

◀ **Fig. 1** Structures of example β -glucosidase substrates. The plant cyanogenic glucosides linamarin, dhurrin, prunasin, and its precursor amygdalin. Other defense-related glycosides include 2-*O*- β -D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOAGlc) and the flavonoids apigenin 7-*O*- β -D-glucoside, the isoflavonoids diadzin and genistin, and phloridzin. Coniferin and coumaryl alcohol represent monolignol β -glucosides, while abscissic acid glucosyl ester is a phytohormone glucoconjugate and salicin and indoxyl β -glucoside are other plant glycosides with similarity to phytohormones. Strictosidine is the metabolic precursor to a wide array of monoterpene alkaloids. Cellobiose and laminaribiose represent plant cell-wall-derived oligosaccharides and can be extended with the same linkage to give the corresponding triose, tetraose, etc. In the lower right is an example of a glucosyl ceramide, one of the substrates for human acid β -glucosidase (GBA1) and other mammalian β -glucosidases

Henrissat developed an alternative classification system for glycoside hydrolases based on amino acid sequence and structural similarity [1–3]. In this system, those enzymes with overall amino acid sequence similarity and well-conserved sequence motifs are grouped into the same family. At this writing, 115 glycoside hydrolase families are listed in the frequently updated Carbohydrate Active enZYme (CAZY) Web site (<http://www.cazy.org>) [3]. The β -glucosidases that have been described in the literature fall in glycoside hydrolase families GH1, GH3, GH5, GH9, and GH30, [1, 3–5]. In addition, the human bile acid β -glucosidase/GBA2 glucocerebrosidase and its relatives are yet to be assigned to a family. The families that have similar catalytic domain structures and conserved catalytic amino acids, suggestive of a common ancestry and catalytic mechanism, are grouped into clans [2, 3]. Of these, clan GH-A has the largest number of families, and it includes the β -glucosidase-containing families GH1, GH5, and GH30.

GH1 includes the largest number of characterized β -glucosidases; therefore, it will be the emphasis of this review, with the other β -glucosidase families receiving brief mention. We will consider the roles of β -glucosidases and related enzymes in animals, in plants, and in microorganisms, followed by the application of these enzymes. Then, we will describe the structure, mechanism, and the resulting general properties of β -glucosidases that have been purified, and end with a perspective on what is known and needs for further study.

β -Glucosidases and their functional roles

Roles of β -glucosidases and their relatives in mammals

Mammals contain several β -glucosidases, including the family GH1 lactase-phloridzin hydrolase and cytoplasmic β -glucosidase, the GH30 human acid β -glucosidase (GBA1) and the bile acid β -glucosidase or GBA2. These

enzymes are thought to play roles in metabolism of glycolipids and dietary glucosides. In addition, a group of related family GH1 proteins is thought to play signaling functions.

Perhaps the best-studied mammalian β -glucosidase is the human acid β -glucosidase, which is generally considered a glucosyl ceramidase. Defects in the function of this enzyme and its transport to the lysosome lead to Gaucher disease, in which glycosphingolipids accumulate in the lysosomes of tissue leukocytes leading to their engorgement and buildup in the tissues [6]. Since enzyme replacement therapy is one way of alleviating the symptoms for this disease, the enzyme has been produced in recombinant mammalian and insect cells, and gene-activated human cells, and the structures of the enzymes determined [7–9]. Other means of treatment for Gaucher disease include use of imino sugars like deoxynojirimycin and isofagomine and their hydrophobic derivatives, which may inhibit synthesis of glycosphingolipids, but also bind to mutant GBA1 in the ER and help it to fold properly for transport to the lysosome [6]. However, these inhibitors also inhibit other β -glucosidases, such as GBA2 [10, 11].

A human bile acid β -glucosidase (GBA2) was found associated with liver microsomes [12]. Immunofluorescence showed a perinuclear reticulated localization [10], consistent with its earlier assignment to the endoplasmic reticulum (ER), while over-expressed green fluorescent protein tagged enzyme was localized near the plasma membrane [11]. When the bile acid β -glucosidase gene was knocked out in mice, little effect was seen on bile acid metabolism, but there was an accumulation of glucosylceramides in the Sertoli cells of the testes, leading to round-headed sperm and decreased fertility [10]. Cells transfected with the gene showed an increased ability to hydrolyze fluorescent glucosylceramides, confirming the protein's identity as GBA2, the nonlysosomal glycosphingolipidase [10, 11].

Currently, five GH1 proteins are known in humans: lactase-phloridzin hydrolase (LPH), cytoplasmic β -glucosidase, Klotho (α -Klotho, KL) β -Klotho (β -KL), and Klotho-LPH-related protein (KLPH) [13]. LPH, an intestinal hydrolase involved in food digestion, has both β -glucosidase activity toward exogenous glucosides, such as phloridzin, and β -galactosidase activity toward lactose. The precursor protein consists of four GH1 domains, the first two of which are removed during maturation, leaving a type I membrane protein with the LPH3 and LPH4 domains bound to the intestinal epithelial cells by a C-terminal transmembrane segment [14]. Flavonoid glucosides appear to be hydrolyzed at both the LPH3 active site and the lactase active site in the LPH4 domain [15, 16]. Deficiency of this enzyme leads to lactose intolerance, one of the most common enzyme deficiencies in humans.

The broad specificity cytoplasmic β -glucosidase has been studied for 30 years [17], and it has recently been given the additional name Klotho-related protein [18]. The cytoplasmic β -glucosidase is found in high levels in hepatocytes and brush border epithelial cells, and it has been shown to hydrolyze plant-derived flavonoid glucosides with high efficiency [19, 20], as shown in Table 1. Recently, the cytoplasmic β -glucosidase was shown to partially account for residual hydrolysis of glucoceramides and galactoceramides in fibroblasts treated with conduritol β -epoxide (CBE), a potent inhibitor of human GBA1 [18]. Although the recombinant enzyme produced in *Escherichia coli* showed extremely slow hydrolysis of natural glyoceramides (Table 1), its structure was solved and found to include lipids in positions suggestive of a glyoceramide binding site and its k_{cat}/K_m values for fluorescently labeled glyoceramides were comparable to those of flavonoid glycosides. The structure of the human cytoplasmic β -glucosidase expressed in *Pichia pastoris* was solved independently, and the residues likely to be involved in its binding of quercetin 4'-O-glucoside were identified by molecular docking and site-directed mutagenesis [13]. No lipids were observed in the active site of the enzyme produced in *P. pastoris*.

The Klotho subfamily of mammalian GH1 proteins (i.e., KL, β -KL, and KLPH) lack a complete active site with both the catalytic acid/base and nucleophile and thus have no β -glucosidase activity [12]. However, KL has been shown to have weak glucuronidase activity and to modify glycosylation of the transient receptor potential ion channel

TRPV5, suggesting they may act as glycoside hydrolases [21]. KL was identified by its induction of rapid aging-like symptoms in knock-out mice [22], and it plays regulatory roles in calcium and phosphate homeostasis [21, 23]. All Klotho subfamily members have N-terminal secretory signal sequences and C-terminal transmembrane domains, and KL has a secretory form as well [24, 25]. The KLPH has a single GH1 domain, while KL and β -KL have two GH1 domains, all of which lack essential catalytic amino acids [24].

Insect β -glucosidases and myrosinases

Although the *Drosophila melanogaster* genome contains only one GH1 gene, suggesting that insects may not have expanded this gene family at an early stage, other insects have adapted glycosides and glycoside hydrolases from the plants on which they feed for protection and digestive purposes [26]. Digestive β -glycosidases from GH1 have been isolated from insect larvae that feed on plants [27, 28]. Similarly, myrosinases have been isolated from specialist insects that feed on crucifers, such as the cabbage aphid, *Brevicoryne brassicae* [29]. The larval β -glycosidases mentioned above can hydrolyze glucoligosaccharides and plant glycosides, such as cellobiose, gentiobiose, and amygdalin [27], in line with their digestive functions. These insect β -glycosidases and myrosinases have sequences most similar to each other, then to vertebrate LPH, suggesting they diverged from the same animal GH1 gene ancestor, after its divergence from plants.

Table 1 Substrate specificity of human cytoplasmic β -glucosidase: substrate K_m and apparent k_{cat} values

Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{mM}^{-1} \text{s}^{-1}$)	Reference
Artificial aryl glycosides				
pNP- β -D-Glc	1,800	12	6.9	Berrin [20]
pNP- β -D-Fuc	370	11	29	Berrin [20]
pNP- α -L-Ara	570	6.0	10	Berrin [20]
pNP- β -D-Gal	3,100	18	5.6	Berrin [20]
Flavonoids and isoflavonoids				
Quercetin 4'-glucoside	34	ND		Day [19]
Quercetin 4'-glucoside	32	1.1	34	Berrin [20]
Quercetin 7-glucoside	42	0.7	16	Berrin [20]
Apigenin 7-glucoside	22	1.5	71	Berrin [20]
Luteolin 4'-glucoside	10	1.2	117	Berrin [20]
Luteolin 7-glucoside	50	3.0	61	Berrin [20]
Genistin	13	ND		Day [19]
Genistin	35	1.5	44	Berrin [20]
Glycosphingolipids				
C6-NBD-GlcCer	4.6	0.121	26	Hayashi [18]
C6-NBD-Gal-Cer	2.0	0.255	128	Hayashi [18]
C18-Glc-Cer	14	0.0072	0.51	Hayashi [18]
C18-Gal-Cer	9.2	<0.0002	<0.02	Hayashi [18]

Berrin et al. [20], recombinant human from *P. pastoris*; Day et al. [19], human from small intestine and liver; Hayashi et al. [18], human expressed in *E. coli*

These genes have since evolved to meet the unique needs of the herbivorous insects in their battle with plant defenses to exploit the plant nutrients.

Roles of GH1 β -glucosidases in plants

Functional diversity and multiplicity

It is in plants that β -glucosidases have been found to play the widest array of biological functions, which include roles in defense, symbiosis, cell wall catabolism and lignification, signalling, and plant secondary metabolism. Several putative β -glucosidase genes have been shown either to be induced by biotic or abiotic stress or to be necessary for successful response to the stress [30–34]. With the advent of genomics, it became clear that about 40 GH1 β -glucosidases are expressed in a typical plant, many in the same tissues [35, 36]. The roles of these enzymes are presumed to be determined by their substrate-specificities, their tissue and subcellular localization, and the conditions under which they come into contact with their physiological substrates.

To match this enormous functional diversity, plants have the largest number of GH1 family proteins. For example, 48 GH1 genes for putative β -glucosidases and thioglucosidases are found in *Arabidopsis thaliana* [35] and 40 GH1 genes are found in rice genome sequences [36]. A number of these represent pseudogenes, and, in the case of rice, two appear to be endophyte genes, but nonetheless both plants appear to express over 30 putative GH1 β -glucosidases. Sequence-based phylogenetic analysis grouped these proteins into eight clusters that include both rice and *Arabidopsis* representatives and two clusters found only in *Arabidopsis* and other plants of the family Caprales, including a cluster of classical thioglucosidases (myrosinases) and a cluster of ER and peroxisomal β -glucosidases and myrosinases. In addition, several groups of enzymes from other plants do not fall into the *Arabidopsis* and rice phylogenetic clusters, including the well-studied chloroplastic β -glucosidases of maize, sorghum, wheat and other cereals, which are not found in rice. Most of these plant GH1 enzymes are closely related to one another, but the lineage of SFR2 [33] shows higher similarity to enzymes from thermophilic bacteria and Archaea than other plant enzymes, and this lineage is thought to be distinct within GH1 [37]. A few of the *Arabidopsis* and rice enzymes have been shown to be primarily β -D-mannosidases [35, 38], so it is possible that some of the others will have different glycone specificities as well, but most are likely to be β -D-glucosidases. Given that plants also contain GH3 and GH5 β -glycosidases with β -glucosidase activity [5, 39], the precise number of β -glucosidase

isoenzymes in a particular plant species has yet to be determined.

Defense and microbial interaction

Plants have long been known to contain glycosides that release toxic compounds, such as cyanide and hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones) [40, 41], and the use of β -glucosidases as “detonators” of these chemical “bombs” has recently been reviewed [42]. In general, the defense glycosides are stored in a different cell or a different cellular compartment from the β -glucosidases that hydrolyze them to release toxic compounds. The defense compounds tend to be stored in the vacuole, while their corresponding β -glucosidases are often found in the apoplast or plastid. Both β -glucosidases and thioglucosidases have been found to play these roles, and specialist insects that feed on these plants have adapted these enzymes to diffuse the glycoside bombs or use them for their own defense [26]. β -Glucosidase-mediated defenses are also required for endophytic fungi to develop symbiotic relationships with plants, evidently by modulating the growth of these microorganisms [43].

Plants have developed a wide range of compounds for defense, examples of which can be seen in Fig. 1. Cyanogenic β -glucosides, including linamarin from clover, cassava and various other plants, dhurrin from sorghum, and prunasin from cherry and other stone fruits, are hydrolyzed to release an α -hydroxynitrile, which then breaks down either enzymatically or spontaneously to release cyanide and an aldehyde [41, 42]. Noncyanogenic defense compounds, such as γ - and β -hydroxynitriles and isoflavones in legumes, other flavonoids, coumarins, hydroxaminic acids, such as 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) in maize and wheat, and saponins are also stored as β -D-glucosides, which are hydrolyzed by specific β -glucosidases [44–54].

Aside from sequestration of the enzyme in the chloroplast or apoplast, several GH1 hydrolases are found in other compartments. The AtBGLU26 (PEN2) myrosinase is found in the pyroxisome [34], while AtBGLU23 (PYK10) is the most abundant protein in an ER-derived compartment called the ER body, which is only found in crucifers [55, 56]. AtBGLU26 has been shown to be critical to the *Arabidopsis* defense against nonspecialist fungi [34, 57]. AtBGLU23 is a β -glucosidase that has been found to be critical for establishment of symbiosis with the endophytic fungus *Piriformospora indica* by preventing it from overgrowing the roots and triggering a defense response [43, 58]. AtBGLU23 and the closely related isoenzymes AtBGLU21 and AtBGLU22 have recently been shown to be specific for scopolin, the most abundant

coumarin glucoside in *Arabidopsis* roots, thereby explaining the antifungal role of AtBGLU23 [54].

Upon cell disruption, plant defensive β -glucosidases and β -thioglucosidases often bind to cytoplasmic aggregating factors, which are thought to help localize the otherwise soluble β -glucosidases and β -thioglucosidases at the site of injury, ensuring a maximal release of defense compounds [53, 59–63]. The functional significance of the interactions of the various defensive β -glucosidases from different cellular compartments and their multiple aggregating factors is an area of active investigation.

Cell wall metabolism

The cell wall of plants is the largest repository of carbohydrates in nature, much of which are β -linked glucosyl residues, so it is not surprising that β -glucosidases should play important roles in cell wall development. β -Glucosidases, in fact, appear to play roles in both the degradation of oligosaccharides generated in cell wall turnover and release of monolignols from their glycosides to allow lignification to stabilize secondary cell walls.

Several β -glucosidases that hydrolyze cell-wall-derived oligosaccharides have been identified over the years and have been studied primarily in monocots. For example, a β -glucosidase in germinating barley seedlings showed activity toward β -1,3- and β -1,4-linked oligosaccharides [64–66]. More recently, it has been shown that this enzyme displays greater preference for manno oligosaccharides, which are also found in barley endosperm cell walls [67]. Rice seedling β -glucosidases have also been shown to hydrolyze oligosaccharides, with varying preferences [36, 38, 68–71]. Rice BGlu1 (Os3BGlu7), Os3BGlu8, and Os7BGlu26 hydrolyzed cellooligosaccharides with increasing efficiency as the degree of polymerization (DP) increased from 2 to 6, while Os4BGlu12 showed little increase in activity with DP and Os3BGlu6 hydrolyzed disaccharides best. The Os3BGlu7 and Os3BGlu8 isoenzymes are widely expressed in rice tissues, so they may be needed for release of glucose from oligosaccharides generated in cell wall remodeling at various stages of plant development. Since the rice β -glucosidase isoenzymes mentioned above also hydrolyze plant-derived glycosides, they may play other roles in the plant as well.

The lignification of secondary cell walls is thought to involve the activation of monolignols by removal of β -glucosyl residues from monolignol glycosides, like cinnamyl alcohol β -glucosides [72]. A coniferin β -glucosidase was identified from lodgepole pine tree xylem [73]. Immunological analysis indicated that this protein was localized to the differentiating region of the xylem, consistent with a role in lignification. More recently, two *Arabidopsis* β -glucosidases (AtBGLU45 and AtBGLU46)

that cluster with lodgepole pine coniferin β -glucosidase in phylogenetic analysis were shown to hydrolyze coniferin, syringin and coumaryl alcohol glucoside [74].

Phytohormone activation

Many phytohormone glucosides are found in plants, and there has been some debate as to whether these are terminal inactivation products or storage forms that can be readily activated by specific β -glucosidases. Partially purified rice β -glucosidases were shown to hydrolyze gibberellin glucosides [75], while maize β -glucosidase (Zm-p60.1, an isoform of ZmGlu1, which hydrolyzes the defense compound DIMBOAGlc) was shown to hydrolyze and activate cytokinin β -glucosides in vitro, as well as in vivo after infusion of the exogenous substrate [76]. Abscissic acid (ABA) glucosyl ester (ABA-GE) was shown to be transported from roots to leaves and be hydrolyzed by extracellular β -glucosidase in the leaves, although free ABA is transported in larger amounts [77]. An enzyme hydrolyzing the auxin glucosyl ester 6-*O*-(4-*O*)-indole-3-ylacetyl- β -D-glucose has also been identified, though the nature of this enzyme remains to be determined [78].

Recently, it has been shown that a specific *Arabidopsis* ER β -glucosidase (AtBGLU18, AtBG1) is activated to hydrolyze ABA-GE in response to drought stress [79]. Mutation of the gene for this enzyme caused early germination and defective stomata closing, which could be rescued by transgenic expression of the gene, but not by a gene encoding an inactive mutant, thereby verifying its role in increasing ABA levels. This is perhaps the clearest demonstration of a physiological role for a β -glucosidase in phytohormone activation and suggests that other phytohormone glucoconjugates also serve as β -glucosidase-activated storage forms.

Secondary metabolism

The monoterpene alkaloid intermediate strictosidine is hydrolyzed by a specific cytoplasmic β -glucosidase to allow metabolism to various monoterpene alkaloids, depending on the plant [80]. This enzyme has been characterized from several plants, and that of *Rauvolfia serpentina* has been structurally characterized [81]. One of the downstream products from strictosidine is raucaffricine, a glucoside that can be deglycosylated by raucaffricine β -glucosidase for further metabolism to ajmaline [82]. Recently, another β -glucosidase hydrolyzing alkaloid glucosides was isolated from *Psychotria ipecacuanha*, further expanding on this theme [83]. As such, β -glucosidases can play metabolic roles to release glucosyl blocking groups from metabolic intermediates and allow

their metabolism to various natural products, many of which are of medicinal importance.

Other functions

Plant β -glucosidases may play a variety of other roles. For instance, they appear to play roles in releasing volatiles from glycoside storage forms. This includes flower scents [84] as well as attractants for parasitic wasps that can attack herbivores feeding on the plant [85]. With the wide variety of glucosides found in plants, it is likely many roles remain to be discovered.

Roles of GH1 β -glucosidases in microorganisms

Although much research has been done on β -glucosidases from microorganisms, most of it has focused on their application rather than their endogenous function. As such, most of the enzymes that have been studied in the context of their natural function are those involved in bioconversion to produce glucose from plant biomass, or in breaking through plant cell walls to establish pathogenic or symbiotic relationships [86]. Bacterial β -glucosidases are often components of large complexes called cellosomes, contain polysaccharide degrading endoglucanases and carbohydrate binding proteins to localize the complex and to the cellulose surface and the cell membrane [87, 88]. Alternatively, some microorganisms secrete soluble endoglycosidases and exoglycosidases for this function, including exoglucanases/ β -glucosidases. Fungi, such as the white rot fungus *Phanerochaete chrysosporium*, may contain cytoplasmic β -glucosidases and extracellular exoglucanases, some of which may act in metabolism of the organism's own cell wall, in addition to plant cell wall biomass metabolism [89–91].

Applications of β -glucosidases

As noted above, β -glucosidases are of interest for biomass conversion, since conversion of β -glucans, the largest source of biomass in the world, is generally accomplished by three enzymes, endo- β -glucanases (e.g., cellulases), exo- β -glucanases (e.g., cellobiosidases) and β -glucosidases [86]. Limiting factors in conversion of cellulose to glucose for fermentation to alcohol include the inhibition of cellulases by oligosaccharides and the lack of adequate β -glucosidase production by certain microorganisms used for biomass breakdown. Thus, the identification and production of β -glucosidases, especially those with high glucose tolerance, has been of interest, and applicable β -glucosidases have been isolated from bacteria and fungi [86, 89–91].

There are hundreds of different β -glucosidic flavor precursors in plants, and their hydrolysis often enhances

the quality of the beverages and foods produced from them [92, 93]. Generally, there are native β -glucosidases in source-plant tissues that hydrolyze these flavor precursors to produce the desired aglycone moiety. Enzymes from the source plants or other sources may be added to foods and beverages before, during, or after processing to enhance food quality. Enzymes with desirable properties may be targeted for breeding programs to increase their abundance in the plants or for overproduction in transgenic microbial or plant hosts, and for engineering to improve their catalytic properties for flavor enhancement and stability.

Aside from flavor enhancement, foods, feeds, and beverages may be improved nutritionally by release of vitamins, antioxidants, and other beneficial compounds from their glycosides. For instance, vitamin B₆ (pyridoxine) can be released from its glucoside by enzymes in rice, in which pyridoxine glucoside is its predominant form [70, 94]. Other vitamins are also found as glucosides in plant sources, and release of their aglycones may improve their nutritional availability, despite the presence of animal and microbial β -glucosidases in the small intestine to aid in this process. Therefore, animal feeds are often treated with crude β -glucosidases. Much work has been done to identify β -glucosidases that can hydrolyze soy isoflavone glycosides, the aglycones of which have antioxidant properties [45–47, 95, 96]. Similarly, the pungent taste of food made from cruciferous vegetables (e.g., broccoli, cabbage, cauliflower, horseradish, kale, mustard, watercress, etc.) is derived from the products of the myrosinase/glucosinolate system, which may also have anticarcinogenic effects, although they may cause endemic goiter in large amounts [93, 97].

The compartmentalized β -glucosidase- β -glucoside defense systems found in such food and feed-plant tissues as cassava roots and leaves, lima beans, flax seeds, and clover leaves produce HCN when the tissue is macerated during preparation or by chewing [93]. The bitterness in almonds is caused by the presence of cyanogenic glucosides [98]. Cassava is highly consumed in parts of Asia, Africa, and South America, and contains the cyanogenic β -glucoside linamarin and its β -glucosidase linamarase. When consumed raw, cyanide poisoning can occur with symptoms of difficulty in breathing, paralysis, convulsion, coma, and even death. Similar symptoms can arise when bitter almonds are eaten raw. Processing of cassava roots by maceration results in the enzyme releasing the HCN, and subsequent aeration and washing removes the products of cyanogenesis. Alternatively, thorough cooking destroys the linamarase enzyme, preventing cyanide release.

In addition to their catalysis of hydrolysis, β -glucosidases also catalyze reverse hydrolysis and transglycosylation reactions, which can be used to synthesize oligosaccharides and glycosides of interest [99]. Various mutations have been

developed to maximize the products of these transglycosylation reactions by manipulating the catalytic mechanism, as described below.

Biochemistry of β -glucosidases

Structures of β -glucosidases

β -Glucosidases have various structures, but the overall fold of the catalytic domain is similar in each GH family. The families GH1, GH5, and GH30 belong to the Clan GH-A, and they all have similar $(\beta/\alpha)_8$ -barrel domains that contain their active site. In contrast, GH3 enzymes have two domains contributing to their active site. GH9 enzymes have $(\alpha/\alpha)_6$ -barrel structures, while the GBA2 family shows weak homology to proteins with this $(\alpha/\alpha)_6$ structure as well (Fig. 2). We will consider these in turn, followed by a more in-depth look at GH1 enzymes, which serve as an excellent model for studying the structural basis for diverse substrate specificities.

The clan GH-A enzymes of families GH1, GH5, and GH30 all have a common $(\beta/\alpha)_8$ -barrel structure and their

active sites contain two conserved carboxylic acid residues on β -strands 4 and 7, serving as the catalytic acid/base and nucleophile, respectively [100, 101]. Although structures are available for all three of these families, our focus will be on GH1 here, since the relatively closely related GH1 plant β -glycosidases show a high diversity of substrate specificities, the basis of which will be considered later. The lengths and subunit masses of these GH1 enzymes vary considerably, depending on the presence of auxiliary domains and redundant GH1 domains (as in human LPH), but the catalytic domain itself ranges from around 440 to 550 residues, depending on the lengths of the variable loops at the C-terminal ends of the β -strands of the $(\beta/\alpha)_8$ -barrel [102]. These monomers form a wide range of quaternary structures, including monomeric enzymes, dimers, tetramers, hexamers, octamers, and large aggregates.

The GH3 β -glucosidases and exoglucanases have a two-domain structure, a $(\beta/\alpha)_8$ -barrel followed by an α/β sandwich comprising a 6-stranded β -sheet sandwiched between three α -helices on either side [103]. The active site of GH3 enzymes is situated between the $(\beta/\alpha)_8$ and $(\alpha/\beta)_6$ sandwich domains, each of which contributes one catalytic carboxylate residue (Fig. 2). The catalytic nucleophile for

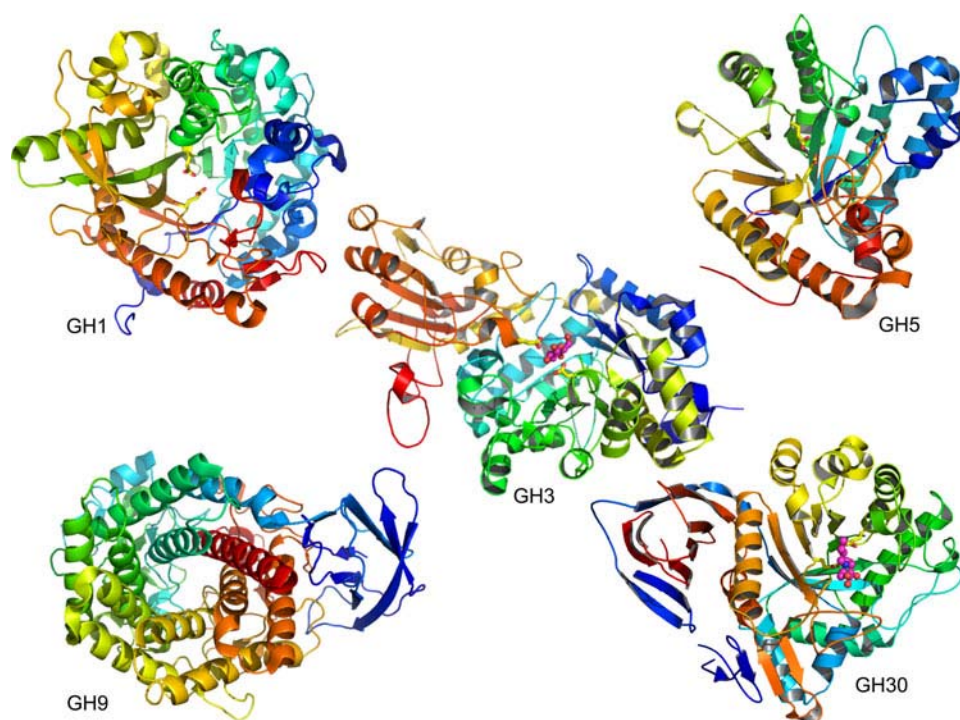


Fig. 2 Structures of β -glucosidases from different GH families. These include β -glucosidases or related enzymes from GH1 (*Zea mays* ZmGlu1, PDB code 1E1E), GH3 (*Hordeum vulgare* Exo I β -glucan glucohydrolase, PDB code 1EX1), GH5 (*Candida albicans* exo- β -(1,3)-glucanase Exg exoglucanase, PDB code 1CZ1), GH30 (*Homo sapiens*, acid β -glucosidase/glucocerebrosidase GBA1, PDB code 2V3D), and GH9 (*Vibrio parahaemolyticus*, putative exoglucanase, PDB code 3H7L). The structural cartoons are colored in a

spectrum from blue to red from their N- to C-termini, with the catalytic nucleophile and acid-base residues shown in stick for those enzymes in which they are known. The ligands shown are glucose in the GH3 barley ExoI and *N*-butyl-deoxynojirimycin in the GH30 human GBA1, both of which are shown with carbons in pink. The human GBA2 (bile acid β -glucosidase) shows low levels of sequence similarity to $(\alpha/\alpha)_6$ enzymes, suggesting its catalytic domain may be similar to the GH9 structure. Drawn with Pymol (DeLano)

barley Exo I is an aspartate at residue D285, which resides on the loop after β -strand 7 of the $(\beta/\alpha)_8$ barrel, while the catalytic acid/base is glutamate E491, which is on a long loop extending from the $(\alpha/\beta)_6$ sandwich domain [104].

Only a few GH9 proteins have been verified to be β -glucosidases [105, 106], as most proteins in this family are endoglycosidases. This family consists of $(\alpha/\alpha)_6$ barrels. Recently, the structure of a *Vibrio parahaemolyticus* protein with 69% amino acid sequence identity over 567 residues with the *Vibrio cholera* β -glucosidase, was determined (PDB accession 3H7L, Fig. 2). These family GH9 enzymes are the first β -glucosidases shown to act through an inverting mechanism (the prevailing mechanism in family GH9), which is unusual, since all other β -glucosidases described so far act through a retaining mechanism [106].

The human GBA2 and its relatives are not related to other β -glycosidases, but show weak similarities to certain $(\alpha/\alpha)_6$ enzymes in homology searches. The GBA2 sequence contains no secretory pathway signal sequence and a single putative transmembrane domain, but was predicted to have its N-terminus in the endoplasmic reticulum and C-terminus in the cytoplasm [11]. The position of this putative transmembrane α -helix falls in the middle of a sequence homologous to soluble $(\alpha/\alpha)_6$ amylohydrolase, chitobiose phosphorylase and α -L-rhamnosidase (3.2.1.40) enzymes, and the low confidence of the transmembrane topology prediction call this topology into question, but the enzyme is clearly associated with membranes by some means [10, 11].

Catalytic mechanisms

Glycoside hydrolases perform catalysis using two mechanisms, one with inversion and one with retention of chirality at the anomeric carbon [105]. Both of these mechanisms use a pair of acidic and nucleophilic residues, usually carboxylic acids, on either side of the sugar, approximately 5 Å apart in the retaining mechanism, and 10 Å apart in the inverting mechanism, in which a water molecule must fit between the catalytic base and the substrate. The GH9 β -glucosidases use an inverting mechanism, in which an activated water molecule makes a direct nucleophilic attack on the anomeric carbon to displace the aglycone in a single step, as shown in Fig. 3a [106]. The catalytic base extracts a proton from the incoming water molecule while the catalytic acid protonates the leaving group aglycone. In contrast, most β -glucosidases that have been characterized (i.e., GH1, GH3, and GH30 enzymes) are retaining enzymes, and they perform catalysis in two steps, glycosylation and deglycosylation (Fig. 3b). In glycosylation, the aglycone departs with the donation of a proton from the catalytic acid/base

and nucleophilic attack of the catalytic nucleophile on the anomeric carbon to yield an α -linked covalent enzyme-glycone intermediate. In the deglycosylation step, the process is reversed, as a water molecule attacks with basic assistance from the catalytic acid/base to displace the catalytic nucleophile from the glucose.

Both the glycosylation and deglycosylation steps are thought to pass through oxocarbenium-ion-like transition states. The glucose of the incoming substrate has sometimes been observed to be distorted into a 1S_3 skew boat as it moves toward the 4H_3 half-chair shape in the first transition state, although in other structures it is poorly defined by the electron density, apparently due to high mobility [108–112]. The structures of certain putative transition state mimics have also been solved in the active site and shown to have a structure close to the 4H_3 half-chair, although others appeared to inhibit by mechanisms other than transition state mimicry [108, 113–118].

The presence of the covalent intermediate was first demonstrated with the GH1 *Agrobacterium* sp. β -glucosidase by covalent labeling with 2,4-dinitrophenyl-2-deoxy-2-fluoroglucoside [119, 120]. In this inhibitor, the electronegative fluoride atom destabilizes the transition state for both half reactions, while the 2,4-dinitrophenylate provides an excellent leaving group to allow the glycosylation step to proceed rapidly. This traps the enzyme in the covalent intermediate and allows the catalytic nucleophile to be identified by tryptic digest and mass spectrometry. This covalent intermediate has also been observed in crystal structures for both the 2-F-glucoside and, in some cases, in the nonfluorinated glucosyl residue in certain acid/base catalyst mutants [104, 112, 121, 122]. The covalent inhibitor CBE has also been used to identify the catalytic nucleophile in some cases [104], but it is less specific and sometimes labels nearby amino acids. The acid/base catalyst of cassava β -glucosidase was also identified with a mechanism-based inhibitor, *N*-bromoacetyl- β -D-glucopyranosylamine [123], but most acid/base residues have been identified through homology, proximity to the glycosidic bond oxygen in crystal structures or site-directed mutagenesis [124].

The double-displacement mechanism for retaining β -glucosidases leads to predictions that mutants of these enzymes in which the acid/base or nucleophile is removed can be rescued by small nucleophiles and utilized for transglycosylation [107, 124–127]. When the acid/base of *Agrobacterium* sp. β -glucosidase was mutated to glycine (E170G), the hydrolysis of 2,4-dinitrophenyl β -D-glucoside (dNPGlc), which has a leaving group that does not require protonation ($pK_a = 3.96$), lost its pH dependence from 7 to 9 and could be rescued by various small nucleophiles, such as azide, which produced β -D-glucosyl azide [124]. This verified E170 as the catalytic acid/base and was consistent

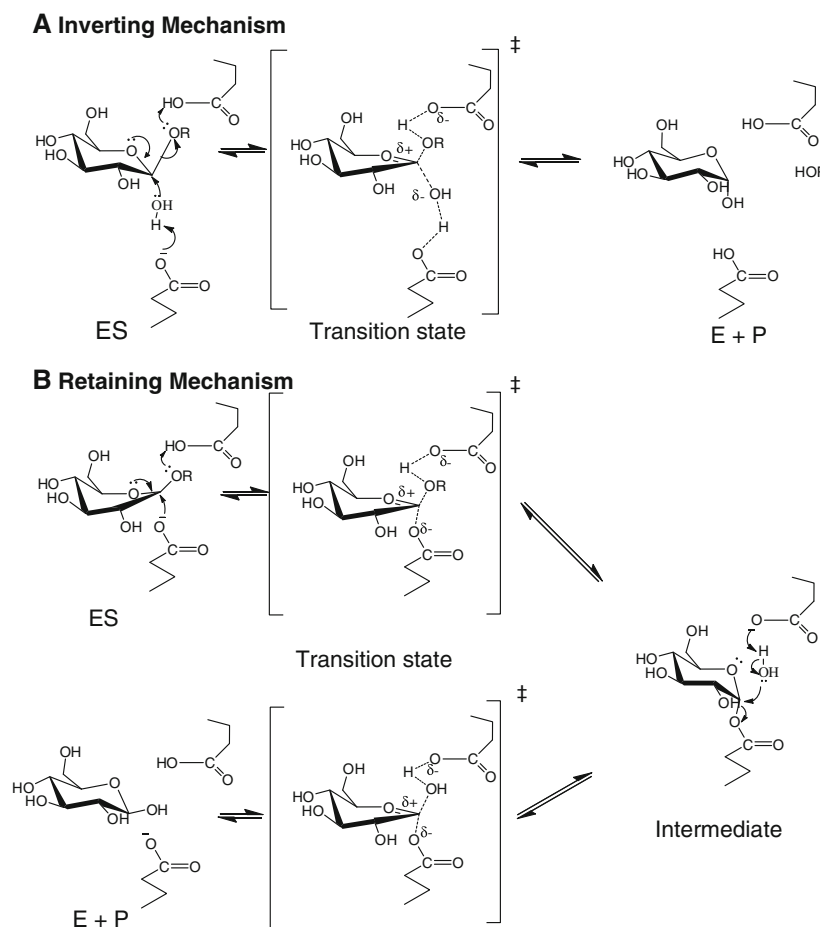


Fig. 3 Retaining catalytic mechanisms of inverting and retaining β -glucosidases. **a** The inverting mechanism that is seen in family GH9 glycoside hydrolases, including β -glucosidases. A single displacement of the aglycone by the water leads to an anomeric carbon with inverted chirality. **b** The commonly accepted mechanism for hydrolysis with retention of anomeric configuration as seen GH Clan A and family GH3 β -glucosidases. The glucosyl moiety is distorted into an 1S_3 skew boat upon binding to the enzyme in preparation to form the 4H_3 half chair conformation of the proposed transition state [107,

108]. The first step is glycosylation, in which the catalytic acid donates a proton to the leaving group, while the catalytic nucleophile attacks from the opposite side to form an α -linked intermediate. In the second, deglycosylation step, the catalytic base (the same carboxylate as the catalytic acid) extracts a proton from a water molecule, improving its nucleophilic power to attack at the anomeric carbon and displace the enzyme. Hydrolysis by either mechanism is equivalent in the organism, since mutarotation of the released glucose will lead to a racemic mixture of glucose in solution after a short time

with the double displacement mechanism in which the role of the catalytic acid has been circumvented. Similarly, conversion of the Abg catalytic nucleophile to a small nonnucleophilic amino acid, i.e., Ala, Ser, or Gly, resulted in an inactive enzyme that could be rescued by azide or fluoride to form α -D-glucosides, thereby converting a retaining enzyme to an inverting enzyme [125, 126]. Alternatively, the use of α -fluoroglucoside, in which the fluoride replaces the enzyme nucleophile in the covalent intermediate, allowed transfer of a β -linked glucosyl moiety onto a sugar or other alcohol. Since these nucleophile mutants have low hydrolytic activity, but relatively high transferase activity, they were designated glycosynthases [127]. Both the acid/base and the nucleophile mutants have potential uses in glycoconjugate synthesis.

