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*

Lionel Verdoucq¶§¶, Jeanne Morinière§¶, David R. Bevan**, Asim Esen‡, Andrea Vasella†‡†, Bernard Henriisst, and Mirjam Czjzek¶§§

*From the Departments of **Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, the †Architectura et Functio des Macromolecules Biologiques, CNRS, 31 Chemin Joseph Aiguier, F13402 Marseille Cedex 20, France, and the ‡‡Laboratory for Organic Chemistry, Eidgenössische Technische Hochschule Zürich, Eidgenössische Technische Hochschule Högerberg, HCI H317, 8093 Zürich, Switzerland

Plant β-glucosidases play a crucial role in defense against pests. They cleave, with variable specificity, β-glucosides to release toxic aglycone moieties. The Sorghum bicolor β-glucosidase isoenzyme Dhr1 has a strict specificity for its natural substrate dhurrin (p-hydroxy-(S)-mandelenitrile-β-D-glucoside), whereas its close homolog, the maize β-glucosidase isoenzyme Glu1, which shares 72% sequence identity, hydrolyzes a broad spectrum of substrates in addition to its natural substrate 2-O-β-D-glucopyranosyl-4-hydroxy-7-methoxy-1,4-benzoxaxin-3-one. Structural data from enzyme-substrate complexes of Dhr1 show that the mode of aglycone binding differs from that previously observed in the homologous maize enzyme. Specifically, the data suggest that Asn259, Phe261, and Ser462, located in the aglycone-binding site of S. bicolor Dhr1, are crucial for aglycone recognition and binding. The tight binding of the aglycone moiety of dhurrin promotes the stabilization of the reaction intermediate in which the glucose moiety is in a deformed 1S conformation within the glycocone-binding site, ready for nucleophilic attack to occur. Compared with the broad specificity maize β-glucosidase, this different binding mode explains the narrow specificity of sorghum dhurrinase-1.

Carbohydrates and their glycoconjugates are one of the most diverse groups of organic molecules in the biosphere. The selective cleavage of glycosidic bonds is crucial in a variety of fundamental biological processes for all living organisms. The large (and growing) number of glycosidase hydrolase families reflects this diversity of substrates and the need for selective cleavage of the glycosidic bond. Ninety-one glycoside hydrolase families are currently available on the continuously updated CAZY web server (1). β-Glucosidases constitute a major group in glycoside hydrolase families 1 and 3 and hydrolyze either O-linked β-glycosidic bonds (β-D-glucoside glucohydrolase, EC 3.2.1.21) or S-linked β-glycosidic bonds (myrosinase or β-D-thioglucoside glucohydrolase, EC 3.2.3.1). More precisely, enzymes of glycoside hydrolase family 1 hydrolyze substrates of the type G-O/S-X, where G indicates the glycosyl residue and X can be either another glycosyl residue or a non-glycosyl aglycone group. In higher plants, the major functions of β-glucosidases are defense against pests with the release of bitter or toxic aglycones and their breakdown products (2, 3), phytohormone activation (4, 5), lignification (6), and cell wall catabolism (7). The nature of the aglycone moiety of substrates is believed to be critical for the specificity and physiological functions of these enzymes.

Because plants possess a large number of β-glucosidases (most of the 48 genes identified as putative β-glucosidase genes in Arabidopsis thaliana do not have a known function), the understanding of the mechanism of substrate specificity is important for accurately predicting the diverse physiological functions of these proteins. The roles of the 2 catalytic glutamates (8, 9) included in the TFINEP and YLV/YTEN peptide motifs of β-glucosidases as the general acid/base catalyst and the nucleophile, respectively (10), are now well understood. These residues are 5.5–6 A apart within the active-site pocket on opposite sides of the glycosidic bond (8, 11) and are required in the two steps of the substrate hydrolysis that results in retention of the anomeric conformation of C-1 at the point of cleavage. The conformational change in the glucose moiety prior to nucleophilic attack and the determinants of binding and deformation of the glucose moiety in subsite −1 are well documented for certain glycoside hydrolases (12–15); but for others, these still need to be determined. However, how β-glucosidases recognize their substrates and interact with them, specifically at the aglycone moiety, have so far been determined only for one enzyme displaying broad substrate specificity (16), even though this mechanism is the basis of subtle substrate specificity occurring in the tremendous diversity of natural substrates. Recent work focused on site-directed mutagenesis of residues found to be crucial for enzyme-substrate interac-
ficity because they represent extremes in substrate specificity.

