic function, and the mutation of this residue was previously studied [17,33]. We systematically mutated for the glucose-binding residues mentioned above and characterized the kinetics of the mutants to elucidate their relative importance and specific roles.

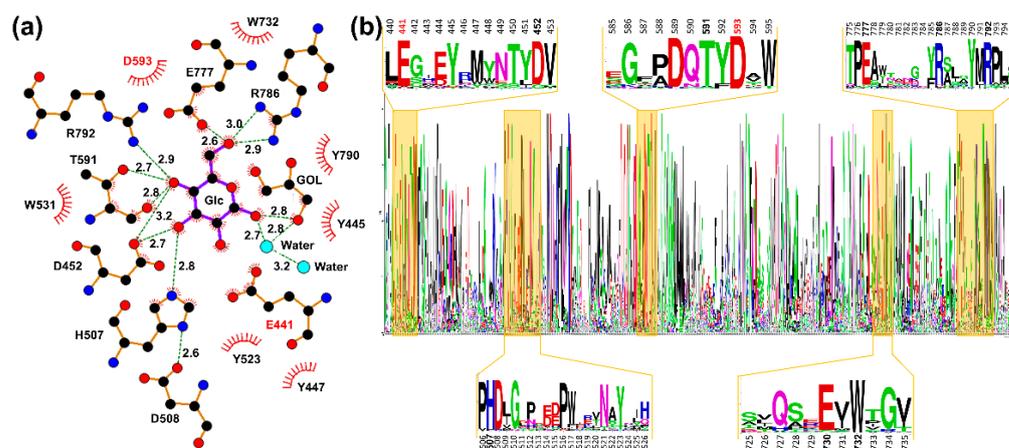


Figure 1. $TxGH116$ glucose-binding residues. (a) Binding sites of $TxGH116$ wild type structural model with glucose ligand. The ligands and protein side chains from wild type $TxGH116$ bound to glucose (PDB: 5BX5) [17] are shown in a ball-and-stick representation, with the ligand bonds colored in purple and protein bonds in brown. Hydrogen bonds are shown as green dotted lines,

while the spoked arcs represent protein residues making nonbonded contacts with the ligand. Water molecules located in the active site are shown as cyan circles. The 2D ligand–protein interaction diagram was made by LigPlot+ [34]. (b) Sequence logo of 34 diverse sequences that are from the GH116 family in different species. The conserved regions around the active site are highlighted, glycone sugar-binding residues' positions in the sequence are shown in bold, and catalytic residues are labeled in red bold print. Amino acid conservation analysis was performed to assess the conserved residues in GH116 enzyme family members.

2.2. Effects of Glucose-Binding Residue Mutations on the Activity of TxGH116

Previously, mutations of D508, which interacts with H507, were shown to cause drastic decreases in TxGH116 activity [17]. The H507A, H507E, and H507Q mutants had similar pH optima at pH 5.0, slightly lower than the wild type at pH 5.5 (Figure 2a), and all showed lower temperature optima, as shown in Figure 2b, suggesting the lower stability of the enzyme or enzyme-substrate complex. The H507 mutations increased the K_M value by 58 to 196-fold compared to the wild type, as seen in Figure 3. While H507A and H507Q decreased the k_{cat} value by less than 10- and 3-fold, H507E decreased the k_{cat} by over 4000-fold and the k_{cat}/K_M value by over 200,000-fold, indicating that the negative charge at this residue is not tolerated.

The interactions of D452 with the glucose C3OH and C4OH are important for catalysis since both D452 mutations, D452A and D452N, greatly decreased activity. The D452A and D452N mutations led to 170- and 70-fold increases in the K_M values, respectively, and they decreased the k_{cat} values by more than 340-fold and 5000-fold, which led to reductions in the k_{cat}/K_M values of more than 56,000- and 352,000-fold (Figure 3). These results are consistent with the much weaker binding of the sugar of the substrate and transition state expected based on this residue's interaction with the glucose C3OH and C4OH (Figures 1a and S1).

In the complex structure, R792 appeared to have hydrogen bond interactions with the Glc C4OH and its neighboring E777 residue (Figures 1a and S1). The R792 residue mutations R792A and R792K had little effect on the pH dependence, as they had broad working pH profiles and pH optima comparable to the wild type (Figure 2a). The R792A and R792K mutants showed similar K_M values that were >25 times higher than the wild type and very low k_{cat} values 5000- and 780-fold lower than the wild type, respectively (Figure 3).

T591 is on the flexible loop containing the catalytic acid/base D593 and forms hydrogen bonds with the Glc C4OH via its backbone carbonyl and sidechain hydroxyl group, which also interacts with D452, as seen in Figures 1a and S1. The T591A mutation eliminated the sidechain interactions, resulting in a 70-fold increase in the K_M and a 10-fold reduction in the k_{cat} value compared to the wild type (Figure 3). At its optimum pH of 4.0, T591A (Figure 2a) had significantly higher activity than at pH 5.5, the pH optimum of the wild type enzyme.

W732 appears to form a CH– π interaction with Glc C6H2 (Figure S1). Although it does not form a hydrogen bond to the sugar, its indole NH group could form one with E730 (Figure S1). The W732F variant had a nearly 3-fold higher K_M as well as a 1.5-fold increase in the k_{cat} value compared to the wild type TxGH116 (Figure 3).

E777 forms a hydrogen bond to the Glc C6OH and also interacts with R792 (Figure 1a). Both E777A and E777Q showed very low activity compared to the wild type (Figure 3). E777A exhibited a 170-fold increased K_M , while E777Q showed only a 40-fold increase compared to the wild type. Both mutants showed very low efficiency, with a reduction around 62,000-fold in the k_{cat}/K_M value for E777A and 87,000-fold for E777Q. E777Q shifted the optimum pH to pH 3.5, as shown in Figure 2a. This suggests that the charge on E777 influenced the pH optimum and was important for catalysis, along with its hydrogen bonding to the C6OH group, which could be maintained in E777Q.

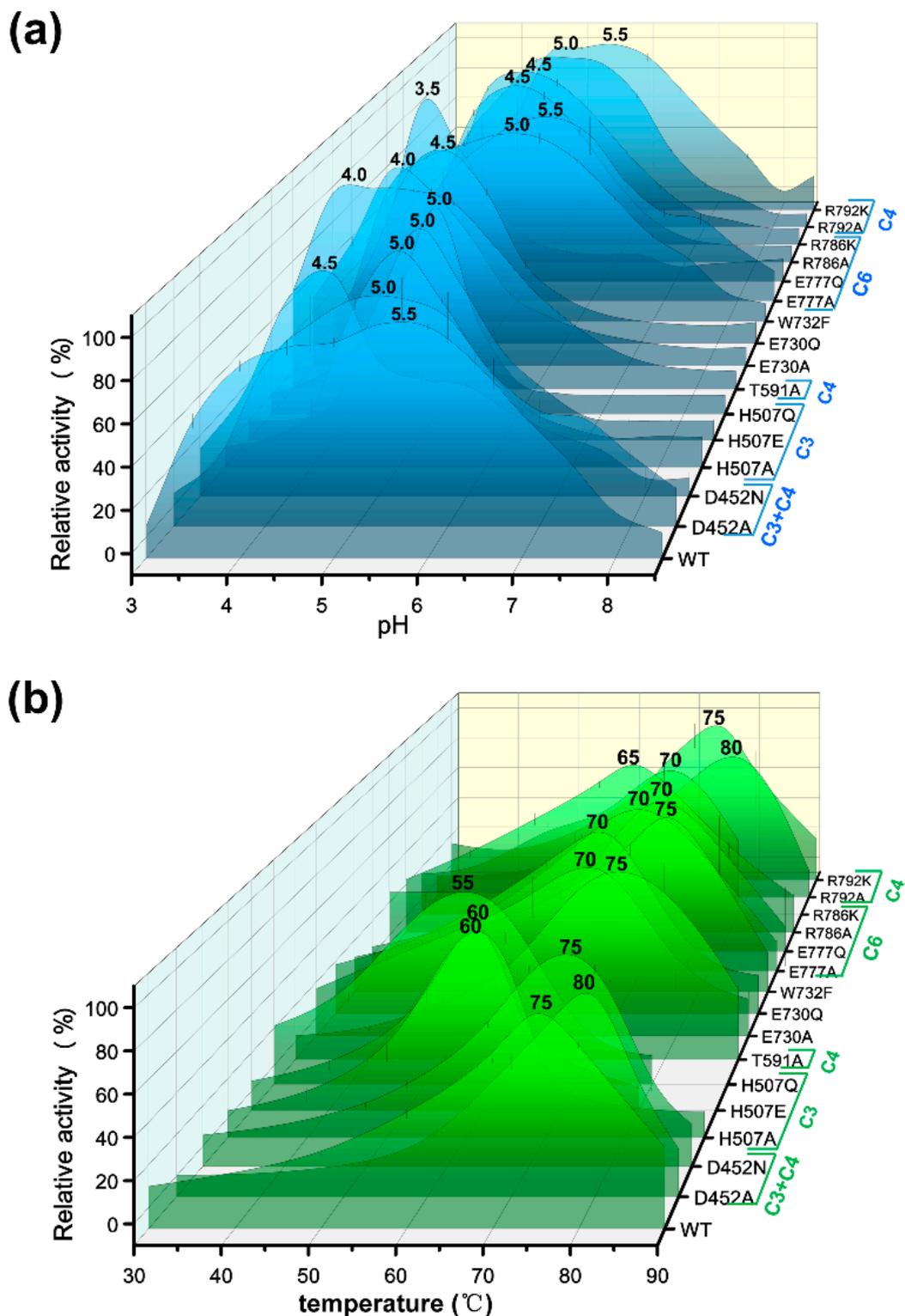


Figure 2. Optimal pH (a) and temperature (b) analysis of mutants and wild type (WT) *TxGH116*. All assays were carried out in triplicate. The 3-D stacked graph was made by Origin 8.0 with Bézier curves. C3, C4, C6, and C3 + C4 represent the glucose carbons carrying hydroxyls interacting with the side chains of the amino acid residues that were mutated.