Mechanism of substrate binding and specificity

Although the residues responsible for the hydrolytic mechanism are well characterized, how β -glucosidases recognize and interact with their substrates, which in large part determines their diverse functions, is less clear. GH1 enzymes are a prime model for these studies and the structures of 23 GH1 enzymes and their variants are currently available, including three from archaea, nine from bacteria, two from animals, one from a fungus, and eight from plants (CAZY website, [5]). The complexes of several of these enzymes with substrates, inhibitors and covalent intermediates are available, allowing in-depth analysis of residues likely to be critical to substrate and transition state binding. Although many of the prokaryotic enzymes show

rather similar and broad substrate specificities, the complexes of β -glycosidase S from the archaeon *Sulfolobus solfataricus* and β -glucosidase A from the eubacteria *Thermotoga maritima* with a range of inhibitors has provided a wealth of information on catalytic and inhibitory mechanisms [113–118]. In addition, site-directed mutagenesis of GH1 enzymes with and without experimentally determined structures has been done to test the roles of various residues.

The GH1 enzymes may have rather broad glycone specificity, for instance one enzyme may hydrolyze β -D-glucosides, β -D-fucosides, β -D-mannosides, β -D-galactosides and α -L-arabinosides, or may be specific for one or a few glycone sugars. Marana [128] analyzed GH1 specificity and concluded that a conserved glutamate, which bridges the glycone hydroxyl groups 4 and 6 in enzymes with β -glucosidase and β -galactosidase activities but is replaced in 6-phosphoglycosidases, is critical for the distinction between enzymes. However, it still remains to be determined how GH1 enzymes can be primarily β -glucosidases or β -mannosidases or show different ranges of allowed glycones, even though they bind the sugar with the same conserved residues [38, 65, 109, 112]. It is worth noting that binding of the aglycone has also been observed to affect the sugar binding position [109, 129], so residues more distant in the substrate binding pocket cannot be excluded from playing roles in glycone specificity.

The basis of the tremendous diversity in function of β -glucosidases, especially in plants, is the substrate aglycone specificity differences that determine their natural substrates. Structures of complexes of enzymes with inhibitors and mutant enzymes with substrates, along with mutagenesis and chimera studies comparing similar enzymes with divergent specificities, have suggested that the basis of aglycone specificity is complex. Although this includes mutagenesis and structural studies of human cytoplasmic β -glucosidase [13, 130], the plant GH1 enzymes have served as the primary model, due to their high diversity in aglycone specificity.

Maize ZmGlu1 and sorghum dhurrinase 1 (SbDhr1) are closely related, displaying 70% amino acid sequence identity, but have distinct specificities. ZmGlu1 has broad specificity, but cannot hydrolyze dhurrin, the natural substrate of SbDhr1, while SbDhr1 hydrolyzes only dhurrin. Studies of reciprocal ZmGlu1/SbDhr1 chimeric enzymes [131] and subsequent structural and site-directed mutagenesis studies [109–111, 129] indicated that the aglycone specificity determining sites are different in ZmGlu1 and SbDhr1. The structures of a catalytically inactive ZmGlu1 mutant (Glu1E191D) in complex with the natural substrate DIMBOAGlc (Fig. 4a), its free aglycone DIMBOA, and the unhydrolyzed competitive inhibitor dhurrin showed that the aglycone moiety of the substrate is sandwiched

between four aromatic residues, W378 on one side and F198, F205, and F466 on the other [109]. The 7-methoxy group of DIMBOA also has a hydrophobic contact with A467. All of these residues, except W378, are variable among β -glucosidases that differ in substrate specificity, which led to the conclusion that these sites and the active-site shape are the basis of aglycone binding specificity in β -glucosidases. In the case of Dhr1, the three phenylalanines are replaced with V196, L203, and S462, and the active site is smaller (Fig. 4b). A water-mediated H-bond between the dhurrin phenolic hydroxyl and Dhr1 S462 provides a more polar and stronger binding interaction than seen in ZmGlu1 [109–111]. This apparently led to a more stable 1S_3 skew boat conformation of the glucosyl residue, whereas in ZmGlu1 the conformation appeared to be variable, leading to poor electron density around the anomeric carbon. Mutagenesis of these aglycone-binding residues in the Zm60.8 isoform of ZmGlu1 confirmed their importance to hydrolysis of synthetic substrates [132].

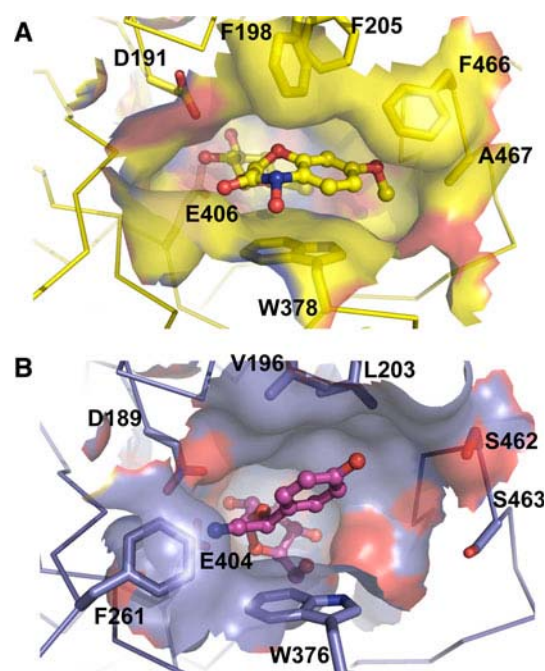


Fig. 4 Active site configurations of maize β -glucosidase 1 (Glu1, **a**) and sorghum dhurrinase 1 (Dhr1, **b**). The active sites of maize Glu1 and sorghum Dhr1 enzymes are shown for the structures of the Glu1 E189D mutant in complex with DIMBOA glucoside (PDB entry 1E56) and Dhr1 E191D mutant in complex with dhurrin (PDB entry 1VO3) [109, 111]. The sidechains of residues noted to interact with the aglycone are shown in stick representation behind the active site surface, which is colored as the underlying residues, which are colored with carbon in yellow for Glu1 and purple for Dhr1, nitrogen in blue, and oxygen in red. The ligands are shown in ball and stick representation with similar coloration. The Phe 261 residue, which narrows the active site in Dhr1, is also shown in front of the catalytic nucleophile Glu 404. Figure created with Pymol

However, the structural investigation of ZmGlu1 to SbDhr1 mutants and the subsequent structure of SbDhr1 and other GH1 hydrolases showed that the above-mentioned variable residues alone cannot designate substrate specificity [111, 131]. Although the Trp corresponding to ZmGlu1 W378 is nearly invariable in other plant GH1 enzymes, its positional variation was found to be critical for binding of substrates like dhurrin and strictosidine [81, 111, 131]. Even the closely related wheat β -glucosidase (TaGlu) was found to have different amino acids at the other aglycone-binding residues found in ZmGlu1, despite the fact that it also hydrolyzes DIMBOAGlc [52]. The oligosaccharide binding site in rice BGlu1 runs out of the active site in a different direction from that of DIMBOAGlc binding in ZmGlu1, so that L442, corresponding to ZmGlu F466, makes no interaction with the substrate, while N190, corresponding to ZmGlu F205, interacts only indirectly ([112], and PDB code 3F5K). Instead, N245 plays a key role in binding to the third glucosyl residue in cellobiosaccharides, while the corresponding residues in SbDhr1 (F261) and rice Os3BGlu6 (M251) appear to block off their active sites, which do not bind to such long substrates [71, 111]. Thus, a different, though overlapping, set of active site residues has been recruited to interact with the aglycone in each GH1 β -glucosidase that has been investigated, in contrast to the highly conserved glycone binding site.

Kinetic parameters for substrate hydrolysis

β -Glucosidases have variable kinetic parameters toward their substrates. The K_m values for natural substrates and other good substrates are typically 1 mM or less, but these values vary roughly 1,000-fold. Similarly, β -glucosidases have relatively low k_{cat} values ($\sim 300\text{ s}^{-1}$ or lower), which may be physiologically beneficial in some roles, but one must suspect the physiological relevance or the enzyme preparation if these values are too low. Comparison of the k_{cat}/K_m (efficiency coefficient) values is generally used to evaluate potential natural substrates, as two substrates with similar K_m values may have vastly different catalytic efficiencies. It is instructive to look at the published values for the hydrolysis of putative natural substrates by natural and recombinant preparations of human cytoplasmic β -glucosidases (Table 1). The plant-derived flavonoid luteolin 4'-glucoside and a synthetic fluorescent glycosceramide, C6-NBD-GalCer, give similar efficiency coefficients, although the former is hydrolyzed approximately five-fold faster than the latter. The natural glycosceramides that have longer acyl chains had much slower hydrolysis rates. Nonetheless, it was shown that RNAi knockdown of the cytoplasmic β -glucosidase resulted in an increase in glycosceramide concentrations, suggesting they may serve as natural substrates in the cell [18].

When β -glucosides with different efficiencies differ in the leaving-group ability of their aglycones, the rate-limiting step will be the glycosylation reaction, while either the glycosylation or deglycosylation step, or both, might be rate-limiting if substrates differ in the glycone. For example, many β -glucosidases hydrolyze *p*-nitrophenyl- β -D-fucoside (*p*NPFuc) with higher efficiency than *p*-nitrophenyl- β -D-glucoside (*p*NPGLc). In the case of maize β -glucosidases [50] and rice Os3BGlu7 [70], the K_m values are similar, and the V_{max} values for *p*NPFuc are clearly higher, however, in *Dalbergia* isoflavonoid β -glucosidases [46, 47, 133] and rice Os3BGlu6 [71], the V_{max} is similar or higher for *p*NPGLc and the K_m is ten-fold lower for *p*NPFuc.

Inhibitors and cofactors

β -Glucosidases are inhibited by transition-state sugar analogues, substrate analogue glycosides, and free aglycones of their substrates, as well as slowly hydrolyzed substrates, such as the fluoroglucosides mentioned earlier. Structural and thermodynamic analysis of 18 putative substrate analogues suggested many may not act as true transition-state analogues, but may nonetheless bind and inhibit *T. maritima* β -glucosidase [118]. Since the aglycone and glycone-binding pockets in the active site are distinct, sugar analogs shaped similar to the half-chair conformation of the transition state can bind to the glycone-binding site and inhibit the enzyme, whereas free aglycones may bind to the aglycone-binding site. Free glucose is a poor inhibitor (typically $K_i = 100\text{--}200\text{ mM}$) because glucose must be distorted toward the half-chair conformation for binding to the glycone-binding site, which is thought to require a portion of the energy of aglycone binding. In contrast, free aglycones can be potent competitive inhibitors because they bind to the aglycone-binding site without energetically unfavorable distortion.

Although most metal ions do not inhibit β -glucosidases, Ag^+ and Hg^{2+} , are potent β -glucosidase inhibitors, and inhibition by Cu^{2+} and Fe^{3+} has also been reported [134]. Although β -O-glucosidases are not known to require any cofactors, ascorbate is known to enhance the activity of β -S-glucosidases (myrosinases) by acting as a surrogate catalytic base [135]. The chelation of Zn^{2+} between the monomers in the biological dimer of myrosinase suggests that metal ions could act in stabilization of some GH1 enzymes.

pH and temperature optima and stability

β -Glucosidases show a range of pH optima and stabilities, depending on their source and amino acid sequence. The pH optima of most β -glucosidases range between pH 4 and 7.5,

depending on their source and cellular location, and they tend to be stable over a range of pH from 4 to 9. It is usually safe to store these enzymes at 0–4°C at pH 7–8, once major protease contaminants have been removed, but this should be tested with each enzyme. As with other proteins, pH extremes, copurifying proteases, and microbial contamination may result in degradation, although many β -glucosidases are resistant to proteases due to their tightly folded core structure. Nonetheless, proteolysis can result in recombinant proteins losing their purification tags and in purified active enzymes appearing to have two subunits on SDS-PAGE due to an internal cleavage that leaves the fold intact, in the authors' experience. Some β -glucosidases are resistant to denaturation by ionic detergents such as SDS, which allows extraction with buffers containing up to 3% SDS and zymogram development after SDS-PAGE when samples are applied without heating.

Thermostability and temperature optima vary greatly among enzymes. Mesophilic β -glucosidases may show highest activity at 30–65°C, but are generally inactivated at and above 55–70°C [46, 53, 58, 73]. High activity at temperatures above the extremes of the enzyme's natural environment is not physiologically relevant and may result in rapid heat denaturation, so assays are often run at 30–40°C. On the other hand, β -glucosidases from thermophilic bacteria, such as *T. maritima* BglA, may have temperature optima of over 100°C [136]. Engineering of a bacterial β -glucosidase to have the same N-terminal and C-terminal residues as *T. maritima* BglA, allowed hydrogen bonding between these termini and stabilized the enzyme, suggesting such interactions may be important for high stability [137]. This thermostability is also thought to be due to an increased number of proline residues, electrostatic bridges, and internal water molecules, and binding of more subunits in the quaternary structure compared to many mesophilic enzymes [138].

Summary and future prospects

The description of β -glucosidases in this review is limited in detail, due to the vast amounts of data that have been generated in the last several years. Nonetheless, we hope the reader will appreciate the wide variety of functions that β -glucosidases play in nature, from biomass degradation by microorganisms, to glycolipid and xenogenic β -glucoside breakdown in animals, to roles in defense, phytohormone regulation, cell wall metabolism, and secondary metabolism in plants, where the β -glucosidases have attained their greatest multiplicity and diversity of function. Although β -glucosidases like sorghum SbDhr1 can be very specific, others show overlapping ranges of activities for multiple substrates, such as the glucocerebrosidase and flavonoid

β -glucosidase activities of human β -glucosidases. It seems likely that β -glucosidases play many as yet undiscovered roles, as well as potential for many applications.

Although the catalytic mechanism is well understood for GH Clan A and GH3 β -glucosidases, the means by which their exact substrate specificity is established has proven to be divergent for even closely related β -glucosidases. This and the high multiplicity of putative β -glucosidase in plants limit the conclusions that can be drawn from genomic sequences as to the putative specificities and functions of new β -glucosidase homologs. Nonetheless, the insights gained from structural and mutagenic studies provide a starting point from which to investigate the functions of new β -glucosidases. As more substrate specificities and structures are determined, it should become more feasible to predict substrate specificity from the sequences of as yet uninvestigated β -glucosidases.

Acknowledgments Rodjana Opassiri and three anonymous reviewers are thanked for useful comments on the manuscript. JRKC was supported by Suranaree University of Technology and the Thailand Research Fund.

References

1. Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280:309–316
2. Henrissat B, Davies G (1997) Structural and sequence-based classification of glycosyl hydrolases. *Curr Opin Struct Biol* 7:637–644
3. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZY): an expert resource for glycogenomics. *Nucleic Acids Res* 37:D233–D238
4. Brunner R, Wirtz W, Rose JKC, Darbill AG, Govers F, Scheel D, Nurnberger T (2002) A β -glucosidase/xylosidase from the phytopathogenic oomycete, *Phytophthora infestans*. *Phytochemistry* 59:689–696
5. Opassiri R, Pomthong B, Akiyama T, Nakphaichit M, Onkoksoong T, Ketudat-Cairns M, Ketudat Cairns JR (2007) A stress-induced rice β -glucosidase represents a new subfamily of glycosyl hydrolase family 5 containing a fascin-like domain. *Biochem J* 408:241–249
6. Butters TD (2007) Gaucher disease. *Curr Opin Chem Biol* 11:412–418
7. Dvir H, Harel M, McCarthy AA, Toker L, Silman I, Futerman AH, Sussman JL (2003) X-ray structure of human acid- β -glucosidase, the defective enzyme in Gaucher disease. *EMBO Rep* 4:704–709
8. Liou B, Kazimierczuk A, Zhang M, Scott CR, Hedge RS, Grabowski GA (2006) Analyses of variant acid β -glucosidases: effects of Gaucher disease mutations. *J Biol Chem* 281:4242–4253
9. Lieberman RL, Wustman BA, Huertas P, Powe AC, Pine CW, Khanna R, Schlossmacher MG, Ringe D, Petsko GA (2007) Structure of acid β -glucosidase with pharmacological chaperone provides insight into Gaucher disease. *Nat Chem Biol* 3:101–107
10. Yildiz Y, Matern H, Thompson B, Allegood JC, Warren RL, Ramirez DMO, Hammer RE, Hamra FK, Matern S, Russell DW

- (2006) Mutation of β -glucosidase 2 causes glycolipid storage disease and impaired male fertility. *J Clin Invest* 116:2985–2994
11. Boot RG, Verhoeck M, Donker-Koopman W, Strijland A, van Marle J, Overkleef HS, Wennekes T, Aerts JM (2007) Identification of the non-lysosomal glucosylceramidase as beta-glucosidase 2. *J Biol Chem* 282:1305–1312
 12. Matern H, Boermans H, Lottspeich F, Matern S (2001) Molecular cloning and expression of human bile acid beta-glucosidase. *J Biol Chem* 276:37929–37933
 13. Tribolo S, Berrin J-G, Kroon PA, Czjzek M, Juge N (2007) The structure of human cytoplasmic β -glucosidase unravels substrate aglycone specificity of a family 1 glycoside hydrolase. *J Mol Biol* 370:964–975
 14. Mantei N, Villa M, Enzler T, Wacker H, Boll W, James P, Hunziker W, Semenza G (1988) Complete primary structure of human and rabbit lactase-phlorizin hydrolase: implications for biosynthesis, membrane anchoring and evolution of the enzyme. *EMBO J* 7:2705–2713
 15. Arribas JC, Herrero AG, Martín-Lomas M, Cañada FJ, He S, Withers SG (2000) Differential mechanism-based labeling and unequivocal activity assignment of the two active sites of intestinal lactase/phlorizin hydrolase. *Eur J Biochem* 267:6996–7005
 16. Day AJ, Canada FJ, Diaz JC, Kroon PA, Mclauchlan R, Faulds CB, Plumb GW, Morgan MR, Williamson G (2000) Dietary flavonoid and isoflavone glycosides are hydrolysed by the lactase site of lactase phlorizin hydrolase. *FEBS Lett* 468:166–170
 17. Daniels LB, Coyle PJ, Chiao YB, Glew RH, Labow RS (1981) Purification and characterization of a cytosolic broad specificity beta-glucosidase from human liver. *J Biol Chem* 256:13004–13013
 18. Hayashi Y, Okino N, Kakuta Y, Shikanai T, Tani M, Narimatsu H, Ito M (2007) Klotho-related protein is a novel cytosolic neutral beta-glycosylceramidase. *J Biol Chem* 282:30889–30900
 19. Day AJ, DuPont MS, Ridley S, Rhodes M, Rhodes MJ, Morgan MR, Williamson G (1998) Deglycosylation of flavonoid and isoflavonoid glycosides by human small intestine and liver beta-glucosidase activity. *FEBS Lett* 436:71–75
 20. Berrin J-G, McLauchlan WR, Needs P, Williamson G, Puigserver A, Kroon PA, Juge N (2002) Functional expression of human liver cytosolic β -glucosidase in *Pichia pastoris*. Insights into its role in the metabolism of dietary glucosides. *Eur J Biochem* 269:249–258
 21. Chang Q, Hoefs S, van der Kemp AW, Topala CN, Bindels RJ, Hoenderop JG (2005) The beta-glucuronidase klotho hydrolyzes and activates the TRPV5 channel. *Science* 310:490–493
 22. Kuro-o M, Matsumura Y, Aizawa H, Kawaguchi H, Suga T, Utsuki T, Ohyama Y, Kurabayashi M, Kaname T, Kume E, Iwasaki H, Iida A, Shiraki-Iida T, Nishikawa S, Nagai R, Nabeshima YI (1997) Mutation of the mouse klotho gene leads to a syndrome resembling ageing. *Nature* 390:45–51
 23. Nabeshima Y, Imura H (2008) α -Klotho: a regulator that integrates calcium homeostasis. *Am J Nephrol* 28:455–464
 24. Ito S, Fujimori T, Hayashizaki Y, Nabeshima Y (2002) Identification of a novel mouse membrane-bound family 1 glycosidase-like protein, which carries an atypical active site structure. *Biochim Biophys Acta* 1576:341–345
 25. Shiraki-Iida T, Aizawa H, Matsumura Y, Sekine S, Iida A, Anazawa H, Nagai R, Kuro-o M, Nabeshima Y (1998) Structure of the mouse klotho gene and its two transcripts encoding membrane and secreted protein. *FEBS Lett* 424:6–10
 26. Zagrobelny M, Bak S, Møller BL (2008) Cyanogenesis in plants and arthropods. *Phytochemistry* 69:1457–1468
 27. Marana SR, Jacobs-Lorena M, Terra WR, Ferreira C (2001) Amino acid residues involved in substrate binding and catalysis in an insect digestive β -glycosidase. *Biochim Biophys Acta* 1545:41–52
 28. Ferreira AHP, Marana SR, Terra WR, Ferreira C (2001) Purification, molecular cloning, and properties of a β -glycosidase isolated from midgut lumen of *Tenebrio molitor* (Coleoptera) larvae. *Insect Biochem Mol Biol* 31:1065–1076
 29. Jones AME, Bridges M, Bones AM, Cole R, Rossiter JT (2001) Purification and characterisation of a non-plant myrosinase from the cabbage aphid *Brevicoryne brassicae*. *Insect Biochem Mol Biol* 31:1–5
 30. Malboobi MA, Lefebvre DD (1997) A phosphate-starvation inducible β -glucosidase gene (*psr3.2*) isolated from *Arabidopsis thaliana* is a member of a distinct subfamily of the BGA family. *Plant Mol Biol* 34:57–68
 31. van de Ven WT, LeVesque CS, Perring TM, Walling LL (2000) Local and systemic changes in squash gene expression in response to silver winged whitefly feeding. *Plant Cell* 12:1409–1423
 32. Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell* 13:889–905
 33. Thorlby G, Fourier N, Warren G (2004) The SENSITIVE TO FREEZING2 gene, required for freezing tolerance in *Arabidopsis thaliana*, encodes a beta-glucosidase. *Plant Cell* 16:2192–2203
 34. Lipka V, Dittgen J, Bednarek P, Bhat R, Wiermer M, Stein M, Landtag J, Brandt W, Rosahl S, Scheel D, Llorente F, Molina A, Parker J, Somerville S, Schulze-Lefert P (2005) Pre- and postinvasion defenses both contribute to nonhost resistance in *Arabidopsis*. *Science* 310:1180–1183
 35. Xu Z, Escamilla-Treviño LL, Zeng L, Lalgondar M, Bevan DR, Winkler BSJ, Mohamed A, Cheng C, Shih M, Poulton JE, Esen A (2004) Functional genomic analysis of *Arabidopsis thaliana* glycoside hydrolase family 1. *Plant Mol Biol* 55:343–367
 36. Opasiri R, Pomthong B, Okoksoong T, Akiyama T, Esen A, Ketudat Cairns JR (2006) Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β -glucosidase. *BMC Plant Biol* 6:33
 37. Marques AR, Coutinho PM, Videira P, Fialho AM, S-Correia I (2003) *Sphingomonas paucimobilis* beta-glucosidase Bgl1: a member of a new bacterial subfamily in glycoside hydrolase family 1. *Biochem J* 370:793–804
 38. Kuntothom T, Luang S, Harvey AJ, Fincher GB, Opasiri R, Hrmova M, Ketudat Cairns JR (2009) Rice family GH1 glycosyl hydrolases with β -D-glucosidase and β -D-mannosidase activities. *Arch Biochem Biophys* 491:84–95
 39. Arthan D, Kittakoop P, Esen A, Svasti J (2006) Furostanol glycoside 26-O- β -glucosidase from the leaves of *Solanum torvum*. *Phytochemistry* 67:27–33
 40. Niemeyer HM (1988) Hydroxamic acids (4-hydroxy-1,4-benzoxazin-3-ones), defense chemicals in the Gramineae. *Phytochemistry* 27:3349–3358
 41. Poulton JE (1990) Cyanogenesis in plants. *Plant Physiol* 94:401–405
 42. Morant AV, Jørgensen K, Jørgensen C, Paquette SM, Sanchéz-Perez R, Møller BL, Bak S (2008) β -Glucosidases as detonators of plant chemical defense. *Phytochemistry* 69:1795–1813
 43. Sherameti I, Venus Y, Drzewiecki C, Tripathi W, Dan VM, Nitz I, Varma A, Grundler F, Oelmüller R (2008) PYK10, a β -glucosidase located in the endoplasmatic reticulum, is crucial for the beneficial interaction between *Arabidopsis thaliana* and the endophytic fungus *Piriformospora indica*. *Plant J* 54:428–439
 44. Morant AV, Bjarnholt N, Kragh ME, Kjaergaard CH, Jørgensen K, Paquette SM, Piotrowski M, Imbert A, Olsen CE, Møller BL, Bak S (2008) The beta-glucosidases responsible for

- bioactivation of hydroxynitrile glucosides in *Lotus japonicus*. *Plant Physiol* 147:1072–1091
45. Suzuki H, Takahashi S, Watanabe R, Fukushima Y, Fujita N, Noguchi A, Yokoyama R, Nishitani K, Nishino T, Nakayama T (2006) An isoflavone conjugate-hydrolyzing β -glucosidase from the roots of soybean (*Glycine max*) seedlings. *J Biol Chem* 281:30251–30259
46. Chuankhayan P, Hua Y, Svasti J, Sakdarat S, Sullivan PA, Ketudat Cairns JR (2005) Purification of an isoflavonoid 7-O- β -apiosyl-glucoside β -glycosidase and its substrates from *Dalbergia nigrescens* Kurz. *Phytochemistry* 66:1880–1889
47. Chuankhayan P, Rimlumduan T, Tantanuch W, Mothong N, Kongsaree PT, Metheenukul P, Svasti J, Jensen ON, Ketudat Cairns JR (2007) Functional and structural differences between isoflavonoid β -glycosidases from *Dalbergia* sp. *Arch Biochem Biophys* 468:205–216
48. Naoumkina M, Farag MA, Sumner LW, Tang Y, Liu CJ, Dixon RA (2007) Different mechanisms for phytoalexin induction by pathogen and wound signals in *Medicago truncatula*. *Proc Natl Acad Sci USA* 104:17909–17915
49. Esen A (1992) Purification and partial characterization of maize (*Zea mays* L.) β -glucosidase. *Plant Physiol* 98:174–182
50. Babcock GD, Esen A (1994) Substrate specificity of maize β -glucosidase. *Plant Sci* 101:31–39
51. Nikus J, Daniel G, Jonsson LM (2001) Subcellular localization of beta-glucosidase in rye, maize and wheat seedlings. *Plant Physiol* 111:466–472
52. Sue M, Yamazaki K, Yajima S, Nomura T, Matsukawa T, Iwamura H, Miyamoto T (2006) Molecular and structural characterization of hexameric beta-D-glucosidases in wheat and rye. *Plant Physiol* 141:1237–1247
53. Nisius A (1988) The stroma centre in *Avena* plastids: an aggregation of β -glucosidase responsible for the activation of oat-leaf saponins. *Planta* 173:474–481
54. Ahn YO, Shimizu B, Sakata K, Gantulga D, Zhou Z, Bevan DR, Esen A (2010) Scopulin-hydrolyzing β -glucosidases in the roots of *Arabidopsis*. *Plant Cell Physiol* 51:131–143
55. Hara-Nishimura I, Matsushima R (2003) A wound-inducible organelle derived from endoplasmic reticulum: a plant strategy against environmental stress? *Curr Opin Plant Biol* 6:538–588
56. Matsushima R, Kondo M, Nishimura M, Hara-Nishimura I (2003) A novel ER-derived compartment, the ER body, selectively accumulates a β -glucosidase with an ER retention signal in *Arabidopsis*. *Plant J* 33:493–502
57. Bednarek P, Pišlewski-Bednarek M, Svatoš A, Schneider B, Doubek J, Mansurova M, Humphry M, Consonni C, Panstruga R, Sanchez-Vallet A, Molina A, Schulze-Lefert P (2009) A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* 323:101–106
58. Nagano AJ, Matsushima R, Hara-Nishimura I (2005) Activation of an ER-body-localized β -glucosidase via a cytosolic binding partner in damaged tissues of *Arabidopsis thaliana*. *Plant Cell Physiol* 46:1140–1148
59. Falk A, Taipalensuu J, Ek B, Lenman M, Rask L (1995) Characterization of rapeseed myrosinase-binding protein. *Planta* 195:387–395
60. Esen A, Blanchard DJ (2000) A specific β -glucosidase-aggregating factor (BGAF) is responsible for the β -glucosidase null phenotype in maize. *Plant Physiol* 122:563–572
61. Blanchard DJ, Cicek M, Chen J, Esen A (2001) Identification of beta-glucosidase aggregating factor (BGAF) and mapping of BGAF binding regions on maize beta-glucosidase. *J Biol Chem* 276:11895–11901
62. Kittur FS, Lalgondar M, Yu HY, Bevan DR, Esen A (2007) Maize β -glucosidase-aggregating factor is a polyspecific jacalin-related chimeric lectin, and its lectin domain is responsible for β -glucosidase aggregation. *J Biol Chem* 282:7299–7311
63. Nagano AJ, Fukao Y, Fujiwara M, Nishimura M, Hara-Nishimura I (2008) Antagonistic jacalin-related lectins regulate the size of ER body-type β -glucosidase complexes in *Arabidopsis thaliana*. *Plant Cell Physiol* 49:969–980
64. Leah R, Kigel J, Svedsen I, Mundy J (1995) Biochemical and molecular characterization of a barley seed β -glucosidase. *J Biol Chem* 270:15789–15797
65. Hrmova M, Harvey AJ, Wang J, Shirley NJ, Jones GP, Stone BA, Hoj PB, Fincher GB (1996) Barley β -D-glucan exohydrolases with β -D-glucosidase activity. *J Biol Chem* 271:5277–5286
66. Hrmova M, MacGregor EA, Biely P, Stewart RJ, Fincher GB (1998) Substrate binding and catalytic mechanism of a barley β -D-glucosidase/(1, 4)- β -D-glucan exohydrolase. *J Biol Chem* 273:11134–11143
67. Hrmova M, Burton RA, Biely P, Lahnstein J, Fincher GB (2006) Hydrolysis of (1,4)- β -D-mannans in barley (*Hordeum vulgare* L.) is mediated by the concerted action of (1,4)- β -D-mannan endohydrolase and β -D-mannosidase. *Biochem J* 399:77–90
68. Akiyama T, Kaku H, Shibuya N (1998) A cell wall-bound β -glucosidase from germinated rice: purification and properties. *Phytochemistry* 48:49–54
69. Opasiri R, Ketudat Cairns JR, Akiyama T, Wara-Aswapati O, Svasti J, Esen A (2003) Characterization of a rice β -glucosidase highly expressed in flower and germinating shoot. *Plant Sci* 165:627–638
70. Opasiri R, Hua Y, Wara-Aswapati O, Akiyama T, Svasti J, Esen A, Ketudat Cairns JR (2004) β -Glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem J* 379:125–131
71. Seshadri S, Akiyama T, Opasiri R, Kuaprasert B, Ketudat Cairns J (2009) Structural and enzymatic characterization of Os3BGlu6, a rice β -glucosidase hydrolyzing hydrophobic glycosides and (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 2)-linked disaccharides. *Plant Physiol* 151:47–58
72. Hösel W, Surholt E, Borgmann E (1978) Characterization of β -glucosidase isoenzymes possibly involved in lignification from chick pea (*Cicer arietinum* L.) cell suspension culture. *Eur J Biochem* 84:487–492
73. Dhamawardhana DP, Ellis BE, Carlson JE (1995) A β -glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. *Plant Physiol* 107:331–339
74. Escamilla-Treviño LL, Chen W, Card ML, Shih MC, Cheng CL, Poulton JE (2006) *Arabidopsis thaliana* β -Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. *Phytochemistry* 67:1651–1660
75. Schliemann W (1984) Hydrolysis of conjugated gibberellins by β -glucosidases from dwarf rice (*Oryza sativa* L. cv. Tan-gin-bozu). *J Plant Physiol* 116:123–132
76. Brzobohatý B, Moore I, Kristoffersen P, Bako L, Campos N, Schell J, Palme K (1993) Release of active cytokinin by a β -glucosidase localized to the maize root meristem. *Science* 262:1051–1054
77. Dietz K-J, Sauter A, Wichert K, Messdaghi D, Hartung W (2000) Extracellular β -glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves. *J Exp Bot* 51:937–944
78. Jakubowska A, Kawalczyk S (2005) A specific enzyme hydrolyzing 6-O(4-O)-indole-3-ylacetyl- β -D-glucose in immature kernels of *Zea mays*. *J Plant Physiol* 162:207–213
79. Lee KH, Piao HL, Kim HY, Choi SM, Jiang F, Hartung W, Hwang I, Kwak JM, Lee IJ (2006) Activation of glucosidase via stress-induced polymerization rapidly increases active pools of abscisic acid. *Cell* 126:1109–1120