Model system to address questions related to substrate specificity in these two enzymes. These two enzymes provide an ideal substrate specificity-determining sites are different, i.e., shown are the structures of DIMBOA-Glc (A), the natural substrate of homologous sorghum dhurrinase-1 (SbDhr1); and the glucotetrazole inhibitor (C), which mimics the H₃ transition-state intermediate.

Fig. 1. Schematic representation of substrate and inhibitor molecules. Shown are the structures of DIMBOA-Glc (A), the natural substrate of the maize β-glucosidase (ZmGlu1); of dhurrin (B), the natural substrate of homologous sorghum dhurrinase-1 (SbDhr1); and the glucotetrazole inhibitor (C), which mimics the H₃ transition-state intermediate.

In the aglycone pocket of the maize β-glucosidase (ZmGlu1) based on the crystal structure of its substrate complexes (17, 18).

Using site-directed reciprocal mutants of the sorghum (SbDhr1) and maize (ZmGlu1) β-glucosidases, which share 70% sequence identity, Verdoucq et al. (17) determined that the aglycone (i.e., substrate) specificity-determining sites are different in these two enzymes. These two enzymes provide an ideal model system to address questions related to substrate specificity because they represent extremes in substrate specificity.

The cyanogenic β-glucosidase SbDhr1 hydrolyzes only its natural substrate dhurrin (Fig. 1), whereas ZmGlu1 hydrolyzes a broad spectrum of artificial and natural substrates, including its natural substrate DIMBOA-Glc, but it does not hydrolyze dhurrin (19, 20). Structural analysis of the inactive double mutant ZmGlu1-E191D/F198V in complex with DIMBOA-Glc showed that the presence of valine instead of phenylalanine at position 198 reduces drastically the activity of the maize enzyme on most substrates tested because of a rearrangement of 3 amino acids, viz. Phe325, Phe466, and Glu189 (21). These residues are involved in either the glycone- or aglycone-binding pocket of the maize β-glucosidase, and their changed position causes binding of the substrate in a different manner compared with the native enzyme (17). The results also show that, whereas the F198V substitution has a drastic effect on ZmGlu1 activity, the reverse substitution has no effect on SbDhr1 activity toward the substrates hydrolyzed by the maize enzyme. Therefore, other amino acids are responsible for the strict specificity of the sorghum enzyme, whereas the corresponding residues have little or no role in ZmGlu1 substrate specificity. Structural analysis of SbDhr1 and its inactive mutant in complex with dhurrin should bring more clarity to this aspect of β-glucosidase specificity.

The purpose of this study is to understand the basis of the strict substrate specificity exhibited by the SbDhr1 β-glucosidase and to gain novel insights into the mechanism of substrate (aglycone) recognition and binding in β-glucosidases. To address these questions, we have produced and purified SbDhr1 and an active mutant of SbDhr1 (referred to as SbDhr1-E189D hereafter) obtained by site-directed mutagenesis for three-dimensional analysis by x-ray crystallography of the native enzyme and mutant enzyme-substrate complexes. To complete the picture of enzyme-substrate or enzyme-inhibitor interactions, this study also describes the crystal structure of inactive ZmGlu1-E191D in complex with a cyclic glucotetrazole inhibitor that mimics a partially planar reaction transition state (21). In this work, we report the identification of the crucial amino acids within the active site of SbDhr1, a β-glucosidase known for its strict specificity for its natural substrate, and compare its mode of substrate binding with that of the homologous maize β-glucosidase (22), an enzyme with broad substrate specificity. This allows us to define determinants responsible for either broad or narrow substrate specificity in these two plant β-glucosidases.