Protein	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (mM ⁻¹ s ⁻¹)
WT	0.170 ± 0.014	41.3 ± 0.63	243
D452A	28.3 ± 1.3	0.120 ± 0.0021	0.0043
D452N	11.7 ± 0.70	0.00810 ± 0.00010	0.00069
H507A	23.0 ± 0.71	4.72 ± 0.051	0.21
H507E	9.93 ± 0.62	0.0100 ± 0.00020	0.0010
H507Q	33.3 ± 4.3	12.2 ± 0.63	0.37
T591A	12.6 ± 0.45	3.63 ± 0.037	0.29
E730A	3.79 ± 0.25	80.8 ± 1.4	21.3
E730Q	7.21 ± 0.49	55.6 ± 1.1	7.72
W732F	0.470 ± 0.010	62.3 ± 0.39	133
E777A	29.9 ± 0.63	0.120 ± 0.0010	0.0039
E777Q	6.97 ± 0.22	0.0200 ± 0.00020	0.0028
R786A	7.89 ± 0.57	171 ± 4.0	21.7
R786K	7.30 ± 0.71	88.9 ± 2.7	12.2
R792A	4.53 ± 0.38	0.00820 ± 0.00010	0.0018
R792K	5.32 ± 0.44	0.0530 ± 0.0013	0.0099

Figure 3. Kinetic constants of the TxGH116 wild type (WT) enzyme and mutants with *p*NPGlc substrate. The heat map represents the fold changes in the apparent k_{cat} , K_M , and k_{cat}/K_M values of each mutant relative to the WT enzyme. Red represents different fold increases and blue represents different fold decreases (from light to dark). Values in white boxes are within a factor of 3 of those for WT.

R786 forms two hydrogen bonds to the Glc C6OH in the -1 subsite and one to the phenolic OH of Y445, in addition to one to E730 (Figures 1a and S1). The R786A mutation eliminated the hydrogen bonds, while R786K diminished the number possible. Both mutants displayed lower temperature optima (Figure 2b), suggesting lower thermostability, especially R786A. These mutants have broad pH profiles, like the wild type, but their pH optimum shifted to pH 4.5 (Figure 2a). Their K_M values were 43–46-fold higher than the wild type, but R786A had a 4-fold and R786K a 2-fold higher k_{cat} value compared to the wild type (Figure 3). At its pH optimum of 4.5, R786K also had a lower k_{cat} (144 s⁻¹) compared to the R786A (171 s⁻¹) (Figures 3 and S4).

E730 interacts with both the W732 and R786 residues, which in turn make direct interactions with the glucose molecule (Figures 1a and S1). Our E730A and E730Q mutants had their temperature optima decrease 5 degrees compared to the wild type and also lower pH optima at pH 4.0 and 4.5, respectively (Figure 2). The E730A mutant should eliminate interactions with W732 and R786, while the E730Q mutant should maintain hydrogen bonding but lose the coulombic component of those interactions. Interestingly, the K_M value for *p*NPGlc increased 20-fold, and its k_{cat} value increased 2-fold in E730A compared to the wild type, while in E730Q, the K_M value increased 40-fold and the k_{cat} value increased 1.5-fold (Figure 3). Given the postulated ability of E730Q to maintain hydrogen bonds with W732 and/or R786, its higher K_M and lower k_{cat} compared to E730A were unexpected.

Because the R786 and E730 mutants were less severely impaired than other mutants, they could be tested with the natural substrate cellobiose, which is less reactive than *p*NPGlc. R786K had lower activity with cellobiose compared to R786A and wild type TxGH116 (Figure 4). R786K and R786A had similar K_M values for the hydrolysis of cellobiose, which

were about 5-fold higher than the wild type, but the k_{cat} of R786K was significantly lower than those of the R786A mutant and wild type TxGH116. As such, the specificity constant (k_{cat}/K_M) of R786K was nearly 2-fold lower than R786A and over 10-fold lower than the wild type enzyme. Thus, the R786K mutation, which was expected to be milder, had a greater defect than R786A. The E730 mutations' effects on the K_M were similar to those of the R786 mutants, with a 4-fold higher K_M for E730A and a nearly 7-fold increase for E730Q. However, these mutants had lower apparent k_{cat} values than the wild type and the R786 mutants, resulting in a 100-fold decrease in the specificity constant (k_{cat}/K_M) for E730A and a 20-fold decrease for E730Q. Thus, the effects of mutations of the E730 residue, which indirectly acts in glucose binding, were stronger than those of mutation of the R786, which directly binds to glucose.

Protein	K_M (mM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (mM ⁻¹ s ⁻¹)
WT	0.360 ± 0.025	48.0 ± 0.40	133
E730A	1.48 ± 0.064	1.99 ± 0.015	1.34
E730Q	2.39 ± 0.12	14.9 ± 0.24	6.21
W732F	0.700 ± 0.036	61.9 ± 0.60	88.4
R786A	1.90 ± 0.17	40.8 ± 1.1	21.5
R786K	2.09 ± 0.057	25.7 ± 0.41	12.3

Figure 4. Kinetic constants of the TxGH116 wild type (WT) enzyme and mutants with cellobiose substrate. The heat map represents the fold changes in each k_{cat} , K_M , and k_{cat}/K_M mutant relative to the WT enzyme. Red represents different fold increases, and blue represents different fold decreases (from light to dark with increasing fold change).

2.3. Structural Investigation of R786 and E730 Mutants

In order to understand the effects of the TxGH116 R786 and E730 mutations, we determined their structures in complexes with glucose. The E730A, R786A, and R786K variant crystals were isomorphous with the original TxGH116 crystals in the $P2_12_12_1$ space group with a single protein molecule in the asymmetric unit [17], while the E730Q variant crystallized in the $P2_12_12_1$ space group with two protein molecules in the asymmetric unit (Table S3). All had clear electron density for glucose molecules in the -1 subsite, as shown in Figure 5a,c. The active sites of the R786 variants were compared to that of wild type TxGH116 in Figure 5. In the crystal without ligand, instead of being in a position to interact with the Glc C6OH, E730, or Y445, as in the wild type, the lysine side chain in R786K was found to turn away, where it displaced the neighboring loop (Figure S10). In the crystal soaked with glucose, about 40% of the density for K786 moved to a position to interact with the Glc C6OH, while the remaining 60% turned away from the sugar toward the outside of the active site, as in the crystal without glucose (Figure 5d,e). The flexible loop in the entrance to the active site also showed a different rotation in the R786K mutant that appeared to make room for the K786 side chain (see Figure 5c–e). The X-ray crystallographic structures of TxGH116 R786A and R786K displayed an enlarged mouth to the active site, as shown in Figures 5f–h and S9, which may have affected the substrate and product movement in and out of the active site (Figure S3). In the E730A mutant structure, the other active site residues and glucose ligand were positioned very similarly to those

in wild type *TxGH116*, while the E730Q variant had changes in the position of the R786 sidechain, which was rotated away from the Q730 residue, and the Y445 side chain also moved, as seen in Figure 6. This change placed the R786 guanidinyll group too close to the glucose C6OH group, with a distance of 2.6 Å in the E730Q mutant, and appeared to cause the displacement of the Y445 side chain in the +1 subsite (Figures 6 and S10).