80. Stöckigt J, Zenk MH (1977) Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpenoid indole alkaloids. *J Chem Soc Chem Commun* 1977:646–648
81. Barleben L, Panjikar S, Ruppert M, Koepke J, Stöckigt J (2007) Molecular architecture of strictosidine glucosidase: the gateway to the biosynthesis of the monoterpenoid indole alkaloid family. *Plant Cell* 19:2886–2897
82. Warzecha H, Gerasimenko I, Kutchan TM, Stöckigt J (2000) Molecular cloning and functional bacterial expression of a plant β -glucosidase specifically involved in alkaloid biosynthesis. *Phytochemistry* 54:657–666
83. Nomura T, Quesada AL, Kutchan TM (2008) The new beta-D-glucosidase in terpenoid-isoquinoline alkaloid biosynthesis in *Psychotria ipecacuanha*. *J Biol Chem* 283:34650–34659
84. Reuveni M, Sagi Z, Eynon D, Hetzroni A (1999) β -Glucosidase activity is involved in scent production in Narcissus flowers. *Plant Sci* 147:19–24
85. Mattiacci L, Dicke M, Posthumus MA (1995) Beta-glucosidase: an elicitor of herbivore-induced plant odor that attracts host-searching parasitic wasps. *Proc Natl Acad Sci USA* 92:2036–2040
86. Gilbert HJ, Stålbrand H, Brumer H (2008) How the walls come tumbling down: recent structural biochemistry of plant polysaccharide degradation. *Curr Opin Plant Biol* 11:338–348
87. Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol* 2:541–551
88. Carvalho AL, Dias FM, Nagy T, Prates JA, Proctor MR, Smith N, Bayer EA, Davies GJ, Ferreira LM, Romão MJ, Fontes CM, Gilbert HJ (2007) Evidence for a dual binding mode of dockerin modules to cohesins. *Proc Natl Acad Sci USA* 2007(104):3089–3094
89. Lyman ES, Li B, Renganathan V (1995) Purification and characterization of a cellulose-binding β -glucosidase from cellulose-degrading cultures of *Phanerochaete chrysosporium*. *Appl Environ Microbiol* 61:2976–2980
90. Igarashi K, Tani T, Kawal R, Samejima M (2003) Family 3 β -glucosidase from cellulose-degrading culture of the white-rot fungus *Phanerochaete chrysosporium*. *J Biosci Bioeng* 95:572–576
91. Tsukada T, Igarashi K, Yoshida M, Samejima M (2006) Molecular cloning and characterization of two intracellular β -glucosidases belonging to glycoside hydrolase family 1 from the basidiomycete *Phanerochaete chrysosporium*. *Appl Microbiol Biotechnol* 73:807–814
92. Günata Z (2003) Flavor enhancement in fruit juices and derived beverages by exogenous glycosidases and consequences of the use of enzyme preparations. In: Whitaker JR, Voragen AGJ, Wong DWS (eds) *Handbook of food enzymology*. Marcel Dekker Inc., New York, pp 303–330
93. Esen A (2003) β -Glucosidases. In: Whitaker JR, Voragen AGJ, Wong DWS (eds) *Handbook of food enzymology*. Marcel Dekker Inc., New York, pp 791–804
94. Yasumoto K, Tsuji H, Iwami K, Mitsuda H (1977) Isolation from rice bran of a bound form of vitamin B6 and its identification as 5'-O- β -D-glucopyranosyl-pyridoxine. *Agric Biol Chem* 41:1061–1067
95. Chuankhayon P, Rimlumduan T, Svasti J, Ketudat Cairns JR (2007) Hydrolysis of soybean isoflavonoid glycosides by *Dalbergia* β -glucosidases. *J Agric Food Chem* 55:2407–2412
96. Ismail B, Hayes K (2005) β -Glycosidase activity toward different glycosidic forms of isoflavones. *J Agric Food Chem* 53:4918–4924
97. Higdon JV, Delage B, Williams DE, Dashwood RH (2007) Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis. *Pharmacol Res* 55:224–236
98. Sánchez-Pérez R, Jørgensen K, Olsen CE, Dicenta F, Möller BL (2008) Bitterness in almonds. *Plant Physiol* 146:1040–1052
99. Crout DH, Vic G (1998) Glycosidases and glycosynthetases in glycoside and oligosaccharide synthesis. *Curr Opin Chem Biol* 2:98–111
100. Henrissat B, Callebaut I, Fabrega S, Lehn P, Mornon JP, Davies G (1995) Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc Natl Acad Sci USA* 92:7090–7094
101. Jenkins J, Lo Leggio L, Harris G, Pickersgill R (1995) Beta-glucosidase, beta-galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold beta/alpha architecture and with two conserved glutamates near the carboxy-terminal ends of beta-strands four and seven. *FEBS Lett* 362:281–285
102. Sanz-Aparicio J, Hermoso JA, Martinez-Ripoll M, Lequerica JL, Polaina J (1998) Crystal structure of beta-glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family 1 glycosyl hydrolases. *J Mol Biol* 275:491–502
103. Varghese JN, Hrmova M, Fincher GB (1999) Three-dimensional structure of a barley β -D-glucan exohydrolase; a family 3 glycosyl hydrolase. *Structure* 7:179–190
104. Hrmova M, Varghese JN, De Gori R, Smith BJ, Driguez H, Fincher GB (2001) Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of a plant beta-D-glucan glucosylhydrolase. *Structure* 9:1005–1016
105. Park JK, Wang L-X, Patel HV, Roseman S (2002) Molecular cloning and characterization of a unique β -glucosidase from *Vibrio cholerae*. *J Biol Chem* 277:29555–29560
106. Qi M, Jun H-S, Forsbert CW (2008) Cel9D, an atypical 1, 4- β -D-glucan glucosylhydrolase from *Fibrobacter succinogenes*: characteristics, catalytic residues, and synergistic interactions with other cellulases. *J Bacteriol* 190:1976–1984
107. Rye CS, Withers SG (2000) Glycosidase mechanisms. *Curr Opin Chem Biol* 4:573–580
108. Davies GJ, Ducros VM-A, Varrot A, Zechel DL (2003) Mapping the conformational itinerary of β -glucosidases by X-ray crystallography. *Biochem Soc Trans* 31:523–527
109. Czjzek M, Cicek M, Zamboni V, Bevan DR, Henrissat B, Esen A (2000) The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, -DIMBOAGlc, and dhurrin complexes. *Proc Natl Acad Sci USA* 97:13555–13560
110. Czjzek M, Cicek M, Zamboni V, Burmeister WP, Bevan DR, Henrissat B, Esen A (2001) Crystal structure of a monocotyledon (maize ZMGLu1) β -glucosidase and a model of its complex with p-nitrophenyl β -D-thiogluconate. *Biochem J* 354:37–46
111. Verdoucq L, Morinière J, Bevan DR, Esen A, Vasella A, Henrissat B, Czjzek M (2004) Structural determinants of substrate specificity in family 1 β -glucosidases: novel insights from the crystal structure of sorghum dhurrinase-1, a plant β -glucosidase with strict specificity, in complex with its natural substrate. *J Biol Chem* 279:31796–31803
112. Chuenchor W, Pengthaisong S, Robinson RC, Yuvaniyama J, Oonanant W, Bevan DR, Esen A, Chen C-J, Opasiri R, Svasti J, Ketudat Cairns JR (2008) Structural insights into rice BGLu1 β -glucosidase oligosaccharide hydrolysis and transglycosylation. *J Mol Biol* 377:1200–1215
113. Zechel DL, Boraston AB, Gloster TM, Boraston CM, Macdonald JM, Tilbrook DMG, Stick RV, Davies GJ (2003) Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β -glucosidases. *J Am Chem Soc* 125:14313–14323
114. Vincent F, Gloster TM, Macdonald J, Morland C, Stick RV, Dias FM, Prates JA, Fontes CM, Gilbert HJ, Davies GJ (2004) Common inhibition of both beta-glucosidases and beta-mannosidases by isofagomine lactam reflects different

- conformational itineraries for pyranoside hydrolysis. *ChemBiochem* 5:1596–1599
115. Gloster TM, Macdonald JM, Tarling CA, Stick RV, Withers SG, Davies GJ (2004) Structural, thermodynamic, and kinetic analyses of tetrahydrooxazine-derived inhibitors bound to β -glucosidases. *J Biol Chem* 279:29236–49242
116. Gloster TM, Madsen R, Davies GJ (2006) Dissection of conformationally restricted inhibitors binding to a beta-glucosidase. *ChemBiochem* 7:738–742
117. Gloster TM, Roberts S, Perugino G, Rossi M, Moracci M, Panday N, Terinek M, Vasella A, Davies GJ (2006) Structural, kinetic, and thermodynamic analysis of glucoimidazole-derived glycosidase inhibitors. *Biochemistry* 45:11879–11884
118. Gloster TM, Meloncelli P, Stick RV, Zechel D, Vasella A, Davies GJ (2007) Glycosidase inhibition: an assessment of the binding of 18 putative transition-state mimics. *J Am Chem Soc* 129:2345–2354
119. Withers SG, Street IP, Bird P, Dolphin DH (1987) 2-Deoxy-2-fluoroglucosides: a novel class of mechanism based inhibitors. *J Am Chem Soc* 109:7530–7531
120. Withers SG, Warren RAJ, Street IP, Rupitz K, Kempton JB, Aebersold R (1990) Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a retaining glycosidase. *J Am Chem Soc* 112:5887–5889
121. Burmeister WP, Cottaz S, Driguez H, Iori R, Palmieri S, Henrissat B (1997) 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* 5:663–675
122. Noguchi J, Hayashi Y, Baba Y, Okino N, Kimura M, Ito M, Kakuta Y (2008) Crystal structure of the covalent intermediate of human cytosolic beta-glucosidase. *Biochem Biophys Res Commun* 374:549–552
123. Keresztessy Z, Kiss L, Hughes MA (1994) Investigation of the active site of the cyanogenic beta-D-glucosidase (linamarase) from *Manihot esculenta* Crantz (cassava). II. Identification of Glu-198 as an active site carboxylate group with acid catalytic function. *Arch Biochem Biophys* 315:323–330
124. Wang Q, Trimbur D, Graham R, Warren RA, Withers SG (1995) Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants. *Biochemistry* 34:14554–14562
125. Wang Q, Graham RW, Trimbur D, Warren RAJ, Withers SG (1994) Changing enzymic reaction mechanisms by mutagenesis: conversion of a retaining glucosidase to an inverting enzyme. *J Am Chem Soc* 116:11594–11595
126. Mackenzie LF, Wang Q, Warren RAJ, Withers SG (1998) Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J Am Chem Soc* 120:5583–5584
127. Ly HD, Withers SG (1999) Mutagenesis of glycosidases. *Annu Rev Biochem* 68:487–522
128. Marana SR (2006) Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUBMB Life* 58:63–73
129. Verdoucq L, Czjzek M, Moriniere J, Bevan DR, Esen A (2003) Mutational and structural analysis of aglycone specificity in maize and sorghum β -glucosidases. *J Biol Chem* 278:25055–25062
130. Berrin J-G, Czjzek M, Kroon PA, McLauchlan WR, Puigserver A, Williamson G, Juge N (2003) Substrate (aglycone) specificity of human cytosolic β -glucosidase. *Biochem J* 373:41–48
131. Cicek M, Blanchard D, Bevan DR, Esen A (2000) The aglycone specificity-determining sites are different in 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)-glucosidase (maize beta-glucosidase) and dhurrinase (sorghum beta-glucosidase). *J Biol Chem* 275:20002–20011
132. Zouhar J, Vévodová J, Marek J, Damborský J, Su X-D, Bryzobohatý B (2001) Insights into the functional architecture of the catalytic center of a maize β -glucosidase Zm-p60.1. *Plant Physiol* 127:973–985
133. Srisomsap C, Svasti J, Surarit R, Champattanachai V, Sawangareetrakul P, Boonpuan K, Subhasitanont P, Chokchai-chamnankit D (1996) Isolation and characterization of an enzyme with beta-glucosidase and beta-fucosidase activities from *Dalbergia cochinchinensis* Pierre. *J Biochem* 119:585–590
134. Zollner H (1989) Handbook of enzyme inhibitors. VCH, Weinheim, pp 94–95
135. Burmeister WP, Cottaz S, Rollin P, Vasella A, Henrissat B (2000) High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and substitutes for the function of the catalytic base. *J Biol Chem* 275:39385–39393
136. Kengen SW, Luesink EJ, Stams AJ, Zehnder AJ (1993) Purification and characterization of an extremely thermostable beta-glucosidase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Eur J Biochem* 213:305–312
137. Nam KH, Kim S-J, Kim M-Y, Kim JH, Yeo Y-S, Lee C-M, Jun H-K, Hwang KY (2008) Crystal structure of engineered beta-glucosidase from soil metagenome. *Proteins* 73:788–793
138. Chi YI, Martinez-Cruz LA, Jancarik J, Swanson RV, Robertson DE, Kim SH (1999) Crystal structure of the beta-glycosidase from the hyperthermophile *Thermosphaera aggregans*: insights into its activity and thermostability. *FEBS Lett* 445:375–383

Article

Crystallisation of Wild-Type and Variant Forms of a Recombinant Plant Enzyme β -D-Glucan Glucohydrolase from Barley (*Hordeum vulgare* L.) and Preliminary X-ray Analysis

Sukanya Luang¹, James R. Ketudat Cairns¹, Victor A. Streltsov² and Maria Hrmova^{3,*}

¹ School of Biochemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand; E-Mail: jeab_bc@hotmail.com (S.L.); cairns@sut.ac.th (J.R.K.C.)

² Molecular and Health Technologies, CSIRO-Commonwealth Scientific Research Organization, Victoria 3052, Australia; E-Mail: victor.streltsov@csiro.au

³ Australian Centre for Plant Functional Genomics, University of Adelaide, Waite Campus, Glen Osmond, South Australia 5064, Australia

* Author to whom correspondence should be addressed; E-Mail: maria.hrmova@adelaide.edu.au; Tel.: +61-8-830-372-80; Fax: +61-8-830-371-02.

Received: 28 May 2010; in revised form: 16 July 2010 / Accepted: 16 July 2010 /

Published: 19 July 2010

Abstract: Wild-type and variant crystals of a recombinant enzyme β -D-glucan glucohydrolase from barley (*Hordeum vulgare* L.) were obtained by macroseeding and cross-seeding with microcrystals obtained from native plant protein. Crystals grew to dimensions of up to $500 \times 250 \times 375 \mu\text{m}$ at 277 K in the hanging-drops by vapour-diffusion. Further, the conditions are described that yielded the wild-type crystals with dimensions of $80 \times 40 \times 60 \mu\text{m}$ by self-nucleation vapour-diffusion in sitting-drops at 281 K. The wild-type and recombinant crystals prepared by seeding techniques achieved full size within 5–14 days, while the wild-type crystals grown by self-nucleation appeared after 30 days and reached their maximum size after another two months. Both the wild-type and recombinant variant crystals, the latter altered in the key catalytic and substrate-binding residues Glu220, Trp434 and Arg158/Glu161 belonged to the $P4_32_12$ tetragonal space group, *i.e.*, the space group of the native microcrystals was retained in the newly grown recombinant crystals. The crystals diffracted beyond 1.57–1.95 Å and the cell dimensions were between $a = b = 99.2$ – 100.8 Å and $c = 183.2$ – 183.6 Å . With one molecule in the asymmetric unit, the calculated Matthews coefficients were between 3.4 – $3.5 \text{ Å}^3 \cdot \text{Da}^{-1}$

and the solvent contents varied between 63.4% and 64.5%. The macroseeding and cross-seeding techniques are advantageous, where a limited amount of variant proteins precludes screening of crystallisation conditions, or where variant proteins could not be crystallized.

Key words: macro- and cross-seeding; wild-type and mutant protein; X-ray diffraction

1. Introduction

Plant β -D-glucan glucohydrolase enzymes are classified in the GH3 family of glycoside hydrolases (<http://www.cazy.org/>) that currently includes nearly 3,000 entries [1]. The majority of these entries represent nucleotide sequences obtained from sequencing of bacterial genomes. Almost all GH3 entries have two, three or more individually folded domains, while the spatial arrangement of the domains varies [2,3]. A large proportion of the GH3 entries are enzymes that are variously annotated as β -D-glucosidases, glucan-1,4- β -D-glucosidases, (1,3)- β -D-glucan exohydrolases, (1,3;1,4)- β -D-glucan exohydrolases, exo-1,3-1,4- β -D-glucanases, N-acetyl- β -D-glucosaminidases, xylan-1,4- β -D-xylosidases and α -L-arabinofuranosidases [1]. Although the GH3 family contains predominantly enzymes, their substrate specificities have not been satisfactorily defined and the most of annotations are based on sequence similarities with enzymes for which biochemical data are available [4].

We have focused our attention in the past upon the biochemical and structural characterization of β -D-glucan glucohydrolases from higher plants, in particular on barley β -D-glucan glucohydrolase, isoenzyme ExoI (designated as HvExoI). This enzyme was crystallized from a native plant source obtained from barley seedlings [5], although the procedure for isolation of a crystallisable quality of HvExoI is both time consuming and technically challenging [6]. Over the past ten years several structures (Protein Data Bank references 1EX1, 1IEQ, 1IEV, 1IEW, 1IEX, 1J8V, 1LQ2, 1X38, 1X39) have been solved from the native HvExoI crystals with substrate analogues and mechanism-based inhibitors to explain the enzyme's catalytic mechanism and substrate specificity [2,7-10]. However, a few facets of catalytic and substrate binding mechanisms remain to be explained. To this end we are interested in the specific roles of amino acid residues in the vicinity of the catalytic pair and how specific structural determinants at the entrance of the catalytic pocket control spatial dispositions of isomeric oligosaccharides entering the catalytic site. Further, we have been intrigued by so-called substrate-product trafficking events at the mouth of the catalytic site [2,7]. Here, the glucose product of the hydrolytic reaction, which is released from the non-reducing termini of substrates, remains bound to the enzyme's active site. However, when a new substrate molecule approaches the enzyme, the glucose product diffuses away from the pocket and the incoming substrate enters the active site [4]. Both events are linked and a precise mechanism, how they proceed, is not known. We presume that the incoming substrate binds in the vicinity of the active site and could, by some delicate mechanism, instigate a product for new substrate interchange. We believe that this substrate/product interchange event represents an ideal model system to study, how products and incoming substrates interact in and/or near enzymes' catalytic sites in general.

Recently, we have reported the high level expression of recombinant HvExoI (designated as rHvExoI) from a codon-optimized HvExoI cDNA in protease-deficient *Pichia pastoris* under low

temperature conditions [11]. We expect that rHvExoI could be used as a suitable enzyme model to study the roles of amino acid residues in catalysis and substrate specificity of this enzyme. To this end, we prepared variants of wild-type rHvExoI, altered (into Ala) in the key catalytic and substrate binding residues Glu220, Trp434 and Arg158/Glu161 by site-directed mutagenesis [12]. From the available structural data [2,7-10] we would expect that Glu220, Trp434 and Arg158 are surface exposed, while Glu161 is buried.

In the current work, we describe the techniques and conditions for preparation of the wild-type and variant rHvExoI crystals that showed excellent diffraction parameters. These crystals were prepared by the macroseeding and cross-seeding techniques with microcrystals obtained from a native plant protein. Seeding techniques have previously been used successfully in bio-macromolecular crystallography [13]. The advantage of these techniques is that they use crystals that provide a preformed, regular surface, onto which new molecules may be added in a regular fashion, generally at a lower degree of supersaturation than is required for nucleation [13].

2. Results and Discussion

2.1. Protein Expression and Purification

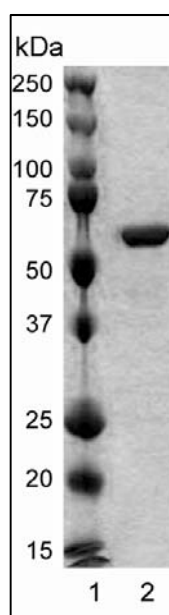
Protein expression in *P. pastoris* at a temperature of 293 K, deglycosylation and purification yielded near-homogenous recombinant rHvExoI (Figure 1) with a molecular mass of 67,169 Da [11]. It is of note that rHvExoI contained at its NH₂-terminus an 8x-His-tag and additional three Ala residues flanking the 8x-His-tag (AHHHHHHHHHAA). The Ala residues resulted from the ligation-based cloning process, while the 8x-His-tag was added for protein purification purposes [11]. This 11-residue tag was not removed from the rHvExoI protein that was subjected to crystallisation trials. Expressed rHvExoI was catalytically competent with the catalytic efficiency value $k_{\text{cat}} \cdot K_M^{-1} = 14 \text{ mM}^{-1} \cdot \text{s}^{-1}$ towards 4-nitrophenyl β -D-glucoside. This second-order rate constant was similar to that reported for native fully N-glycosylated HvExoI [14] and a recombinant N-deglycosylated form [11]. Further, the biophysical properties of rHvExoI such as pH optimum and thermostability were also similar to those reported for native or recombinant HvExoI [11,14].

2.2. Protein Crystallisation

The concentrated rHvExoI was subjected to four types of crystallisation trials (Figures 2 and 3). We firstly set-up the microbatch under paraffin oil trials at 277 K and 287 K with the goal to screen approximately 130 independent conditions, including those that were previously found to be successful with a native plant enzyme [5]. These conditions (1.7 M ammonium sulfate, 75 mM HEPES-NaOH buffer, pH 7, containing 7.5 mM sodium acetate and 1.2% (w/v) PEG 400) previously produced crystals of native plant HvExoI that belonged to the primitive tetragonal space group $P4_32_12$ and yielded high-resolution diffraction patterns [2,5,7]. Nevertheless, despite intensive effort, no diffraction-quality crystals grew, and after approximately 14 days only the formation of a highly intricate interlaced web of thin needles of approximately 40-60 μm in their longest dimensions was observed. Figure 2A and its inset show the appearance of these needles that grew from 1.6 M magnesium sulfate at pH 6.5 and 277 K.

Our second approach to growing the diffraction quality crystals led us to set-up close to 400 trials at the Bio21 Collaborative Crystallisation Centre at 281 K and 293 K by a sitting-drop vapour-diffusion method. After approximately 21 days at 281 K, we could observe formation of short thin needle-shaped crystals that either remained dispersed throughout the droplets (Figure 2B) or formed well-organised round balls (Figure 2C). However, after about a month, in some of the droplets with needles (Figures 2B and 2C) truncated bi-pyramidal crystals formed, which reached dimensions of $80 \times 40 \times 60 \mu\text{m}$ after about 97 days (Figure 2B). It was of note that these crystals were only observed at 281 K in droplets with 1.6 to 2.2 M ammonium sulfate, containing 10 mM malate-MES-Tris buffer, pH 5. These conditions were similar to the conditions that we found previously for native HvExoI [5], except that here the pH value of 5 was more acidic than that used previously (pH 7), and that the protein concentration was almost twice as high ($12.5 \text{ mg}\cdot\text{mL}^{-1}$ versus $6.8 \text{ mg}\cdot\text{mL}^{-1}$). We expected that these truncated bi-pyramidal crystals could belong to the tetragonal $P4_32_12$ space group found for the native protein crystals [5], although at this stage we were not able to collect their diffraction patterns. Notably, in an identical screen at a higher temperature of 293 K the bi-pyramidal crystals were not formed, so it seemed that the temperature was a critical factor for crystal formation of rHvExoI. Thirdly, we also attempted to grow crystals by self-nucleation at 277 K by hanging-drop vapour diffusion and using the conditions developed for native HvExoI [5]. However, we could not observe crystal formation within 180 days and rHvExoI mostly precipitated or the drops remained clear.

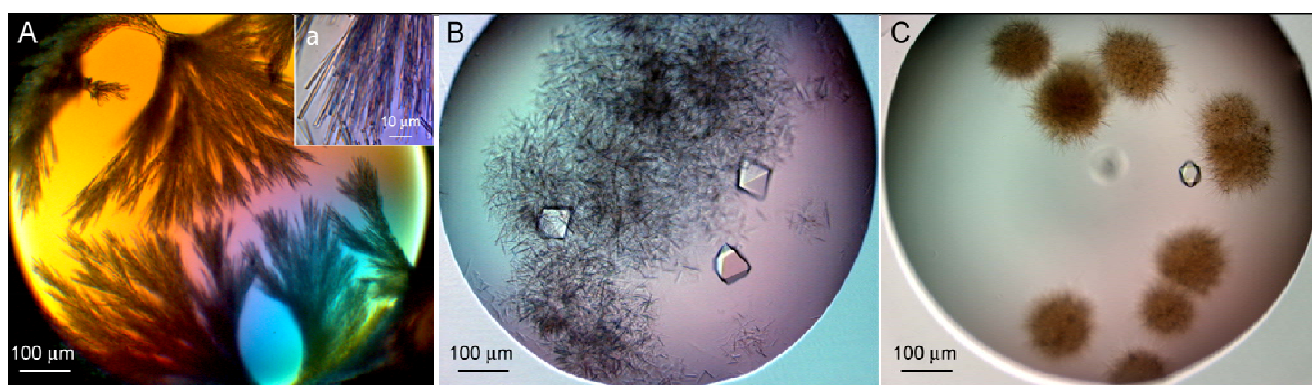
Figure 1. SDS-PAGE profile of recombinant wild-type rHvExoI (lane 2) used for crystallisation by self-nucleation and macro- and cross-seeding using hanging-drop and sitting-drop vapour-diffusion methods. The rHvExoI protein (25 μg) was visualised with a Coomassie Brilliant Blue stain. Protein standards (5 μL of the Precision Plus Protein Standards' stock) are indicated in lane 1.



As no diffraction-quality crystals were obtained with the three crystallisation approaches described above, we turned our attention to seeding at 277 K using a hanging-drop vapour-diffusion method that was used successfully for preparation of large well-diffracting crystals of native HvExoI [7-10]. We

first examined, if we could use microcrystals obtained from native HvExoI to seed the wild-type recombinant rHvExoI protein, despite the differences between the two proteins. These differences included the 11 additional residues of the affinity tag at the NH₂-terminus, as described above. Further, the rHvExoI protein was N-deglycosylated by endoglycosidase H, such that only one N-linked N-acetyl- β -D-glucosaminyl residue [12] remained attached to each of the three N-glycosylation sites Asn221, Asn498 and Asn600 [2,11]. Approximately 48 h after rHvExoI was macroseeded with the native microcrystals of the sizes between $10 \times 5 \times 7.5 \mu\text{m}$ and $20 \times 10 \times 15 \mu\text{m}$, the original native microcrystals started growing in size. The fully grown crystals of wild-type rHvExoI reached dimensions that varied between $100 \times 50 \times 75 \mu\text{m}$ (Figure 3B) and $500 \times 250 \times 375 \mu\text{m}$ after 5 to 7 days and these crystals had a similar bi-pyramidal morphology as the native microcrystals (Figures 3A). Having succeeded in growing wild-type crystals, using native microcrystal seeds, we prepared large recombinant crystals of three rHvExoI variants, specifically Glu220Ala (Figure 3C), Trp434Ala and of the double mutant Arg158Ala/Glu161Ala, also using native seed crystals. The newly grown variant crystals reached sizes of $100 \times 50 \times 75 \mu\text{m}$ (Figure 3C) to $500 \times 250 \times 375 \mu\text{m}$ and showed a similar bi-pyramidal morphology as the native microcrystals. The variant crystals grew slightly slower and reached their maximum dimensions after 10-14 days.

Figure 2. Crystals of recombinant wild-type rHvExoI grown by self-nucleation. Various forms of crystals grown for 14 (A) and 97 (B-C) days using a microbatch-under-paraffin-oil technique from 1.6 M magnesium sulfate, pH 6.5 at 277 K (A) or obtained using sitting-drop vapour-diffusion from 1.8 (B) or 2.2 M (C) ammonium sulfate, both at pH 5 and 281 K. The crystals in (A) formed thin needles of approximately 40-50 μm in their longest dimensions (inset), while in B and C short thin needles are shown that appeared within 14 days and after another 30 days the truncated bi-pyramidal crystals grew that reached dimensions of $80 \times 40 \times 60 \mu\text{m}$ after 97 days.

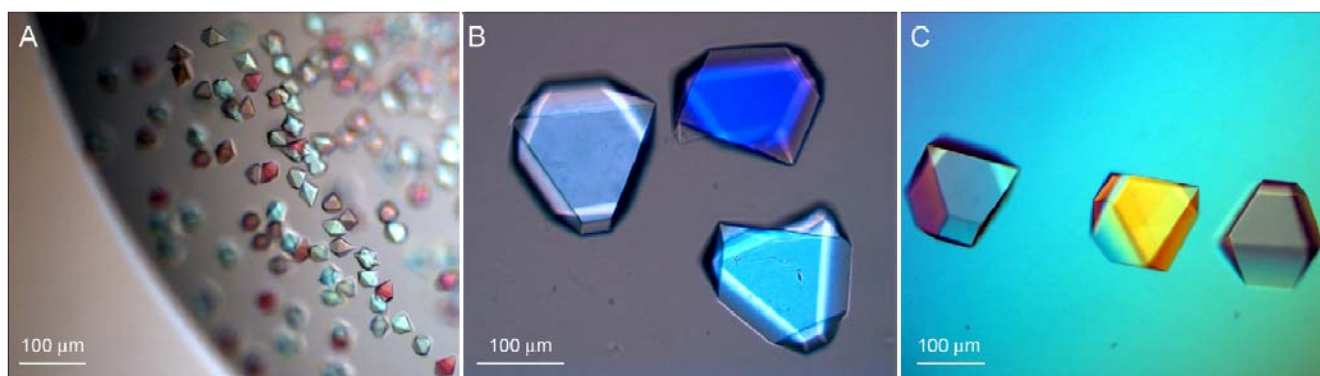


2.3. X-ray Diffraction

Single wild-type and variant rHvExoI crystals were cryo-protected and subjected to diffraction at the MX1 beamline of the Australian Synchrotron. All of the X-ray diffraction data sets were virtually complete beyond 1.57-1.95 Å (Figure 4; Table 1). The HKL2000 indexing and systematic absences calculated that the space groups of the wild-type and variant rHvExoI crystals were consistent with a

primitive tetragonal space group $P4_32_12$ (Figure 4 and Table 1), and as expected, these space group characteristics were similar to their native counterparts [2,7]. In line with these observations was our previous finding that the native HvExoI microcrystals of the sizes $20 \times 10 \times 15 \mu\text{m}$ diffracted beyond $2.80 \text{ \AA}$ on an in-house rotating anode X-ray source and belonged to a primitive tetragonal space group $P4_32_12$ [15]. Thus, our data are in agreement with other reports, where the space group characteristics of macroseeds and fully grown crystals were identical [16–18], or that as a bonus, the resolution of the newly grown crystals has improved [18]. On the contrary, other authors have reported that during seeding trials mutant crystals often crystallized in different space groups than their macroseeds [19].

Figure 3. Microcrystals ($10 \times 5 \times 7.5 \mu\text{m}$ to $20 \times 10 \times 15 \mu\text{m}$) of native HvExoI (A) used to grow the recombinant wild-type (B) and variant Glu220Ala (C) rHvExoI crystals that grew to their full-sizes within 5–14 days. The crystals obtained by seeding in hanging-drops by vapour-diffusion grew to dimensions of up to $500 \times 250 \times 375 \mu\text{m}$ at 277 K .



The best diffraction data, to $1.57 \text{ \AA}$ were collected from the wild-type crystals, followed by the Arg158Ala/Glu161Ala, Glu220Ala and Trp434Ala variants that diffracted to $1.65 \text{ \AA}$, $1.90 \text{ \AA}$ and $1.95 \text{ \AA}$, respectively (Figure 4, Table 1). The lattice dimensions of wild-type and variant crystals varied between $a = b = 99.2\text{--}100.8 \text{ \AA}$ and $c = 183.2\text{--}183.6 \text{ \AA}$, and there appeared to be one molecule in the asymmetric units, according to the Matthews coefficient calculation [20]. The Matthews coefficients of recombinant crystals were between $3.4\text{--}3.5 \text{ \AA}^3 \cdot \text{Da}^{-1}$ with solvent contents of 63.4% to 64.5% . The R_{merge} values of 5.6% to 10.1% were obtained with $\langle I/\sigma(I) \rangle$ of 37.0 to 71.2 , whereas the multiplicity of the individual datasets was well above 20 and varied between 26 to 29 (Table 1). It was of note that the completeness for the highest resolution shells of the wild-type and Glu220Ala datasets was lower, despite the high multiplicity and crystal symmetry, because the data were integrated into the corners of the detector. Also, the mean $\langle I/\sigma(I) \rangle$ values for these datasets indicated that they actually diffracted to higher resolution than that stated in Table 2. It was not surprising that the most favourable diffraction statistics was obtained with the wild-type crystals that were seeded with the wild-type native macroseeds (Figure 4 and Table 1).

Table 1. Data collection statistics from the wild-type and variant rHvExoI crystals, calculated by the HKL2000 suite of programs.

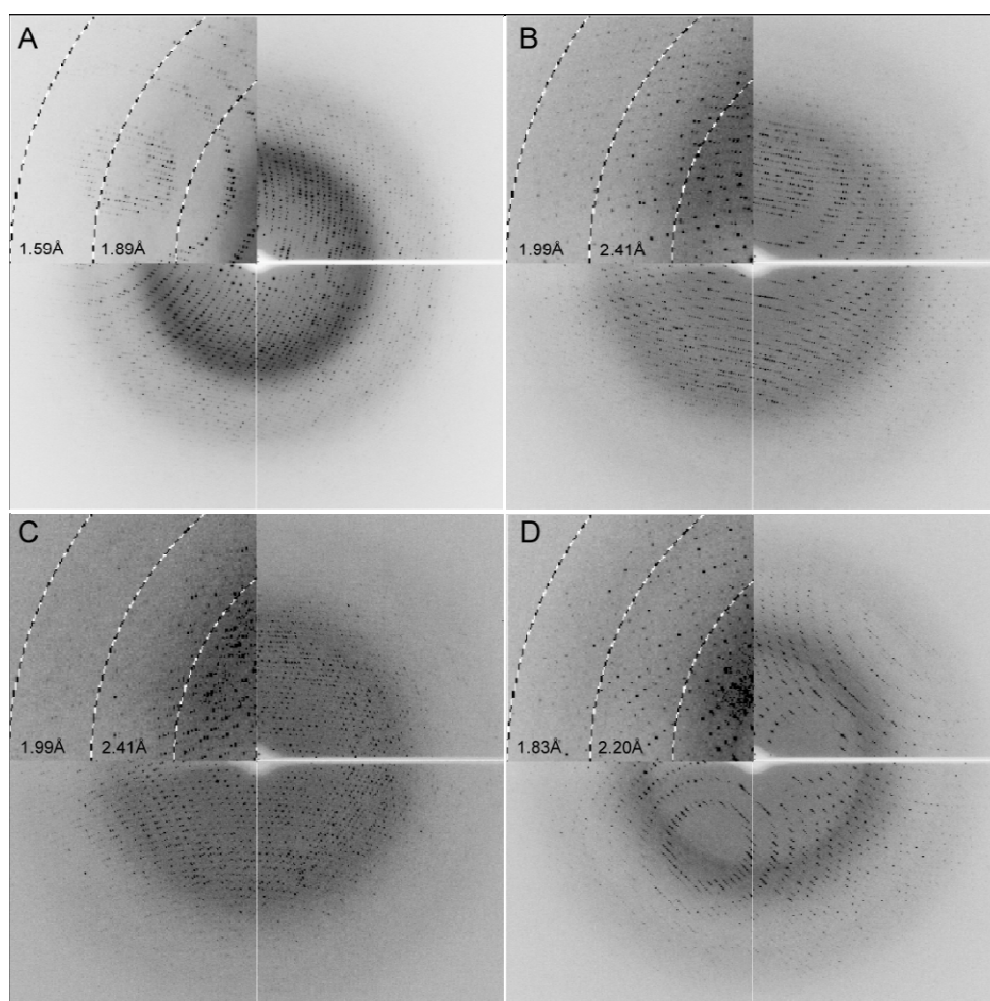
	Wild-type	Glu220Ala	Trp434Ala	Arg158Ala/Glu161Ala
Unique reflections	119968	74171	64460	107601
Resolution ^a (Å)	1.57 (1.6-1.57)	1.90 (1.94-1.90)	1.95 (1.98-1.95)	1.65 (1.68-1.65)
Mean multiplicity ^{a,b}	29 (26)	27 (16)	27 (16)	26 (12)
Completeness ^{a,b} (%)	99 (86)	99 (88)	100 (100)	99.8 (98)
Mean $\langle I/\sigma(I) \rangle$ ^a	71.2 (5.6)	54.8 (4.3)	37.0 (2.2)	58.7 (2.3)
R_{merge} ^{a-c} (%)	6.7 (47)	8.9 (64)	10.1 (87)	5.6 (82)
$a = b$ (Å) ^d	99.2	100.2	100.1	100.8
c (Å) ^d	183.5	183.2	183.6	183.2

^a Numbers in parenthesis represent the values in the highest resolution shell.

^b The numbers were rounded to no decimal place.

^c $R_{\text{merge}} = 100[\sum(I_i - \langle I \rangle)^2 / \sum I_i^2]$ is summed over all independent reflections.

^d The numbers were rounded to the 1st decimal place.

Figure 4. X-ray diffraction patterns of the recombinant wild-type (A), and variant Glu220Ala (B), Trp434Ala (C) and Arg158Ala/Glu161Ala (D) rHvExoI crystals. The left top insets show diffraction intensities to 1.59-1.89 Å (A), 1.99-2.41 Å (B and C) and 1.83-2.20 Å (D).

3. Experimental Section

3.1. Expression and Purification of Wild-Type and Variant Forms of rHvExoI

Wild-type (GenBank accession No. GU441535) and variant codon-optimized cDNAs, encoding a mature barley β -D-glucan exohydrolase I (HvExoI) inserted in pPICZ α BNH₈ expression vectors, were expressed in *P. pastoris*, strain SMD1168H and purified by ion exchange, immobilized metal affinity chromatography (IMAC), N-deglycosylation by endoglycosidase H and a second round of IMAC, as described previously [11]. The variant forms included Glu220Ala, Trp434Ala and a double mutant Arg158Ala/Glu161Ala, for which the constructions of the cDNA fusions will be described elsewhere [12]. At the final purification step before crystallisation, the N-deglycosylated rHvExoI wild-type and variant enzymes were eluted from a BioGel-P100 size-exclusion column with 50 mM sodium acetate buffer, pH 5.25 containing 200 mM NaCl and 1 mM dithiothreitol at a linear flow rate of 0.5 cm·h⁻¹. The protein purities of the rHvExoI fractions were analyzed by SDS-PAGE, using 12.5% w/v polyacrylamide and bis-polyacrylamide gels and standard techniques [6]. The protein concentration was estimated with a Bio-Rad protein assay kit (Bio-Rad Laboratories, Gladesville, New South Wales, Australia) using bovine serum albumin (Sigma Chemical Company, St. Louis, MO, USA) as a standard. The protein standards 'Precision Plus Protein Standards' used for SDS-PAGE were from Bio-Rad Laboratories.

3.2. Enzyme Assays

The activities of pooled and concentrated (using 10 kDa cut-off centrifugal filter units (Millipore, Bedford, MA, USA) wild-type and variant enzymes were assayed against 4-nitrophenyl β -D-glucopyranoside (Sigma) in 50 mM sodium acetate buffer, pH 5.25.