EXPERIMENTAL PROCEDURES

General Methods—Cloning and expression of the dhr1 cDNA in Escherichia coli were performed using standard procedures (23). DNA synthesis by PCR was performed with high fidelity Pfu Turbo DNA polymerase (Stratagene). DNA was visualized with SYBR Gold (Molecular Probes, Inc.) and purified from excised gel bands with the QIAquick gel extraction kit (QIAGEN Inc.). All restriction enzymes were from Stratagene, and T4 DNA ligase was from Promega. All constructs were sequenced for confirmation in the Virginia Bioinformatics Institute Core Laboratory Facility (Blacksburg, VA).

Production of SbDhr1 and SbDhr1-E189D in E. coli and Purification—The mutant SbDhr1-E189D was overexpressed and purified for structural analysis in complex with the natural substrate. Therefore, it was necessary to mutate the acid/base catalyst residue Glu189 to Asp189 to obtain a catalytically inactive enzyme. Mutated cDNAs were constructed by the PCR-mediated overlap extension technique using a pair of specific complementary primers (EDrev, 5’-GACCTTTAATGACCCGGAG-3’; and EDfor, 5’-CTCCGGGTCATTTAAAGGTC-3’) and a pair of vector-specific primers (T7prom, 5’-TAATAGCCTACGCTAT-AGG-3’; and T7term, 5’-ATGCTTTTAGTGGTCACGGGTTG-3’) as described (24). Because of a codon usage bias for E. coli expression with two consecutive rare codons (AGG) in the cDNA of SbDhr1, the cDNA was mutated using the complementary primer pair SbRfor (5’-CCTGGAAATCCCTCCGGCCGGTAGCTTGG-3’) and SbRrev (5’-CCGATTGACCCGGGAGGGTTTCCGGAGGCGGTTTCCGAGG-3’) to have the best E. coli codon usage for Arg195 and Arg240. The DNA template used in this study was pET28a-SbDhr1, which was previously used for expression in E. coli and subsequent purification in our laboratory (17). The T7prom, EDrev, and T7term, EDfor PCR products were gel-purified and combined in the second PCR step using the T7prom, and T7term oligonucleotides to obtain the full-length mutated cDNA. The resulting PCR product was gel-purified, digested with NheI and XhoI restriction enzymes, and cloned into the expression plasmid pET28a (Novagen). This plasmid allows the fusion of a His₆ extension at the N-terminal part of the mutated β-glucosidases for purification on a Ni²⁺-agarose column as described previously (17).

The final product was submitted to dialysis and subsequent concentration to obtain a pure 8 mg/ml solution in 20 mM Tris buffer, pH 7.0.

Crystallization of SbDhr1 and SbDhr1-E189D in Complex with Dhurrin—Crystallization conditions were first investigated using four sparse matrix sampling kits (Molecular Dimensions I and II and Wizard 1 and II). Of these, two distinct conditions led to crystals for both the native and inactive mutant enzymes and were chosen for further optimization by the hanging drop vapor diffusion method. The first optimized crystallization condition led to a cubic crystal form, and it involved mixing 2 µl of protein solution (8 mg/ml in 20 mM Tris, pH 7) with 2 µl of reservoir solution containing 0.1 M Hepes, pH 7.5, 2 M (NH₄)₂SO₄, and 50 mM NaCl. The crystals belonged to space group P4₁32 and had unit cell parameters of a = b = c = 195.19 Å. The second crystallization condition led to a hexagonal crystal form in a mixture containing 1 µl of protein solution in 0.1 M Tris, pH 7, and 14–16%
The crystals were soaked in these two solutions for a total of 30 min and diffracted in a cryoprotective solution supplemented with ligand. First, the crystals of the inactive mutant enzyme were transferred into a crystallization solution supplemented with 5% glycerol and ~5 mmol of ligand. They were then transferred into a second solution containing the crystallization solution supplemented with 10% glycerol and ~5 mmol of ligand prior to freezing in a cold nitrogen stream at 100 K. The total soaking time was ~15 min. Because the hexagonal crystal form had a large unit cell parameter, the maximum resolution attainable was 1.9 Å. The cubic crystals did not diffract at a resolution better than 2.0 Å. The ShbDhr1-E189D-dhurrin complex was obtained with crystals of the cubic form.