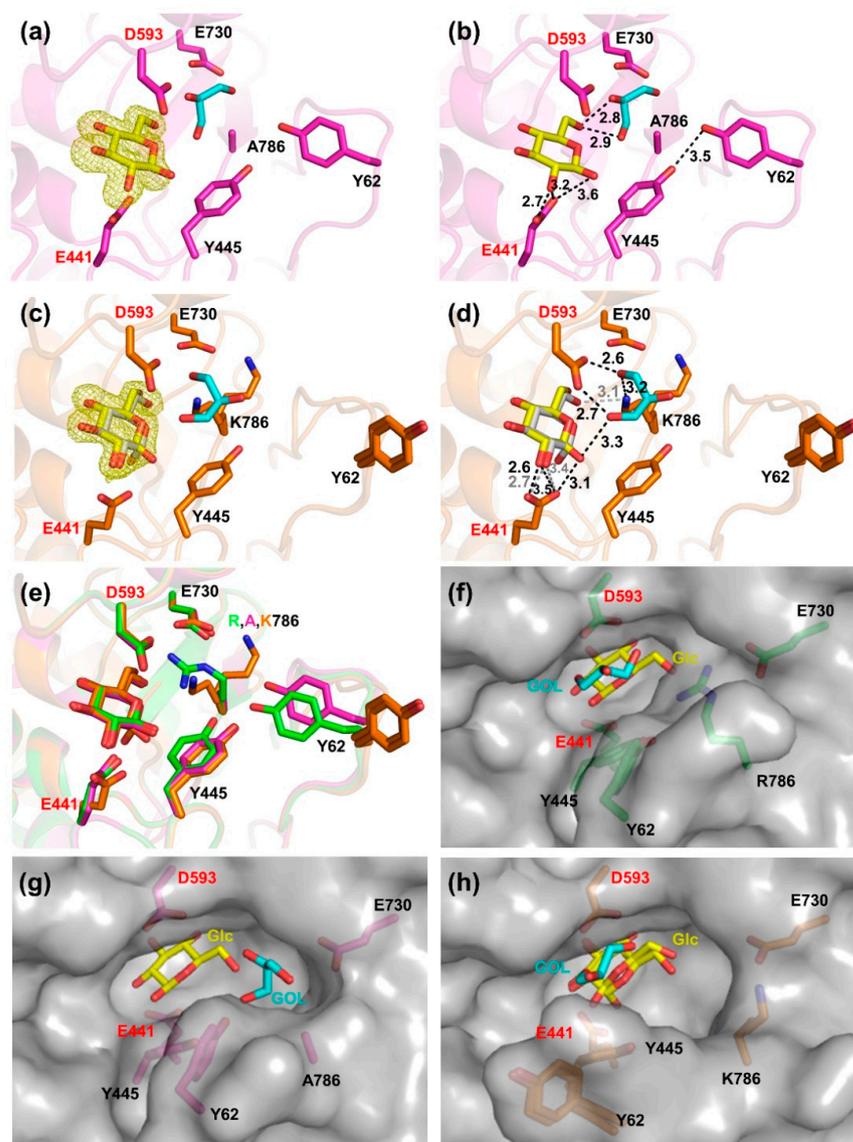


Figure 5. Active site comparison between *TxGH116* R786A and R786K and the wild type (WT) with glucose ligand. The glucose ligand weighted $|F_O - F_C|$ electron density maps calculated with the ligands omitted and contoured at 3σ are shown in the active sites of the glucose complexes of R786A (a) and R786K (c). Active site interactions with glucose and glycerol molecules are shown for R786A (b) and R786K (d), in which, with two anomers, glucose ligands were found (alpha 80%, in grey and beta 20%, in yellow). (e) Superposition of active site residues of WT (PDB: 5BX5, green [17]), R786A (magenta), and R786K (orange). Slot-like active site entrance comparison between WT (f), R786A (g) and R786K (h). R786A and R786K have an enlarged entrance space, with an extra glycerol molecule in the place of the original arginine side chain. The potential hydrogen bonds between the glucose (Glc) ligands (yellow sticks) and surrounding residues are displayed as black dashed lines. Glycerol (GOL) molecules are shown as cyan sticks, and catalytic residues are labeled in red.

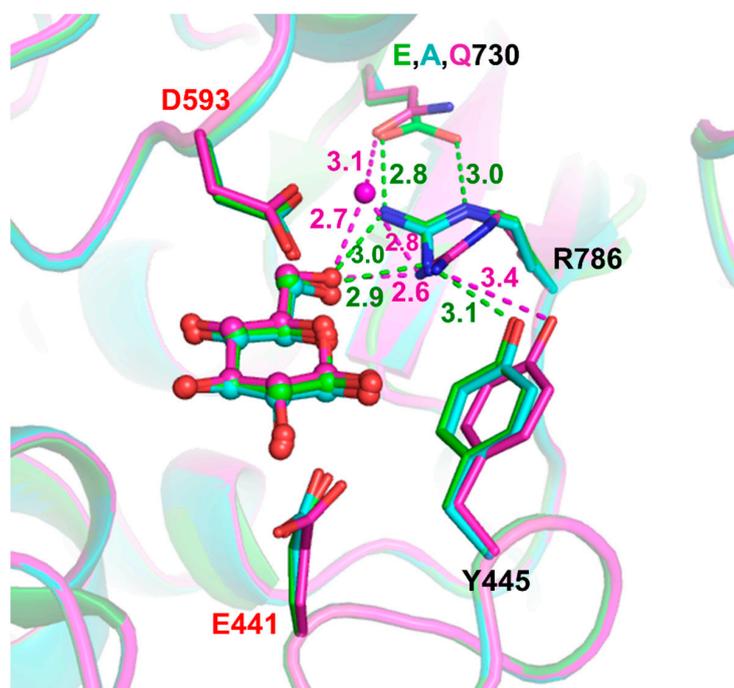


Figure 6. Structures of E730A (cyan) and E730Q (magenta) with glucose ligand. The glucose ligand molecules are positioned similar to that in the wild type (green, PDB: 5BX5 [17]). One water molecule that takes the position near that vacated by the R786 residue guanidiny group in E730Q mutant is shown as a magenta ball. Hydrogen bonds and close polar contacts are shown as dashed lines with distances marked (green for wild type and magenta for E730Q).

2.4. Inhibition of *p*NPGlc Hydrolysis by Glucose and Cellobiose

To investigate whether the more open active site and weaker binding of glucose would lead to lower product and substrate inhibition, we characterized the inhibition by glucose and cellobiose. Both glucose and cellobiose displayed competitive inhibition of *TxGH116* hydrolysis of *p*NPGlc, as seen in Table 1 and Figures S6 and S7, with the wild type and R786K appearing more sensitive than *TxGH116* R786A. Cellobiose gives K_i values of 0.94 mM for the wild type, 17 mM for R786A, and 8.3 mM for R786K, which are higher than their respective K_M values for cellobiose in Figure 4, especially for the mutants. This behavior is more complex than simply acting as a competitive substrate and led us to investigate the transglycosylation of cellobiose. Transglycosylation products of both cellobiose and *p*NPGlc were observed during the reactions (see Figure S8).

Table 1. Glucose and cellobiose inhibition kinetics.

Protein	Glucose K_i (mM)	Cellobiose K_i (mM)
WT	4.5 ¹	0.94
R786A	6.5	16.9
R786K	3.4	8.3
E730A	19.0	N.D. ²
E730Q	4.4	N.D.

¹ Value from [17]. ² N.D.—not determined.

2.5. Effect of R786 Mutations on Sugar Specificity

Since R786 interacts with the glucose C6OH group, we tested whether loss of this interaction would affect the sugar specificity, especially for β -D-glucopyranoside versus β -D-xylopyranoside, which lacks this hydroxymethyl group. We compared the activity of the C6OH-related mutants to the wild type on *p*NPGlc, *p*NP- β -D-xylopyranoside, *p*NP- β -D-galactopyranoside, and *p*NP- β -cellobioside. All the mutants had reduced activity

on *p*NP-Glc compared to the wild type, but with *p*NP-xylopyranoside, R786A and R786K showed high hydrolysis activity compared to the wild type, as shown in Figure 7.

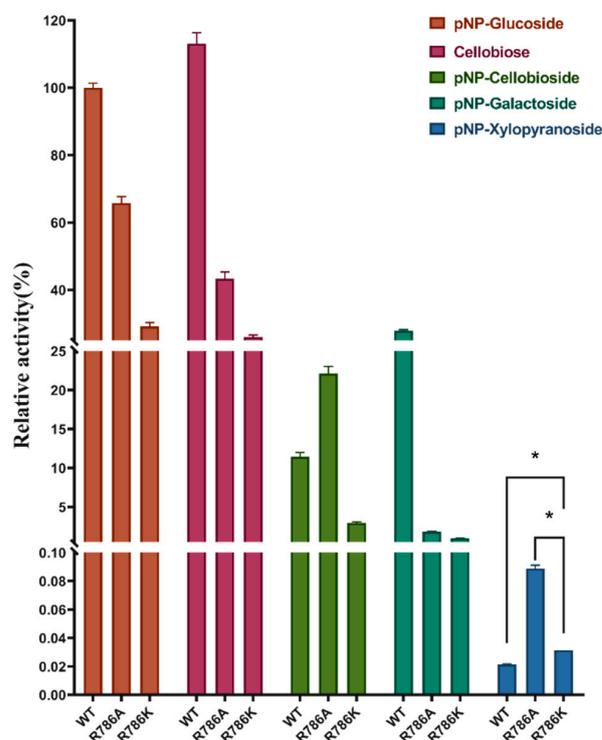


Figure 7. Relative activity on different substrates, *p*NP- β -D-glucopyranoside, cellobiose, *p*NP- β -D-cellobioside, *p*NP- β -D-galactopyranoside, and *p*NP- β -D-xylopyranoside. The assay contained 1 mM substrate in 50 mM sodium acetate buffer, pH 5.5, at 60 °C. Relative activity is based on *p*NP released from *p*NP- β -glycosides or glucose release from cellobiose. Data are expressed as means of three independent reactions \pm SD. Stars represent one-way ANOVA significance at $p < 0.05$.

All mutants had very low activity on *p*NP- β -D-galactoside, which differed from the glucose in the chirality of the C4OH group, consistent with R786 not interacting with this group; however, it was previously noted that C6 sidechain binding is different between β -glucosidases and β -galactosidases [35]. Surprisingly, the R786A mutation increased the rate of hydrolysis of *p*NP- β -cellobioside, which was expected to be hydrolyzed in a two-step process beginning with the removal of the nonreducing glucosyl residue, followed by the release of the detected *p*NP from the remaining glucose.

2.6. Computational Analysis of Glucose-Binding Residues

To elucidate the mutation-induced effect on glucose-binding residues, particular interaction analysis, which describes the contribution of each residue in the total binding energy, was conducted for wild type and 10 mutant models (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, R786K), based on the ONIOM method [36–38]. The ONIOM-optimized structure of wild type *Tx*GH116-glucose complex and the particular interaction energy (IE) between glucose and the *Tx*GH116 binding site for the wild type and mutants are illustrated in Figures 8 and S11–S14. The BE values of all complexes are plotted in Figure S12, which are in the ranges of -94.8 to -131.7 kcal mol $^{-1}$ for the mutants and -134.1 kcal mol $^{-1}$ for the wild type. This result clearly shows that all mutations decreased the binding affinity of glucose to the *Tx*GH116 active site, suggesting the important roles of the *Tx*GH116 binding residues.