3.3. Crystallisation of Wild-Type and Variant Forms of rHvExoI

Near-homogenous N-deglycosylated wild-type and variant rHvExoI proteins were concentrated to 12.5 mg·mL⁻¹ in 20 mM sodium acetate pH 5.25 and filtered through a 0.22 μ m filter (Millipore). Screening of crystallisation conditions was performed by four experimental approaches. Firstly, initial crystallisation conditions were screened using a microbatch under paraffin oil technique. Here, 1 μ L of precipitant solutions and 1 μ L of the solution containing 12.5 mg·mL⁻¹ of rHvExoI were added into 10 μ L of 100% Paraffin oil (Hampton Research, Aliso Viejo, CA, USA) that was previously dispensed in microbatch 72 well Greiner (Terasaki style) plates (Hampton Research). The formulations of Crystal Screen 2, Crystal Screen Lite and Grid Screen Ammonium Sulfate (Hampton Research) were used as precipitants and the crystals grew at 277 K or 287 K in a vibration-free crystallographic cabinet (Molecular Dimensions, Suffolk, UK). Secondly, crystallisation trials were set-up robotically (Phoenix Nano-Dispenser, Art Robbins Instruments, Sunnyvale, CA, USA) in sitting drops, in which 300 nL droplets of rHvExoI were mixed with the same volumes of precipitants and crystal growth proceeded at 281 K and 293 K at the Bio21 Collaborative Crystallisation Centre (CSIRO, Parkville, Australia) [21]. The precipitants from the PSS_1_Com5 and PS gradient-mid range formulation screens were used, whereas both screens were prepared in-house at the Bio21 Centre

(http://www.csiro.au/c3/Facility/c3_centre_robotic_crystal.htm), following the recommendation from Emerald BioSystems (Bainbridge Island, WA, USA) for the first screen and those reported by Newman [22] for the second screen. The PSS_1_Com5 screen uses the inorganic precipitants such as sulfates, chlorides, citrates and phosphates, but also 2-methyl-2,4-pentanediol, glycerol and polyethyleneglycols (PEGs) in the pH ranges of 5–8.7. On the other hand, the PS gradient-mid range formulation relies on ammonium sulfate and sodium malonate as precipitants in a 10 mM malate-MES-Tris buffer system in the pH ranges of 4.5–9 [22]. Thirdly, attempts were made to grow crystals under the conditions developed for native HvExoI at 277 K using a hanging-drop vapour diffusion method [5]. Lastly, and most importantly, the rHvExoI crystals were grown in hanging drops at 277 K that were seeded with the native HvExoI microcrystals prepared as described previously [5]. The sizes of the microcrystals for the latter conditions varied between $10 \times 5 \times 7.5 \mu\text{m}$ and $20 \times 10 \times 15 \mu\text{m}$. The hanging drops were prepared at 277 K as follows. The volume of 4 to 6 μL of rHvExoI at $12.5 \text{ mg}\cdot\text{mL}^{-1}$ was added to 4 μL of the precipitant solution A (100 mM HEPES-NaOH buffer pH 7, 2.4% (w/v) PEG 400, 1.6 M ammonium sulfate) on 22-mm siliconized circular glass cover slips (Hampton Research). A few microcrystals of native HvExoI were transferred into the hanging drop with a cat whisker. Here, the whisker gently touched the surface of a macroseed stock of native HvExoI and subsequently the whisker was swiftly immersed into a new rHvExoI drop. The cover slips with the seeded hanging drops were placed over 1 mL of reservoir solutions (1.7 M ammonium sulfate in 50 mM HEPES-NaOH buffer, pH 7) contained in the 24 well Linbro plates (Hampton Research), and the wells were sealed with vacuum grease (Dow Corning Corporation, Midland, MI, USA). Crystals from the seeded drops appeared within 5–14 days and were suitable for X-ray data collection. The crystals were photographed through a Leica Laser Microdissection microscope (Leica, North Ryde, Australia) equipped with fluorescence and differential interference contrast.

3.4. X-ray Data Collection and Processing

Single enzyme crystals with the longest dimensions of 100 to 500 μm were cryo-protected in 20% (v/v) glycerol concentration in solution A (as specified above in Section 3.3.) [10] and flash cooled in the cold N_2 stream at the beamline MX1 of the Australian Synchrotron. X-ray diffraction data sets were collected at 0.5° oscillations (1 sec exposures) through 360° on the ADSC Quantum 210r Detector [23]. The data were processed with the HKL2000 suite of programs [24].

4. Conclusions

In summary, excellent X-ray diffraction data were obtained from the recombinant wild-type and variant rHvExoI crystals grown by seeding from a native plant source protein in hanging-drops by vapour-diffusion. The recombinant crystals grew relatively fast and within 5–14 days reached dimensions of up to $500 \times 250 \times 375 \mu\text{m}$. The fully grown recombinant crystals retained the space group characteristics of their native macroseeds and diffracted beyond $1.57 \text{ \AA}$ to $1.95 \text{ \AA}$. As reported for other proteins, this technique could be valuable, where a limited amount of variant proteins is available precluding crystallisation trials, or where variant protein forms could not be crystallized. We project that cross-seeding using native protein as a source of microcrystals could be successfully used for generation of large recombinant wild-type and variant crystals that could potentially yield high

resolution diffraction patterns. Lastly, the diffraction data collected from the wild-type and variant rHvExoI crystals reported here could be used successfully for structure solution. The structural data are currently being prepared for publication [12].

Acknowledgements

The authors acknowledge the MX1 beamline at the Australian Synchrotron (Victoria, Australia) and Janet Newman from the Bio21 Collaborative Crystallisation Centre (Melbourne, Australia). We are also grateful to Gwenda Mayo (Australian Centre for Plant Functional Genomics) for taking photographs of crystals. This work was supported by grants from the Australian Research Council to M.H., the Thailand Research Fund (TRF) to J.R.K.C. and a TRF Royal Golden Jubilee PhD Fellowship to S.L. The constructive help of both anonymous referees is also acknowledged.

References

1. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **2008**, *37*, D233-D238.
2. Varghese, J.N.; Hrmova, M.; Fincher, G.B. Three-dimensional structure of a barley β -D-glucan exohydrolase; a family 3 glycosyl hydrolase. *Struct. Fold. Des.* **1999**, *7*, 179-190.
3. Pozzo, T.; Pasten, J.L.; Karlsson, E.N.; Logan, D.T. Structural and functional analyses of beta-glucosidase 3B from *Thermotoga neapolitana*: A thermostable three-domain representative of glycoside hydrolase 3. *J. Mol. Biol.* **2010**, *397*, 724-739.
4. Hrmova, M.; Fincher, G.B. Dissecting the catalytic mechanism of a plant β -D-glucan glucosylhydrolase through structural biology using inhibitors and substrate analogues. *Carbohydr. Res.* **2007**, *305*, 209-221.
5. Hrmova, M.; Varghese, J.N.; Høj, P.B.; Fincher, G.B. Crystallization and preliminary X-ray analysis of β -glucan exohydrolase isoenzyme ExoI from barley (*Hordeum vulgare*). *Acta Cryst.* **1998**, *D54*, 6887-689.
6. Hrmova, M.; Harvey, A.J.; Wang, J.; Shirley, N.J.; Jones, G.P.; Høj, P.B.; Fincher, G.B. Barley β -D-glucan exohydrolases with β -D-glucosidase activity. Purification and determination of primary structure from a cDNA clone. *J. Biol. Chem.* **1996**, *271*, 5277-5286.
7. Hrmova, M.; Varghese, J.N.; DeGori, R.; Smith, B.J.; Driguez, H.; Fincher, G.B. Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of plant β -D-glucan glucosylhydrolase. *Struct. Fold. Des.* **2001**, *9*, 1005-1016.
8. Hrmova, M.; DeGori, R.; Smith, B.J.; Driguez, H.; Varghese, J.N.; Fincher, G.B. Structural basis for a broad specificity in higher plant β -D-glucan glucosylhydrolases. *Plant Cell* **2002**, *14*, 1033-1052.
9. Hrmova, M.; De Gori, R.; Smith, B.J.; Vasella, A.; Varghese, J.N.; Fincher, G.B. Three-dimensional structure of the barley β -D-glucan glucosylhydrolase in complex with a transition state mimic. *J. Biol. Chem.* **2004**, *279*, 4970-4980.

10. Hrmova, M.; Streltsov, V.A.; Smith, B.J.; Vasella, A.; Varghese, J.N.; Fincher, G.B. Structural rationale for low nanomolar binding of transition state mimics to a family GH3 β -D-glucan glucohydrolase from barley. *Biochemistry* **2005**, *44*, 16529-16539.
11. Luang, S.; Hrmova, M.; Ketudat Cairns, J.R. High-level expression of barley β -D-glucan exohydrolase HvExo I from a codon-optimized cDNA in *Pichia pastoris*. *Prot. Expr. Purif.* **2010**, *73*, 90-98.
12. Luang, S.; Streltsov, V.A.; Ketudat Cairns, J.R.; Hrmova, M. Characterization of variant forms of β -D-glucan glucohydrolase from barley (*Hordeum vulgare* L). *Biochem. J.* **2010**, to be submitted.
13. Stura, E.; Wilson, I.A. Analytical and production seeding techniques. *Methods* **2000**, *1*, 38-49.
14. Hrmova, M.; Fincher, G.B. Barley β -D-glucan exohydrolases. Substrate specificity and kinetic properties. *Carbohydr. Res.* **1998**, *305*, 209-221.
15. Varghese, J.N.; van Donkelaar, A.; Balaic, D.X.; Barnea, Z. CSIRO-Commonwealth Scientific Research Organization. Personal communication, Melbourne, Victoria, 2001.
16. Radaev, S.; Agniswamy, J.; Sun, P.D. A case of structure determination using pseudosymmetry. *Acta Cryst.* **2009**, *D65*, 1334-1340.
17. Stec, B.; Holtz, K.M.; Wojciechowski, C.L.; Kantrowitz, E.R. Structure of the wild-type TEM-1 β -lactamase at 1.55 Å and the mutant enzyme Ser70Ala at 2.1 Å suggest the mode of noncovalent catalysis for the mutant enzyme. *Acta Cryst.* **2005**, *D61*, 1072-1079.
18. Heaslet, H.; Rosenfeld, R.; Giffin, M.; Lin, Y.-C.; Tam, K.; Torbett, B.E.; Elder, J.H.; McRee, D.E.; Stout, C.D. Conformational flexibility in the flap domains of ligand free HIV protease. *Acta Cryst.* **2007**, *D63*, 866-875.
19. Mizutani, H.; Saraboji, K.; Malathy Sony, S.M.; Ponnuswamy, M.N.; Kumarevel, T.; Krishna Swamy, B.S.; Simanshu, D.K.; Murthy, M.R.; Kunishima, N. Systematic study on crystal-contact engineering of diphthine synthase: Influence of mutations at crystal-packing regions on X-ray diffraction quality. *Acta Cryst.* **2008**, *D64*, 1020-1033.
20. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33*, 491-497.
21. Newman, J.; Pham, T.M.; Peat, T.S. Phoenix experiments: combining the strengths of commercial crystallization automation. *Acta Cryst.* **2008**, *F64*, 991-996.
22. Newman, J. Optimization of buffer solutions for protein crystallization. *Acta Cryst.* **2004**, *D60*, 610-612.
23. McPhillips, T.M.; McPhillips, S.E.; Chiu, H.J.; Cohen, A.E.; Deacon, A.M.; Ellis, P.J.; Garman, E.; Gonzalez, N.K.; Sauter, N.K.; Phizackerley, R.P.; Soltis, S.M.; Kuhn, P. Blu-ice and the distributed control system: Software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **2002**, *9*, 401-406.
24. Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307-326.

Binding of β -D-Glucosides and β -D-Mannosides by Rice and Barley β -D-Glycosidases with Distinct Substrate Specificities[†]

Teerachai Kuntothom,^{‡,§,▽} Michal Raab,^{‡,§} Igor Tvaroška,[§] Sebastien Fort,^{||} Salila Pengthaisong,[‡] Javier Cañada,[⊥] Luis Calle,[⊥] Jesús Jiménez-Barbero,[⊥] James R. Ketudat Cairns,[‡] and Maria Hrmova^{*,@,●}

[‡]School of Biochemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima, Thailand,

[§]Department of Structure and Function of Saccharides, Institute of Chemistry, Center for Glycomics, Slovak Academy of Sciences, Bratislava, Slovak Republic, ^{||}Centre de Recherches sur les Macromolécules Végétales, Grenoble, France, [⊥]Centro de Investigaciones Biológicas, CSIC, Madrid, Spain, and [@]Australian Centre for Plant Functional Genomics, University of Adelaide, Glen Osmond, Australia. [▽]These authors made equal contributions to this work. [●]Current address: Department of Chemistry,

Maharakham University, Maharakham, Thailand. [●]Also affiliated with Institute of Chemistry, Slovak Academy of Sciences, Bratislava, Slovak Republic.

Received July 12, 2010; Revised Manuscript Received September 7, 2010

ABSTRACT: Predominantly, rice Os3BGlu7 operates as a β -D-glucosidase (EC 3.2.1.21), while barley HvBII acts as a β -D-mannosidase (EC 3.2.1.25). Saturation transfer difference nuclear magnetic resonance (STD NMR) and transferred nuclear Overhauser effect (trNOE) spectroscopy in conjunction with quantum mechanics/molecular mechanics (QM/MM) modeling and docking at the 6-31+G* level were used to investigate binding of S- and O-linked *gluco*- and *manno*-configured aryl- β -D-glycosides to Os3BGlu7 and HvBII. Kinetic analyses with 4-nitrophenyl β -D-thioglucoside (4NP-S-Glc) and 4-nitrophenyl β -D-thiomannoside (4NP-S-Man) indicated that the inhibitions were competitive with apparent K_i constants of 664 and 710 μ M for Os3BGlu7 and 95 and 266 μ M for HvBII, respectively. The STD NMR and trNOESY experiments revealed that 4NP-S-Glc and 4NP-S-Man bound weakly in ⁴C₁ conformations to Os3BGlu7; 4NP-S-Glc adopted ³S₅ (B_{3,0}) or ¹S₃ (^{1,4}B) conformations, and 4NP-S-Man preferred ⁴C₁ geometry, when bound to HvBII. The QM modeling and docking, based on GLIDE scores, predicted that 4NP-O-Glc, 4NP-O-Man, and 4NP-S-Man bound preferentially in ¹S₃ geometries to both enzymes, contrary to 4NP-S-Glc that could also adopt a ⁴C₁ conformation, although in a “flipped-down” ring position. The experimental and computational data suggested that in glycoside recognition and substrate specificity of Os3BGlu7 and HvBII, a combination of the following determinants is likely to play key roles: (i) the inherent conformational and spatial flexibilities of *gluco*- and *manno*-configured substrates in the enzymes’ active sites, (ii) the subtle differences in the spatial disposition of active site residues and their capacities to form interactions with specific groups of substrates, and (iii) the small variations in the charge distributions and shapes of the catalytic sites.

The glycoside hydrolase GH1 family includes enzymes with approximately 20 known substrate specificities (1), including β -D-glucosidases (EC 3.2.1.21), 6-phospho- β -D-glucosidases (EC 3.2.1.86), β -D-mannosidases (EC 3.2.1.25), β -D-galactosidases (EC 3.2.1.23), β -D-glucuronidases (EC 3.2.1.31), and others (2). A detailed examination of the substrate specificity of the barley β -D-glucosidase isoenzyme β II (HvBII) (3, 4), also designated Hv β MannosI (5), revealed that it exhibits a marked preference for manno-oligosaccharides and that the rate of hydrolysis increases with the degree of polymerization of both cello- and manno-oligosaccharides (3–6). Hence, the substrate specificity and action patterns of HvBII are characteristic of an oligosaccharide

exohydrolase, rather than of an enzyme with a preference for low-molecular mass cello-oligosaccharides. Similar conclusions were drawn for an Os3BGlu7 β -D-glucosidase from rice (also called BGlu1), although binding energies at individual subsites differ somewhat (6, 7). Both plant enzymes are capable of catalyzing transglycosylation reactions with 4NP-O-Glc¹ (3, 4, 6, 7), but not with 4NP-O-Man. The presence of an extended series of subsites in these two plant β -D-glycosidases indicates that their biological

[†]This work was supported by grants from the Australian Research Council to M.H. and from the Thailand Research Fund (BRG5080007) to J.R.K.C. T.K. was sponsored by The Institute for the Promotion of Teaching Science and Technology of Thailand. M.R. and I.T. are thankful for support from the Science and Technology Assistance Agency under Contract APVV-0607-07. J.C., L.C., and J.J.-B. thank the Ministerio de Ciencia e Innovación for support (CTQ2009-08536).

*To whom correspondence should be addressed. Telephone: +61 8 8303 7160. Fax: +61 8 8303 7102. E-mail: maria.hrmova@adelaide.edu.au.

¹Abbreviations: CAZy, Carbohydrate-Active enZymes; DP, degree of polymerization; CORCEMA, Complete Relaxation and Conformational Exchange Matrix; DPGFSE, double pulse field-gradient spin-echo; EA, catalytic nucleophile; EB, catalytic acid/base; ESP, electrostatic potential; GH, glycoside hydrolase; HDO, hydrogen/deuterium water; ISPA, isolated spin-pair approximation; NOEs, nuclear Overhauser effects; PDB, Protein Data Bank; rmsd, root-mean-square deviation; QM/MM, quantum mechanics/molecular mechanics; STD NMR, saturation transfer difference nuclear magnetic resonance; trNOESY, transferred nuclear Overhauser effect spectroscopy; 1D, one-dimensional; 2D, two-dimensional; 3D, three-dimensional; 4NP, 4-nitrophenyl; 4NP-O-Glc, 4-nitrophenyl β -D-glucopyranoside; 4NP-O-Man, 4-nitrophenyl β -D-mannopyranoside; 4NP-S-Glc, 4-nitrophenyl β -D-thioglucopyranoside; 4NP-S-Man, 4-nitrophenyl β -D-thiomannopyranoside.

functions could lie in the hydrolysis of longer oligosaccharides, possibly derived from cell wall (1,3;1,4)-glucans and (1,4)- β -D-(gluco)mannans (3–6, 8).

Substrate hydrolysis in the GH1 family proceeds with retention of anomeric configuration in a two-step double-displacement catalytic mechanism (9). The catalytic event advances with participation of a pair of glutamate residues, a catalytic acid/base and a catalytic nucleophile. The exceptions to this rule in the GH1 family are plant myrosinases, for example, that from *Sinapis alba* (10), which do not have proton donors. Most information about the catalytic mechanism of the GH1 β -D-glucosidases has been derived from chemical modification (11, 12), mutagenesis (13), molecular modeling (3, 14), or X-ray diffraction studies (e.g., refs 15 and 16). The catalysts in a range of β -D-glucosidases were determined using mechanism-based inhibitors; e.g., the catalytic nucleophile in the *Agrobacterium* sp. β -D-glucosidase (17) and rice Os3BGlu7 (16) were identified with 2',4'-dinitrophenyl 2-deoxy-2-fluoro- β -D-glucoside, while conduritol B epoxide was used with the *Aspergillus wentii* β -D-glucosidase (18) and barley HvBII (3). In the HvBII enzyme, the two catalytic amino acid residues, Glu179 and Glu386, are positioned near the bottom of the substrate-binding pocket and are separated by approximately 5–6 Å, although it is not clear how the nonreducing end of the substrate is preferably recognized versus the reducing end by the active site residues (3). It is also not apparent how the remainder of the substrate dissociates from the active site after glucose is cleaved, to provide sufficient space for the product to diffuse out of the pocket.

Approximately 30 unique three-dimensional (3D) structures or molecular models (1) have been reported for the GH1 enzymes that fold into (β/α)₈ barrel projections. In contrast to the open cleft structures of endohydrolases, plant and microbial β -D-glucosidases align their glycosidic substrate in dead-end funnels (3, 19). The substrates are brought into juxtaposition with the catalytic pairs close to a glycosidic linkage at the nonreducing terminal residue (8, 15, 20). Much attention has been devoted to dissecting structural features that are responsible for glycon and aglycon substrate specificities in the GH1 group of enzymes (e.g., refs 8, 12, 15, and 20). However, despite this effort, so far we have not been able to dissect what precise conformations the *gluco*- and *manno*-configured substrates adopt in bound states in the active sites of plant GH1 β -D-glucosidases and β -D-mannosidases, what transition states develop during their catalytic cycles, and what structural determinants underlie the substrate specificities of these two closely related β -D-glycosidases.

One approach that has previously been fruitful for descriptions of hydrolase–glycoside interactions is to use thio analogues that mimic natural substrates (e.g., refs (21–23)). In the work presented here, we have used the S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides and the GH1 enzymes Os3BGlu7 from rice and HvBII from barley. These enzymes represent two types of β -D-glycosidases with distinct substrate specificities; i.e., these catalysts prefer to hydrolyze β -D-glucosides and β -D-mannosides, respectively (3, 4, 6, 7, 24). But contrary to employing X-ray crystallography as an experimental tool that has been used in the majority of previous studies, here we interchangeably used saturation transfer difference nuclear magnetic resonance (STD NMR) with transferred nuclear Overhauser effect (trNOE) spectroscopy, in conjunction with predictive computational quantum mechanics/molecular mechanics (QM/MM) calculations. The approaches specified above determined that O- and S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides

adopted a range of conformations. The resultant data are discussed in relation to the substrate preferences of the Os3BGlu7 and HvBII enzymes.

EXPERIMENTAL PROCEDURES

Materials. 4-Nitrophenyl β -D-glucopyranoside (4NP-O-Glc) and 4-nitrophenyl β -D-mannopyranoside (4NP-O-Man) were obtained from Sigma Chemical Co. (St. Louis, MO), and 4NP-S-Glc and 4NP-S-Man were prepared through the published organo-synthetic procedures for 1-thioglycosides (25). The source of all other chemicals was specified elsewhere (4, 6).

Cloning, Heterologous Expression, and Purification. The cDNAs encoding Os3BGlu7 and HvBII were subcloned into the pET32a expression vector encoding a thioredoxin folding partner protein (Novagen, Madison, WI), at the *Nco*I and *Xho*I restriction sites as described previously (6). The HvBII and Os3BGlu7 cDNAs encoding mature proteins corresponded to the EU807965 and U28047 sequences, respectively. Protein expression was conducted after chemical transformation in *Escherichia coli* Origami(DE3) cells (Novagen) and selection on Luria-Bertani agar containing 50 μ g/mL ampicillin, 15 μ g/mL kanamycin, and 12.5 μ g/mL tetracycline. The selected clones were induced for 8 h at 20 °C with 0.3 mM isopropyl β -D-thiogalactopyranoside as previously described (24). The Os3BGlu7 and HvBII enzymes were synthesized as His₆-tagged fusion proteins that were purified by immobilized metal affinity chromatography using a Talon resin (Invitrogen Life Technologies, Carlsbad, CA). The NH₂-terminal thioredoxin-His tags were removed by digestion with enterokinase (New England Biolabs, Beverly, MA) (6, 26). The recombinant enzymes were reconstituted in 50 mM sodium acetate buffer at their pH optima, i.e., at pH 5.0 for Os3BGlu7 and pH 4.0 for HvBII (6, 7).

Kinetic Constants and Inactivation by Thio Inhibitors. Kinetic constants of Os3BGlu7 and HvBII were determined as previously detailed (6). The reaction mixtures contained 0.038–1.5 mM (0.2–3 times the K_m value) 4NP-O-Glc or 0.042–3.81 mM (0.2–3 times the K_m value) 4NP-O-Man, 0.016% (w/v) BSA, and 2–4 pmol (24–48 nM) of Os3BGlu7 or HvBII in 50 mM sodium acetate buffer at pH 5.0 or 4.0, respectively. The enzyme inactivation constants were determined using 0–600 μ M 4NP-S-Glc or 4NP-S-Man. Each inhibitor was tested at four concentrations that were 0.4–3 times the individual K_i (dissociation constant of the enzyme–inhibitor complex) values. The reactions were developed with 2 M Na₂CO₃, and the amount of 4-nitrophenol (4NP) released was measured in microtiter plates using a PolarStar Optima Plate Reader (BMG LabTech, Offenburg, Germany). The enzyme activity was monitored at 405 nm, and the K_i values were determined from Dixon plots (inhibitor concentrations vs 1/v). Substrate hydrolyses during incubations never exceeded 10% of their initial concentrations, and kinetic constants were determined in triplicate at 30 °C. Hydrolytic and inhibition parameters were calculated by non-linear regression of the Michaelis–Menten curves with Graft version 5.0 (Erithacus Software, Horley, Surrey, U.K.).

STD NMR. STD NMR experiments were performed in a phosphate buffer of two different pH values as specified below (D₂O, uncorrected for isotope effects), between 290 and 300 K without saturation of a residual HDO (hydrogen/deuterium water) signal at 4NP-S-Glc or 4NP-S-Man to enzyme molar ratios between 20:1 and 50:1. A train of Gaussian-shaped pulses of 50 ms for each measurement was employed, with a total saturation time of the protein envelope of 2 s and a maximum B1

field strength of 50 Hz. An off-resonance frequency (δ) of 40 ppm and an on-resonance frequency (δ) of -1.0 ppm (protein aliphatic signal region) were applied. In all cases, the line broadening of inhibitor protons was monitored. In each case, the intensities were normalized with respect to the strongest response, which always corresponded to one of the H2 protons. The STD NMR experiments were repeated twice and the results averaged.

For the nuclear Overhauser effects (NOEs) in free states, selective experiments that employed a double-pulse field-gradient spin-echo (DPFGSE) module were conducted. NOE intensities were normalized with respect to the diagonal peak at zero mixing time. Selective T_1 measurements were performed on anomeric and several other protons to obtain the values mentioned above. Experimental NOEs were fitted to the double-exponential function $f(t) = p_0(e^{-p_1 t})(1 - e^{-p_2 t})$, as described previously (27, 28), where p_0 , p_1 , and p_2 are adjustable parameters. The initial slope was determined from the first derivative at time zero [$f'(0) = p_0 p_2$]. Interproton distances were obtained from the initial slopes by employing the isolated spin-pair approximation (ISPA).

The trNOESY experiments with HvBII were performed with freshly prepared mixtures containing HvBII and 4NP-S-Glc or 4NP-S-Man inhibitors, at approximately 30:1 molar ratios, in 20 mM phosphate buffer (pH 5.0) (near the pH optimum), with mixing times of 50, 100, 150, and 200 ms (28), resulting in final thio inhibitor concentrations of approximately 2–3 mM. No purging spin-lock period was employed to remove the NMR signals arising from the HvBII background, because typically the background signals are not observed with large proteins such as HvBII. First, line broadening of the inhibitor protons was monitored after the addition of the enzyme. The theoretical analysis of the trNOEs of 4NP-S-Glc or 4NP-S-Man protons was performed using Complete Relaxation and Conformational Exchange Matrix (CORCEMA), and a relaxation matrix with exchange as described previously (28). Varying exchange rate constants were employed to obtain optimal matches between the experimental and theoretical results of the methylene protons of 4NP-S-Glc or 4NP-S-Man at the C6 atoms, which could be separated by a fixed distance of 1.8 Å. The overall correlation time (τ_c) for the free states of inhibitors was always set to 50 ps for the monosaccharides, while τ_c was set to 10 ns for bound states. To fit the experimental trNOE intensities, off-rate constants between 100 and 1000 s $^{-1}$ were tested. An optimal agreement was achieved for a k_{off} of 100–300 s $^{-1}$.

Interactions of Os3BGlu7 with 4NP-S-Glc and 4NP-S-Man were investigated as described above, except that 20 mM Tris-HCl buffer (pH 8.0) or 20 mM phosphate buffer (pH 6.0) (near the pH optimum) was used, to avoid protons in NMR spectra in the latter case. The other experimental details used were as specified for HvBII.

All spectra were recorded between 290 and 300 K on a Bruker AVANCE 500 MHz spectrometer, equipped with a triple-channel ^1H , ^{13}N , ^{15}N gradient probe and processed with the Topspin Bruker software.

Homology Modeling of HvBII. A rice β -D-glucosidase Os3BGlu7 [PDB entry 2RGL (16)] was used as a structural template. The template sequence was aligned with that of HvBII using ClustalW (29), and the alignment was checked manually to maintain the integrity of secondary structure elements (30). The levels of sequence similarity and identity between HvBII and 2RGL were 92 and 66%, respectively. The structurally aligned sequences were used as input parameters to generate 200 models of HvBII with Modeler 9v2 (31). The top five models with the

lowest value of the Modeler 9v2 Objective Function were assessed by the DOPE module of Modeler 9v2, and the model with the most favorable value was chosen for evaluation. The overall G factors (estimates of stereochemical parameters) in both coordinate files, assessed by PROCHECK (32), were 0.31 and 0.04 for 2RGL and HvBII, respectively, with 99.8 and 0.2% of the residues in the favored, additional, and generously allowed regions and disallowed regions, respectively. WHATIF (33) evaluated the packing environment in the structures and returned quality control values of -0.97 and -0.77 for 2RGL and HvBII, respectively, while Verify 3D (34), which scores the fitness of protein sequences in their 3D environment, showed that 98.3% (2RGL) and 97.7% (HvBII) of the residues had an acceptable 3D–1D score. The z score values (35) reflecting combined statistical potential energy for 2RGL and HvBII were -11.59 and -12.38 , respectively. The rmsd value in the C α positions between the model and its template was 0.147 Å over 470 residues, from a total of 474 and 479 residues in 2RGL and HvBII, respectively, as determined with Stamp (36). Surfaces of Os3BGlu7 and HvBII were calculated with PyMol (37) with a probe radius of 1.4 Å, electrostatic potentials with the Adaptive Poisson–Boltzmann Solver (the dielectric constants of the solvent and solute were 80 and 2, respectively) (<http://apbs.sourceforge.net/>) implemented in PyMol as a plugin, and mapped on the protein molecular surfaces. The overall buriedness parameters of the HvBII and 2RGL active site funnels were calculated by the PocketPicker plugin (38) in PyMol. Molecular graphics were generated with PyMol and the Maestro graphical interface within the Schrödinger software suite (39).

Computational Docking of 4NP-O- and 4NP-S-Glycosides. Twelve starting structures for 4NP-Glc, 4NP-Man, 4NP-S-Glc, and 4NP-S-Man were constructed (39). For 1-O- and 1-S-glucos- and -mannopyranosides, three ring conformations were considered, namely, $^1\text{S}_3$ (^1B) (1B, 2B, 3B, and 4B), $^3\text{S}_5$ (B_3O) (1S, 2S, 3S, and 4S), and $^4\text{C}_1$ (1C, 2C, 3C, and 4C). All 12 structures were fully optimized using the DFT/M05-2X method (40–42) at the 6-31+G* basis. The electrostatic potential (ESP) charges fitted to the atom centers were calculated at the same level of theory that was used for the optimized structures. Calculations were conducted using Jaguar (39).

The 12 optimized structures of β -D-glycosides were docked in the active sites of the crystal structure of Os3BGlu7 (PDB entry 2RGM) and the HvBII model. The preparation of proteins for docking consisted of the following steps. First, both protein structures were reduced to single monomeric units, and water and bound inhibitor molecules were removed. The protein preparation utility within the Schrödinger suite was used to add hydrogen atoms, explicitly define disulfide linkages, and assign protonation states of residues. The structures of both proteins were minimized through 1000 cycles with MacroModel (39) using the OPLS-AA force field (43). The last structures from minimizations were selected for the calculation of the ESP charges for the residues in the active sites utilizing the hybrid QM/MM approach, as implemented in the QSite subroutine (39). For this purpose, the DFT/M05-2X method with the lacvp basis set was used for calculations, over the QM region that consisted of residues within an ~ 4 Å radius around the -1 subsite. The MM part contained the remainder of the proteins that were treated with an OPLS-AA all-atom force field approximation. The QM/MM boundaries were treated with a hydrogen capping approach. Both proteins were assigned charges from the QM/MM calculations that were used to construct the grid models for docking. The docking sites

Table 1: Kinetic Constants for Hydrolysis and Inhibition by Os3BGlu7 and HvBII

	K_m^a (mM)	k_{cat}^a (s ⁻¹)	$k_{cat}K_m^{-1b}$ (s ⁻¹ mM ⁻¹)	K_i (μM)	ΔG^c (kJ mol ⁻¹)
Os3BGlu7					
4NP-O-Glc	0.23 ± 0.02	7.90 ± 0.4	35.1 ± 1.0		
4NP-O-Man	1.30 ± 0.1	1.32 ± 0.05	1.01 ± 0.02		
4NP-S-Glc				664.3 ± 64.4	-18.4
4NP-S-Man				710.3 ± 115.1	-18.3
HvBII					
4NP-O-Glc	0.50 ± 0.03	0.50 ± 0.07	1.0 ± 0.05		
4NP-O-Man	0.25 ± 0.01	3.10 ± 0.02	12.7 ± 0.2		
4NP-S-Glc				94.7 ± 16.8	-23.3
4NP-S-Man				265.9 ± 57.2	-20.7

^aRounded to a single decimal digit, when higher than 1.0. ^bData from refs 6 and 7. ^cCalculated according to the equation $\Delta G = -RT \ln(1/K_i)$ (59).

were restricted to a cube with a 12 Å radius around the active site; no additional constraints were defined, and all 12 β-D-glycosides structures were docked into the active sites of both enzymes. The GLIDE program from the Schrödinger suite (39, 44) was used to calculate and evaluate descriptors for docking fitness. The resultant complexes were inspected for proper orientations and distances of bound substrates with respect to the corresponding active site residues.

RESULTS

Kinetic Parameters for Inhibition by β-D-Thioglycosides.

The parameters K_i and ΔG resulting from measurements of inhibition kinetics of rice Os3BGlu7 and barley HvBII with 4NP-S-Glc and 4NP-S-Man are listed in Table 1. The values of the K_i constants varied between 95 and 710 × 10⁻⁶ M with both enzymes, and the thioglycosides exhibited competitive-type inhibition (data not shown). Hence, the $v = (V_{max}S)/(K_m(1 + I/K_i) + S)$ equation was used for calculations of K_i constants. The ΔG values of inactivation for HvBII with 4NP-S-Glc and 4NP-S-Man were approximately 2–5 kJ/mol more favorable than the respective values for Os3BGlu7. This difference represented a small but significant variation between the two enzymes (Table 1). Table 1 also contains the k_{cat} and K_m values for hydrolysis of 4NP-O-Glc and 4NP-O-Man that were previously determined (4, 6, 7), but for comparison to our work, these constants are included. The catalytic efficiency factor for Os3BGlu7 with 4NP-O-Glc was ~35-fold higher than that for 4NP-O-Man (6, 7), while for HvBII, this second-order rate constant ($k_{cat}K_m^{-1}$) was ~13-fold higher with 4NP-O-Man than with 4NP-O-Glc (Table 1).

STD and *tr*NOESY NMR Investigations. As a first step, the conformations of free thio inhibitors 4NP-S-Glc and 4NP-S-Man were investigated (Figure 1). The analysis of vicinal proton–proton coupling constants permitted deduction of the exclusive existence of ⁴C₁(D) chair conformations for both inhibitors in solution, as expected. In these cases, NOE experiments were performed to measure the cross relaxation rates that could be correlated with interproton distances. Positive NOEs were obtained at a variety of mixing times from 200 to 1000 ms (Figure 1), as expected for small molecules that display fast tumbling in solution. Excellent agreement was found between the experimentally estimated distances from NOE signals and those measured for the MM3*-minimized ⁴C₁(D) 4NP-S-Glc and 4NP-S-Man chair geometries (data not shown).