**Crystallization of ZmGlu1-E191D in Complex with Glucotetrazole**—The production and purification of the inactive mutant ZmGlu1-E191D were described previously (17). The ZmGlu1-E191D-glucotetrazole complex was obtained in soaking experiments in which 92 μl of crystallization solution was mixed with 5 μl of glycerol (the cryoprotectant) and 3 μl of inhibitor (10 mM) solution. Subsequently, 45.5 μl of this mixture was supplemented with 2.5 μl of glycerol and 2 μl of inhibitor solution. The crystals were soaked in these two solutions for a total of 30 min and then frozen in a stream of N₂ at 100 K. They belonged to space group P3₁, and had unit cell parameters of \( a = b = 100.12 \), \( c = 279.54 \) Å for the hexagonal crystal form, and \( a = 96.18, b = 104.80, c = 119.34 \) Å for the cubic crystal form. The unit cell contains three dimers, whereas in the cubic crystal form, the dimer is obtained by the crystallographic 2-fold symmetry. The cubic crystals did not diffract at a resolution better than 2.0 Å. The ShbDhr1-E189D-dhurrin complex was obtained with crystals of the cubic form.

**Crystallographic Data Refinement**—The refinement of all structures was performed using the program REFMAC5 (26), and details of the refinement statistics are presented in Table 1. The overall structure of ShbDhr1 in the two crystal forms is identical, having the classical (βα)₉ fold representative of glucoside hydrolase family 1 enzymes. The hexagonal asymmetric unit contains three dimers, whereas in the cubic crystal form, the dimer is obtained by the crystallographic 2-fold symmetry operation. The structure of ShbDhr1 shows remarkably high similarity to the homologous maize β-glucosidase, reflecting the high sequence similarity (>70% identity) between them (Fig. 2). Consequently, the dimer interface of ShbDhr1 is also very similar to that of the maize enzyme. The only difference in the dimer interface concerns a threonine in ZmGlu1 that is replaced by a more hydrophobic valine in ShbDhr1. Otherwise, the same interfacial hydrogen bonds involving residues of the same nature are found in the dimer of ShbDhr1 as in ZmGlu1. The root mean square deviation of all C-α atoms between the dimers of ShbDhr1 and ZmGlu1 is 0.62 Å, and that between the dimers of ShbDhr1-E189D and ZmGlu1 is 0.56 Å. The only notable differences are two surface loops surrounding the active-site cleft, loops B and D (for loop definitions, see Ref. 22), which display higher root mean square deviations (1.03 and 0.89 Å, respectively) than the overall mean.

The superimposition of the two homologous enzyme structures revealed that, as expected, the major difference between the two enzymes resides in the shape and nature of the active-site cleft (Fig. 3, A and B). This cleft has been described to be a flattened crater or slot in ZmGlu1 (16, 22), whereas it appears to be considerably wider and smaller in ShbDhr1. This difference in shape is due mainly to substitutions F196V, F203L, and F462S (residue numbers are for ShbDhr1), the 3 residues particularly important in determining aglycone specificity in ZmGlu1 (16), and is due also to substitution M261F, which covers one-third of the cleft opening in ShbDhr1.

**Analysis of the Crystal Structure of ShbDhr1-E189D in Complex with Dhurrin**—The structure of ShbDhr1-E189D in complex with its natural substrate dhurrin was determined at a previous study (17). No significant differences in stability and kinetic parameters were observed between the two forms, with a pH optimum of 6.2, similar to the one found by Hösel et al. (29), and a strict specificity for dhurrin hydrolysis (\( K_m = 61 \pm 10 \mu M \) and \( k_{cat} = 14.3 \pm 1.7 \) s⁻¹) (data not shown). In comparison, the pH optimum of ZmGlu1 is 5.8, and the hydrolysis of the natural substrate DIMBOA-β-glucoside is performed with \( K_m = 98 \mu M \) and \( k_{cat} = 29.2 \) s⁻¹ (19).