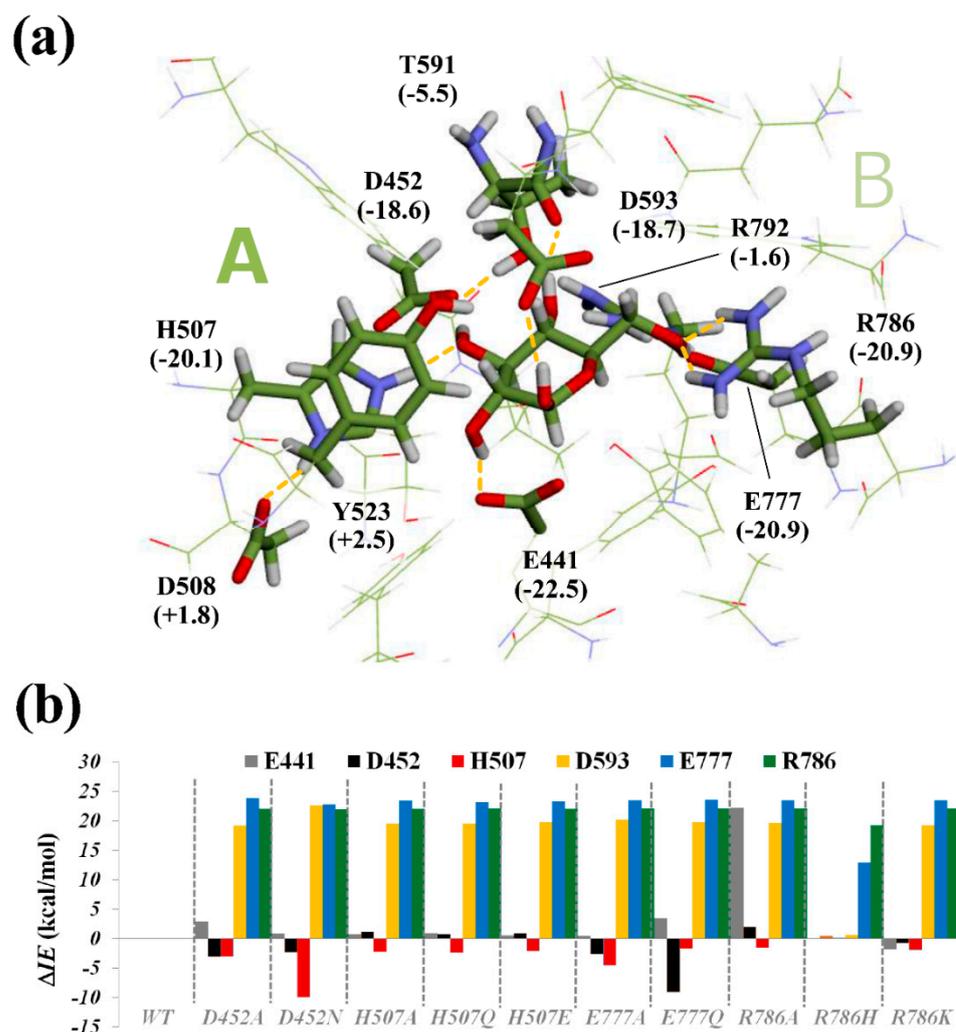


Figure 8. ONIOM2-optimized structure of wild type (WT) TxGH116-glucose complex (a) and change in particular interaction energy (ΔIE) between glucose and six important residues for different mutations (b). The residues analyzed were those with $|IE| > 15 \text{ kcal mol}^{-1}$ in Figure S13, and the mutations analyzed were D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K (b). Geometries were obtained at ONIOM2(B3LYP/6-31G(d,p):PM3) level of theory. All energies in the bar plot for the mutants are relative to the IE values of WT (shown in parentheses in (a) in kcal mol^{-1}). The energies were calculated at the B3LYP/6-31G(d,p) level with BSSE-CP. Negative and positive ΔIE values mean the particular residue stabilizes and destabilizes the glucose binding, respectively. The hydrogen bonds are displayed as orange dash lines.

The particular interaction energies were calculated to identify the important residues that contributed to the binding of glucose. They can also be used as a simple and qualitative indicator for understanding the mutation effect on the electronic structure and protein–ligand interaction. The interaction energies for the individual residues of both the wild type and mutants are shown in Figure S13. It was calculated that six residues, E441 (nucleophile), D452, H507, D593 (acid/base), E777, and R786, are major (electrostatic) contributors for glucose binding to TxGH116, as evidenced by the large negative IE values ($|IE| > 15 \text{ kcal mol}^{-1}$). Further protein–ligand analysis illustrated in Figure S11 also showed that the hydrogen bond and van der Waals interactions were the dominant stabilization factors of the non-bonded interactions between the glucose ligand and the TxGH116 active site.

The geometries of the wild type and the mutated residues (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K) and their relevant H-bond inter-

actions with the C2OH, C3OH, C4OH, or C6OH of the glucose are visualized in Figures S15–S18. Noticeably, the mutations caused significant changes in the binding energy contribution, for which TxGH116 D452 and E777 mutations reduced the favorable contribution of the residues E441, D593, E777, and R786, and increased the favorable contributions of D452 and H507 (Figures 8 and S13). Furthermore, the mutations also resulted in a reduction in the total ONIOM binding energy compared to the wild type, as seen in Figure S12, with D452N, E777Q, and R786K exhibiting a small effect ($\sim 2\text{--}7$ kcal mol $^{-1}$) compared to the alanine mutations ($\sim 10\text{--}40$ kcal mol $^{-1}$). These relative binding energy changes were consistent with the relative increases in the K_M values and decreases in the k_{cat}/K_M described above, in most cases (Figure S14).

3. Discussion

In this work, we performed a systemic functional analysis of TxGH116 active-site glycone sugar-binding residues in substrate binding and catalysis via mutational, kinetic, structural, and quantum mechanical analysis. Due to the low activities of the mutations of the sugar-binding residues, most studies have simply reported a lack of activity for such mutants, but the activities were high enough to detect with elevated enzyme concentrations. In each kinetic assessment, the substrate concentration was always greater than 40 times that of the enzyme so that the formation of the enzyme–substrate (ES) complex would not significantly deplete the substrate concentration. From enzyme–substrate kinetics and structural analysis, we clarified the roles of each amino acid residue. In general, simple binding interactions with the glycone of the substrate cannot fully explain the activity changes in the mutation of the glucose-binding residues, and their interactions, charge distribution, and the stabilization of the transition state appear to be more critical, based on larger effects on the k_{cat} than the K_M .

H507A removed the imidazole ring, thereby eliminating the interaction with the 3OH group, while in principle, H507Q and H507E could still interact. However, H507Q may have also lost the interaction with the 3OH and instead interacted with D508 since it has similar kinetic parameters to H507A. The greater effect of H507E suggests that the previously reported large effects of D508 mutants on TxGH116 catalysis [17] are likely due to the stabilization of a positive charge on H507 rather than an inductive charge transfer to produce a partial negative charge. This would balance the negative charge on D452, which also hydrogen-bonds to the glucose C3OH and the catalytic nucleophile, which bonds to the neighboring C2OH, as seen in Figures 1a and S1. The mutation of a corresponding conserved histidine residue in the GH1 *Sulfolobus solfataricus* β -glycosidase was found to interact with changes in the C2 and C4 positions of glucose, in addition to hydrogen bonding to the C3OH, although with milder effects on activity [39].

The interactions at the Glc C4OH made small to moderate contributions to glucose binding in the ONIOM calculations in Figure 8 but had relatively large effects on the enzyme kinetics. The hydrogen bonds of R792 to the glucose C4OH and E777 seen in Figures 1a and S1 appear to be essential to transition-state binding, based on the large decrease in the k_{cat}/K_M . The importance of guanidino group participation in hydrogen bond interactions with substrate glucosyl hydroxyl groups was previously noted in a GH1 β -glucosidase [39]. T591 also interacts with the C4OH group via its main chain carbonyl oxygen, and potentially through its side chain OH (see Figures 1a and S1). Although the side chain hydroxyl group oxygen is 2.7 Å from the D452 sidechain oxygen and 2.8 Å from Glc O4, ONIOM calculations suggested that it mainly interacts with D452 (Figure S15C). The 74-fold increase in the K_M in T591A implies this interaction is important in substrate binding, and the 10-fold drop in the k_{cat} in T591A suggests this interaction may also help stabilize the flexible loop containing the catalytic acid/base, D593, and/or the transition state conformation.

It was reported that β -glucosidases have one common phenomenon that the binding interactions at the glucose C3 and C4 positions individually contribute 3–10 kJ mol $^{-1}$ to each transition state, whereas contributions in the ground state are much weaker

(≤ 3 kJ mol⁻¹) [40]. However, our calculations in Figure 8 and Table S3 show larger contributions to the glucose binding of residues interacting at the C3OH, and compared to the structure with glucoimidazole (PDB: 5BX4 [17]), a putative transition-state mimic, showed no hydrogen bond lengths differing by more than 0.1 Å from those with glucose. Based on the change in the k_{cat}/K_M , which is related to the binding energy of the glycosylation half-reaction transition state [40], these mutations resulted in $\Delta\Delta G = -RT\ln[(k_{\text{cat}}/K_M)_{\text{mut}}/(k_{\text{cat}}/K_M)_{\text{WT}}]$ of 7.2 kcal mol⁻¹ for D452A and 4.7 kcal mol⁻¹ for H507A.

The interactions with E777 and R786 held the glucose C6H₂OH side chain in a *gauche,trans* (*gt*) conformation with respect to the C5-O5 and C5-C4 bonds, which is unusual for β -glucosidases, which usually bind in the *gauche,gauche* (*gg*) [35,41,42]. Although a *gg* conformation provides more stabilization to β -glucosidase reactions, this would place the O6 very close to the catalytic acid/base due to its unusual position above the pyranose ring. This would perturb the acid/base interaction with the glycosidic oxygen atom, thereby explaining the C6H₂OH side chain being held in the second-most stabilizing *gt* conformation. The large decrease in the hydrolytic rate in E777 mutations suggests that this positioning makes a strong contribution to catalysis, while the negative charge on E777 may also help stabilize the oxocarbenium ion transition state. In contrast, the mutation of W732 to F did not change the turnover rate but led to an increased K_M value compared to the wild type (see Figures 3 and 4), which suggests that the W732 CH- π interaction with C6 makes a small contribution to the substrate binding.