The STD NMR experiments with HvBII permitted the unambiguous deduction of that the enzyme recognized both inhibitor molecules, because clear STD signals were observed for all the protons in 4NP-S-Glc and 4NP-S-Man (Figure 2; the data

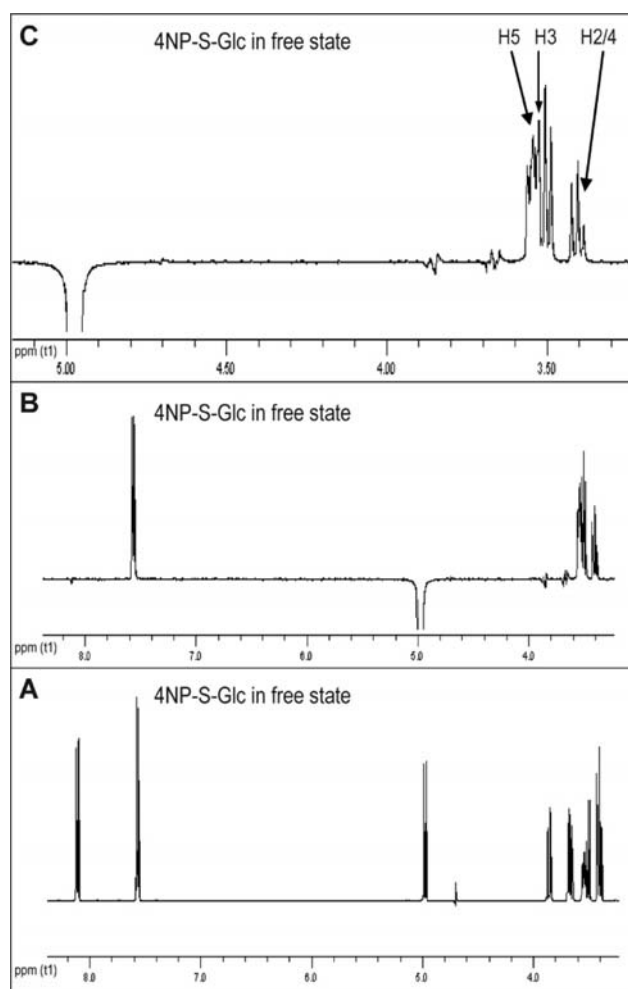


FIGURE 1: Nuclear Overhauser effects observed for free 4NP-S-Glc. (A) The 500 MHz NMR spectrum of 4NP-S-Glc in a free state. Overlapping of H2 and H4 occurs at the high field signal (δ 3.40). (B) NOE enhancements (mixing time of 600 ms) measured upon inversion of an anomeric signal (δ 4.98). Positive NOEs at the ortho-4NP, H5, H3, and H2/4 protons are clearly observed. (C) Expansion of the ring sugar protons, showing that the signal for the H1–H5 NOE was slightly larger than that for the H1–H3 proton pair. The latter pair was significantly larger than that for the H1–H2/4 proton pair. The observations were valid for all mixing times between 200 and 1000 ms.

shown for 4NP-S-Glc). Because of the small size of the thio inhibitors, a clear quantitative definition of the inhibitor epitopes was not attempted, as it was expected that the entire ring of thioglycosides would be recognized by HvBII. Here, a small transfer of saturation to the hydroxymethyl protons was observed with

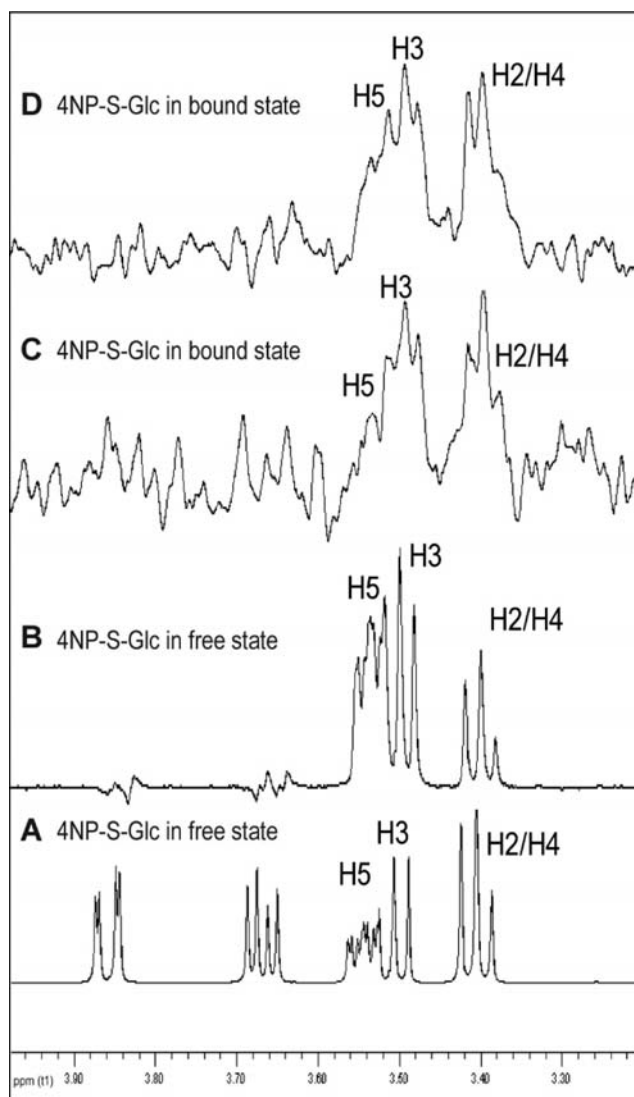


FIGURE 2: Binding of 4NP-S-Glc to HvBII. (A) The 500 MHz NMR spectrum of 4NP-S-Glc in the free state. The arrows indicate the positions of the H1–H5, H1–H3, and H1–H2/4 proton pairs. The overlapping of H2 and H4 occurred at a high field signal (δ 3.40). (B) Section of the NOESY spectrum showing the NOE enhancements (mixing time of 600 ms) measured upon inversion of the anomeric signal of free 4NP-S-Glc (δ 4.98). Positive NOEs at the H1–H5, H1–H3, and H1–H2/4 proton pairs were observed. The NOE signal for the H1–H5 proton pair was slightly larger than that for the H1–H3 proton pair, which was significantly larger than that for the H1–H2 proton pair. (C) Section of the trNOESY spectrum (4NP-S-Glc:HvBII molar ratio of 30:1, mixing time of 100 ms) showing the trNOE enhancements measured upon inversion of the anomeric signal of free 4NP-S-Glc (δ 4.98). (D) Negative NOEs at the H1–H3 and H1–H2/4 proton pairs. The signal for the H1–H5 proton pair was weaker and the signal for the H1–H3 proton pair similar to that of the H1–H2/4 proton pair.

both thio inhibitors, while strong STD signals were observed for the 4NP protons, indicating that this moiety also binds to the HvBII active site (Figure 2A,B; the data shown for 4NP-S-Glc).

Further, strong and negative trNOESY cross-peaks were observed for both inhibitors at a 30:1 ligand:enzyme molar ratio (Figure 2C) with the 100 ms NOE mixing time, as well as at other regimes of 50–200 ms (data not shown). The data for 4NP-S-Glc contrasted with those of 4NP-S-Glc in a free state, where the NOE cross-peaks were positive (cf. Figures 1 and 2). This change from positive to negative cross-peaks upon addition of the

enzyme again indicated binding and that, importantly, these trNOESY experiments could be used to deduce the conformations of bound 4NP-S-Glc and 4NP-S-Man. A CORCEMA-based full-relaxation matrix analysis of the cross-peaks was performed to deduce the experimental proton–proton distances in the bound state, and these cross-peaks were then compared with those that were estimated for all possible chair and skew boat conformers of both inhibitors (Table S1 of the Supporting Information). It was evident that for 4NP-S-Glc, the relative cross-relaxation rates (σ) were rather different from those observed in a free state. Here, the σ values were measured for the H1–H2 pair (σ_{12}), and the data indicated that these values were similar to σ_{13} but were larger than σ_{15} (Table S1 of the Supporting Information). These observations contrasted with those obtained for free 4NP-S-Glc. In this case, σ_{15} was larger than σ_{13} , which in turn was 2-fold larger than σ_{12} . On the basis of these data, we concluded that, in the state bound to HvBII, 4NP-S-Glc could adopt a geometry for which the interproton H1–H2 distance was similar to the H1–H3 distance but was shorter than that between the H1 and H5 protons. These analyses suggested that a conformational change of the six-member ring of 4NP-S-Glc had occurred upon binding to HvBII, and thus, the 4NP-S-Glc inhibitor in the bound state no longer adopted a 4C_1 chair conformation (Table S1 of the Supporting Information).

In contrast, 4NP-S-Man bound to HvBII had very similar relative cross-relaxation rates for the key proton pairs σ_{12} , σ_{13} , and σ_{15} (Table S2 of the Supporting Information). Essentially, identical cross-relaxation rates were observed for both free and bound states, although they had different signs, a negative one for a bound state and a positive one for a free state. This quantitative data suggested that, in principle, no changes in the interproton distances occurred in 4NP-S-Man, when bound to HvBII, and thus no significant conformational geometry changes from the ground 4C_1 conformation occurred during the molecular recognition by HvBII (Table S2 of the Supporting Information).

The nature of the distortion of the pyranose ring of 4NP-S-Glc upon binding to HvBII was assessed by comparison of the experimental interproton distances of bound 4NP-S-Glc with those measured in the MM3*-minimized geometries of the pseudorotational itinerary of 4NP-S-Glc. Here, we considered two extreme chair geometries, ${}^4C_1(D)$ and ${}^1C_4(D)$, and five skew boat conformers, 0S_2 , 1S_3 , 3S_5 , 4S_0 , and 1S_5 . Obviously, a chair conformation could not explain the observed NOE enhancements and the corresponding cross-relaxation rates (Table S1 of the Supporting Information). The best match between the observed and computed distances was found for the 3S_5 skew boat conformer (Table S1 of the Supporting Information and Figure 3A), although given that there were intrinsic uncertainties for the trNOESY experiments and for the conversion of the cross-relaxation rates into distances, the existence of the 1S_3 geometry for bound 4NP-S-Glc could not completely be excluded (Figure 3B). It was characteristic to observe that, for the 3S_5 and 1S_3 geometries, the anomeric C–S bond adopted a pseudoaxial orientation with a proper orientation for the aglycone moiety to depart. The same pseudoequatorial arrangement was also adopted by the hydroxyl groups at the C2, C3, and C4 atoms of the glucose ring. The major difference between the 1S_3 and 3S_5 conformers of 4NP-S-Glc is in the orientation of the C6 hydroxymethyl group, which is pseudoaxial for the 3S_5 conformer and pseudoequatorial for the 1S_3 geometry (Figure 3A,B).

The analogous STD NMR and trNOESY analyses for 4NP-S-Man showed that no major distortion occurred upon binding

to HvBII, because the estimated interproton distances were in full agreement with those computed for the ${}^4\text{C}_1(\text{D})$ chair conformer (Figure 3C). Our findings indicated that the data obtained for 4NP-S-Man could function as an “additional internal reference point” for validating the conclusions deduced for the conformational changes of bound 4NP-S-Glc, given that both experiments were performed under identical experimental conditions.

The STD experiments with Os3BGlu7 interacting with 4NP-S-Glc provided weak transferred signals (data not shown). These experiments proceeded in 20 mM Tris-HCl buffer at a pH value of 8.0 that was similar to the pH value at which the crystal structure of Os3BGlu7 was determined [PDB entry 2RGL (16)]. A high magnetization transfer was observed only for the aromatic ring protons, while the peaks in the carbohydrate part of 4NP-S-Glc were barely visible. Also, the two-dimensional (2D) trNOESY method did not provide cross-peaks for the 4NP-S-Glc bound to Os3BGlu7 at a molar ratio of 20:1 or 50:1. We have further investigated 1D selective NOESY on H1 of the glucose residue using different mixing times; however, no NOESY peaks were observed in the carbohydrate or aromatic regions of the spectra (data not shown). Hence, we concluded that a rather weak binding of 4NP-S-Glc to Os3BGlu7 could be responsible for a lack of measurable signals. These conclusions were supported by the weak K_i constant (664 μM) for binding of 4NP-S-Glc to Os3BGlu7 (Table 1).

One of the reasons that explains why we did not detect the measurable NOE signals with Os3BGlu7 and 4NP-S-Glc at pH 8.0 could be that at this particular pH value, the enzyme's affinity for the thio inhibitor could be low. Hence, the measurements were repeated at pH 6.0 in 20 mM phosphate buffer, whereas under these conditions, the pH value was closer to the pH optimum of 5.0 of Os3BGlu7 (7). Nevertheless, under these conditions, Os3BGlu7 unexpectedly hydrolyzed 4NP-S-Glc to a

small extent, as evidenced by the appearance of new spectral peaks that corresponded to hydrolytic products of 4NP-S-Glc (data not shown). Hence, it could be concluded that with Os3BGlu7, 4NP-S-Glc should be cautiously used as a substrate mimic for investigations of binding interactions.

To investigate binding of 4NP-S-Man to Os3BGlu7, the experimental approach used for binding of 4NP-S-Glc to Os3BGlu7 was adopted, where we first used 20 mM Tris-HCl buffer (pH 8.0) and then 20 mM phosphate buffer (pH 6.0). The STD and transferred NOESY experiments conducted at pH 8.0 with 4NP-S-Man and Os3BGlu7 produced better signals at a molar ratio of 50:1 than at 20:1 (data not shown). A major saturation transfer was again observed on the aromatic protons, followed by the H1 and H2 protons of the mannosyl residue (data not shown). As indicated in Figure 4, in 20 mM phosphate buffer (pH 6) using 1D selective trNOESY conditions, the cross-peaks from H1 to H2, H3, and H5 exhibited the same intensity, as seen for the free inhibitor. These data suggested that no major distortion of the mannosyl ${}^4\text{C}_1$ chair ring had occurred.

QM Modeling and Docking Simulations. The geometry optimizations of the starting ${}^1\text{S}_3$, ${}^3\text{S}_5$, and ${}^4\text{C}_1$ conformers of the O- and S-linked 4NP-Glc and 4NP-Man were conducted at the M05-2X/6-31+G* level of theory, and the calculations yielded the β -D-glycoside conformers in their local minima (Figure 5). The optimized values of selected bond lengths, bond angles, and dihedral angles and ESP partial charges of conformers are summarized in Table S3 of the Supporting Information. A comparison of the structural features of the O- and S-linked 4NP-Glc and 4NP-Man conformers revealed that the glycosides displayed the following structural characteristics. The ESP partial charge on the anomeric carbon of the O-linked glycosides was less negative than the charges on the S-linked glycosides, ranging from 0.36 to 0 for the O-linked versus 0.05 to -0.51 for the S-linked glycosides. Further, the ESP charge on the glycosidic linkage atoms was less negative for the S-linkage (-0.01 to -0.24) than for the O-linkage (-0.29 to -0.54). The partial charge of the endocyclic oxygen was similar for both types of glycosides (-0.36 to -0.53 and -0.22 to -0.40 for the O- and the S-linked glycosides, respectively). The lengths of the C1–O1 linkages in the O-linked glycosides were approximately 0.4–0.5 Å shorter than the lengths of the C1–S1 linkages. Hence, the distances in the O-linked glycosides were between 1.41 and 1.44 Å, while the same distances in the S-linked glycosides were between 1.85 and 1.92 Å (Table S3 of the Supporting Information). The same trend was observed for the lengths of the O1–C7 linkages, which were shorter by approximately 0.44 Å than the lengths of the S1–C7 linkages (Table S3 of the Supporting Information). The calculated C1–X1–C7 angles

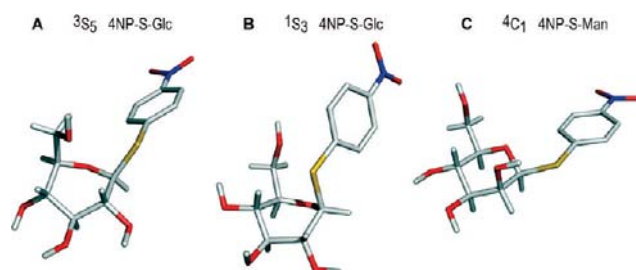


FIGURE 3: Geometries of 4NP-S-Glc and 4NP-S-Man bound to HvBII, as identified by STD NMR. 4NP-S-Glc is shown in the ${}^3\text{S}_5$ (A) and ${}^1\text{S}_3$ (B) conformations, while 4NP-S-Man is in the ${}^4\text{C}_1$ conformation (C).

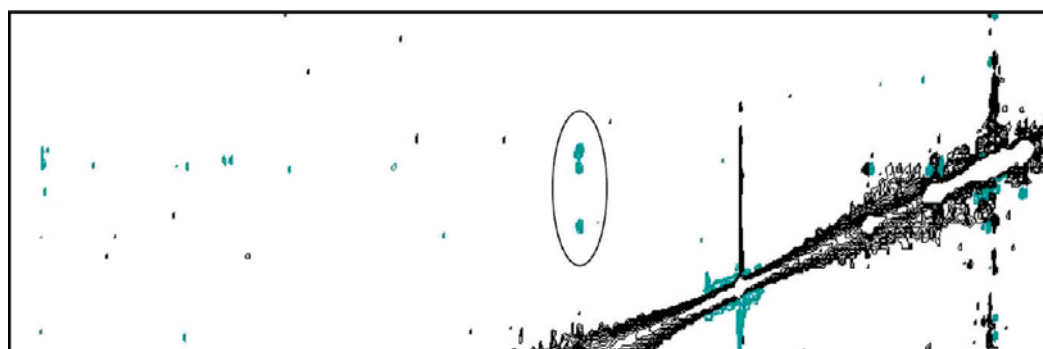


FIGURE 4: Binding of 4NP-S-Man to Os3BGlu7 at a molar ratio of 10:1. In the 2D spectrum, the circled cross-peaks between the anomeric proton and the H2, H3, and H5 protons of 4NP-S-Man showed similar intensities, suggesting that the glycoside was bound in the ${}^4\text{C}_1$ conformation.

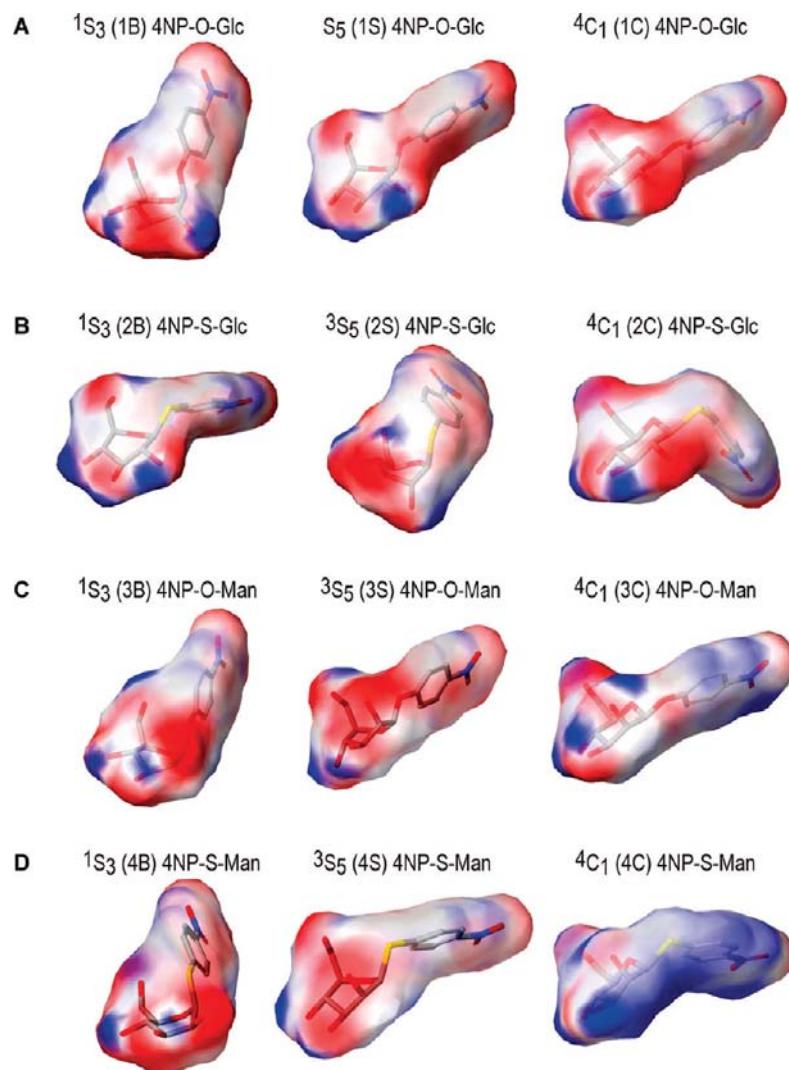


FIGURE 5: Optimized pyranose ring conformations of the *gluco*- and *manno*-configured aryl-glycosides in their local minima that were used for docking: (A) 4NP-O-Glc in 1S_3 (1B), 3S_5 (1S), and 4C_1 (1C); (B) 4NP-S-Glc in 1S_3 (2B), 3S_5 (2S), and 4C_1 (2C); (C) 4NP-O-Man in 1S_3 (3B), 3S_5 (3S), and 4C_1 (3C); and (D) 4NP-S-Man in 1S_3 (4B), 3S_5 (4S), and 4C_1 (4C). The contours denote surfaces and their electrostatic potentials derived from the ESP charges.

of the glycosidic linkages also varied significantly, whereby the C1–S1–C7 bond angles were sharper by as much as 21–30° than the C1–O1–C7 angles. The values were between 87° and 95° for the C1–S1–C7 angle and between 116° and 118° for the C1–O1–C7 angle (Table S3 of the Supporting Information). The minima around the C1–X1 linkages of all aryl-glycoside conformers, predefining their dihedral angles Θ , Φ , and Ψ , corresponded to the synclinal (sc) orientations of the phenyl groups with respect to the pyranose rings, and to the antiperiplanar (ap) orientations at the C2 atoms, in agreement with the exoanomeric effect (45).

Twelve optimized aryl-glycosides in 1S_3 , 3S_5 , and 4C_1 geometries, that is, 4NP-O-Glc in 1B, 1S, and 1C, 4NP-S-Glc in 2B, 2S, and 2C, 4NP-O-Man in 3B, 3S, and 3C, and 4NP-S-Man in 4B, 4S, and 4C, were docked in the active sites of Os3BGlu7 and HvBII, using GLIDE (Table 2 and Figures 6–8). As expected for a *gluco*-O-configured aryl-glucoside, the 1S_3 conformer, also known as the 1,4B (1B) skew boat conformer, fitted most favorably in Os3BGlu7 and HvBII, as supported by the GLIDE score values of –7.2 kcal/mol in both instances; these scores were the highest from all conformers for both enzymes (Table 2). The predicted affinities for the 3S_5 conformer, also known as $B_{3,O}$ (1S)

and 4C_1 (1C) conformers of 4NP-O-Glc, were between –5.2 and –6.0 kcal/mol for Os3BGlu7 and HvBII (Table 2). However, a similar GLIDE score value was predicted for the 1S_3 (2B) (–6.7 kcal/mol) and 4C_1 (2C) (–6.8 kcal/mol) conformers of 4NP-S-Glc docked in Os3BGlu7. Here the difference between the 1S_3 and 4C_1 conformers of 4NP-S-Glc was smaller than the precision of the scoring function. Therefore, on the basis of the GLIDE scores, we concluded that the 1S_3 (2B) or 4C_1 (2C) conformers of 4NP-S-Glc could equally be fitted in Os3BGlu7 (Table 2). On the other hand, for 4NP-S-Glc bound to HvBII, the higher scores were predicted for the 1S_3 and 4C_1 conformers (–5.8 and –5.7 kcal/mol, respectively) than for the 3S_5 conformer (–4.9 kcal/mol). Similarly, for the *manno*-configured aryl-glycosides, the best predicted score values were obtained for the 1S_3 conformations. In the case of 4NP-O-Man in 1S_3 (3B), the binding to Os3BGlu7 was slightly more favorable than that to HvBII, with scores of –7.3 kcal/mol versus –6.9 kcal/mol, while 4NP-O-Man in the 3S_5 and 4C_1 conformers bound less favorably, with affinities of –6.3 kcal/mol for Os3BGlu7 versus –5.6 kcal/mol for HvBII, and –5.5 kcal/mol for Os3BGlu7 versus –5.4 kcal/mol for HvBII (Table 2). The predicted affinities for 4NP-S-Man

Table 2: Summary of Docking of 4NP-O-Glc (¹S₃, ³S₅, and ⁴C₁ or 1B, 1S, and 1C), 4NP-S-Glc (¹S₃, ³S₅, and ⁴C₁ or 2B, 2S, and 2C), 4NP-O-Man (¹S₃, ³S₅, and ⁴C₁ or 3B, 3S, and 3C), and 4NP-S-Man (¹S₃, ³S₅, and ⁴C₁ or 4B, 4S, and 4C) in Os3BGlu7 and HvBII, Illustrating the Relative Conformational “Strain” Energy (Δ*E*) of Docked Conformers at the M05-2X/6-31+G* Level of Theory

	GLIDE score		<i>d</i> ^a (C1⋯O _{EA} ^b)		<i>d</i> ^a (O1/S1⋯O _{EB} ^b)		<i>d</i> ^a (HO ₂ ⋯O _{EB} ^b)		strain energy (Δ <i>E</i> ^a)	
	Os3BGlu7	HvBII	Os3BGlu7	HvBII	Os3BGlu7	HvBII	Os3BGlu7	HvBII	Os3BGlu7	HvBII
1B	−7.24	−7.15	3.19	2.89	3.64	3.42	1.61	3.20	17.21	18.31
1S	−6.02	−5.34	3.51	3.33	3.47	3.16	1.60	1.46	30.90	22.72
1C	−5.82	−5.20	5.13	5.02	4.38	4.02	7.27	3.39	24.05	16.45
2B	−6.66	−5.77	4.67	4.66	4.32	4.03	4.37	3.14	26.30	26.83
2S	−6.43	−4.86	4.64	3.48	3.40	3.37	7.08	1.50	26.58	33.80
2C	−6.76	−5.74	3.84	4.82	4.37	4.90	7.25	3.85	23.30	24.29
3B	−7.28	−6.89	3.33	3.00	4.22	3.69	4.10	3.01	30.44	26.10
3S	−6.33	−5.57	3.87	3.40	3.89	3.23	4.49	3.12	21.21	14.84
3C	−5.51	−5.36	4.39	4.92	3.51	3.87	4.28	2.46	25.61	21.80
4B	−7.56	−7.26	3.10	2.95	3.59	3.27	1.47	1.56	15.31	22.46
4S	−5.97	−5.70	4.96	4.40	3.13	4.46	4.21	5.20	26.39	15.66
4C	−6.61	−5.68	4.42	4.72	5.93	3.14	5.01	1.70	27.35	26.38

^aLengths (*d*) in angstroms and GLIDE scores and strain energies (Δ*E*) in kilocalories per mole. ^bEA denotes the catalytic nucleophile residues Glu386 in Os3BGlu7 and Glu389 in HvBII, and EB denotes the catalytic acid/base residues Glu176 in Os3BGlu7 and Glu179 in HvBII.

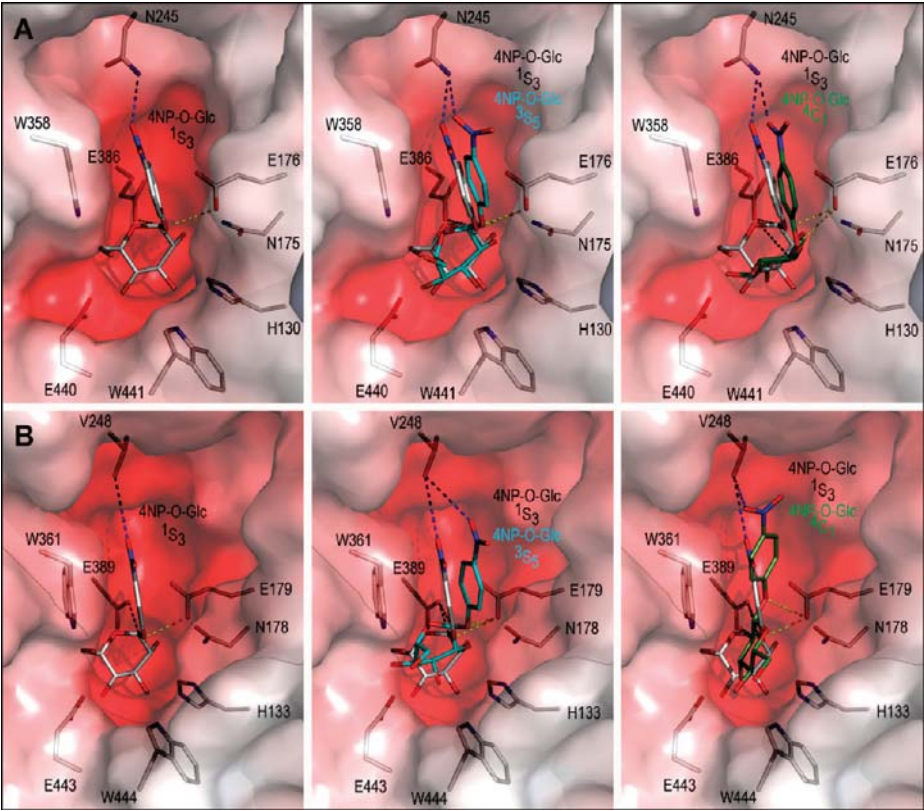


FIGURE 6: Predicted binding positions of 4NP-O-Glc in the active sites of Os3BGlu7 (A) and HvBII (B). The positions of 4NP-O-Glc (A) [left panel, in ¹S₃ (cpk); middle panel, in ¹S₃ (cpk) and ³S₅ (cpk cyan); right panel, in ¹S₃ (cpk) and ⁴C₁ (cpk green)] were predicted by GLIDE. The separations between the catalytic nucleophiles E386 (Os3BGlu7) and E389 (HvBII) and acid/base residues E176 (Os3BGlu7) and E179 (HvBII), and the O1 and C1 atoms in 4NP-O-Glc, respectively, are shown as yellow and black dashes. Blue dashes indicate separations between the O atom of the 4NP group and the nearest amino acid residue. Blue and red patches indicate electropositive and electronegative areas contoured at +5 and −10*kT*/e, respectively. Electrostatic potentials were calculated with Adaptive Poisson–Boltzmann Solver implemented in PyMol.

in the ⁴C₁ (4B) conformer were very similar to those of 3B with both enzymes, whereas the highest values of −7.6 kcal/mol (Os3BGlu7) and −7.3 kcal/mol (HvBII) were calculated for 4NP-S-Man in the ¹S₃ conformer. The GLIDE scores of 4NP-S-Man in the ³S₅ conformer with Os3BGlu7 (−6.0 kcal/mol) and HvBII (−5.7 kcal/mol) ranked relatively close to each other, although 4NP-S-Man bound in the ⁴C₁ conformer

was more favored for Os3BGlu7 (−6.6 kcal/mol) than HvBII (−5.7 kcal/mol).

To evaluate the predicted orientations of docked *gluco*- and *manno*-configured O- and S-linked 4NP-glycosides, closer inspections that are summarized in Figures 6–8 were conducted. The best docking positions of 4NP-O-Glc in the ¹S₃ (1B) geometry indicated that the pyranose rings were located in the intermediate

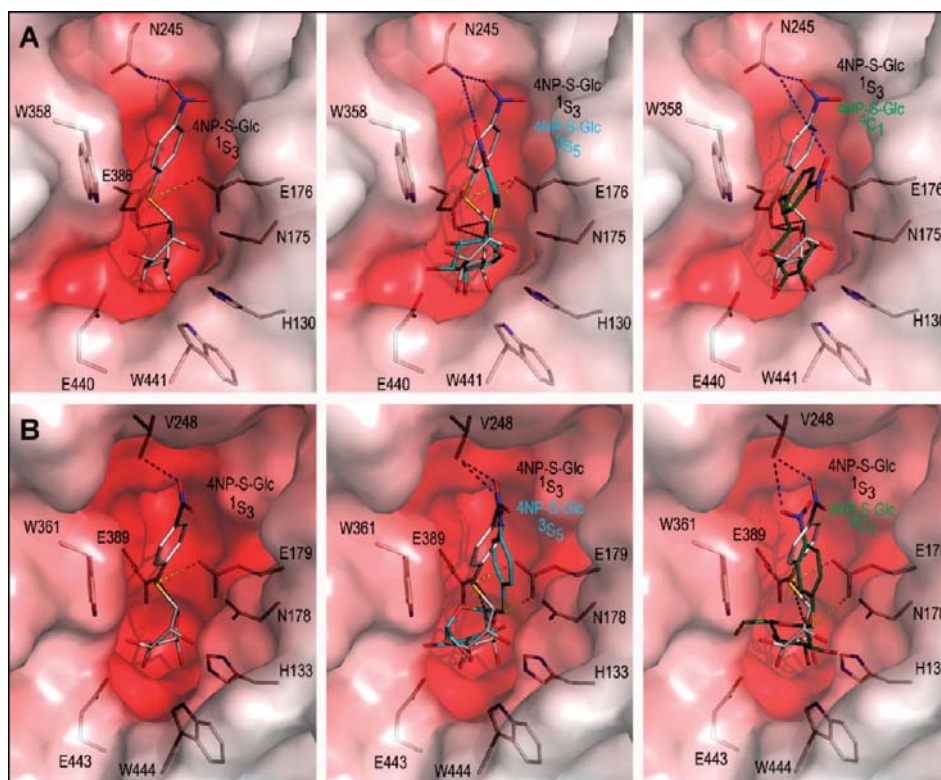


FIGURE 7: Predicted binding positions of 4NP-S-Glc in the active sites of Os3BGlu7 (A) and HvBII (B). In panel A, the positions of 4NP-S-Glc in (left) 1S_3 (cpk), (middle) 1S_3 (cpk) and 3S_5 (cpk cyan), and (right) 1S_3 (cpk) and 4C_1 (cpk green) were predicted by GLIDE. The separations between the catalytic nucleophiles E386 (Os3BGlu7) and E389 (HvBII) and acid/base residues E176 (Os3BGlu7) and E179 (HvBII), and the S1 and C1 atoms in 4NP-S-Glc, are shown as yellow and black dashes, respectively. Blue dashes indicate the separations between the O atom of the 4NP group and the nearest residue.

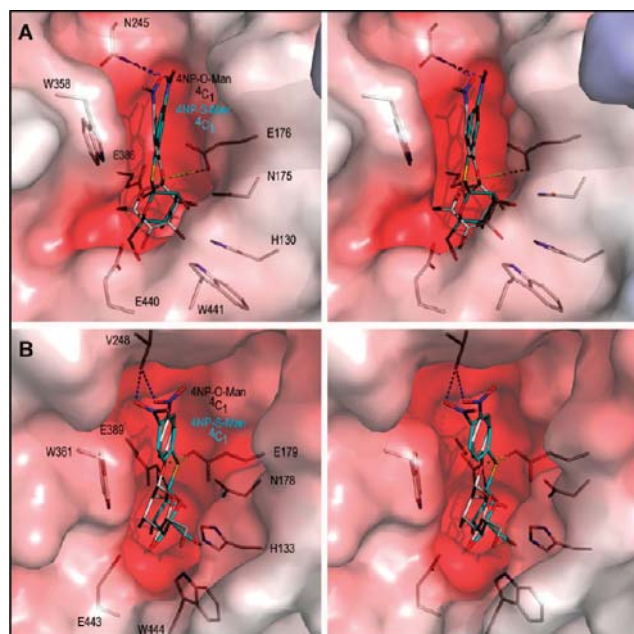


FIGURE 8: Stereoviews of predicted binding positions of 4NP-O-Man and 4NP-S-Man in the active sites of Os3BGlu7 (A) and HvBII (B). The positions of 4NP-O-Man (cpk) and 4NP-S-Man in the 4C_1 conformer (cpk cyan) were predicted by GLIDE. The separations between the catalytic nucleophiles E386 (Os3BGlu7), and E389 (HvBII) and acid/base residues E176 (Os3BGlu7) and E179 (HvBII), and the O1 (S1 for 4NP-S-Man) and C1 atoms in 4NP-mannosides, respectively, are shown as yellow and black dashes. Blue dashes indicate the separations between the O atom of the 4NP group and the nearest residue.

positions in the active sites of both enzymes and formed networks of hydrogen bonds with the surrounding residues (Figure 6). The 4NP aglycon of 4NP-O-Glc in 1S_3 was in a favorable orientation in the aromatic/hydrophobic funnel, although it did not stretch through its entire length (Figure 6A, right panels). The distances between the anomeric carbons of 4NP-O-Glc and one of the oxygen atoms of the catalytic nucleophiles (EA) Glu386 and Glu389 were 3.2 and 2.9 Å in Os3BGlu7 and HvBII, respectively. The predicted C1–O_{EA} distances in the complexes of Os3BGlu7 and HvBII with 4NP-O-Glc in the 1S_3 conformer (black dashes in Figure 6A,B, left panels) suggested that the anomeric carbons were positioned optimally for nucleophilic attacks by the oxygen atoms of catalytic nucleophiles. Further, the glycosidic oxygens in 1S_3 were in the reasonable proximities of the catalytic acid/base residues (EB) Glu176 and Glu179, with O1–O_{EB} distances of 3.6 and 3.4 Å in Os3BGlu7 and HvBII, respectively. In addition, the hydrogen atoms of the C2–OH groups pointed to the oxygen atoms of the catalytic acid/base residues. The separations between the protons and the oxygen atoms of EB (HO2–O_{EB}) were 1.6 and 3.2 Å for Os3BGlu7 and HvBII, respectively. However, in the case of HvBII, the hydrogen atom was rotated toward Glu389, forming a strong hydrogen bond that was shorter than the HO2–O_{EB} distance of 3.2 Å. These rather short distances suggested that exceptionally strong hydrogen bonds were formed between the participating atoms that may potentially stabilize the enzyme–substrate complex and contribute to the higher affinity of the 1S_3 conformer of 4NP-O-Glc to both enzymes, as predicted by GLIDE (Table 2). It was also clear from Table 2 that the affinities of 4NP-O-Glc in the 3S_5 and 4C_1 geometries were less favorable. Here, the distances between the catalytic residues and C1, O1, and

C2-OH were significantly longer (Table 2). It was remarkable to observe that the 4NP aglycons of 4NP-O-Glc in the 3S_5 and 4C_1 geometries stretched further through the pocket than that of the 1S_3 conformer; in particular, the 4NP aglycon of the 4C_1 conformer stretched all the way through the length of the HvBII active site (cf. Figure 6A,B, middle and right panels). While the conformational penalty of the 1S_3 conformer was by only 1 kcal/mol lower for Os3BGlu7 than for HvBII, the penalties of the 3S_5 and 4C_1 conformers of 4NP-O-Glc were clearly higher for HvBII, with energy differences of 8.2 and 7.6 kcal/mol, respectively (cf. ΔE values in Table 2, rows 3 and 4).