**Overall Structure of ShbDhr1**—The structure of ShbDhr1 was solved by molecular replacement, and the solutions obtained had an overall R-factor of 35% and a correlation coefficient of 62. The overall structure of ShbDhr1 in the two crystal forms is identical, having the classical (βα)₉ fold representative of glucoside hydrolase family 1 enzymes. The hexagonal asymmetric unit contains three dimers, whereas in the cubic crystal form, the dimer is obtained by the crystallographic 2-fold symmetry operation. The structure of ShbDhr1 shows remarkably high similarity to the homologous maize β-glucosidase, reflecting the high sequence similarity (>70% identity) between them (Fig. 2). Consequently, the dimer interface of ShbDhr1 is also very similar to that of the maize enzyme. The only difference in the dimer interface concerns a threonine in ZmGlu1 that is replaced by a more hydrophobic valine in ShbDhr1. Otherwise, the same interfacial hydrogen bonds involving residues of the same nature are found in the dimer of ShbDhr1 as in ZmGlu1. The root mean square deviation of all C-α atoms between the dimers of ShbDhr1 and ZmGlu1 is 0.62 Å, and that between the dimers of ShbDhr1-E189D and ZmGlu1 is 0.56 Å. The only notable differences are two surface loops surrounding the active-site cleft, loops B and D (for loop definitions, see Ref. 22), which display higher root mean square deviations (1.03 and 0.89 Å, respectively) than the overall mean.

The superimposition of the two homologous enzyme structures revealed that, as expected, the major difference between the two enzymes resides in the shape and nature of the active-site cleft (Fig. 3, A and B). This cleft has been described to be a flattened crater or slot in ZmGlu1 (16, 22), whereas it appears to be considerably wider and smaller in ShbDhr1. This difference in shape is due mainly to substitutions F196V, F203L, and F462S (residue numbers are for ShbDhr1), the 3 residues particularly important in determining aglycone specificity in ZmGlu1 (16), and is due also to substitution M261F, which covers one-third of the cleft opening in ShbDhr1.
Table II

Refinement statistics of native ShDhr1, ShDhr1-E189D in complex with dhurrin, and ZmGlu1-E191D in complex with the inhibitor glucotetrazole

|                  | ShDhr1       | ShDhr1-E189D-dhurrin | ZmGlu1-E191D-glucotetrazole |
|------------------|--------------|----------------------|-----------------------------|
| Resolution       | 1.9–30.0     | 2.0–30.0             | 1.95–30.0                   |
| R_factor (%)     | 17.3         | 18.7                 | 19.6                        |
| R_work (%)       | 20.9         | 21.2                 | 26.1                        |
| No. non-H protein atoms | 28,327   | 4742                 | 7934                        |
| No. ligand atoms | 0            | 21                   | 28                          |
| No. water O atoms | 5062      | 842                  | 494                         |
| Mean deviation from ideal values | | | |
| Bond lengths (Å) | 0.015        | 0.017                | 0.021                       |
| Bond angles      | 1.34°        | 1.53°                | 1.94°                       |
| Torsion angles   | 5.18°        | 5.85°                | 6.89°                       |
| Ramachandran plot | 89.5, 9.8, 0.5 | 88.8, 10.4, 0.8 | 89.8, 9.4, 0.8             |
| No. residues in unallowed regions (%) | 0.2        | 0                    | 0                           |