The elimination of R786 H bonds with the C6OH in the R786A and R786K mutations resulted in higher K_M and k_{cat} values for *p*NPGlc hydrolysis than the wild type, which are seen in Figures 3 and 4. This is consistent with our previous work in which the TxGH116 R786H mutation (corresponding to a pathogenic human GBA2 mutation) [43] showed a 90-fold increase in the K_M value, which led to a reduction in the k_{cat}/K_M of more than 20-fold [17]. For TxGH116 wild type and W732F, the turnover numbers for *p*NPGlc and cellobiose were nearly equivalent, suggesting that the deglycosylation step, which was the same with both substrates, is largely rate limiting, since $k_{\text{cat}} = k_2k_4/(k_2 + k_4)$ is equal to k_4 when $k_2 \gg k_4$ (see Figure S3 for the kinetic scheme) [17]. This was not the case for the E730A, E730Q, R786A, and R786K mutants, however, suggesting that the glycosylation step becomes at least partially rate limiting for at least one of these substrates in these variants. This is consistent with the downward shift in the pH optima of the E730 mutants seen in Figure 2a since the protonation of the catalytic acid that is necessary for the initial glycosylation step was decreased. Although the X-ray crystal structures of R786A and R786K had a similar active-site structure to that of the wild type, the active-site entrance or mouth was wider and the nucleophile residue, E441, was closer to the glycoside bond (Figure 5c,d), which may have also affected the relative rates of the glycosylation and deglycosylation half-reactions.

Despite the loss of contacts, the glucose competitively inhibited R786A and R786K, with K_i values similar to the wild type (Table 1). Both glucose and cellobiose provided competitive inhibition of the hydrolysis *p*NPGlc substrate, as seen in Figures S6 and S7. Remarkably, despite the loss of interactions at the C6OH, the calculated losses in the binding energy seen in Figure 8, Figures S12 and S13, and the 25-fold increase in the K_M values for *p*NPGlc, the K_i for glucose for the R786 mutants was similar to that of wild type TxGH116. This suggests the R786 interactions actually contribute relatively little to glucose-binding in its inhibitory mode, although they have a larger effect on the K_M of *p*NPGlc and cellobiose, for which the contributions of R786 to the +1 subsite must be considered. This may reflect a contribution of the α -anomer of glucose to inhibition, which could be easier in the mutant with its wider entrance. Although the crystal structure shows the β -anomer, this structure is stabilized by glycerol from the cryosolution, which would not be present in the catalytic reaction. However, the calculated binding energies of the C6OH-binding residues, especially R786A, were relatively poorly correlated to the kinetic parameters relative to the mutations at the residues binding other glucose hydroxyls (Figure S14), possibly reflecting the greater potential entropy of this position and its strained binding orientation.

Cellobiose also acted as a competitive substrate, but the K_i was higher than the K_M for cellobiose in all cases, especially for the mutant enzymes. This may have been due to the cellobiose acting as a transglycosylation acceptor since a slight increase in the oligosaccharides was observed in the mutant enzyme reactions compared to wild type TxGH116 (Figure S8).

This R786A mutant also had high hydrolysis activity with *p*NP-xylopyranoside substrate, which lacks the C6OH group, reflecting the role of R786 in establishing specificity for glucopyranoside substrates (Figure 7). Notably, the structurally similar family GH52 β -xylosidase from *Geobacillus thermoglucosidasius* (PDB: 4C1P [44]) lacks residues corresponding to E777 and R786, but has a glutamine residue (Q701) in a position too close to the sugar to allow the Glc C6OH side chain to fit in the active position for hydrolysis [17,44]. The GH116 Ss01353 β -xylosidase/ β -glucosidase has a conserved glutamate at the E777 position, but is not conserved at the TxGH116 R786 position [17,45], consistent with loss of R786 allowing higher activity on xyloside, while E777 is necessary for activity on β -glucoside and helping to stabilize the cationic transition state.

One unexpected finding was that converting D452 and E777 to their amines had greater effects on catalysis than the elimination of the functional groups by mutation to alanine (Figure 3). The ONIOM calculations suggested possible explanations for the displacement of the sugar atoms. In each case, the donation of the hydrogen bond from the amino acid residue side chain amide caused a rotation of the hydroxyl to place an electron pair in a position to make the hydrogen bond. This also resulted in the displacement of the pyranose ring relative to its position in the wild type and alanine mutants (Figures S15–S18). In the case of D452N, the C3OH made an alternative hydrogen bond with the E441 nucleophile residue, contributing to the shift in position (Figure S15). In the case of E777Q, the ring displacement was less obvious, but the pyranose ring was distorted to a 3B_O boat conformation, which was further from the expected 4H_3 conformation that puts C5, O5, C1, and C2 in the same plane in the oxocarbenium ion-like transition state (Figure S16). Thus, it would appear that subtle shifts in hydrogen bonding caused by carboxylate to amide conversions can have significant effects on the ability of the sugar ring to reach the transition state.

4. Materials and Methods

4.1. Conserved Residue Alignment and Site-Directed Mutagenesis of TxGH116 Amino Acid Residues Acting in Glucose Binding

The crystal structure of TxGH116 (PDB ID: 5BVU [17]) was used as the initial input to a Pfam [46] search to identify its protein family. The enzyme TxGH116 is assigned to the glycoside hydrolases family with 2263 sequence homologs residing in the clan. To filter out redundant catalytic domain sequences, 34 seed sequences were taken from the Pfam noline alignment. After aligning the remaining sequences with MEGAX, they were further curated to remove the proteins that do not possess the catalytic residues associated with enzyme function. This series of bioinformatics operations produced the final set of 34 protein sequences for analysis in Table S1. The active-site sequence logo of the GH116 family was generated by WebLogo (<http://weblogo.berkeley.edu/>, accessed on 24 March 2021). The amino acid residues that interact with glucose were identified from the known structure of the TxGH116 glucose complex. Site-directed mutagenesis was performed on the previously reported recombinant pET30a(+)/TxGH116 plasmid [17] by the QuikChange method (Stratagene, Agilent Corp, Santa Clara, CA, USA). The mutagenic primers for the corresponding amino acid residue mutations are shown in Table S2.

4.2. Protein Expression and Purification and Enzymatic Characterization

The wild type and mutant pET30a(+)/TxGH116 plasmids were transformed into the *Escherichia coli* strain BL21(DE3) and the proteins were expressed and purified as previously described [17].

The purified TxGH116 proteins, the SDS-PAGE results for which are shown in Figure S2, were used for enzymatic characterization. The hydrolytic activity of the purified en-

zymes from *E. coli* was determined as a release of *p*-nitrophenol from *p*NPGlc and other *p*-nitrophenyl glycosides. The enzyme activity was assayed against these synthetic substrates in 50 mM sodium acetate buffer, pH 5.5, at 60 °C in 140 µL reactions. Time courses were used to assess the amount of protein and incubation time that gave the apparent initial velocities. The enzyme concentrations and incubation times for each mutant were H507A 0.074 µM, 20 min; H507E 1.85 µM, 30 min; H507Q 0.037 µM, 20 min; D452A 1.85 µM, 30 min; D452N 7.4 µM, 30 min; R792A 3.7 µM, 30 min; R792K 1.85 µM, 15 min; T591A 0.037 µM, 30 min; W732F 0.0037 µM, 20 min; E730A 0.0074 µM, 20 min; E730Q 0.0037 µM, 15 min; R786A 0.0037 µM, 15 min; R786K 0.0037 µM, 15 min; E777A 1.85 µM, 30 min; and E777Q 3.7 µM, 30 min. Three independent reactions were set up for each reaction condition. The reactions were stopped by alkalization with 70 µL 2 M Na₂CO₃, and *p*-nitrophenol was quantified by the 405 nm absorbance of its phenolate ion, in comparison to a *p*-nitrophenolate standard curve in the same solution. The kinetic parameters, including the k_{cat} , K_M , and k_{cat}/K_M of purified TxGH116 mutants from *E. coli*, were calculated by the nonlinear regression of the Michaelis–Menten plots of initial velocity data at concentrations covering at least 0.3 to 3-fold the apparent K_M value (except in the cases of the mutants D452A, H507A, H507Q, E777A, and E777Q, for which substrate concentrations of only 2 times the K_M could be achieved due to the solubility and substrate inhibition) with the Grafit 5.0 computer program (Erithacus Software, Horley, UK) and checked with linear (Hanes–Wolff $[S]/v_0$ vs. $[S]$) plots (Figure S4). For the standard activity assays, 1 mM *p*NPGlc was incubated with enzyme at 60 °C for 15 min [17]. Unless otherwise stated, all assays were carried out in triplicate.

4.3. Peroxidase/Glucose Oxidase (PGO) Assay of Oligosaccharide Substrates

The activity of enzyme with cellobiose was assayed in 50 mM sodium acetate buffer, pH 5.5, at 60 °C in 50 µL reactions. The enzyme concentrations and incubation times for each mutant were wild type 0.0037 µM, 15 min; W732F 0.0037 µM, 20 min; E730A 0.0037 µM, 20 min; E730Q 0.0037 µM, 20 min; R786A 0.0037 µM, 15 min; and R786K 0.0037 µM, 15 min, in order to obtain a detectable signal. The reactions were stopped by boiling for 5 min after incubation, followed by centrifugation at 4 °C 12,000 rpm for 30 min to remove the enzyme pellets. Then, 40 µL supernatant was pipetted into a microplate well that contained 10 µL reaction buffer and 50 µL 1 mg mL⁻¹ ABTS (2,2'-Azinobis-(3-ethylbenzthiazolin-6-sulfonic acid), Sigma-Aldrich/Merck), and 100 µL 0.01 g mL⁻¹ peroxidase/glucose oxidase (PGO) enzyme (Sigma-Aldrich/Merck). The PGO reaction was incubated at 37 °C for 30 min, and the absorbance was measured at 405 nm and compared to a glucose standard curve.