In Os3BGlu7, GLIDE predicted a different binding mode for 4NP-S-Glc in the 1S_3 geometry (Figure 7). In this instance, the predicted position of the pyranose ring of 4NP-S-Glc overlapped well with that of 2-deoxy-2-fluoroglucoside, observed in the crystal structure of Os3BGlu7 (16). However, upon comparison of the positions of the pyranose rings of 4NP-S-Glc with that of 4NP-O-Glc in the 1S_3 geometries in both enzymes, the pyranose rings of the S-linked 4NP-glycoside were slightly "out of alignment" in the active sites (Figures 6 and 7, right panels). Moreover, it could be observed that unlike the 1S_3 geometry of 4NP-O-Glc, the 1S_3 conformer of 4NP-S-Glc in Os3BGlu7 adopted a more "equatorial-like" orientation for the glycosidic linkage with respect to the pyranose ring. However, in HvBII, the 3S_5 and 4C_1 geometries of 4NP-S-Glc had the glycosidic linkages in equatorial orientations with respect to the pyranose rings, and these linkage orientations seemed to be optimally predisposed for hydrolysis (Figure 7). Here, the distances between C1 of 4NP-S-Glc in 1S_3 and O atoms of the catalytic nucleophiles stretched to 4.7 Å in both enzymes, and the distances between S1 of 4NP-S-Glc and O atoms of catalytic acid/base residues were 4.3 and 4.0 Å for Os3BGlu7 and HvBII, respectively. Further, the distances between the C2-OH groups of 4NP-S-Glc and the O atoms of catalytic acid/base residues were 4.4 and 3.1 Å in Os3BGlu7 and HvBII, respectively (Table 2). In summary, the distances between the C1 and S1 atoms of 4NP-S-Glc in 1S_3 (2B) conformer and the catalytic residues in both enzymes were significantly longer than those between the C1 and O1 atoms of 4NP-O-Glc in the 1S_3 (1B) conformer (Table 2).

Clear differences in the binding of 4NP-S-Glc in the 3S_5 geometry were observed between Os3BGlu7 and HvBII (Table 2 and Figure 7, middle panels). In Os3BGlu7, the participating respective distances of 4.6 and 3.4 Å between the EA and EB catalysts and the C1 atoms of 4NP-S-Glc in 3S_5 were slightly shorter than those for the 1S_3 conformer. Further, the pyranose ring of 4NP-S-Glc in the 3S_5 conformer adopted a "flipped-down" orientation compared to that of 4NP-O-Glc in 1S_3 , as demonstrated by the HO2–O_{EB} distance of 7.1 Å (Table 2). Notably, this conformation posed steric limitations for the interactions between the C1 atom of 4NP-S-Glc in 3S_5 and the O1 atom of Glu386 in Os3BGlu7. On the other hand, the C1 atom in the HvBII complex with 4NP-S-Glc in the 3S_5 geometry showed a better associative power with EA and EB than 4NP-S-Glc in the 1S_3 geometry. Here, the corresponding C1–O_{EA}, S1–O_{EB}, and HO2–O_{EB} distances were 3.5, 3.4, and 1.5 Å, respectively (Table 2). However, the large conformational penalty of 34 kcal/mol for the HvBII complex with 4NP-S-Glc in the 3S_5 geometry most likely contributed to a low GLIDE score of –4.9 kcal/mol for this docking position (Table 2). Finally, binding of 4NP-S-Glc in the 4C_1 geometry to both enzymes resulted in scores similar to those that were calculated for 4NP-S-Glc in the 1S_3 conformer, although these conformations again adopted the flipped-down pyranose ring

orientations. The latter orientations produced longer separations between the C1 atom of 4NP-S-Glc and the O atoms of the EA and EB catalysts and also generated steric hindrances between C1 of 4NP-S-Glc and EA in both enzymes (Figure 7, right panels).

The binding modes of the 4C_1 conformers of 4NP-O-Man and 4NP-S-Man in Os3BGlu7 and HvBII that were predicted by GLIDE are illustrated in Figure 8. The stereoviews of binding modes revealed that the two *manno*-configured O- and S-linked 4NP-glycosides adopted very different positions in both enzymes, although the 4NP aglycon moieties of 4NP-O-Man and 4NP-S-Man overlapped reasonably well. As illustrated in Figure 8B, the binding modes of *manno*-configured O- and S-linked 4NP-glycosides in HvBII were almost identical, while in Os3BGlu7, the 4NP aglycons adopted different orientations (Figure 8A). Compared to the 1S_3 mannose geometry, the pyranose rings of 4NP-O-Man in the 4C_1 geometry in both enzymes adopted flipped-down orientations, where the C1–O_{EA}, O1–O_{EB}, and HO2–O_{EB} distances were 4.4 and 4.9 Å, 3.5 and 3.9 Å, and 4.3 and 2.5 Å in Os3BGlu7 and HvBII, respectively. The conformational penalty of 4NP-O-Man in the 4C_1 conformer favored the HvBII complex over that of Os3BGlu7, as illustrated by the energy difference of nearly 4 kcal/mol (Table 2). On the other hand, the 4C_1 conformer of 4NP-S-Man scored slightly better with both enzymes, although the binding modes were dissimilar. In Os3BGlu7, 4NP-S-Man in the 4C_1 conformer adopted an orientation that was similar to that of 4NP-O-Glc in the 1S_3 conformer, and where the C1 and S1 atoms interacted weakly with EA and EB, as denoted by the C1–O_{EA}, S1–O_{EB}, and HO2–O_{EB} distances of 4.4, 5.9, and 5.0 Å, respectively (Table 2). However, in HvBII, although the C1 atom of 4NP-S-Man in 4C_1 interacted weakly with EA at a C1–O_{EA} distance of 4.7 Å, the S1 atom and the C2-OH group of 4NP-S-Man formed strong interactions with EB with S1–O_{EB} and HO2–O_{EB} distances of 3.1 and 1.7 Å, respectively. Hence, in the HvBII complexes with 4NP-O-Man and 4NP-S-Man in the 4C_1 geometry, the axial hydroxyl groups in the C2 positions and the oxygens of EB formed strong interactions, as indicated by their respective separations of 1.7 and 2.5 Å (Table 2 and Figure 8). Noteworthy is the fact that the conformational penalty with 4NP-S-Man in 4C_1 was similar for both Os3BGlu7 and HvBII, regardless of the separations between the axial hydroxyl groups in the C2 positions and the EB oxygens (Table 2).

DISCUSSION

A detailed knowledge of binding of β -D-glycosides has paramount importance for understanding of hydrolytic processes. Naturally, the deduction of molecular features of glycosidase–substrate complexes is a complex task as substrates are transformed into products at fast rates. To dissect how binding of β -D-glycosides proceeds at atomistic levels and in particular what specific sugar geometries participate during hydrolysis, mechanism-based inhibitors and substrate analogues have been used to mimic natural glycoside substrates (46). In other approaches, variant enzymes containing mutated active site residues have been employed (e.g., ref 47). The information gained from these studies is critical for glycosidase inhibitor design that has applications in a wide range of biotechnologies, including food and dietary fiber processing and production of pharmaceuticals, nutraceuticals, paper, pulp, wood, and biofuels.

There is mounting evidence that enzymes recognize substrates with a preferential conformational selection mechanism and that

in the absence of substrates, a dynamic equilibrium exists between substrate-free and substrate-bound configurations (48). It has been suggested that the most realistic information for substrate binding events is obtained with probes resembling putative transition states (46, 49). Here, binding of transition-state probes is linked to energy gains, resulting from binding of high-energy conformers on a transition-state itinerary of hydrolytic enzymes, where oxocarbenium ion-like-shaped intermediates have been suggested to play key roles (11). These events have been investigated predominantly by X-ray diffraction (e.g., refs 15 and (5052)), although STD NMR techniques have contributed significantly (27, 28, 53).

Against this background, the goal of our study was to shed light on the structural basis of binding of S- and O-linked *gluco*- and *manno*-configured aryl- β -D-glycosides to the Os3BGlu7 and HvBII enzymes that preferentially hydrolyze β -D-glucoside and β -D-mannoside substrates, respectively (3–7). Our aim was also to understand how chemical and structural features in substrates and enzymes' active sites underlie the substrate specificity of these enzymes that are classified in the GH-A clan of CAZY classification (1).

We first determined and compared kinetic parameters of inhibition (K_i) and hydrolysis (K_m) with S- and O-linked *gluco*- and *manno*-configured aryl- β -D-glycosides (Table 1). The comparisons of the kinetic data of HvBII and Os3BGlu7, using the corresponding O-glycoside substrates, revealed that the observed trends in K_i values of the thio analogues did not necessarily follow those in the K_m values. For example, with Os3BGlu7, where both thio inhibitors were almost equally effective, the K_m values for 4NP-O-Glc and 4NP-O-Man differed by a factor of ~ 6 . This conclusion was also valid for HvBII, where, while 4NP-S-Glc was nearly 3-fold more potent than 4NP-S-Man, the K_m value for the O-linked glucoside was 2 times higher than that for 4NP-O-Man (Table 1). Overall, the inhibition experiments clearly demonstrated that HvBII was far more sensitive to the inhibition by 4NP-S-Glc and 4NP-S-Man than Os3BGlu7 and that the K_m and K_i values with O- and S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides could not be correlated.

To reconcile the differences in binding of the O- and S-linked aryl glycosides that serve as respective substrates and inhibitors in Os3BGlu7 and HvBII, STD NMR and trNOESY investigations coupled with QM/MM modeling and docking experiments were undertaken. Here, we aimed to investigate precise conformations of S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides in bound states, and if the distortions from the 4C_1 ground states could be observed by STD NMR and trNOESY spectroscopy. These studies were complemented with predictive molecular modeling approaches. We surmised that these two approaches could provide useful insights, in particular, when crystal structures of inhibitor–ligand complexes are not available. The former technique has become a powerful tool in recent years for the characterization of binding of mechanistic probes to proteins (e.g., refs 27 and 28) and relies on saturation transfer from protein to ligand molecules. Thus, when a protein is selectively irradiated, the ligand between bound and free forms also becomes saturated in a bound state. Subtraction of the spectra acquired with off-resonance irradiation readily reveals conformations of a bound ligand, and the differential STD effects within a ligand molecule provide the information about the proximity of individual protons present in the vicinity of a protein. The advantage of this technique is that it is rapid and does not require isotope labeling of proteins or excessively large quantities of protein.

The STD NMR and trNOESY experiments were set up in several stages. As previously shown (28), when ligands bind weakly and when exchange rates between free and bound states occur at reasonably fast rates, trNOESY signals could provide adequate means for determining conformations of glycosides. After the S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides were incubated with HvBII, strong negative NOE cross-peaks were observed, indicating binding of inhibitors to the enzymes, in contrast to the free states of inhibitors. Via the trNOESY experiments with HvBII, we could unambiguously deduce conformational changes of the six-member rings of 4NP-S-Glc and 4NP-S-Man. Upon comparison of the experimental interproton distances of bound 4NP-S-Glc with those measured in the MM3*-minimized geometries, we concluded that 4NP-S-Glc adopted the 3S_5 or 1S_3 skew boat geometry, while 4NP-S-Man did not deviate from a starting low-energy 4C_1 geometry. The analogous STD NMR and trNOESY analyses for 4NP-S-Glc and 4NP-S-Man bound to Os3BGlu7 were also performed and showed that both thioglycosides remained in the ground 4C_1 conformations and that major saturation transfers were observed only on aromatic protons. The latter conclusions were supported by inhibition kinetics indicating that both β -D-thioglycosides were rather weak inhibitors (Table 1). The existence of no major contacts with the enzyme, in principle, suggests that no distortion of the ring is taking place upon binding. Further, the lack of observable NOESY cross-peaks for the sugar part during binding of 4NP-S-Glc and 4NP-S-Man to Os3BGlu7 may also suggest that future experiments with other inhibitors such as S-alkyl or S-benzyl derivatives are necessary and will ultimately clarify if the distortions of 4NP-S-glycosides upon binding to Os3BGlu7 occur.

The STD NMR and trNOESY experiments were complemented with QM/MM docking computations. During the docking procedure, the geometry of active site side chains was optimized, and this approach coupled with the QM-derived partial charges provided a basis for binding predictions by molecular docking. For these computations, we used a crystal structure of Os3BGlu7 (PDB entry 2RGM), while a homology model was constructed for HvBII, in the absence of its 3D structure. During the docking procedure, the orientations of the key amino acid residues in both proteins were sufficiently relaxed for the docked molecules to adjust. Although caution should be exercised in interpreting the computational data using homology models, as subtle errors in modeled HvBII could lead to erroneous miscalculations of docked positions and docking energies, in our experience the docking procedure using reliable homology models leads to reasonable predictions. Nevertheless, the use of the same models of, e.g., variant enzymes, where large conformational changes or variations in partial charges of amino acid residues occur, must be interpreted carefully because they could lead to erroneous prediction of binding affinities.

For docking of β -D-glycosides, we used as a guide an experimentally determined position of 2-deoxy-2-fluoroglucoside in Os3BGlu7 (16). When the characteristics of the active sites of Os3BGlu7 and HvBII were compared, it became obvious that they contained a plethora of charged and hydrophobic residues, where the catalytic nucleophiles Glu386 (Os3BGlu7) and Glu389 (HvBII) (4, 16) at subsite -1 were expected to form covalent linkages with the anomeric carbons of the substrates during formation of glycosyl–enzyme intermediates. The other important catalysts in the active sites of Os3BGlu7 and HvBII are the catalytic acid/base residues Glu176 (Os3BGlu7) and Glu179

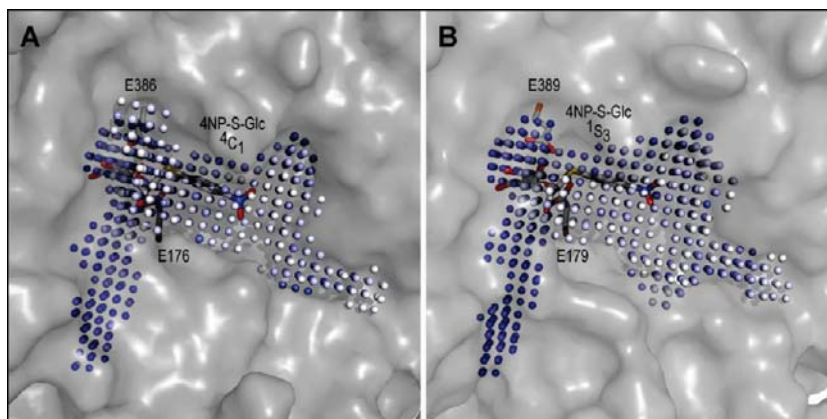


FIGURE 9: Shapes and buriedness of active site funnels of Os3BGlu7 and HvBII. The positions of 4NP-S-Glc (cpk) are shown in preferred geometries, 4C_1 for Os3BGlu7 (A) or 1S_3 for HvBII (B), as identified by STD NMR. The catalytic nucleophiles E386 (Os3BGlu7) and E389 (HvBII) and acid/base residues E176 (Os3BGlu7) and E179 (HvBII) are shown in cpk. Binding sites are given in the PocketPicker descriptors (38), with darker spheres indicating greater buriedness.

(HvBII) that, according to the canonical mechanism, should protonate glycosidic oxygens during hydrolysis. Moreover, in Os3BGlu7, the catalytic site is surrounded by a series of bulky hydrophobic/aromatic residues, namely, Trp433 and Trp441, and other residues such as Trp358 and Tyr315 that form a hydrophobic funnel (16). Here, a particular emphasis is placed on Trp433 positioned at the bottom of the active site pocket that mediates stacking interactions with the pyranose ring of 2-deoxy-2-fluoroglucoside (16).

One would expect that these interactions might be one of the major forces that keep substrates in favorable orientations for a nucleophilic attack. Further, at subsite 1, residues Ile179, Leu183, Tyr315, and Trp358 formed a hydrophobic funnel that presumably could accommodate the aglycon moieties of substrates. Thus, it would be expected that a specific and characteristic shape, buriedness, and geometry of active sites are required for particular conformations of substrates to bind in the active site regions. Here, a comparison of shape properties, using PocketPicker descriptors (38), indicated that both enzymes had very similar shapes, depths, and geometries of active site funnels (Figure 9). One would also expect that significant differences in conformations and charges in O- and S-linked 4NP-Glc and 4NP-Man might have serious consequences for the shapes of electrostatic surfaces of glycosides that could influence binding modes of substrates in enzymes' active sites, in particular those in subsites -1 and 1. These features are summarized in Figures 5 and 9 and in Table S3 of the Supporting Information, collectively indicating that O- and S-linked aryl- β -D-glycosides exhibit certain characteristic electrostatic potentials, and that these potentials are dependent on geometries of the individual pyranose rings. In particular, the data in Table S3 of the Supporting Information indicated that significant differences in the ESP charge distributions in the O- and S-linked glycosides occurred in the vicinity of their anomeric C1 atoms. It was therefore not unexpected that, after the glycosidic oxygens were replaced with sulfurs, the ESP charges on the C1 atoms changed from the positive values to the negative ones in all conformers (Table S3 of the Supporting Information). On the other hand, the overall ESP charges on the X1 and O4 atoms became less negative (Table S3 of the Supporting Information).

A few observations from the modeling experiments are worth analyzing. It seemed that in addition to characteristic positions of the S-linked *gluco*- and *manno*-configured aryl- β -D-glycoside

conformers in the active site funnels, defined by their separations from the respective active site residues, the overall sidedness of the pyranose ring orientations varied. The so-called flipped-down 4C_1 conformations of 4NP-S-Glc and 4NP-O-Man were adopted in Os3BGlu7 and HvBII (Figures 6 and 8), and a similar orientation was also seen in the 3S_5 conformer of 4NP-S-Glc that bound to Os3BGlu7 (Figure 7A, middle panel), compared to a 2-deoxy-2-fluoroglucoside moiety in the covalent complex of Os3BGlu7 (16). These flipped-down conformations resulted in unusually long C2-OH–O_{EB} separations of more than 7 Å in Os3BGlu7 (Table 2), except for that of 4NP-O-Man, and were also associated with longer separations between β -D-glycosides and respective active site residues EA and EB. This observation indicated that both enzymes could select incorrect orientations of incoming substrates that bind less favorably in their catalytic funnels. The importance of hydrogen bonds between the C2-OH group and catalytic residues was emphasized in a theoretical study of the catalytic mechanisms of β -D-glycoside hydrolases, where Bras and co-workers (52) concluded that the C2-OH–O_{EA} hydrogen bond is responsible for a low-energy activation barrier in the *E. coli* β -galactosidase catalytic mechanism.

Further, scoring algorithms that evaluate the accuracy of docking (Figures 6–8 and Table 2) and that could distinguish between various binding modes of ligands deserve mentioning, although the accuracy of predicting precise affinities in quantitative terms could be relatively low (44, 54). Nevertheless, the data from our modeling experiments clearly demonstrated that binding behavior of 4NP-S-Glc and 4NP-S-Man differed between the two plant enzymes. It was noteworthy that the bound O- and S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides did not adopt necessarily their lowest-energy conformations in bound states (Table 2). To estimate the strain energy imposed by the active site environments, we calculated the energy differences of bound aryl- β -D-glycosides at the M05-2X/6-31+G* level of theory. These energy differences arose from the differences between their local minima in free and bound states (Table 2). From these values, one could infer how binding of O- and S-linked *gluco*- and *manno*-configured aryl- β -D-glycosides in the higher-energy states influences the hydrolytic activation barrier.

Finally, it has been suggested for the members of the GH26 group of β -D-mannanases and (1,3;1,4)- β -D-glucan endohydrolases, which like the β -D-glucosidases and β -D-mannosidases

studied here are the members of clan GH-A, that while the interactions around the C1 and C2-OH groups were similar, environments around C3-OH groups could contribute to the specificity of the reaction coordinate (55). The authors further suggested that although the chemistry of the bound sugar is important, it does not dominate substrate specificity; whereas for *manno*-configured substrates, both B_{2,5} and ³H₄ transition states participate, for *gluco*-configured substrates, the ⁴H₃ and B^{2,5} transition states participate (55). Subsequently, a Michaelis complex of a GH2 β -D-mannosidase BtMan2A was published, which suggested that a boatlike transition state develops on a ¹S₅–B_{2,5}–⁰S₂ itinerary in this class of hydrolases (56). However, in the work described here, we concluded that other conformations, such as those participating in the ¹S₅–B_{2,5}–⁰S₂ itinerary for BtMan2A, are highly unlikely.

Our computational data indeed showed that the so-called flipped-down ⁴C₁ conformation of 4NP-O-Man was adopted in Os3BGlu7 and HvBII. On the other hand, the QM/MM modeling experiments indicated that the most favorable geometry for 4NP-O-Man with both Os3BGlu7 and HvBII was ¹S₃ (Table 2). Here we did not test the significance of ¹S₅ geometry, although we presume that geometrical interconversions between ¹S₅ and ¹S₃ should be less significant than those between ³S₅ and ¹S₃, and thus, the ¹S₃ conformer could be considered as an intermediate between the ³S₅ (B_{3,0}) and ¹S₅ conformers. The ¹S₃ geometry was also identified in 4NP-O-Glc with both β -D-glycosidases through the QM modeling and docking experiments, although STD NMR and trNOESY experiments confirmed its existence only with 4NP-S-Glc. To complicate the matters even further, it has recently been reported with a β -D-glucosidase from a soil metagenome library (57) that in addition to the glycon, the nature of the aglycon could contribute to substrate specificity. Similar observations were previously reported with a maize β -glucosidase (20) and diverse β -D-mannanases (58).

In summary, STD NMR and trNOESY experiments in conjunction with molecular docking simulations provided information about binding of S- and O-linked *gluco*- and *manno*-configured aryl- β -D-glycosides to rice Os3BGlu7 and barley HvBII that operate predominantly as a β -D-glucosidase and β -D-mannosidase, respectively. Kinetic analyses with 4NP-S-Glc and 4NP-S-Man indicated that the inhibitions were competitive. The STD NMR and trNOESY experiments revealed that 4NP-S-Glc and 4NP-S-Man bound weakly in ⁴C₁ conformations to Os3BGlu7, 4NP-S-Glc adopted the ³S₅ or ¹S₃ conformation, and 4NP-S-Man preferred ⁴C₁ geometry with HvBII. The docking and modeling studies predicted that 4NP-O-Glc, 4NP-O-Man, and 4NP-S-Man bound preferentially in the ¹S₃ geometries to both enzymes, contrary to 4NP-S-Glc that could adopt a range of geometries (¹S₃, ³S₅, or ⁴C₁). Although NMR, kinetic, and QM/MM examinations presented in this study provide valuable information about the conformations of sugars bound by the two plant β -D-glycosidases with distinct substrate specificities, we further surmise that the depth of information could be further enhanced via X-ray crystallography, and thus the amalgamation of the four techniques would ideally represent the most powerful approach for evaluation of binding affinities (60).

We conclude that in the plant β -D-glucoside hydrolase Os3BGlu7 and β -D-mannoside hydrolase HvBII a combination of determinants is likely to play roles in substrate recognition, such as (i) the inherent conformational and spatial flexibilities of

gluco- and *manno*-configured substrates in the enzymes' active sites, (ii) the subtle spatial differences in the disposition of active site residues and their capacities to form interactions with specific groups of substrates, and (iii) the small variations in charge distributions and shapes of the catalytic sites. We anticipate that the combination of these determinants may collectively drive the key interactions between the active site residues and *gluco*- and *manno*-configured substrates and underlie the enzymes' substrate specificities.

ACKNOWLEDGMENT

J.C., L.C., and J.J.-B. thank J. J. Hernández-Gay for helpful discussions.

SUPPORTING INFORMATION AVAILABLE

Supplementary Tables 1–3. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The carbohydrate-active enzymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238.
- Hrmova, M., and Fincher, G. B. (2009) Depolymerizing enzymes. In *Chemistry, Biochemistry and Biology of (1,3)- β -D-Glucans and Related Polysaccharides* (Bacic, T., Fincher, G. B., and Stone, B. A., Eds.) pp 119–170, Academic Press and Elsevier Inc., San Diego.
- Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Høj, P. B., and Fincher, G. B. (1996) Barley β -D-glucan exohydrolases with β -D-glucosidase activity. Purification and determination of primary structure from a cDNA clone. *J. Biol. Chem.* 271, 5277–5286.
- Hrmova, M., MacGregor, E. A., Biely, P., Stewart, R. J., and Fincher, G. B. (1998) Substrate binding and catalytic mechanism of a barley β -D-glucosidase/(1,4)- β -D-glucan exohydrolase. *J. Biol. Chem.* 273, 11134–11143.
- Hrmova, M., Burton, R. A., Biely, P., Lahnstein, J., and Fincher, G. B. (2006) Hydrolysis of (1,4)- β -D-mannans in barley (*Hordeum vulgare* L.) is mediated by a concerted action of a (1,4)- β -D-mannan endohydrolase and a β -D-mannosidase. *Biochem. J.* 399, 77–90.
- Kuntthom, T., Luang, S., Fincher, G. B., Opasiri, R., Hrmova, M., and Ketudat Cairns, J. R. (2008) Closely related barley and rice family GH1 glycosyl hydrolases with β -D-glucosidase and β -D-mannosidase activities. *Arch. Biochem. Biophys.* 491, 85–95.
- Opasiri, R., Hua, Y., Wara-aswapati, O., Akiyama, T., Svasti, J., Esen, A., and Ketudat Cairns, J. R. (2004) β -Glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem. J.* 379, 125–131.
- Marana, S. R. (2006) Molecular basis of substrate specificity in family 1 glycoside hydrolases. *IUMB Life* 58, 6–73.
- Vocadlo, D. J., and Davies, G. J. (2008) Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* 12, 539–555.
- Burmeister, W. P., Cottaz, S., Driguez, H., Iori, R., Palmieri, S., and Henrissat, B. (1997) 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* 5, 663–675.
- Zechel, D. L., and Withers, S. G. (2001) Dissection of nucleophilic and acid-base catalysis in glycosidases. *Curr. Opin. Chem. Biol.* 5, 643–649.
- Vallmitjana, M., Ferrer-Navarro, M., Planell, R., Abel, M., Querol, E., Planas, A., and Pérez-Pons, J.-A. (2001) Mechanism of the family β -glucosidase from *Streptomyces* sp.: Catalytic residues and kinetic studies. *Biochemistry* 40, 5975–5982.
- Ly, H. D., and Withers, S. G. (1999) Mutagenesis of glycosidases. *Annu. Rev. Biochem.* 68, 487–522.
- Cicek, M., Blanchard, D., Bevan, D. R., and Esen, A. (2000) The aglycone specificity-determining sites are different in 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)-glucosidase (maize β -glucosidase) and dhurrinase (sorghum β -glucosidase). *J. Biol. Chem.* 275, 20002–20011.

15. Czjzek, M., Cicek, M., Zamboni, V., Bevan, D. R., Henrissat, B., and Esen, A. (2000) The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, -DIMBOAGlc, and -dhurrin complexes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13555–13560.
16. Chuenchor, W., Pengthaisong, S., Robinson, R. C., Jirundon, Y., Esen, A., Chen, C.-J., Opassiri, R., Svasti, J., and Ketudat Cairns, J. R. (2008) Structural insights into rice BGLu1 β -glucosidase oligosaccharide hydrolysis and transglycosylation. *J. Mol. Biol.* 377, 1200–1215.
17. Street, I. P., Kempton, J. B., and Withers, S. G. (1992) Inactivation of a β -glucosidase through the accumulation of a stable 2-deoxy-2-fluoro- β -D-glucopyranosyl-enzyme intermediate: A detailed investigation. *Biochemistry* 31, 9970–9978.
18. Bause, E., and Legler, G. (1980) Isolation and structure of a tryptic glycopeptide from the active site of β -glucosidase A₃ from *Aspergillus wentii*. *Biochim. Biophys. Acta* 626, 459–465.
19. Aguilar, C. F., Sanderson, I., Moracci, M., Ciaramella, M., Nucci, R., Rossi, M., and Pearl, L. H. (1997) Crystal structure of the β -glycosidase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: Resilience as a key factor in thermostability. *J. Mol. Biol.* 271, 789–802.
20. Verdoucq, L., Moriniere, J., Bevan, D. R., Esen, A., Vasella, A., Henrissat, B., and Czjzek, M. (2004) Structural determinants of substrate specificity in family 1 β -glucosidases: Novel insights from the crystal structure of sorghum dhurrinase-1, a plant β -glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem.* 279, 31796–31803.
21. Sulzenbacher, G., Driguez, H., Henrissat, B., Schülein, M., and Davies, G. J. (1996) Structure of the *Fusarium oxysporum* endoglucanase I with a nonhydrolyzable substrate analogue: Substrate distortion gives rise to the preferred axial orientation for the leaving group. *Biochemistry* 35, 15280–15287.
22. Hrmova, M., De Gori, R., Smith, B. J., Fairweather, J. K., Driguez, H., Varghese, J. N., and Fincher, G. B. (2002) Structural basis for a broad specificity in higher plant β -D-glucan glucohydrolases. *Plant Cell* 14, 1033–1052.
23. Money, V., Cartmell, A., Guerreiro, C. I. P. D., Ducros, V. M.-A., Fontes, C. M. G. A., Gilbert, H. J., and Davies, G. J. (2008) Probing the β -1,3:1,4 glucanase, C_LLic26A, with a thio-oligosaccharide and enzyme variants. *Org. Biomol. Chem.* 6, 851–853.
24. Opassiri, R., Ketudat Cairns, J. R., Akiyama, T., Wara-aswapati, O., Svasti, J., and Esen, A. (2003) Characterization of a rice β -glucosidase highly expressed in flower and germinating shoot. *Plant Sci.* 165, 627–638.
25. Blanc-Muesser, M., Defaye, J., and Driguez, H. (1978) Syntheses stereoselectives de 1-thiogluco-sides. *Carbohydr. Res.* 67, 305–328.
26. Chuenchor, W., Pengthaisong, S., Yuvaniyama, J., Opassiri, R., Svasti, J., and Ketudat Cairns, J. R. (2006) Purification, crystallization and preliminary X-ray analysis of rice BGLu1 β -glucosidase with and without 2-deoxy-2-fluoro- β -D-glucoside inhibitor. *Acta Crystallogr. F* 62, 798–801.
27. Vogther, M., and Peters, T. (2000) Application of NMR based binding assays to identify key hydroxyl groups for intermolecular recognition. *J. Am. Chem. Soc.* 122, 6093–6099.
28. Clavel, C., Canales, A., Gupta, G., Cañada, F. J., Penadés, S., Suriola, A., and Jiménez-Barbero, J. (2007) NMR investigation of the bound conformation of natural and synthetic oligomannosides to banana lectin. *Eur. J. Org. Chem.* 10, 1577–1585.
29. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 24, 4876–4882.
30. Kim, D. E., Chivian, D., and Baker, D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32 (Suppl. 2), W256–W261.
31. Sali, A., and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.
32. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
33. Vriend, G. (1990) WHAT IF: A molecular modeling and drug design program. *J. Mol. Graphics* 8, 52–56.
34. Lüthy, R., Bowie, J. E., and Eisenberg, D. (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356, 83–85.
35. Sippl, M. J. (1993) Recognition of errors in 3-dimensional structures of proteins. *Proteins: Struct., Funct., Bioinf.* 17, 355–362.
36. Sumathi, K., Ananthalakshmi, P., Md Roshan, M. B. A., and Sekar, K. (2006) 3dSS: 3D structural superposition. *Nucleic Acids Res.* 34, W128–W134.
37. DeLano, W. (2009) The PyMOL Molecular Graphics System, DeLano Scientific LLC, San Carlos, CA.
38. Weisel, M., Proschak, E., and Schneider, G. (2007) PocketPicker: Analysis of ligand-binding sites with shape descriptors. *Chem. Cent. J.* 1, 7.
39. MacroModel, version 9.6 (2008) Schrödinger, LLC, New York.
40. Zhao, Y., Schultz, N. E., and Truhlar, D. G. (2005) Exchange-correlation functional with broad accuracy for metallic and nonmetallic compounds, kinetics, and noncovalent interactions. *J. Chem. Phys.* 123, 161103.
41. Zhao, Y., Schultz, N. E., and Truhlar, D. G. (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* 2, 364–382.
42. Zheng, J., Zhao, Y., and Truhlar, D. G. (2007) Representative benchmark suites for barrier heights of diverse reaction types and assessment of electronic structure methods for thermochemical kinetics. *J. Chem. Theory Comput.* 3, 569–582.
43. Jorgensen, W. L., Maxwell, D. S., and Tirado-Rives, J. (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118, 11225–11236.
44. Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* 47, 1739–1749.
45. Tvaroška, I., and Carver, J. P. (1996) *Ab initio* molecular orbital calculation of carbohydrate model compounds. 5. Anomeric, exoanomic, and reverse anomeric effects in C-, N-, and S-glycosyl compounds. *J. Phys. Chem.* 100, 11305–11313.
46. Rempel, B. P., and Withers, S. G. (2008) Covalent inhibitors of glycosidases and their applications in biochemistry and biology. *Glycobiology* 18, 570–586.
47. Blanchard, J. E., and Withers, S. G. (2001) Rapid screening of the aglycone specificity of glycosidases: Applications to enzymatic synthesis of oligosaccharides. *Chem. Biol.* 8, 627–633.
48. Davulcu, O., Flynn, P. F., Chapman, M. S., and Skaliky, J. J. (2009) Intrinsic domain and loop dynamics commensurate with catalytic turnover in an induced-fit enzyme. *Structure* 17, 1356–1367.
49. Ermert, P., Rupitz, K., Vasella, A., Weber, M., and Withers, S. G. (1993) Transition state analogue inhibitors of glycosidases are configurationally selective: A study with nojiritetrazoles, a new class of glycosidase inhibitors. *Carbohydr. Res.* 250, 113–128.
50. Hrmova, M., Streltsov, V. A., Smith, B. J., Vasella, A., Varghese, J. N., and Fincher, G. B. (2005) Structural rationale for low nanomolar binding of transition state mimics to a family GH3 β -D-glucan glucohydrolase from barley. *Biochemistry* 44, 16529–16539.
51. Gloster, T. M., Meloncelli, P., Stick, R. V., Zechel, D., Vasella, A., and Davies, G. J. (2007) Glycosidase inhibition: An assessment of the binding of 18 putative transition-state mimics. *J. Am. Chem. Soc.* 129, 2345–2354.
52. Bras, N. F., Moura-Tameres, A., Fernandes, P. A., and Ramos, M. J. (2008) Mechanistic studies on the formation of glycosidase-substrate and glycosidase-inhibitor covalent intermediates. *J. Comput. Chem.* 29, 2565–2574.
53. Martin-Pastor, M., Vega-Vázquez, M., De Capua, A., Canales, A., Andre, S., Gabius, H. J., and Jiménez-Barbero, J. (2006) Enhanced signal dispersion in saturation transfer difference experiments by conversion to a 1D-STD-homodecoupled spectrum. *J. Biomol. NMR* 36, 103–109.
54. Englebienne, P., Fiaux, H., Kuntz, D. A., Corbeil, C. R., Gerber-Lemaire, S., Rose, D. R., and Moitessier, N. (2007) Evaluation of docking programs for predicting binding of Golgi α -mannosidase II inhibitors: A comparison with crystallography. *Proteins* 69, 160–176.
55. Money, V. A., Smith, N. L., Scaffidi, A., Stick, R. V., Gilbert, H. J., and Davies, G. J. (2006) Substrate distortion by a lichenase highlights the different conformational itineraries harnessed by related glycoside hydrolases. *Angew. Chem., Int. Ed.* 45, 5136–5140.

56. Offen, W. A., Zechel, D. L., Withers, S. G., Gilbert, H. J., and Davies, G. J. (2009) Structure of the Michaelis complex of β -mannosidase, Man2A, provides insight into the conformational itinerary of mannoside hydrolysis. *Chem. Commun.*, 2484–2486.
57. Nam, K. H., Sung, M. W., and Hwang, K. Y. (2010) Structural insights into the substrate specificity of β -glucosidase. *Biochem. Biophys. Res. Commun.* 391, 1131–1135.
58. Tailford, L. E., Ducros, V. M., Flint, J. E., Roberts, S. M., Morland, C., Zechel, D. L., Smith, N., Bjørnvad, M. E., Borchert, T. V., Wilson, K. S., Davies, G. J., and Gilbert, H. J. (2009) Understanding how diverse β -mannanases recognize heterogeneous substrates. *Biochemistry* 48, 7009–7018.
59. Fersht, A. (1999) Structure and Mechanism in Protein Science, pp 1–631, W. H. Freeman and Co., New York.
60. Hrmova, M., and Fincher, G. B. (2009) Functional genomics and structural biology in the definition of gene function. In *Methods in Molecular Biology: Plant Genomics* (Somers, D., Langridge, P., and Gustafson, P., Eds.) Vol. 513, pp 175–198, Humana Press Inc., Totowa, NJ.