2.0-Å resolution from cubic crystals soaked in a ligand-containing solution (see “Experimental Procedures” for details). The electron density was clear for both the glycone moiety (binding subsite -1) and the aglycone moiety (binding subsite -1) (for subsite nomenclature, see Ref. 30) of the substrate molecule (Fig. 4A). This is in contrast to the crystal structure of the ZmGlu1-E191D-DIMBOA-Glc complex, where certain atoms around the scissile C-1-O bond were not well defined. The lack of definition of the electron density in that case had been interpreted as the result of flexibility of the substrate at the level of glycosidic linkage (16). It is known that the glucose unit is deformed prior to nucleophilic attack (14, 31); and for this reason, it is plausible that the active-site pocket of ZmGlu1 adapts to both the non-deformed substrate, present in solution, and the deformed substrate, prior to nucleophilic attack. The lack of electron density is therefore a consequence of disordered substrate molecules in a range of possible conformations.

In ShDhr1, however, the situation is entirely different in that the enzyme binds the substrate directly in the necessary conformational state. In the structure of the complex described here, we observed that the glucose moiety assumes the 1S3 skew boat conformation as opposed to the half-chair conformation in ZmGlu1. The stabilization of the reaction intermediate in the active site of ShDhr1 is due to tighter binding of the aglycone moiety. In contrast to the pure hydrophobic interactions between ZmGlu1 and its natural substrate DIMBOA-Glc, we found 2 residues forming indirect (mediated through water molecules) hydrogen bonds with the substrate pockets of dhurrin in ShDhr1. These are Ser462, which undergoes a conformational rotation to form an indirect hydrogen bond with the hydroxyl group in the para-position of the phenyl ring with respect to the glycoside, and Asn463, which binds via a water molecule to the cyano group. Furthermore, the cyano group points into a polar pocket, covered by Phe261. It is interesting to note that ShDhr2, which has a methionine at this position, as does ZmGlu1 (Fig. 2), displays a broader substrate specificity with ShDhr1 since it cleaves, albeit at slow rate, 4-methylumbelliferyl-β-D-glucoside and nitrophenyl-β-D-glucosides in addition to the natural substrate dhurrin (20). Calculation of the dipole moment of this pocket and that of the cyano group showed that their directions coincide (Fig. 4B), indicating that even though no direct hydrogen bond is formed, binding of the cyano group is dominated by polar interactions. These two polar interactions together with hydrophobic interactions of the phenyl ring with Val196 and Leu203 determine a single position of the aglycone moiety within the aglycone-binding pocket. Consequently, the position of the glucosidic moiety is equally constrained to one single conformation, i.e. the skew boat conformation necessary for catalysis to take place.

Analysis of the Crystal Structure of ZmGlu1-E191D in Complex with the Inhibitor Glucotetrazole—The structure of ZmGlu1-E191D in complex with the glucotetrazole inhibitor (Fig. 1C) was determined at 1.95-Å resolution from crystals that were soaked in the inhibitor solution. This type of transition-state analog inhibitor has specifically been developed for family 1 β-glucosidases and typically has inhibition constants (Ki) in the range of 1 μM (32, 33). The electron density of the Fourier difference map, calculated including only the protein atoms, clearly defined all atoms of the inhibitor molecule (Fig. 4C). This is in contrast to the natural substrate in complex with ZmGlu1-E191D, described elsewhere (16). Interestingly, it is perfectly superimposable with the same inhibitor in complex with myrosinase (Fig. 4D) (33). The glucotetrazole binds in a pure 4H2 half-chair conformation. However, glucotetrazole does not adopt the same position as the glucose in the complex with the natural substrate DIMBOA-Glc (Fig. 5). The hydrogen bonds formed with the surrounding protein residues are summarized in Table III. It must be noted that the residues that bind to groups O-3, O-4, and O-6 of the glucose moiety in the DIMBOA-Glc complex with ZmGlu1-E191D (16) now bind to O-2, O-3, and O-4 of the glucotetrazole inhibitor molecule. The only residue that does not change its hydrogen bonding partner is Glu264, which binds to O-6 in all complexes. Consequently, Glu264 undergoes a conformational change upon binding in this complex.