4.4. pH and Temperature Dependence

The pH dependence of the enzymatic activity was determined over the pH range of 2.5–8.0 at 0.5 pH unit intervals, with 1 mM *p*NPGlc substrate at 60 °C for 30 min in 100 mM McIlvaine universal (citrate-phosphate) buffer [47]. The optimal temperature was evaluated at the optimal pH in assays at temperatures over the range of 10–90 °C at 5–10 °C intervals.

4.5. Glucose and Cellobiose Inhibition Kinetics of Different Mutants

The activity against *p*NPGlc was assayed in a range of glucose and cellobiose concentrations (final concentrations: 0 to 200 mM) in triplicate reactions at five *p*NPGlc substrate concentrations bracketing the apparent K_M value. Glucose was preincubated with 0.004 µM enzyme at 37 °C for 10 min before the addition of *p*NPGlc and incubation at 60 °C for 15 min. Because cellobiose is hydrolyzed, no preincubation was done for the cellobiose inhibition kinetics. The K_i values were calculated from the x-intercept of the derivative plot of the slopes of Lineweaver–Burk reciprocal plots ($1/v$ vs. $1/[pNPGlc]$) vs. the inhibitor concentration (Figures S6 and S7) [48].

4.6. Protein Crystallization

Based on our previous crystallization of wild type TxGH116 by hanging drop vapor diffusion [49], the crystallization conditions were optimized for the mutant proteins. The purified TxGH116 was adjusted to 1, 2, 3, 4, and 10 mg mL⁻¹ in 20 mM Tris-HCl, 150 mM NaCl, pH 8.0, by centrifugal filtration [49]. The hanging drop vapor diffusion plate contained 500 µL precipitant in the reservoir, and the drops consisted of 2 µL protein solution and 1 µL precipitant solution. All crystallization experiments were carried out at 288 K. The optimized conditions for crystal incubation consisted of 0.15–0.3 M ammonium sulfate, 0.1 M MES buffer, pH 5.5, with 19–24% polyethylene glycol (PEG) 3000 for the reservoir solution.

4.7. Synchrotron X-ray Diffraction and Structure Solution

Crystals of TxGH116 mutants were soaked briefly in cryosolution that increased each component 10% from the precipitant solution and contained 20% glycerol. For the glucose complexes, the soaking solution contained 100 mM D-glucose. The crystals were flash vitrified in liquid nitrogen and shipped to the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan) for data collection. Diffraction data were collected on beamline BL13B1 at the NSRRC with an ADSC Quantum-315r CCD detector. The wavelength was set at 1.000 Å and the crystals were maintained at 100–105 K during diffraction with a nitrogen cold stream. The data were processed and scaled with the HKL-2000 package [50]. The initial phases for the mutants and their complexes were determined by molecular replacement with the MOLREP program [51] in the CCP4 suite [52] using the wild type native crystal structure (PDB: 5BVU [17]) with solvent and glycerol removed as the template. The structure was adjusted with the COOT graphic program [53] in alternation with refinement with REFMAC5 [54]. The final models were analyzed with MolProbity [55] and validated on the PDB Website. The data collection and refinement statistics parameters are shown in Table S3. The mutant protein structures were visualized and compared with the wild type TxGH116 structure in PyMol (Schrödinger LLC, Portland, OR, USA) and Discovery Studio 3.1 (BioVia, San Diego, CA, USA).

4.8. Computational Procedures: Binding Energy and Particular Interaction Analysis

Our own two-layered N-layered integrated molecular orbital and molecular mechanics (ONIOM) [56] calculations (ONIOM2(B3LYP/6-31G(d,p):PM3)) were performed to examine the mutation-induced effects on the binding affinity and stability of glucose ligand in the TxGH116 active site. Here, eleven protein–ligand complexes for TxGH116 wild type and mutants (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K) were modeled and optimized using the X-ray structure of the TxGH116-glucose complex (PDB: 5BX5 [17]) because of their potential hydrogen bond interactions. The ONIOM2 calculations were set up in a similar manner as in the previously published papers [36,57], and were performed on the TxGH116 active site residues in a sphere of 5 Å around the glucose. In this complex, the high-level region (region A in Figure 8) comprises the atoms in the region of the complex, including the glucose and the residues E441 (the catalytic nucleophile), H507, D452, D508, Y523, T591, D593 (the catalytic acid/base), E777, R786, and R792, and were treated to a high-level calculation, in this case, B3LYP/6-31G(d,p). These residues are, from the crystallographic point of view, involved mainly in hydrogen-bonding interactions with the glucose [17]. The low-level region (region B in Figure 8) includes all the atoms not previously selected in the high-level region, which were treated at the semi-empirical level PM3. The ONIOM2 optimized structures, as well as their binding energies, BE, were obtained from the extrapolated energy of the ONIOM2 approach, which is defined as follows:

$$\begin{aligned} \text{BE(ONIOM2)} &= E[\text{Cpx}] - E[\text{P}] - E[\text{L}] \\ &= \Delta E(\text{high, A}) + [\Delta E(\text{low, AB}) - \Delta E(\text{low, A})] \end{aligned}$$

where $E[\text{Cpx}]$ is the total energy of the protein–ligand complex, $E[\text{P}]$ is the total energy of the protein pocket, $E[\text{L}]$ is the total energy of the ligand (i.e., glucose), ΔE (high, A) is the energy of the region A calculated using a high level of calculations, ΔE (low, AB) is the energy of the whole model system (regions A and B in Figure 8) calculated using a low level of calculations, and ΔE (low, A) is the energy of the region A calculated using a low level of calculations.

To further identify the important contribution of individual residues in the TxGH116 active site, we have performed a particular interaction analysis on those mutated amino acids by calculating the interaction energies, $\text{IE}(\text{L} + \text{X}_i)$, between ligand (glucose) and individual residues, X_i , at the B3LYP/6-31G(d,p) level of theory using the ONIOM geometry described above. Counterpoise correction for the basis set superposition error (BSSE-CP) was also calculated to correct the interaction energy. The total interaction energy, IE, can be expressed as follows:

$$\text{IE}(\text{L} + \text{X}_i) = E(\text{L} + \text{X}_i) - E(\text{L}) - E(\text{X}_i)$$

where $E(\text{L})$ and $E(\text{X}_i)$ are the energies of the ligand (i.e., glucose) and each individual residue, respectively. All computations were performed using the Gaussian09 program [58].

5. Conclusions

The systematic investigation of the GH116 indicates that a delicate balance of charges and hydrogen-bonding interactions enable the binding of the sugar, stabilization of the transition state, and activation of the catalytic residues at the optimal pH. Even relatively subtle changes from a hydrogen bond acceptor to a hydrogen bond donor can have significant effects on the glycone/transition state optimization, pointing to subtle steric effects [2] in addition to charge effects. Thus, the investigated residues play important stabilization, steric, and activation roles, which critically complement the proton shuffling and covalent roles played by the catalytic acid/base and nucleophile, respectively. Although GH116 appears to have relatively little contribution from aromatic electron clouds stacking the sugars via interactions with the polarized sugar ring hydrogens compared to other β -glycosidase families, the polar interactions appeared to be similar, suggesting these lessons likely apply to other β -glycosidase families as well. The critical importance of these stabilizing, steric, and activating residues must be considered when engineering β -glucosidases for various applications.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/catal12030343/s1>. Supplementary Materials file 1: Table S1: Proteins from GH116 family used for sequence alignment; Table S2: Glycone sugar-binding interactions, related residues, and their response mutations; Table S3: Data collection and refinement statistics; Figure S1: WT TxGH116 glycone-binding residues and their interactions with glucose ligands; Figure S2: SDS gel after protein expression and 1st IMAC purification of all mutants; Figure S3: Schematic view of Ping-Pong Bi Bi reaction of TxGH116 with pNPGlc substrate; Figure S4: pNPGlc kinetics of glycone sugar-binding residues mutants; Figure S5: Cellobiose kinetics of TxGH116 variants with mutations of glycone sugar-binding residues; Figure S6: Competitive inhibition of glucose to TxGH116 and its mutants; Figure S7: Competitive inhibition of cellobiose to TxGH116 and its mutants; Figure S8: Transglycosylation analysis between WT and R786 mutation variants on hydrolysis of pNPGlc and cellobiose substrate; Figure S9: Slot-like active site entrance comparison between TxGH116 R786K with (a) and without (b) glucose ligands; Figure S10. E730A and E730Q active-site interactions; Figure S11: 2D protein–ligand interaction map (left) and 3D (optimized) structure (right) of WT TxGH116–glucose (Glc) complex obtained at ONIOM2 (B3LYP/6-31G(d,p):PM3) level; Figure S12: Binding interaction energies (BE, kcal/mol) of glucose for WT and mutant (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, R786K) models; Figure S13: Particular interaction energy (IE, kcal/mol) between the glucose and individual residues for wild type and mutant TxGH116 models; Figure S14: Correlation of calculated binding energy (BE) with kinetic constants; Figure S15: Active site comparison of ONIOM2 models of wild type, D452A, and D452N; Figure S16: Active site comparison of ONIOM2 models of wild type, E777A, and E777N; Figure S17: Active site comparison of ONIOM2 models of wild type, R786H, R786A, and R786K; Figure S18:

Active site comparison of ONIOM2 models of wild type, H507Q, H507A, and H507E. Supplementary Materials file 2: Cartesian coordinates of ONIOM2(B3LYP/6-31G(d,p):PM3) optimized geometries for all glucose (Glc)-binding structures of wild type (WT) and mutant systems (D452A, D452N, H507A, H507Q, H507E, E777A, E777Q, R786A, R786H, and R786K).