Contents lists available at ScienceDirect

Journal of Structural Biology

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

The structural basis of oligosaccharide binding by rice BGlu1 beta-glucosidase

Watchalee Chuenchor^{a,1}, Salila Pengthaisong^a, Robert C. Robinson^b, Jirundon Yuvaniyama^c, Jisnusun Svasti^c, James R. Ketudat Cairns^{a,*}

^aSchools of Biochemistry and Chemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand

^bInstitute of Molecular and Cell Biology, 61 Biopolis Dr., Proteos, Singapore 138673, Singapore

^cDepartment of Biochemistry and Center of Excellence in Protein Structure and Function, Faculty of Science, Mahidol University, Bangkok 10400, Thailand

ARTICLE INFO

Article history:

Received 22 June 2010

Received in revised form 27 August 2010

Accepted 22 September 2010

Available online xxxx

Keywords:

Glycoside hydrolase

Acid–base mutant

Cello-oligosaccharides

X-ray crystallography

Enzyme–substrate complex

ABSTRACT

Rice BGlu1 β -glucosidase is an oligosaccharide exoglucosidase that binds to six β -(1 $\rightarrow$ 4)-linked glucosyl residues in its active site cleft. Here, we demonstrate that a BGlu1 E176Q active site mutant can be effectively rescued by small nucleophiles, such as acetate, azide and ascorbate, for hydrolysis of aryl glycosides in a pH-independent manner above pH 5, consistent with the role of E176 as the catalytic acid–base. Cellotriose, cellotetraose, cellopentaose, cellohexaose and laminaribiose are not hydrolyzed by the mutant and instead exhibit competitive inhibition. The structures of the BGlu1 E176Q, its complexes with cello-tetraose, cellopentaose and laminaribiose, and its covalent intermediate with 2-deoxy-2-fluoroglucoside were determined at 1.65, 1.95, 1.80, 2.80, and 1.90 Å resolution, respectively. The Q176 N ϵ was found to hydrogen bond to the glycosidic oxygen of the scissile bond, thereby explaining its high activity. The enzyme interacts with cellooligosaccharides through direct hydrogen bonds to the nonreducing terminal glucosyl residue. However, interaction with the other glucosyl residues is predominantly mediated through water molecules, with the exception of a direct hydrogen bond from N245 to glucosyl residue 3, consistent with the apparent high binding energy at this residue. Hydrophobic interactions with the aromatic sidechain of W358 appear to orient glucosyl residues 2 and 3, while Y341 orients glucosyl residues 4 and 5. In contrast, laminaribiose has its second glucosyl residue positioned to allow direct hydrogen bonding between its O2 and Q176 O ϵ and O1 and N245. These are the first GH1 glycoside hydrolase family structures to show oligosaccharide binding in the hydrolytic configuration.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

β -Glucosidases (E.C. 3.2.1.21, β -D-glucopyranosidases) hydrolyze β -glycosidic bonds to release nonreducing terminal D-glucopyranosyl residues from oligosaccharides and glycosides (Ketudat Cairns and Esen, 2010). As such, they have become much studied due to their potential use in biomass conversion of cellulosic waste, together with cellulases and exoglucanases, which break the long cellulose chains into shorter (1 $\rightarrow$ 4)- β -linked glucooligosaccharides

that are favored substrates for certain β -glucosidases (Sticklin, 2006). Microbial β -glucosidases, which have evolved to take advantage of the vast cellulosic energy pool of plant cell wall materials, are typically used for biomass conversion applications. However, plants have separately evolved their own β -glucosidases for (1 $\rightarrow$ 4)- β -linked oligosaccharides generated in the process of cell wall recycling within the plant. One of the most well-studied of these plant enzymes is rice (*Oryza sativa* L.) BGlu1, systematically named Os3BGlu7, which was identified as a highly expressed isoenzyme in rice seedlings and subsequently noted to be the most abundantly expressed glycoside hydrolase family GH1 (Henrissat, 1991) member in rice (Opassiri et al., 2003; Opassiri, 2006).

GH1 proteins, including rice BGlu1, are members of glycoside hydrolase clan A, which consists of several families of glycoside hydrolases with catalytic domains that have a (β/α)₈ barrel structure and two conserved carboxylic acid residues on β -strands 4 and 7 of the β -barrel, which serve as the catalytic acid–base and nucleophile, respectively (Fig. 1A) (Jenkins et al., 1995; Henrissat et al., 1995). These enzymes are thought to act through a double displacement mechanism (Fig. 1B), whereby the glycosidic oxygen is protonated by the catalytic acid–base and the catalytic

Abbreviations: Abg, *Agrobacterium* sp. β -glucosidase; dNPGlc, 2,4-dinitrophenyl β -D-glucopyranoside; G2F, 2-deoxy-2-fluoro-glucoside; GH1, glycoside hydrolase family 1; Glc, glucosyl residue; pNPGlc, para-nitrophenyl β -D-glucopyranoside.

* Corresponding author. Address: Institute of Science, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima 30000, Thailand. Fax: +66 44 224185.

E-mail addresses: watchalee_lee@hotmail.com (W. Chuenchor), salila01@yahoo.com (S. Pengthaisong), rrobinson@icmb.a-star.edu.sg (R.C. Robinson), scjyv@mahidol.ac.th (J. Yuvaniyama), scjsv@mahidol.ac (J. Svasti), cairns@sut.ac.th (James R. Ketudat Cairns).

¹ Present Address: Dept. of Chemistry & Biochemistry, University of Maryland, College Park, MD 20742, USA

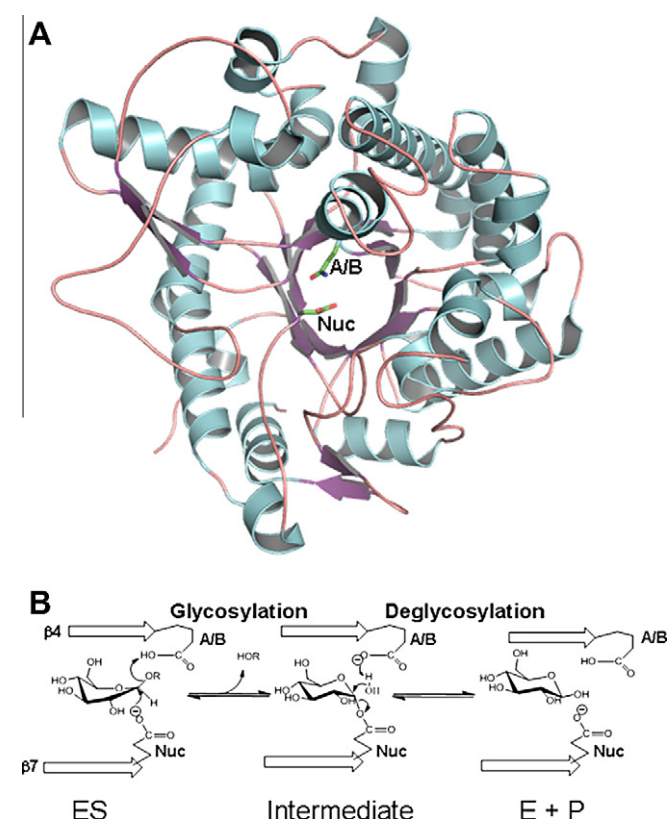


Fig. 1. Overall structure and double displacement mechanism proposed for glycosyl hydrolase family 1 β -D-glucosidases. (A). Cartoon diagram of the rice BGLu1 E176Q structural model (PDB entry 3F4V) with the catalytic acid–base (A/B) and nucleophile (Nuc) shown in stick representation. Note that in this acid/base mutant, the catalytic acid/base has been substituted with Gln, but it should normally be Glu. (B). In the double displacement mechanism, the β -glucoside substrate binds with the glycone in a distorted (1S_3 skew boat) conformation with the anomeric carbon between the acid–base glutamic acid (A/B) and the nucleophile glutamate (Nuc), located on beta-strands $\beta 4$ and $\beta 7$ of the (β/α)₈ barrel, respectively. Protonation of the leaving group allows the nucleophile to displace the aglycone from the glucosyl residue via an oxocarbenium ion-like transition state to form a covalent intermediate in the glycosylation step. Deglycosylation occurs when the intermediate is hydrolyzed by water, which is activated by extraction of a proton by the catalytic acid–base, essentially reversing the glycosylation step.

nucleophile attacks the anomeric carbon in the glycosylation step to form an α -linked covalent intermediate, followed by deglycosylation, in which a water or another nucleophile attacks at the anomeric carbon with basic assistance from the catalytic acid–base to displace the nucleophile and release the β -D-glucopyranose. Evidence for the formation of the covalent intermediate and the identity of the nucleophile was first provided by the trapping the GH1 *Agrobacterium* sp. β -glucosidase (Abg) linked to a 2-fluoroglucosyl residue, in which the electron withdrawal by the fluorine atom destabilizes the oxocarbenium cation-like transition state, greatly slowing both stages of the reaction (Withers et al., 1990). The use of a highly ionizable 2,4-dinitrophenolate as the leaving group of the glycoside substrate allowed the glycosylation to proceed relatively rapidly, while the deglycosylation half-reaction was slowed to trap the covalent intermediate. This covalent intermediate has since been observed in a number of GH1 enzyme crystal structures (Burmeister et al., 1997; Zechel et al., 2003; Gloster et al., 2004; Isorna et al., 2007; Chuenchor et al., 2008).

Mutagenesis to replace the catalytic nucleophile with small non-nucleophilic residues changed the GH1 mechanism to allow direct attack on β -linked substrates by small nucleophiles to form α -linked products or the use of α -fluoroglucoside as a surrogate for

the covalent intermediate in transglycosylation reactions (Mackenzie et al., 1998). Similarly, mutations of the catalytic acid–base validate its identity by providing catalysis when substrates with good leaving groups, like 2,4-dinitrophenolate, that do not require acid-assistance are used. These substrates allow the glycosylation step to proceed, and small anionic nucleophiles and sulfhydryls are provided to displace the enzyme from the glucosyl group in the deglycosylation step (Wang et al., 1995). The use of such acid–base mutants has allowed the trapping of unmodified glucosyl residues on the enzyme, as recently shown for the GH1 human cytoplasmic β -glucosidase E165Q mutant (Noguchi et al., 2008). Interestingly, a natural example of the rescue of an acid–base mutant occurs in the GH1 myrosinases (β -D-thioglycosidases), where the acid–base is replaced with glutamine and ascorbate seems to serve in the catalytic base function (Burmeister et al., 2000).

Recently, crystal structures have been determined for bacterial, fungal and plant GH1 enzymes that act on cellooligosaccharides (Isorna et al., 2007; Nijkken et al., 2007; Chuenchor et al., 2008). Rice BGLu1 has evolved separately from the microbial enzymes and is more closely related to plant enzymes that act on other substrates than to the microbial enzymes that act on cellooligosaccharides. X-ray crystal structures have been determined for a number of plant GH1 enzymes (Barrett et al., 1995; Burmeister et al., 1997; Czjzek et al., 2000, 2001; Verdoucq et al., 2004; Sue et al., 2006; Barleben et al., 2007; Seshadri et al., 2009), but none of the plant GH1 β -glucosidases with known structures is known to hydrolyze cellooligosaccharides efficiently, other than rice BGLu1. The aglycone specificity of the maize, sorghum and wheat β -glucosidases has been extensively studied by mutagenesis and structural studies of mutants with substrates or inhibitors (Czjzek et al., 2000, 2001; Zouhar et al., 2001; Verdoucq et al., 2003, 2004; Sue et al., 2006; Dopitova et al., 2008). These studies suggest that a conserved Trp, corresponding to rice BGLu1 W358, and several variable residues around the active site pocket are important for substrate aglycone binding. However, no general rule for how to identify the substrate specificity from the active site residues has emerged.

BGLu1, like a closely related β -mannosidase/ β -glucosidase from barley and its closely related rice isoenzymes Os3BGLu8 and Os7BGLu26, has been shown to hydrolyze cellooligosaccharides with up to 6 glucosyl residues in length with increasing efficiency, although the barley enzyme hydrolyzed cellobiose more rapidly than cellotriose (Hrmova et al., 1996; Hrmova et al., 1998; Opassiri et al., 2004; Kuntothom et al., 2009). Kinetic subsite mapping suggested that this difference was primarily due to the relative affinities at subsites +1 and +2 (which bind cellooligosaccharide glucosyl residues 2 and 3, respectively). Both enzymes hydrolyze laminaribiose with high efficiency, but show decreasing rates of hydrolysis of longer β -(1 $\rightarrow$ 3)-linked oligosaccharides as the degree of polymerization increases. The structures of BGLu1 and its covalent complex with 2-fluoroglucose showed a wide slot running out of the active site, which could act to bind oligosaccharides (Chuenchor et al., 2008). Mutagenesis of I179, N190 and N245, which line this slot affected the efficiency of hydrolysis of cellobiose and cellotriose, but computational docking of a cellotriose molecule into the active site was not precise enough to allow the interactions involved to be defined. To better understand the basis for hydrolysis of oligosaccharides, we have identified a catalytic acid–base mutant of rice BGLu1 (E176Q) that has relatively high activity for *p*-nitrophenyl- β -D-glucopyranoside (pNPGlc) substrate compared to other catalytic residue mutants, but cannot hydrolyze oligosaccharides. This mutant enzyme allowed binding of oligosaccharides to be studied by inhibition kinetics and X-ray crystallography.

2. Materials and methods

2.1. Mutagenesis and protein expression

The E176Q mutation of the rice *bglu1* cDNA was done by the overlap extension method (Wurch et al., 1998) with the T7 promoter (5'-TAATACGACTCACTATAGGG-3'), E176Qr (5'-CCTTGGCTGATTAATGTA-3'), E176Qf (5'-TTTAATCAGCAAGGATAGTA-3'), and EcoRIr (5'-TTCGATGAATTCTCAGTGCTT-3') primers. The cDNA was then subcloned into pET32a(+) as previously described (Opassiri et al., 2003). All other mutations of the rice *bglu1* cDNA were made in the pET32/Bglu1 expression plasmid by the Quikchange method (Stratagene), as described by the supplier. The oligonucleotides for mutagenesis had the following sequences and their reverse complements, for E176A: 5'-CACTGGTTTACATTTAATGCGCAAGGATAGTAGCAG-3', for E176D: 5'-GCACTGGTTTACATTTAATGACCAAGGATAGTAGCA-3', for E386D: 5'-GACAGTCGTCATACTGACAACGGAATGGATCAAC-5', and for E386Q: 5'-CCGACAGTCGTCATACTCAGACGGAATGGATCAACCT-5'. All mutations were verified by sequencing the full cDNA to ensure that each had only the desired mutation. The proteins were expressed and purified as previously described (Opassiri et al., 2003; Chuenchor et al., 2006).

Enzyme rates were determined in triplicate reactions, as previously described (Chuenchor et al., 2008), with a time course to ensure linearity. The pH dependence of the wildtype Bglu1 and the E176 mutants was determined with *p*-nitrophenol β -D-glucopyranoside (pNPGlc) as a substrate in universal pH buffer (0.2 M boric acid, 0.05 M citric acid mixed with 0.05 M tri-sodium phosphate to achieve the pH values). Relative activities of wildtype Bglu1 and Bglu1 E176 mutants were determined in 50 mM buffer: sodium acetate, MES, phosphate or universal buffer, and 50 mM nucleophiles after adjustment of the reaction solution (without enzyme and substrate) to pH 5.0. Kinetic parameters (K_m and apparent k_{cat}) of wildtype Bglu1 and Bglu1 E176 mutants in the presence of 50 mM sodium azide were calculated from nonlinear regression of Michaelis–Menten curves with Grafit 5.0 (Leatherbarrow et al., 2001). The competitive inhibition constants (K_i) were determined from the plots of the apparent K_m/V_{max} vs. the inhibitor (oligosaccharide) concentration in the presence of 50 mM sodium azide, and were taken to be the apparent dissociation constants (K_D) of the inhibitors. The association constants were calculated as $K_A = 1/K_D$; and Gibbs free energies of binding were calculated as $\Delta G_{binding} = -RT \ln K_A$, where R is the gas constant (8.314 J K⁻¹ mol⁻¹) and T is the absolute temperature (303 K). Relative inhibitor potencies were calculated by setting that of cellohexaose at 100%, dividing the association constants of the oligosaccharides by that of cellohexaose and multiplying by 100%.

2.2. X-ray crystallography

Prior to crystallization trials, Bglu1 E176Q protein was released from its N-terminal thioredoxin tag by enterokinase digestion and further purified by adsorption of the tag to immobilized cobalt resin, followed by S200 gel filtration chromatography of the tag-free Bglu1 E176Q protein, as previously described for wildtype Bglu1 (Chuenchor et al., 2006, 2008). The protein was crystallized alone and in the presence of 1 mM 2,4-dinitrophenyl 2-deoxy-2-fluorogalactoside, up to 100 mM cellobiose, 2 mM cellotetraose, 1 mM cellopentaose or 0.5 mM cellohexaose under conditions similar to those reported for wildtype Bglu1 (2 mg/mL protein, 20–23% PEG 5000 MME, 0.18–0.22 M (NH₄)₂SO₄, 0.1 M MES, pH 6.7, with microseeding). Apo Bglu1 E176Q crystals were also soaked with 1 and 10 mM laminaribiose (crystals soaked with 1 mM laminaribiose diffracted to 1.37 Å resolution, but had low occupancy of laminaribiose, while those soaked with 10 mM diffracted to 2.8 Å and had

higher occupancy, so it was used for the ligand structure building). The crystals were flash vitrified in liquid nitrogen after soaking in precipitant solution supplemented by 18% glycerol with the same concentrations of ligands. Data were collected with the crystals cooled by a 105 K nitrogen stream at the BL13B and BL13C beamlines of the National Synchrotron Radiation Research Center, Hsinchu, Taiwan, with 1.0 Å wavelength X-rays and ADSC Quantum 210 (BL13C) and Quantum 315 (BL13B) detectors. Data were indexed, refined and scaled with HKL2000 (Otwinowski and Minor, 1997).

The crystals of Bglu1 E176Q were isomorphous with those of previously reported wildtype Bglu1 crystals, allowing the structures to be solved by simple rigid body refinement with the wildtype structure (PDB code 2RGL) in Refmac 5.0 (Collaborative Computing Project Number 4, 1994; Murshudov et al., 1999). The structures were refined by restrained refinement with the same program and rebuilt with O (Jones et al., 1991). The occupancy of two possible modes of binding of cellopentaose in the active site was refined by setting the total occupancy at 1.0 and testing different fractions of each ligand position to determine which gave the lowest ligand temperature (B) factor and R_{free} value. In doing so, a glycerol molecule was included at either the –1 or +5 subsite, for the +1 to +5 and –1 to +4 binding modes, respectively. The structural parameters of the final models were validated with PROCHECK (Laskowski et al., 1993), and are shown in Tables 3 and 4. Superimposed structures were viewed and structural figures were drawn with Pymol (Schrodinger, LLC). Pymol was also used to measure sugar torsion angles, while Cremer–Pople analysis of sugar puckering to identify the ring form (Cremer and Pople, 1975) was done via the Cremer–Pople calculator website of Shinya Fushinobu (Univ. of Tokyo).

2.3. Accession numbers

Coordinates and structure factors have been deposited in the Protein Data Bank under the accession numbers 3F4 V (apo Bglu1 E176Q), 3AHV (Bglu1 E176Q with G2F), 3F5 J (Bglu1 E176Q with cellotetraose), 3F5 K (Bglu1 E176Q with cellopentaose), and 3AHT and 3F5L (Bglu1 E176Q with laminaribiose).

3. Results

3.1. Catalytic residue mutation activities

In order to identify rice Bglu1 mutants appropriate for producing structures in complex with oligosaccharides, three mutations of the putative acid–base, E176A, E176D and E176Q were generated, along with two mutations of the catalytic nucleophile, E386D and E386Q. The mutant enzymes were produced as soluble proteins and their circular dichroism spectra were nearly identical to wildtype (data not shown). Under standard assay conditions, the activity of the nucleophile mutants was undetectable, and assays at increased enzyme concentrations and times showed that E386D and E386Q mutants were at least 3000- and 60,000-fold less active than wildtype Bglu1 in pNPGlc hydrolysis, respectively.

The acid–base mutations also had decreased activity, however, for E176Q, hydrolysis of pNPGlc was still significant in the standard assay. As acid–base mutants of retaining β -glycosidases are expected to be rescued by small nucleophiles (Wang et al., 1995), the activity of the enzymes in the presence of several buffers and nucleophiles was tested, as shown in Table 1. The E176Q basal activity in MES buffer was over 3-fold and 6-fold higher than the E176D and E176A mutants, respectively, and the E176Q activity increased greater than 7-fold and 13-fold in the presence of acetate and azide, respectively. The E176A activity also increased 2–3-fold in the presence of acetate and azide, but E176D was not rescued by

Table 3

Data collection statistics.

	3F4 V	3AHV	3F5 J	3F5 K	3AHT
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameter (Å, °)	$a = 79.8$ $b = 101.0$ $c = 127.6$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.2$ $b = 100.4$ $c = 127.4$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.7$ $b = 100.8$ $c = 127.5$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.8$ $b = 101.2$ $c = 127.4$ $\alpha = \beta = \gamma = 90^\circ$	$a = 79.7$ $b = 101.5$ $c = 128.3$ $\alpha = \beta = \gamma = 90^\circ$
No. of molecules per ASU	2	2	2	2	2
Resolution range (Å)	30–1.65 (1.71–1.65)	30–1.90 (1.97–1.90)	30–1.95 (2.02–1.95)	30–1.80 (1.86–1.80)	30–2.80 (2.90–2.80)
No. of observations	636,554	560,709	462,300	429,830	158,410
No. of unique observations	124,448	79,635	73,603	94,774	24,681
Redundancy	5.1 (5.0)	7.0 (6.7)	6.3 (5.8)	4.5 (3.6)	6.1 (6.3)
Completeness (%)	99.9 (99.9)	99.1 (98.9)	97.4 (93.7)	98.9 (93.3)	99.8 (99.8)
$I/\sigma(I)$	24.5 (4.5)	23.3 (6.1)	19.2 (6.9)	17.9 (2.7)	16.1 (5.8)
$R_{\text{sym}} (\%)^a$	6.2 (36.5)	8.3 (35.2)	8.3 (29.2)	8.3 (45.6)	11.7 (34.1)

Numbers in parentheses are outer shell parameters.

^a $R_{\text{sym}} = \sum_{\text{hkl}} \sum_i |I_i(\text{hkl}) - \langle I(\text{hkl}) \rangle| / \sum_{\text{hkl}} \sum_i I_i(\text{hkl})$.**Table 4**

Refinement statistics.

	3F4 V	3F5I	3F5 J	3F5 K	3AHT
Resolution range (Å)	30–1.65	26.29–1.90	30–1.95	30–1.80	30–2.80
No. of amino-acid residues	944	944	944	944	942
No. of protein atoms	7618	7618	7618	7618	7602
No. of water molecules	901	668	727	801	71
Refined carbohydrate	None	G2F	Cellotetraose	Cellopentaose	Laminaribiose
No. of carbohydrate atoms	None	22	90	112	46
No. of other hetero atoms	53	53	35	47	35
$R_{\text{factor}} (\%)^b$	17.5	18.6	18.0	17.8	20.7
$R_{\text{free}} (\%)^c$	19.6	18.9	21.4	21.0	24.8
Ramachandran statistics					
Favored region (%)	97.4	97.6	97.6	98.0	95.6
Allowed region (%)	2.6	2.4	2.4	2.0	4.4
Outlier region (%)	0	0	0	0	0
R.m.s.d. from ideality					
Bond distances (Å)	0.011	0.014	0.018	0.015	0.008
Bond angle (°)	1.357	1.461	1.585	1.512	1.057
Mean B factors (Å²)					
All protein atoms	12.03	14.43	9.15	14.74	21.874
Waters	24.43	27.35	19.10	26.88	31.87
Hetero atoms	19.54	28.44	20.59	34.49	46.97
Carbohydrate atoms	None	9.84	24.29	32.71	43.16
Subsite –1 (A/B)	None	9.67/10.01	20.32/19.58	26.09/23.99	41.20/41.93
Subsite +1 (A/B)	None	None	21.73/21.57	29.75/29.06	44.69/44.59
Subsite +2 (A/B)	None	None	21.94/21.82	26.97/26.94	None
Subsite +3 (A/B)	None	None	32.87/32.90	32.91/32.66	None
Subsite +4 (A/B)	None	None	None	49.01/46.49	None

^b $R_{\text{factor}} = (\sum |F_o| - |F_c|) / \sum |F_o|$.^c Based on 5% of the unique observations not included in the refinement.

any anionic nucleophiles. Formate and cyanate could also slightly rescue E176Q activity. Ascorbate was found in the active site of white mustard myrosinase, which also has Gln in the place of the catalytic acid–base (Burmeister et al., 2000), so rescue with 50 mM ascorbate was attempted. The E176 activity increased 15–18-fold in the presence of ascorbate, while E176A activity was increased 3–4-fold, E176D was not affected, and wildtype BGlu1 appeared to be inhibited. However, due to the instability of ascorbate, azide was used for further kinetic analysis.

In the presence of 50 mM azide, the apparent k_{cat} values of the E176Q and E176A mutants were nearly independent of the pH from pH 5–10, while those of the wildtype and E176D enzymes decreased 30- and 8-fold, respectively, as the pH increased from 5 to 10 (Fig. 2). The pH dependence over the acidic range of 3–5 was maintained in the mutants and wildtype enzyme. These data confirmed that E176 is the catalytic acid–base, and suggest that E176Q is unusually active with a high capacity to transglycosylate anionic

nucleophiles, whereas E176D appears to maintain an acid catalyst that is weakened by its greater distance from the substrate due to the shortened sidechain.

While wildtype BGlu1 hydrolyzed 2,4-dinitrophenyl- β -D-glucoside (dNPGlc), which has an aglycone with a pK_a of 3.96, with an apparent k_{cat} of $9.1 \pm 0.7 \text{ s}^{-1}$ compared to $4.7 \pm 0.3 \text{ s}^{-1}$ for pNPGlc (pK_a 7.18), the BGlu1 E176Q mutant had essentially the same apparent k_{cat} for both substrates in the presence of azide ($1.32 \pm 0.01 \text{ s}^{-1}$ for dNPGlc with 0.2 M azide and $1.84 \pm 0.05 \text{ s}^{-1}$ for pNPGlc with 0.32 M azide). In contrast, hydrolysis of oligosaccharides by BGlu1 E176Q was not observed.

3.2. Oligosaccharide binding by BGlu1 E176Q

The availability of a semi-active mutant that could be efficiently rescued by small nucleophiles for aryl glycosides, but not oligosaccharides, allowed kinetic investigation of the binding of

Table 1

Relative activities of wildtype and acid–base mutants of BGlu1 in the presence and absence of nucleophiles.

Buffer	Specific activity ($\mu\text{mol}/\text{mg}/\text{min}$)			
	WT	E176Q	E176A	E176D
MES	n.d.	0.070	0.011	0.022
Acetate/MES	6.8	0.51	0.028	0.029
Azide/MES	7.2	0.89	0.031	0.028
Ascorbate/MES	5.6	1.06	0.040	0.029
Formate/MES	7.7	0.111	0.014	0.026
TFA/MES	7.5	0.046	0.005	0.028
KF/MES	5.8	0.072	0.008	0.022
Cyanate/MES	7.2	0.036	0.007	0.027
Acetate/UB	7.3	0.33	0.021	0.024
Azide/UB	7.3	0.84	0.030	0.025
Ascorbate/UB	5.1	0.95	0.036	0.026
UB	7.9	0.052	0.013	0.028

* All anionic nucleophiles and MES buffer, were prepared 50 mM, pH 5.0. Universal buffer (UB) was 0.2 M boric acid, 0.05 mM citric acid mixed with 0.05 M tri-sodium phosphate to attain pH 5.0. Assays were performed with 1 mM pNPGlc, as described in the methods. “n.d.” means not determined.

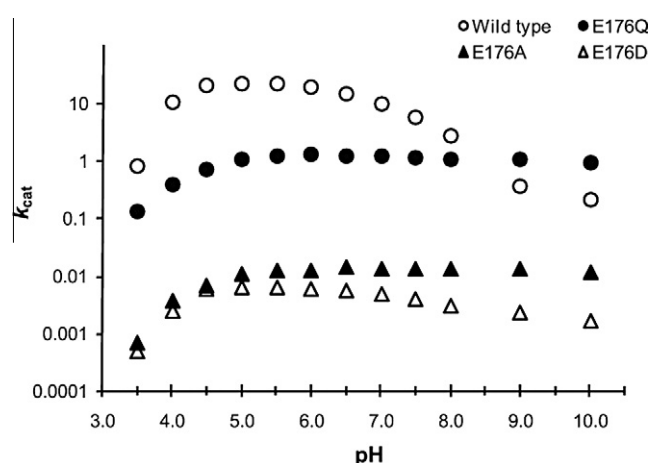


Fig. 2. Comparison of pH dependence of the BGlu1 wild type, E176Q, E176A, and E176D enzymes in 50 mM azide in universal pH buffer. The semi-log plot of the k_{cat} against the pH for each of the four enzymes to show that the E176Q and E176A mutants show little dependence of the rate on the pH between six and ten, whereas the wildtype and E176D mutant proteins have 10–100-fold decreases in k_{cat} over this range.

oligosaccharides to the BGlu1 E176Q active site by inhibition studies. Cellotriose, cellotetraose, cellopentaose, cellohexaose, and laminaribiose gave predominantly competitive inhibition (Table 2). From the competitive inhibition constants, apparent association constants for each oligosaccharide encompassing all possible binding modes could be calculated, along with the apparent overall

energy of binding. However, cellobiose was a very weak inhibitor that showed mixed inhibition, so only maximal values could be estimated for the competitive inhibition constant and association constant.

3.3. X-ray crystallographic structures of E176Q complexes

Binding of BGlu1 E176Q to oligosaccharides was also investigated by X-ray crystallography. Five structures were determined, including the apo BGlu1 E176Q mutant protein (PDB code 3F4V), its covalent complex with 2-fluoroglucoside (G2F, PDB code 3AHV), and its noncovalent complexes with cellotetraose (3F5 J), cellopentaose (3F5 K), and laminaribiose (3AHT). Attempts to cocrystallize or soak the crystals with cellobiose concentrations up to 100 mM resulted in no significant density in the active site and crystals with cellohexaose could not be obtained by co-crystallization. Diffraction and model parameters for the 5 structures can be found in Tables 3 and 4.

The apo BGlu1 E176Q (3F4 V) and BGlu1 E176Q covalent intermediate with G2F (3AHV) structures were nearly identical to the structures of wildtype BGlu1 (2RGL) and its complex with G2F (2RGM) (12), with an RMSD of 2RGL to 3F4 V of 0.163 Å over 926 C α atoms and for 2RGM to 3AHV of 0.245 Å over 925 atoms. For comparison, the wildtype structures (2RGL and 2RGM) had RMSD of 0.121 over 881 atoms, and the E176Q mutant structures (3F4 V and 3AHV) also had RMSD of 0.121 over 881 atoms.

3.4. Structural analysis of oligosaccharide binding

Fig. 3 shows the electron density omit maps (Fo–Fc maps calculated with the ligands removed) for the active site glycerol in the apo enzyme (A), and α -2-deoxy-2-F-glucoside (B), laminaribiose (C) cellotetraose (D) and cellopentaose (E) in their complexes with BGlu1 E176Q. Although the electron density for the reducing and nonreducing end glucosyl residues in the cellotetraose and cellopentaose complexes is less well defined, that for residues 2, 3 and 4 (in the cellopentaose structure) is clear. The nonreducing terminal residue was clearly found in the –1 subsite, since the density was more extensive than the glycerol in the apo enzyme active site, which superimposes on C2, O2, C3, O3, and C4, but not O4, C5, O5, C6 and O6 (Fig. 3 and Supplementary Fig. 1). The density for C1 was very weak, but the nonreducing glucosyl residue could be refined best when a glucosyl residue in a 1S_3 skew boat conformation was inserted and the C2–C1–O5–C5 dihedral angles refined to values of +3.5° for laminaribiose, +29° for cellotetraose and +24° for cellopentaose. This puts one O ϵ of E386 nearly directly in line with the glycosidic bond for its nucleophilic attack, at a distance of 3.15 Å in the cellopentaose structure, while the other E386 O ϵ makes a strong H-bond (2.53 Å) with the neighboring C2 OH (Fig 5 A). The cellopentaose nonreducing glucosyl residue with a C2–C1–O5–C5 angle of +24° appears to be somewhere between the

Table 2

Inhibition of BGlu1 E176Q hydrolysis of pNPGlc by gluco-oligosaccharides.

Oligosaccharide	Type of inhibition	Apparent K_i (mM)	Apparent K_A (M^{-1})	Apparent $\Delta G_{\text{binding}}$ *	Relative inhibitor potency (percent)
Cellobiose	Mixed	44 + 3	23	–7.9**	0.04
Cellotriose	Competitive	0.48 + 0.03	2,100	–19.3	3.8
Cellotetraose	Competitive	0.052 + 0.006	19,200	–24.8	34.6
Cellopentaose	Competitive	0.024 + 0.002	41,700	–26.8	75.0
Cellohexaose	Competitive	0.018 + 0.001	56,000	–27.5	100
Laminaribiose	Competitive	0.191 + 0.02	5,200	–21.6	9.4

* The apparent K_A and $\Delta G_{\text{binding}}$ are calculated based on the assumption of purely competitive inhibition. As such, they represent maximal values. The equations used to calculate these values were: $K_A = 1/K_D = 1/K_i$; $\Delta G_{\text{binding}} = -RT \ln K_A$, where R is the gas constant ($8.314 \text{ J K}^{-1} \text{ mol}^{-1}$) and T is the absolute temperature (303 K). The relative inhibitor potency was calculated by dividing the apparent K_A by that of the most potent inhibitor (cellohexaose) and multiplying by 100%.

** Since cellobiose gave mixed inhibition, the apparent K_A is expected to be an overestimate of the maximal value. The $\Delta G_{\text{binding}}$ is therefore more energetically favorable than is realistic.

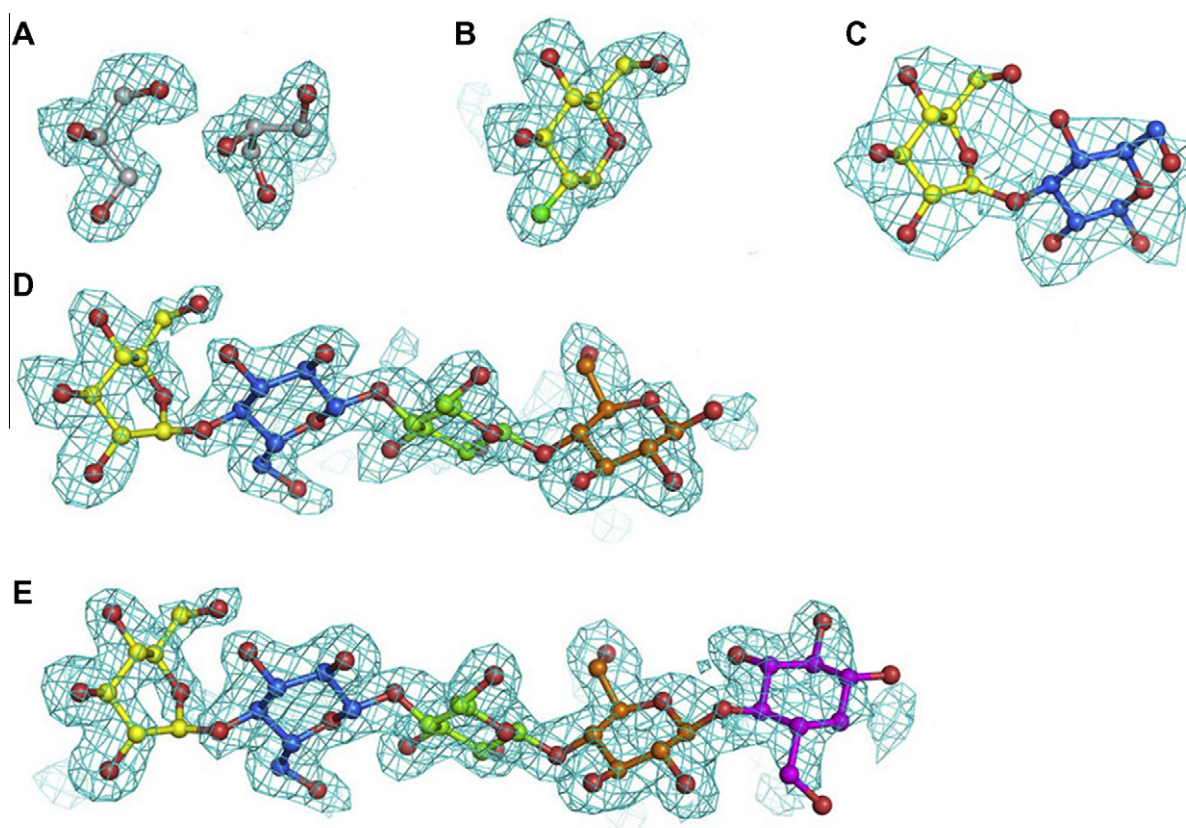


Fig. 3. The electron densities in the active sites of BGLu1 E176Q and its ligand complexes. The electron density (Fo–Fc) maps were calculated for each structure with all heteroatoms omitted from the active site, and are shown in blue mesh at $+2\sigma$. The ligands from the final structures are superimposed on these electron density maps and shown in ball and stick representation. (A) shows the BGLu1 E176Q apo enzyme to which no ligands were added, but in which 2 glycerol (carbons in silver) are seen in the active site. (B) shows the BGLu1 E176Q complex with 2-deoxy-2-fluoroglucoside, which is found in an alpha-linkage with the catalytic nucleophile. (C) shows the BGLu1 E176Q complex with laminaribiose, (D) shows the complex with cellotetraose, and (E) shows the complex with cellopentaose. The ligand glucosyl residues are color coded with carbons in yellow for the nonreducing glucosyl residue, blue for the second, green for the third, orange for the fourth and pink for the fifth. Oxygen is shown in red and fluoride in green. Direct overlays of the glycerol densities on those of cellopentaose and laminaribiose are shown in [Supplementary Fig. 1](#).