Author Contributions: Conceptualization, J.R.K.C., S.P., and J.J.; methodology, M.H., S.P., R.C., and J.J.; software, J.J.; validation, M.H., S.P., R.C., J.R.K.C., and J.J.; formal analysis, M.H., S.P., R.C., N.T., J.J., and J.R.K.C.; investigation, M.H. and S.P.; resources, J.J. and J.R.K.C.; data curation, S.P., R.C., and J.J.; writing—original draft preparation, M.H.; writing—review and editing, M.H., S.P., R.C., N.T., J.J., and J.R.K.C.; visualization, M.H., S.P., N.T., and J.J.; supervision, J.J. and J.R.K.C.; project administration, J.R.K.C.; funding acquisition, J.J. and J.R.K.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by grant RSA6280073 from the Suranaree University of Technology (SUT), the Synchrotron Light Research Institute (Public Organization), and the Thailand Research Fund. J.J. is supported by the Thailand Research Fund and the University of Phayao (UP) (grants RSA6280104 and FF64-RIB002), and the School of Science, UP (grant PBTSC65006) and thanks the National e-Science Infrastructure Consortium for computing resources (www.e-science.in.th). M.H. and S.P. are supported by SUT, Thailand Science Research and Innovation (TSRI), and the National Science, Research and Innovation Fund (NSRF) (project code 90464). The National Synchrotron Radiation Research Center was supported by the National Science Council of Taiwan, ROC.

Data Availability Statement: The X-ray crystal structures and structure factors obtained in this work, PDB 7W2S, 7W2T, 7W2U, 7W2V, and 7W2X, are available from the worldwide Protein Data Bank (wwPDB.org) via links to its affiliated members, the PDBj, PDBe, and RCSB PDB. Enzymatic kinetic data is available in the Supplementary Materials and any additional data can be provided by the authors. Cartesian coordinates for all ONIOM optimized structures are provided in the Supplementary Materials.

Acknowledgments: The NSRRC beamline staffs are thanked for their assistance in the data collection and processing.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Holliday, G.L.; Bartlett, G.J.; Almonacid, D.E.; O'Boyle, N.M.; Murray-Rust, P.; Thornton, J.M.; Mitchell, J.B. MACiE: A database of enzyme reaction mechanisms. *Bioinformatics* **2005**, *21*, 4315–4316. [[CrossRef](#)]
2. Holliday, G.L.; Mitchell, J.B.; Thornton, J.M. Understanding the functional roles of amino acid residues in enzyme catalysis. *J. Mol. Biol.* **2009**, *390*, 560–577. [[CrossRef](#)] [[PubMed](#)]
3. Ketudat Cairns, J.R.; Mahong, B.; Baiya, S.; Jeon, J.S. Beta-Glucosidases: Multitasking, moonlighting or simply misunderstood? *Plant Sci.* **2015**, *241*, 246–259. [[CrossRef](#)] [[PubMed](#)]
4. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P.M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490–D495. [[CrossRef](#)] [[PubMed](#)]
5. McCarter, J.D.; Withers, S.G. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885–892. [[CrossRef](#)]
6. Zechel, D.L.; Withers, S.G. Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.* **2000**, *33*, 11–18. [[PubMed](#)]
7. McCarthy, J.K.; Uzelac, A.; Davis, D.F.; Eveleigh, D.E. Improved catalytic efficiency and active site modification of 1,4-beta-D-glucan glucohydrolase A from *Thermotoga neapolitana* by directed evolution. *J. Biol. Chem.* **2004**, *279*, 11495–11502. [[CrossRef](#)] [[PubMed](#)]
8. Burmeister, W.P.; Cottaz, S.; Driguez, H.; Iori, R.; Palmieri, S.; Henrissat, B. The crystal structures of *Sinapis alba* myrosinase and a covalent glycosyl-enzyme intermediate provide insights into the substrate recognition and active-site machinery of an S-glycosidase. *Structure* **1997**, *5*, 663–675. [[CrossRef](#)]
9. Kurakata, Y.; Uechi, A.; Yoshida, H.; Kamitori, S.; Sakano, Y.; Nishikawa, A.; Tonzuka, T. Structural insights into the substrate specificity and function of *Escherichia coli* K12 YgjK, a glucosidase belonging to the glycoside hydrolase family 63. *J. Mol. Biol.* **2008**, *381*, 116–128. [[CrossRef](#)]
10. Vocadlo, D.J.; Davies, G.J. Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539–555. [[CrossRef](#)]
11. Chuenchor, W.; Pengthaisong, S.; Robinson, R.C.; Yuvaniyama, J.; Svasti, J.; Cairns, J.R. The structural basis of oligosaccharide binding by rice BGLu1 beta-glucosidase. *J. Struct. Biol.* **2011**, *173*, 169–179. [[CrossRef](#)]

12. Geronimo, I.; Payne, C.M.; Sandgren, M. The role of catalytic residue pKa on the hydrolysis/transglycosylation partition in family 3 beta-glucosidases. *Org. Biomol. Chem.* **2018**, *16*, 316–324. [[CrossRef](#)] [[PubMed](#)]
13. Dopitova, R.; Mazura, P.; Janda, L.; Chaloupkova, R.; Jerabek, P.; Damborsky, J.; Filipi, T.; Kiran, N.S.; Brzobohaty, B. Functional analysis of the aglycone-binding site of the maize beta-glucosidase Zm-p60.1. *FEBS J.* **2008**, *275*, 6123–6135. [[CrossRef](#)] [[PubMed](#)]
14. Pengthaisong, S.; Chen, C.F.; Withers, S.G.; Kuaprasert, B.; Ketudat Cairns, J.R. Rice BGLu1 glycosynthase and wild type transglycosylation activities distinguished by cyclophellitol inhibition. *Carbohydr. Res.* **2012**, *352*, 51–59. [[CrossRef](#)]
15. Lucas, J.E.; Siegel, J.B. Quantitative functional characterization of conserved molecular interactions in the active site of mannitol 2-dehydrogenase. *Protein Sci.* **2015**, *24*, 936–945. [[CrossRef](#)] [[PubMed](#)]
16. Zhang, X.; Wang, S.; Wu, X.; Liu, S.; Li, D.; Xu, H.; Gao, P.; Chen, G.; Wang, L. Subsite-specific contributions of different aromatic residues in the active site architecture of glycoside hydrolase family 12. *Sci. Rep.* **2015**, *5*, 18357. [[CrossRef](#)] [[PubMed](#)]
17. Charoenwattanasatien, R.; Pengthaisong, S.; Breen, I.; Mutoh, R.; Sansenya, S.; Hua, Y.; Tankrathok, A.; Wu, L.; Songsirittitthigul, C.; Tanaka, H.; et al. Bacterial beta-Glucosidase Reveals the Structural and Functional Basis of Genetic Defects in Human Glucocerebrosidase 2 (GBA2). *ACS Chem. Biol.* **2016**, *11*, 1891–1900. [[CrossRef](#)] [[PubMed](#)]
18. Jeng, W.-Y.; Wang, N.-C.; Lin, C.-T.; Chang, W.-J.; Liu, C.-I.; Wang, A.-J. High-resolution structures of *Neotermes koshunensis* β -glucosidase mutants provide insights into the catalytic mechanism and the synthesis of glucoconjugates. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 829–838. [[CrossRef](#)] [[PubMed](#)]
19. Woeste, M.A.; Wachten, D. The enigmatic role of GBA2 in controlling locomotor function. *Front. Mol. Neurosci.* **2017**, *10*, 386. [[CrossRef](#)] [[PubMed](#)]
20. Sultana, S.; Stewart, J.; van der Spoel, A.C. Truncated mutants of beta-glucosidase 2 (GBA2) are localized in the mitochondrial matrix and cause mitochondrial fragmentation. *PLoS ONE* **2020**, *15*, e0233856. [[CrossRef](#)] [[PubMed](#)]
21. van Weely, S.; Brandsma, M.; Strijland, A.; Tager, J.M.; Aerts, J.M. Demonstration of the existence of a second, non-lysosomal glucocerebrosidase that is not deficient in Gaucher disease. *Biochim. Biophys. Acta* **1993**, *1181*, 55–62. [[CrossRef](#)]
22. Boot, R.G.; Verhoek, M.; Donker-Koopman, W.; Strijland, A.; van Marle, J.; Overkleeft, H.S.; Wennekes, T.; Aerts, J.M. Identification of the non-lysosomal glucosylceramidase as beta-glucosidase 2. *J. Biol. Chem.* **2007**, *282*, 1305–1312. [[CrossRef](#)]
23. Jatoorathawichot, P.; Talabnin, C.; Ngiwsara, L.; Rustam, Y.H.; Svasti, J.; Reid, G.E.; Ketudat Cairns, J.R. Effect of expression of human glucosylceramidase 2 isoforms on lipid profiles in COS-7 cells. *Metabolites* **2020**, *10*, 488. [[CrossRef](#)]
24. Mistry, P.K.; Liu, J.; Sun, L.; Chuang, W.-L.; Yuen, T.; Yang, R.; Lu, P.; Zhang, K.; Li, J.; Keutzer, J. Glucocerebrosidase 2 gene deletion rescues type 1 Gaucher disease. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 4934–4939. [[CrossRef](#)]
25. Bouscary, A.; Quessada, C.; René, F.; Spedding, M.; Henriques, A.; Ngo, S.; Loeffler, J.-P. Drug repositioning in neurodegeneration: An overview of the use of ambroxol in neurodegenerative diseases. *Eur. J. Pharmacol.* **2020**, *884*, 173446. [[CrossRef](#)] [[PubMed](#)]
26. Kuo, C.L.; Kallemeijn, W.W.; Lelieveld, L.T.; Mirzaian, M.; Zoutendijk, I.; Vardi, A.; Futerman, A.H.; Meijer, A.H.; Spaink, H.P.; Overkleeft, H.S. In vivo inactivation of glycosidases by conduritol B epoxide and cyclophellitol as revealed by activity-based protein profiling. *FEBS J.* **2019**, *286*, 584–600. [[CrossRef](#)] [[PubMed](#)]
27. Artola, M.; Wu, L.; Ferraz, M.J.; Kuo, C.L.; Raich, L.; Breen, I.Z.; Offen, W.A.; Codee, J.D.C.; van der Marel, G.A.; Rovira, C.; et al. 1,6-Cyclophellitol cyclophosphates: A new class of irreversible glycosidase inhibitor. *ACS Cent. Sci.* **2017**, *3*, 784–793. [[CrossRef](#)]
28. Lahav, D.I.; Liu, B.; van den Berg, R.J.; van den Nieuwendijk, A.M.; Wennekes, T.; Ghisaidoobe, A.T.; Breen, I.; Ferraz, M.J.; Kuo, C.-L.; Wu, L. A fluorescence polarization activity-based protein profiling assay in the discovery of potent, selective inhibitors for human nonlysosomal glucosylceramidase. *J. Am. Chem. Soc.* **2017**, *139*, 14192–14197. [[CrossRef](#)]
29. Artola, M.; Kuo, C.-L.; Lelieveld, L.T.; Rowland, R.J.; van der Marel, G.A.; Codée, J.D.; Boot, R.G.; Davies, G.J.; Aerts, J.M.; Overkleeft, H.S. Functionalized cyclophellitols are selective glucocerebrosidase inhibitors and induce a bona fide neuropathic Gaucher model in zebrafish. *J. Am. Chem. Soc.* **2019**, *141*, 4214–4218. [[CrossRef](#)] [[PubMed](#)]
30. Gorantla, J.N.; Pengthaisong, S.; Choknud, S.; Kaewpuang, T.; Manyum, T.; Promarak, V.; Ketudat Cairns, J.R. Gram scale production of 1-azido- β -D-glucose via enzyme catalysis for the synthesis of 1,2,3-triazole-glucosides. *RSC Adv.* **2019**, *9*, 6211–6220. [[CrossRef](#)]
31. Gorantla, J.N.; Maniganda, S.; Pengthaisong, S.; Ngiwsara, L.; Sawangareetrakul, P.; Chokchaisiri, S.; Kittakoop, P.; Svasti, J.; Ketudat Cairns, J.R. Chemoenzymatic and Protecting-Group-Free Synthesis of 1, 4-Substituted 1, 2, 3-Triazole- α -D-glucosides with Potent Inhibitory Activity toward Lysosomal α -Glucosidase. *ACS Omega* **2021**, *6*, 25710–25719. [[CrossRef](#)] [[PubMed](#)]
32. Tankrathok, A.; Iglesias-Fernández, J.; Williams, R.J.; Pengthaisong, S.; Baiya, S.; Hakki, Z.; Robinson, R.C.; Hrmova, M.; Rovira, C.; Williams, S.J. A single glycosidase harnesses different pyranoside ring transition state conformations for hydrolysis of mannosides and glucosides. *ACS Catal.* **2015**, *5*, 6041–6051. [[CrossRef](#)]
33. Pengthaisong, S.; Hua, Y.; Ketudat Cairns, J.R. Structural basis for transglycosylation in glycoside hydrolase family GH116 glycosynthases. *Arch. Biochem. Biophys.* **2021**, *706*, 108924. [[CrossRef](#)] [[PubMed](#)]
34. Laskowski, R.A.; Swindells, M.B. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. *J. Chem. Inf. Model.* **2011**, *51*, 2778–2786. [[CrossRef](#)] [[PubMed](#)]
35. Quirke, J.C.K.; Crich, D. Glycoside hydrolases restrict the side chain conformation of their substrates to gain additional transition state stabilization. *J. Am. Chem. Soc.* **2020**, *142*, 16965–16973. [[CrossRef](#)] [[PubMed](#)]
36. Saen-oon, S.; Kuno, M.; Hannongbua, S. Binding energy analysis for wild-type and Y181C mutant HIV-1 RT/8-CI TIBO complex structures: Quantum chemical calculations based on the ONIOM method. *Proteins* **2005**, *61*, 859–869. [[CrossRef](#)] [[PubMed](#)]