1S_3 skew boat (C2–C1–O5–C5 dihedral angle of $+45^\circ$ (Biarnes et al., 2006) and the expected 4H_3 half chair transition state or the closely related 4E envelope (C2–C1–O5–C5 dihedral angle of 0°), whereas this dihedral angle in other glucosyl residues was close to the relaxed 4C_1 chair value of -45° . Calculation of the Cremer–Pople parameters (Cremer and Pople, 1975) for these glucosyl rings gave $\varphi = 240^\circ$, $\theta = 58^\circ$ and $Q = 0.605$ for laminaribiose, $\varphi = 227^\circ$, $\theta = 70^\circ$ and $Q = 0.723$ for cellotetraose, and $\varphi = 231^\circ$, $\theta = 58^\circ$ and $Q = 0.675$ for cellopentaose, which were all closest to the values for a 4E envelope, though with higher puckering amplitudes (Q). This suggested the initially inserted 1S_3 skew boats had been substantially moved toward a 4E structure, which is closely related to the 4H_3 half chair and might act as a transition state, by the refinement. The weak density for C1 is consistent with movement between these ring structures in the complex.

The structure of BGLu1 with cellotetraose was close to that with cellopentaose with an RMSD of 0.078 Å (911 residues). The glucosyl residues of cellotetraose and cellopentaose bound to BGLu1 E176Q in the same positions for the first four residues (Fig. 4A), in line with the subsite model of oligosaccharide binding. The interactions between the five glucosyl residues of cellopentaose with the surrounding amino acids are mapped on the interaction diagram shown in Fig. 5A. Aside from the interaction of the catalytic nucleophile, mentioned above, several other direct interactions with Glc1 (the nonreducing terminal glucosyl residue) are found, in common with other GH1 structures (Sanz-Aparicio et al., 1998). No water molecules in position to interact with Glc1 were found in the -1 subsite, in contrast to all other subsites,

where nearly all polar interactions are mediated by water molecules. As seen in Fig. 5A, the only direct hydrogen bond between cellopentaose (and cellotetraose) Glc2 to Glc5 and BGLu1 E176Q is that between N δ of N245 and O2 of Glc3, and the Glc3 O3 atom is also 3.3 Å from this nitrogen, suggesting it may also form a hydrogen bond to N245. These interactions may help explain the strong binding at this subsite (Opassiri et al., 2004) (demonstrated in Table 2). Glucose residues Glc2 and Glc3 are aligned to stack onto the indole ring of W358, while the oligosaccharide twists nearly 90° between Glc3 and Glc4 to stack Glc4 and Glc5 along the face of the Y341 phenol ring (Fig. 4B). It is notable that in the apo enzyme structure Y341 and its neighbors in the flexible C loop, V342 and F343, are found in two possible positions, while they are stabilized in the one position by interactions with either cellotetraose or cellopentaose.

The poor density of the nonreducing end glucosyl residue of cellopentaose in the -1 subsite, together with some diffuse density in the apparent $+5$ site of the cellopentaose structure (Fig. 3E), led us to suppose that the cellopentaose was found in two binding modes in the crystal. Therefore, the cellopentaose model was built in both the productive position occupying subsites -1 to $+4$ and in the unproductive position occupying subsites $+1$ to $+5$, with a glycerol in the -1 subsite, as seen in the apo protein crystal structure. However, when the occupancy was refined to find the percent binding in each position that gave the lowest R_{free} and ligand temperature (B) factor, the best solution was one in which the occupancy in the productive position was at least 95% (with 5% binding from the $+1$ to $+5$ subsites). In fact, the density in the $+5$ site could be fit to a

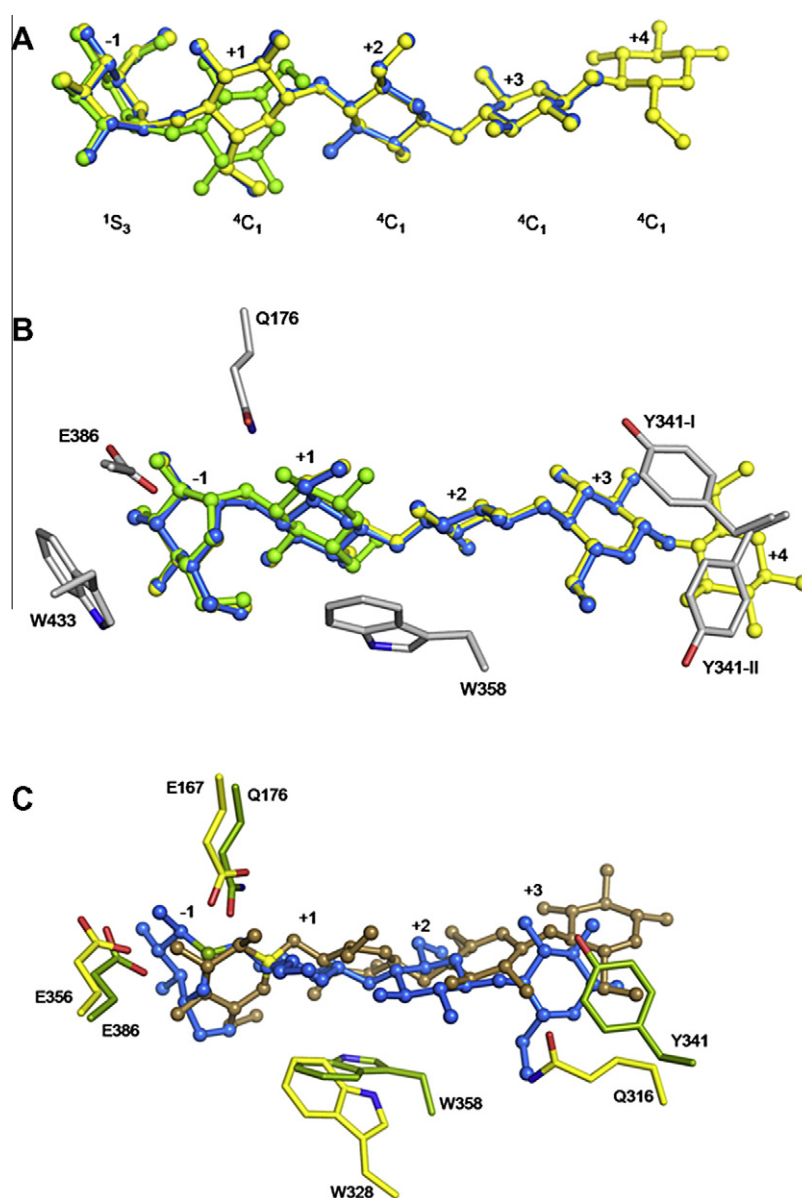


Fig. 4. Superposition of oligosaccharides in the active site. (A). Relative positions of cellotetraose, cellopentaose, and laminaribiose in the superimposed active sites of their E176Q mutant complexes. The sugar conformations fit into the density for each residue (4C_1 chair or 3S_1 skew boat) and the C2–C1–O5–C5 dihedral angles for glucosyl residues 1 and 2 from the nonreducing end in cellotetraose (Cel4), cellopentaose (Cel5) and laminaribiose (Lam2) are shown below that residue. Cremer–Pople analysis indicated that the refinement moved all 3S_1 skew boats to be best fit to 4E envelope rings in the final structures, while those residues in relaxed chairs stayed in that conformation. Cellotetraose (blue) and cellopentaose (yellow) were nearly completely superimposed, while the laminaribiose (red) Glc2 was flipped to lie on the opposite face, in addition to being connected to Glc1 C1 by O3 instead of O4. (B) Aromatic-sugar interactions most critical for stacking the sugars in their places. The Cel4, Cel5 and Lam2 are drawn as in A, while the residues W433 upon which the -1 subsite glucosyl residue stacks, W358, upon which the +1 subsite, and to a lesser extent +2 subsite, glucosyl residues stack, and Tyr341, upon which the +3 subsite, and to a lesser extent +4 subsite, glucosyl residues stack, along with the catalytic residues, are drawn with gray carbons. Two alternate conformations of Y341 (I and II) were seen in the electron densities generated from the high resolution data of the apo enzyme and its complexes with G2F and low occupancy Lam2, but Y341 is locked into position I by interactions with cellotetraose and cellopentaose. The +1 subsite glucosyl residue of Lam2 is flipped 180° compared to the cellooligosaccharides, thereby allowing the C1, C3, and C5 hydrogens to interact with the indole ring of W358. (C) Superimposition of rice BGLu1 (green) and *Paenibacillus polymyxa* β -glucosidase B (PpBGLuB, yellow) structures in complex with cellotetraose. Cellotetraose in complex with rice BGLu1 is colored blue, while it is sand colored in its complex from PpBGLuB. It can be seen that, whereas BGLu1 E386 is 3.2 Å from the C1 (colored green) of the nonreducing glucosyl residue of cellotetraose in its complex, the catalytic nucleophile of PpBGLuB, E356, is 6.3 Å away and blocked from making a direct attack on the anomeric carbon (C1, colored yellow) by C3. In addition to the catalytic acid–base and nucleophile, the aromatic amino acid side chains that stack the 2nd and 3rd glucosyl residues (Trp258) and 4th and 5th glucosyl residues (Tyr341) in BGLu1 and the corresponding residues in PpBGLuB are shown. The glucosyl residue-binding subsites (-1 to +3) are marked above the sugars.

glycerol better than a glucosyl residue, so little evidence was seen for unproductive binding. Therefore, it was deduced that, under the conditions in the crystals at the time of flash cooling, at least 95% of the cellopentaose bound in the -1 to +4 position. Cellotetraose was similarly largely in the productive position.

In the complex of BGLu1 E176Q with laminaribiose, the nonreducing sugar is in nearly the same position as in those with cello-

tetraose and cellopentaose, with the largest differences occurring at the positions of O5 and C1. The positions of these atoms and its glycosidic linkage to the C3 of the reducing glucosyl residue placed this reducing end residue in a distinct position (Fig. 4A) and the ring was flipped 180° compared to the Glc2 of the cellooligosaccharides. This orientation positioned the C1, C3 and C5 protons, which are expected to have a weak positive dipole due to

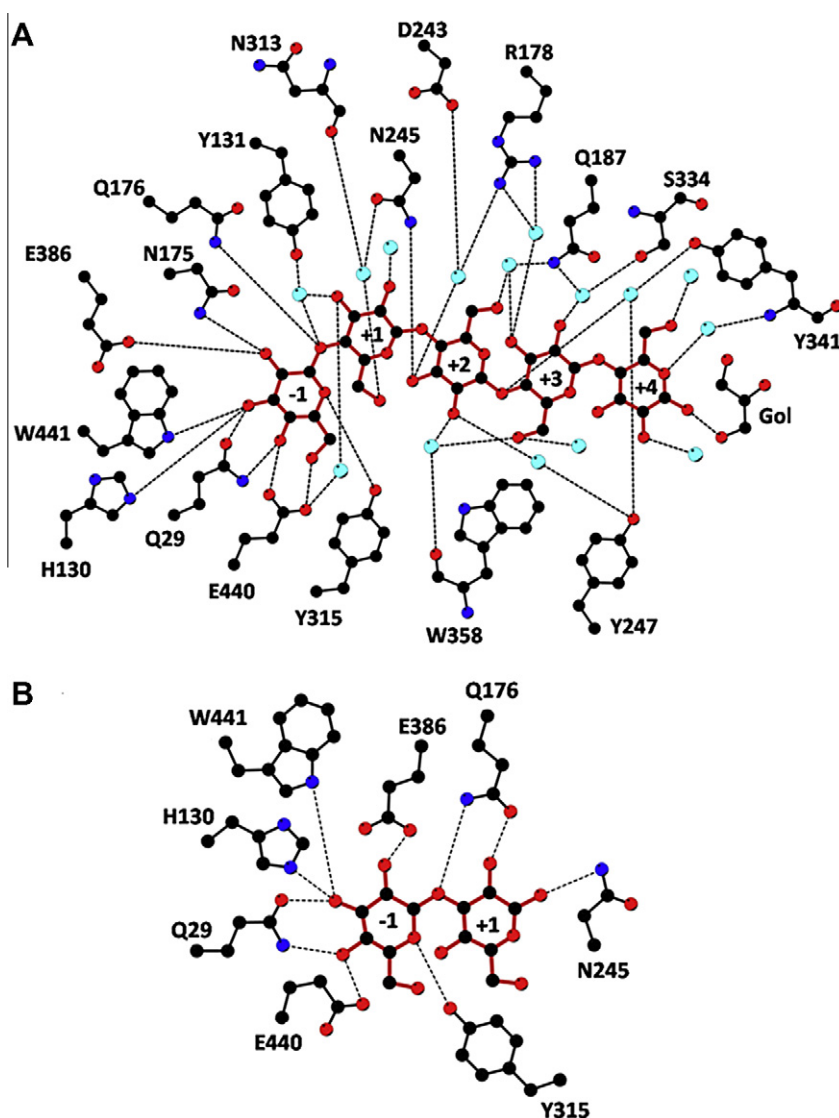


Fig. 5. Polar interactions between the BGluc1 E176Q mutant active site residues and cellopentaose (A) and laminaribiose (B). The diagram in A shows that all apparent hydrogen bonds (dotted lines shown where the distance is <3.2 Å) are directly between the cellopentaose Glc1 and the protein, but are mediated through water for all other residues, except for one direct hydrogen bond between N245 and Glc3 O3. In contrast, (B) shows that, in addition to water-mediated hydrogen bonds, laminaribiose forms three direct hydrogen bonds to the protein at Glc2, while Glc1 makes the same direct hydrogen bonds to the protein as does cellopentaose Glc1.

the sigma withdrawal of electrons to the sugar oxygen atoms, to make weakly polar interactions (sometimes called soft hydrogen bonds (Weiss et al., 2001)) with the aromatic electron cloud of the Trp358 indole ring to stack this sugar on that ring (Fig. 4B). The β -conformation required for this positioning of the C1 hydrogen, also facilitated the formation of a direct hydrogen bond between the C1 OH and the N245 sidechain, as well as a water-mediated hydrogen bond between the Glc2 O1 and the same sidechain (Fig. 5B).

4. Discussion

4.1. Analysis of catalytic residue mutants

The identity of the catalytic acid–base and nucleophile were already known, due to their conservation with those of other GH1 β -glucosidases and their positions in the structure. Additional evidence for the catalytic nucleophile was provided by the glycosynthase activity of E386A, E386G and E386S mutants (Hommalai et al., 2007), and the covalent labeling of the E386 with G2F

(Chuenchor et al., 2008), so loss of activity in the E386D and E386Q mutants was expected. The loss of activity in the E176 mutants, the rescue of E176A and E176Q by anionic nucleophiles, and the loss of the basic shoulder of the pH-activity profiles of these two mutants provided support for the identification of E176 as the acid–base catalyst, as noted previously for the corresponding *Agrobacterium* sp. β -glucosidase Abg mutants (Wang et al., 1995; Mülleger et al., 2005). It is notable that these mutations were most highly active in the presence of ascorbate, a cofactor for the hydrolysis of thioglucosides by myrosinase, which also has a Gln in the position corresponding to the GH1 catalytic acid–base (Burmeister et al., 2000). The hydrolysis of thioglucoside by almond β -glucosidase has also been shown to be promoted by low concentrations of ascorbate, but inhibited by high concentrations (Shen and Byers, 2007). Wildtype BGluc1 also shows inhibition by ascorbate (Table 1) and higher concentrations of ascorbate inhibited BGluc1 E176Q, but hydrolysis of pNP- β -D-thioglucoside was not detected under standard assay conditions (data not shown).

A similarly high activity to that seen with the BGluc1 E176Q mutant was seen with the corresponding mutation of Abg,

although that enzyme appeared to show more rapid hydrolysis of dNPGlc ($k_{\text{cat}} = 16 \text{ s}^{-1}$) than pNPGlc ($k_{\text{cat}} = 1 \text{ s}^{-1}$) (Mülegger et al., 2005). It has been noted previously that the *Agrobacterium* enzyme seemed to have a higher requirement for acid-assistance than other enzymes, such as *Cellulomas fimi* Cex (Wang et al., 1995), and, as shown here, rice BGlu1. The occurrence of the G2F covalent intermediate confirms that the 2,4-dinitrophenolate does not require acid-assistance in its departure, since glycosylation takes place despite the destabilization of the presumed oxycarbenium cation-like transition state by the electron withdrawal by the fluorine at C2. Recently, a similar covalent intermediate with an unmodified glucosyl residue (without a fluorine atom) with the acid–base E to Q mutant of human cytosolic β -glucosidase has been reported (Noguchi et al., 2008). Attempts by our group to produce a similar covalent intermediate with BGlu1 E176Q have so far failed to show noticeable density in the active site, perhaps due to the unusually high activity of this mutant of BGlu1.

Mülegger et al. (2005) suggested that the high activity of the Abg acid–base E to Q mutant may have resulted from a hydrogen bond made by the glutamine. Indeed, the Q176 N ϵ and the glycosidic oxygen of the scissile bond are within hydrogen bonding distance in both the covalent G2F intermediate and Michaelis-complex-like oligosaccharide complexes with BGlu1 E176Q (Fig. 4), as was seen the corresponding mutants of human cytoplasmic and strictosidine β -glucosidases (Noguchi et al., 2008; Barleben et al., 2007). In fact, despite not being hydrolyzed, the oligosaccharide complexes showed distortion of the nonreducing glucosyl residue between the $^1\text{S}_3$ skew boat and a planar transition state most similar to the ^4E envelope, based on the Cremer–Pople parameters and C2–C1–O5–C5 torsion angles for the refined structures. Movement between these structures may account for the poor electron density for the anomeric carbon (C1).

4.2. Comparison of substrate binding to other GH1 structures

The glycone moiety is bound by the conserved residues found in other GH1 β -glucosidases (Czjzek et al., 2000; Verdoucq et al., 2004; Sanz-Aparicio et al., 1998), as seen in the G2F-labeled wild-type enzyme (Chuenchor et al., 2008), but the aglycone-binding residues are variable among these enzymes, reflecting their different substrate specificities. Previously, the conserved Trp corresponding to W358 has been noted to be critical for binding of aglycones, and variability in its position has been noted to reflect substrate specificity (Czjzek et al., 2000, 2001; Zouhar et al., 2001; Verdoucq et al., 2003, 2004; Barleben et al., 2007). Other residues found to interact with the aglycone in maize Glu1 included F198, F205, and F466, which correspond to L183, N190, and L442 in rice BGlu1. In sorghum SbDhr1, V196 and L203, corresponding to the former two of these residues, formed hydrophobic interactions with the dhurriin phenolic ring, but S462, corresponding BGlu1 L442, was most critical in forming a water-mediated hydrogen bond to the aglycone phenolic oxygen (Czjzek et al., 2000; Verdoucq et al., 2004). Similar to SbDhr1, rice BGlu1 residue L183 appears to form a hydrophobic interaction with Glc residue 3. However, N190 is over 6 Å away and interacts only indirectly, while the substrate turns away from L442. Mutagenesis of L442 had no effect on catalytic activity (Chuenchor et al., 2008). F261 in SbDhr1 appears to restrict the active site, while M251 in this position has a similar effect on the rice Os3BGlu6 active site (Seshadri et al., 2009). In contrast, the corresponding BGlu1 N245 forms two hydrogen bonds to Glc residue 3, and its polar sidechain has been shown to contribute to hydrolysis of pNPGlc, as well as cellooligosaccharides (Chuenchor et al., 2008). Since the cellooligosaccharide substrates of BGlu1 are longer, the BGlu1 active site is more extensive than the other plant enzymes and includes residues further

out, such as Y341 and S334, which interact with the last two glucosyl residues.

4.3. Analysis of oligosaccharide binding

The increase in the value of the apparent association constant (K_A) for up to 6 residues, supports the previous prediction of 6 glucosyl residue-binding subsites based on subsite mapping by kinetic analysis of hydrolysis of oligosaccharides by BGlu1 (Opassiri et al., 2004). Since cellotetraose and cellopentaose appeared to bind predominantly in the productive positions with the four common glucosyl residues in the same positions, it is likely that the differences in binding energies for, at least, the $\Delta G_{\text{binding,cellohexaose}} - \Delta G_{\text{binding,cellopentaose}}$ and $\Delta G_{\text{binding,cellopentaose}} - \Delta G_{\text{binding,cellotetraose}}$ are reasonable estimates of the energy of binding subsites +5 and +4, respectively. In general, the changes of free energy of binding upon addition of a glucosyl residue correlate well with the number of direct interactions and water-mediated hydrogen bonds at the subsite to which that residue binds when the nonreducing end is bound to subsite –1. If the assumption is made that the binding of cellotriose to cellohexaose is overwhelmingly in the position that would be productive for hydrolysis in the wildtype enzyme with the nonreducing glucosyl residue in the –1 site, then subsite affinities of >11.4 kJ/mol for the +2 subsite, 4.5 kJ/mol for the +3 subsite, 2 kJ/mol for the +4 subsite and 0.7 kJ/mol for the +5 subsite can be calculated for binding of the (1 $\rightarrow$ 4)- β -linked D-glucosyl residues of cellooligosaccharides. The relative energies of interaction for the different subsites are similar to those calculated from subsite mapping (Opassiri et al., 2004).

The strong interactions between the reducing end glucosyl residue and N245 may explain why the binding of laminaribiose is more than 2 kJ/mol more energetically favorable than cellotriose. These interactions could only be formed for the β -anomer of the reducing glucosyl residue, demonstrating an apparent preference for the chirality of this residue. An initial high resolution (1.37 Å) dataset from a crystal soaked with 1 mM laminaribiose showed partial occupancy for the laminaribiose (PDB code 3F5L). Since the protein without ligands (3F4V) had 2 glycerol molecules in subsites –1 and +1 that overlapped with the laminaribiose position, their strong partial occupancies may have biased the laminaribiose placement. However, the dataset from a crystal that had been soaked in 10 mM laminaribiose collected to 2.8 Å resolution (PDB code 3AHT) had more substantial laminaribiose density.

4.4. Comparison to the complex of cellotetraose with *Paaenobacillus polymyxa* β -glucosidase B

Isorna et al. (2007) reported the structure of *P. polymyxa* β -glucosidase B with cellotetraose bound into the active site (PDB code 3Z1S), based on the partial electron density observed in the active site after a short incubation and flash freezing. In that case, the enzyme was fully active, unlike BGlu1 E176Q, which is active toward aryl β -glucosides with good leaving groups, but inactive toward oligosaccharides. In that structure, the catalytic nucleophile was 6.3 Å away from the anomeric carbon of the nonreducing sugar, which was in a relaxed chair conformation and out of position for a direct nucleophilic attack (Fig 4C). In contrast, in the BGlu1 E176Q oligosaccharide complexes, E386 is within 3.2 Å of the anomeric carbon, and directly opposite and inline with the glycosidic bond due to the distorted conformation of the nonreducing glucosyl residue. The position and distortion of cellotetraose and cellopentaose in the BGlu1 E176Q complexes appear to be indicative of a Michaelis complex, while the cellotetraose in the *P. polymyxa* BgluB model appears to bind with the glucosyl residues half way between the productive binding subsites, which is likely why it could be observed in an enzyme that is competent to hydrolyze

it. The fact that it could be observed may indicate that binding in this position is somewhat stable, which suggests that it may be a staging site for binding of substrates before insertion into the catalytic pocket and may have implications for the mechanism of such insertion of oligosaccharides. We have yet to observe such a staging position in rice BGLu1, but it may occur transiently in BGLu1 and other GH1 β -glucosidases.

5. Conclusions

The complexes of rice BGLu1 E176Q with gluco-oligosaccharide appear to be the first for a family GH1 hydrolase with oligosaccharides in productive positions for hydrolysis. The determination of the glucosyl residue binding sites in BGLu1 is critical for understanding how plant β -glucosidases have adapted to the job of cellobiosaccharide binding and hydrolysis. The demonstration that BGLu1 binds to the nonreducing terminal residue directly, yet interacts with the remaining glucosyl residues predominantly through water-mediated hydrogen bonds and weakly polar stacking interactions between aromatic residues and glucosyl CH groups, indicates the interaction involves dehydration of Glc1, while other glucosyl residues remain hydrated. It is possible that the transition from hydrated to hydrogen bonded requires a bound intermediate, which may have been observed in *P. polymyxa* BgluB (Isorna et al., 2007), but further investigation is needed to see if such an intermediate can be observed in plant family 1 glycoside hydrolases.

Acknowledgements

Chartchai Krittanai and Apichai Bourchookarn are thanked for assistance in circular dichroism of the wildtype and mutant enzymes. Chun-Jung Chen, Rodjana Opasiri, Phimonphan Chuankhayan, Worrapoj Oonant, Catleya Rojviriyi, and the staff of the NSRRC PX beamlines are thanked for helpful suggestions and assistance in data collection, processing and interpretation. Shinya Fushinobu is thanked for the use of his website for Cremer–Pople calculations. This work was supported by the Thailand Research Fund grant BRG5080007 and Suranaree University of Technology. R.C.R. was supported by A*STAR, Singapore. Portions of this research were carried out at the National Synchrotron Radiation Research Center, a national user facility supported by the National Science Council of Taiwan, ROC. The Synchrotron Radiation Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jsb.2010.09.021](https://doi.org/10.1016/j.jsb.2010.09.021).

References

- Barleben, L., Panjikar, S., Ruppert, M., Koepke, J., Stockigt, J., 2007. Molecular and structural characterization of hexameric β -D-glucosidases in wheat and rye. *Plant Cell* 19, 2886–2897.
- Barrett, T., Suresh, C.G., Tolley, S.P., Dodson, E.J., Hughes, M.A., 1995. The crystal structure of a cyanogenic β -glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure* 3, 951–960.
- Biarnes, X., Nieto, J., Planas, A., Rovira, C., 2006. Substrate distortion in the Michaelis complex of *Bacillus* 1, 3–1, 4- β -glucanase: Insight from first principles molecular dynamics simulations. *J. Biol. Chem.* 281, 1432–1441.
- Burmeister, W.P., Cottaz, S., Driguez, H., Palmieri, S., Henrissat, B., 1997. 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* 5, 663–675.
- Burmeister, W.P., Cottaz, S., Rollin, P., Vasella, A., Henrissat, B., 2000. High resolution X-ray crystallography shows that ascorbate is a cofactor for myrosinase and

- substitutes for the function of the catalytic base. *J. Biol. Chem.* 275, 39385–39393.
- Chuenchor, W., Pengthaisong, S., Yuvaniyama, J., Opasiri, R., Svasti, J., Ketudat Cairns, J.R., 2006. Purification, crystallization and preliminary X-ray analysis of rice BGLu1 β -glucosidase with and without 2-deoxy-2-fluoro- β -D-glucoside. *Acta Crystallogr. Sect. F* 62, 798–801.
- Chuenchor, W., Pengthaisong, S., Robinson, R.C., Yuvaniyama, J., Oonant, W., Bevan, D.R., Esen, A., Chen, C.-J., Opasiri, R., Svasti, J., Ketudat Cairns, J.R., 2008. Structural insights into rice BGLu1 β -glucosidase oligosaccharide hydrolysis and transglycosylation. *J. Mol. Biol.* 377, 1200–1215.
- Collaborative Computing Project Number 4, 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D* 50 760–763.
- Cremer, D., Pople, J.A., 1975. A general definition of ring puckering coordinates. *J. Am. Chem. Soc.* 97, 1354–1358.
- Czjzek, M., Cicek, M., Zamboni, V., Bevan, D.R., Henrissat, B., Esen, A., 2000. The mechanism of substrate (aglycone) specificity in β -glucosidases is revealed by crystal structures of mutant maize β -glucosidase-DIMBOA, -DIMBOAGlc, and-dhurrin complexes. *Proc. Natl. Acad. Sci. USA* 97, 13555–13560.
- Czjzek, M., Cicek, M., Zamboni, V., Burmeister, W.P., Bevan, D.R., Henrissat, B., Esen, A., 2001. Crystal structure of a monocotyledon (maize ZMGLu1) β -glucosidase and a model of its complex with *p*-nitrophenyl β -D-thiogluconate. *Biochem. J.* 354, 37–46.
- Dopitová, R., Mazura, P., Janda, L., Chaloupková, R., Jeřábek, P., Damborský, J., Filipi, T., Kiran, N.S., Bryzobohatý, B., 2008. Functional analysis of the aglycone-binding site of maize β -glucosidase Zm-p60.1. *FEBS J.* 275, 6123–6135.
- Gloster, T.M., Roberts, S., Ducros, V.M., Perugino, G., Rossi, M., Hoos, R., Moracci, M., Vasella, A., Davies, G.J., 2004. Structural studies of the beta-glycosidase from *Sulfolobus solfataricus* in complex with covalently and noncovalently bound inhibitors. *Biochemistry* 43, 6101–6109.
- Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280, 309–316.
- Henrissat, B., Callebaut, I., Fabrega, S., Lehn, P., Mornon, J.P., Davies, G., 1995. Conserved catalytic machinery and the prediction of a common fold for several families of glycosyl hydrolases. *Proc. Natl. Acad. Sci. USA* 92, 7090–7094.
- Hommel, G., Withers, S.G., Chuenchor, W., Ketudat Cairns, J., Svasti, J., 2007. Enzymatic synthesis of cellobio-oligosaccharides by mutated rice β -glucosidases. *Glycobiology* 17, 744–753.
- Hrmova, M., Harvey, A.J., Wang, J., Shirley, N.J., Jones, G.P., Stone, B.A., Høj, P.B., Fincher, G.B., 1996. Barley β -D-glucan exohydrolases with β -D-glucosidase activity. *J. Biol. Chem.* 271, 5277–5286.
- Hrmova, M., MacGregor, E.A., Biely, P., Stewart, R.J., Fincher, G.B., 1998. Substrate binding and catalytic mechanism of a barley β -D-glucosidase/(1, 4)- β -D-glucan exohydrolase. *J. Biol. Chem.* 273, 11134–11143.
- Isorna, P., Polaina, J., Latorre-Garcia, L., Canada, F.J., Gonzalez, B., Sanz-Aparicio, J., 2007. Crystal structures of *Paenibacillus polymyxa* β -glucosidase B complexes reveal the molecular basis of substrate specificity and give new insights into the catalytic machinery of family I glycosidases. *J. Mol. Biol.* 371, 1204–1218.
- Jenkins, J., Lo Leggio, L., Harris, G., Pickersgill, R., 1995. β -glucosidase, β -galactosidase, family A cellulases, family F xylanases and two barley glycanases form a superfamily of enzymes with 8-fold β/α architecture and with two conserved glutamates near the carboxy-terminal ends of β -strands four and seven. *FEBS Lett.* 362, 281–285.
- Jones, T.A., Zou, J.Y., Cowan, S.W., Kjeldgaard, M., 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. Sect. A* 47, 110–119.
- Ketudat Cairns, J.R., Esen, A., 2010. β -Glucosidases. *Cell. Mol. Life Sci.* doi 10.1007/s00018-010-0399-2.
- Kuntothom, T., Luang, S., Harvey, A.J., Opasiri, R., Fincher, G.B., Hrmova, M., Ketudat Cairns, J.R., 2009. Rice family GH1 glycoside hydrolases with β -D-glucosidase and β -D-mannosidase activities. *Arch. Biochem. Biophys.* 491, 85–95.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291.
- Leatherbarrow, R.J., 2001. GraFit, 5th ed. Erithacus Software Ltd, Horley, UK.
- Mackenzie, L.F., Wang, Q., Warren, R.A.J., Withers, S.G., 1998. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* 120, 5583–5584.
- Mülegger, J., Jahn, M., Chen, H.-M., Warren, R.A.J., Withers, S.G., 2005. Engineering of a thioglycoligase: randomized mutagenesis of the acid-base residue leads to improved catalysts. *Prot. Engineer. Design Select.* 18, 33–40.
- Murshudov, G.N., Lebedev, A., Vagin, A.A., Wilson, K.S., Dodson, E.J., 1999. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr. Sect. D* 55, 247–255.
- Nijkken, Y., Tsukada, T., Igarashi, K., Samejima, M., Wakagi, T., Shoun, H., Fushinobu, S., 2007. Crystal structure of intracellular family 1 β -glucosidase BGL1A from the basidiomycete *Phanerochaete chrysosporium*. *FEBS Lett.* 581, 1514–1520.
- Noguchi, J., Hayashi, Y., Baba, Y., Zkino, O., Kimura, M., 2008. Crystal structure of the covalent intermediate of human cytosolic beta-glucosidase. *Biochem. Biophys. Res. Comm.* 374, 549–552.
- Opasiri, R., Ponthong, B., Onkokoosong, T., Akiyama, T., Esen, A., Ketudat Cairns, J.R., 2006. Analysis of rice glycosyl hydrolase family I and expression of Os4bglu12 β -glucosidase. *BMC Plant Biol.* 6, 33.
- Opasiri, R., Ketudat Cairns, J.R., Akiyama, T., Wara-Aswapati, O., Svasti, J., Esen, A., 2003. Characterization of a rice β -glucosidase highly expressed in flower and germinating shoot. *Plant Sci.* 165, 627–638.

- Opassiri, R., Hua, Y., Wara-Aswapati, O., Akiyama, T., Svasti, J., Esen, A., Ketudat Cairns, J.R., 2004. β -Glucosidase, exo- β -glucanase and pyridoxine transglucosylase activities of rice BGlu1. *Biochem. J.* 379, 125–131.
- Otwinowski, Z., Minor, W., 1997. Processing of X-ray diffraction data collected in oscillation mode. In: Carter, C.W., Sweet, R.M. (Eds.), *Meth. Enzymol.*, Vol. 276. Academic Press, New York, pp. 307–326.
- Sanz-Aparicio, J., Hermoso, J.A., Martínez-Ripoll, M., Lequerica, J.L., Polaina, J., 1998. Crystal structure of β -glucosidase A from *Bacillus polymyxa*: insights into the catalytic activity in family1 glycosyl hydrolases. *J. Mol. Biol.* 275, 491–502.
- Seshadri, S., Akiyama, T., Opassiri, R., Kuaprasert, B., Ketudat Cairns, J.R., 2009. Structural and enzymatic characterization of Os3BGlu6, a rice β -glucosidase hydrolyzing hydrophobic glycosides and (1 $\rightarrow$ 3)- and (1 $\rightarrow$ 2)-linked disaccharides. *Plant Physiol.* 151, 47–58.
- Shen, H., Byers, L.D., 2007. Thioglycoside hydrolysis catalyzed by β -glucosidase. *Biochem. Biophys. Res. Commun.* 362, 717–720.
- Sticklin, M., 2006. Plant genetic engineering to improve biomass characteristics for biofuels. *Curr. Opin. Biotechnol.* 17, 315–319.
- Sue, M., Yamazaki, K., Yajima, S., Nomura, T., Matsukawa, T., Iwamura, H., Miyamoto, T., 2006. Molecular and structural characterization of hexameric β -D-glucosidases in wheat and rye. *Plant Physiol.* 141, 1237–1247.
- Verdoucq, L., Czjzek, M., Moriniere, J., Bevan, D.R., Esen, A., 2003. Mutational and structural analysis of aglycone specificity in maize and sorghum β -glucosidases. *J. Biol. Chem.* 278, 25055–25062.
- Verdoucq, L., Moriniere, J., Bevan, D.R., Esen, A., Vasella, A., Henrissat, B., Czjzek, M., 2004. Structural determinants of substrate specificity in family 1 β -glucosidases: novel insights from the crystal structure of sorghum dhurrinase-1, a plant β -glucosidase with strict specificity, in complex with its natural substrate. *J. Biol. Chem.* 279, 31796–31803.
- Wang, Q., Trimbur, D., Graham, R., Warren, R.A.J., Withers, S.G., 1995. Identification of the acid/base catalyst in *Agrobacterium faecalis* beta-glucosidase by kinetic analysis of mutants. *Biochemistry* 34, 14554–14562.
- Weiss, M.S., Brandl, M., Sühnel, J., Pal, D., Hilgenfeld, R., 2001. More hydrogen bonds for the (structural) biologist. *Trends Biochem. Sci.* 26, 521–523.
- Withers, S.G., Warren, R.A.J., Street, I.P., Rupitz, K., Kempton, J.B., Aebersold, R., 1990. Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a “retaining” glycosidase. *J. Am. Chem. Soc.* 112, 5887–5889.
- Wurch, T., Lestienne, F., Pauwels, P.J., 1998. A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes. *Biotechnol. Techniq.* 12, 653–657.
- Zechel, D.L., Boraston, A.B., Gloster, T., Boraston, C.M., Macdonald, J.M., Tilbrook, D.M., Stick, R.V., Davies, G.J., 2003. Iminosugar glycosidase inhibitors: structural and thermodynamic dissection of the binding of isofagomine and 1-deoxynojirimycin to β -glucosidases. *J. Am. Chem. Soc.* 125, 14313–14323.
- Zouhar, J., Vévodová, J., Marek, J., Damborský, J., Su, X.-D., Bryzobohatý, B., 2001. Insights into the functional architecture of the catalytic center of a maize beta-glucosidase Zm-p60.1. *Plant Physiol.* 127, 973–985.