37. Nunriam, P.; Kuno, M.; Saen-oon, S.; Hannongbua, S. Particular interaction between efavirenz and the HIV-1 reverse transcriptase binding site as explained by the ONIOM2 method. *Chem. Phys. Lett.* **2005**, *405*, 198–202. [[CrossRef](#)]
38. Kuno, M.; Hannongbua, S.; Morokuma, K. Theoretical investigation on nevirapine and HIV-1 reverse transcriptase binding site interaction, based on ONIOM method. *Chem. Phys. Lett.* **2003**, *380*, 456–463. [[CrossRef](#)]
39. Romaniouk, A.; Vijay, I.K. Structure-function relationships in glucosidase I: Amino acids involved in binding the substrate to the enzyme. *Glycobiology* **1997**, *7*, 399–404. [[CrossRef](#)] [[PubMed](#)]
40. Namchuk, M.N.; Withers, S.G. Mechanism of *Agrobacterium* beta-glucosidase: Kinetic analysis of the role of noncovalent enzyme/substrate interactions. *Biochemistry* **1995**, *34*, 16194–16202. [[CrossRef](#)] [[PubMed](#)]
41. Zhao, H.; Pan, Q.; Zhang, W.; Carmichael, I.; Serianni, A.S. DFT and NMR studies of 2JCOH, 3JHCOH, and 3JCCOH spin-couplings in saccharides: C-O torsional bias and H-bonding in aqueous solution. *J. Org. Chem.* **2007**, *72*, 7071–7082. [[CrossRef](#)] [[PubMed](#)]
42. Quirke, J.C.K.; Crich, D. Side chain conformation restriction in the catalysis of glycosidic bond formation by Leloir glycosyltransferases, glycoside phosphorylases, and transglycosidases. *ACS Catal.* **2021**, *11*, 5069–5078. [[CrossRef](#)] [[PubMed](#)]
43. Sultana, S.; Reichbauer, J.; Schule, R.; Mochel, F.; Synofzik, M.; van der Spoel, A.C. Lack of enzyme activity in GBA2 mutants associated with hereditary spastic paraplegia/cerebellar ataxia (SPG46). *Biochem. Biophys. Res. Commun.* **2015**, *465*, 35–40. [[CrossRef](#)] [[PubMed](#)]
44. Espina, G.; Eley, K.; Pompidor, G.; Schneider, T.R.; Crennell, S.J.; Danson, M.J. A novel β -xylosidase structure from *Geobacillus thermoglucosidasius*: The first crystal structure of a glycoside hydrolase family GH52 enzyme reveals unpredicted similarity to other glycoside hydrolase folds. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 1366–1374. [[CrossRef](#)] [[PubMed](#)]
45. Cobucci-Ponzano, B.; Aurilia, V.; Riccio, G.; Henrissat, B.; Coutinho, P.M.; Strazzulli, A.; Padula, A.; Corsaro, M.M.; Pieretti, G.; Pocsfalvi, G.; et al. A new archaeal beta-glycosidase from *Sulfolobus solfataricus*: Seeding a novel retaining beta-glycan-specific glycoside hydrolase family along with the human non-lysosomal glucosylceramidase GBA2. *J. Biol. Chem.* **2010**, *285*, 20691–20703. [[CrossRef](#)] [[PubMed](#)]
46. Finn, R.D.; Mistry, J.; Schuster-Bockler, B.; Griffiths-Jones, S.; Hollich, V.; Lassmann, T.; Moxon, S.; Marshall, M.; Khanna, A.; Durbin, R.; et al. Pfam: Clans, web tools and services. *Nucleic Acids Res.* **2006**, *34*, D247–D251. [[CrossRef](#)] [[PubMed](#)]
47. McIlvaine, T. A buffer solution for colorimetric comparison. *J. Biol. Chem.* **1921**, *49*, 183–186. [[CrossRef](#)]
48. Dixon, M. The determination of enzyme inhibitor constants. *Biochem. J.* **1953**, *55*, 170–171. [[CrossRef](#)]
49. Sansenya, S.; Mutoh, R.; Charoenwattanasatien, R.; Kurisu, G.; Ketudat Cairns, J.R. Expression and crystallization of a bacterial glycoside hydrolase family 116 beta-glucosidase from *Thermoanaerobacterium xylanolyticum*. *Acta Crystallogr. F Struct. Biol. Commun.* **2015**, *71*, 41–44. [[CrossRef](#)]
50. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzym.* **1997**, *276*, 307–326.
51. Vagin, A.; Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 22–25. [[CrossRef](#)]
52. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)] [[PubMed](#)]
53. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)]
54. Murshudov, G.N.; Skubak, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 355–367. [[CrossRef](#)]
55. Chen, V.B.; Arendall, W.B., 3rd; Headd, J.J.; Keedy, D.A.; Immormino, R.M.; Kapral, G.J.; Murray, L.W.; Richardson, J.S.; Richardson, D.C. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12–21. [[CrossRef](#)] [[PubMed](#)]
56. Chung, L.W.; Sameera, W.M.C.; Ramozzi, R.; Page, A.J.; Hatanaka, M.; Petrova, G.P.; Harris, T.V.; Li, X.; Ke, Z.; Liu, F.; et al. The ONIOM method and its applications. *Chem. Rev.* **2015**, *115*, 5678–5796. [[CrossRef](#)] [[PubMed](#)]
57. Boonsri, P.; Kuno, M.; Hannongbua, S. Key interactions of the mutant HIV-1 reverse transcriptase/efavirenz: An evidence obtained from ONIOM method. *Med. Chem. Commun.* **2011**, *2*, 1181–1187. [[CrossRef](#)]
58. Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Scuseria, G.E.; Robb, M.A.; Cheeseman, J.R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G.A.; et al. *Gaussian 09, Revision A.02*; Gaussian, Inc.: Wallingford, CT, USA, 2009.