**DISCUSSION**

Catalysis by glycosidases is driven to a large extent by harnessing noncovalent interactions all along the reaction pathway (34–36). Analyses of enzyme-substrate or enzyme-inhibitor complexes provide an experimental probe of such interactions and allow detailed dissection of hydrogen bonds involving reaction intermediates and/or putative transition-state mimics. Here, the crystal structures of inactive ShDhr1 and ZmGlu1 in complex with the natural substrate for the former and an inhibitor mimicking a transition-state analog for the latter not only shed light on the hydrogen bonding network and further noncovalent interactions within the active site, but also allowed us to identify the determinants responsible for a highly specific enzyme in the case of ShDhr1 and an enzyme with broad specificity in the case of ZmGlu1.

The first stable state along the reaction coordinate is the so-called “Michaelis complex” of unhydrolyzed substrate (Fig. 4A). This complex has now been observed in several glycosidases (32) and involves in all cases a distorted species in which the glycosyl moiety is twisted out of its relaxed 4C1 (chair)

---

3 N. Vicitphan and A. Esen, personal communication.
FIG. 2. Sequence alignment of sorghum dhurrinase SbDhr1 and maize β-glucosidase ZmGlu1 and their isoenzymes SbDhr2 and ZmGlu2. The four mature protein sequences were aligned with the program ESPRRIPT and are presented with Boxshade. The regions of sequence identity are white with a red background. The 2 catalytic glutamates are indicated by blue arrowheads. Colored arrowheads highlight the amino acids located within the glycone-binding pocket (green) and those not conserved between SbDhr1 and ZmGlu1 as well as those located within the aglycone-binding pocket (purple).
Analyses of such complexes for different glycoside hydrolase families (15) revealed that the type of deformation is characteristic for an enzyme family, and different conformational ensembles have been observed for \( H_{9252} \)-mannanases (glycoside hydrolase family 26) (37) and for retaining cellulases (families 5 and 7) (13, 38). In general, for \( H_{9252} \)-retaining enzymes that hydrolyze gluco-configured substrates, 1S3 skew boat Michaelis complexes imply catalysis via the adjacent 4H3 half-chair transition state (36). The conformations of the reaction intermediates of the glycosyl moiety are achieved by “electrophilic migration” of the C-1 atom along a reaction coordinate (39). Halfway between the Michaelis and covalent intermediate conformations, the anomeric carbon is coplanar with C-5, O-5, and C-2. For this reason, half-chair transition-state mimics such as the inhibitor glucotetrazole display tight binding (40, 41). For glycoside hydrolase family 1, several covalent glycosyl-enzyme intermediates have already been experimentally observed in myrosinase (33). In the structure of SbDhr1-E189D in complex with dhurrin described here, we have reported for the first time the experimental observation of a Michaelis complex in this family in which the substrate molecule adopts a distorted 1S3 conformation of the glucosyl moiety. The position of the aglycone of dhurrin in the aglycone pocket determines the binding of glucose in the distorted 1S3 configuration, ready for catalysis to take place. In ZmGlu1 and myrosinase (33), the same position of the glycosyl moiety has been observed only for

**Figure 3.** Aglycone-binding pockets of ZmGlu1 and SbDhr1. A, surface representation of the active-site slot of ZmGlu1. The location of the natural substrate molecule DIMBOA-Glc is shown in stick representation. The residues forming the aglycone pocket are visible underneath the surface. B, surface representation of the active-site funnel of SbDhr1. The location of the natural substrate molecule dhurrin is shown in stick representation. The residues responsible for the different form of the active-site pocket are visible underneath the surface.

**Figure 4.** Dhurrin bound in the active sites to SbDhr1-E189D and glucotetrazole bound to ZmGlu1-E191D. A, electron density surrounding the dhurrin molecule in the active site of SbDhr1-E189D. The 2Fo - Fc Fourier difference maps at the final stage of refinement are shown contoured at 1σ above the mean density. B, a slice in the surface representation of SbDhr1 in complex with dhurrin showing the cyano group-binding pocket. The dipole moment of the polar pocket, calculated with GRASP (28), coincides with that of the cyano group. C, electron density around the glucotetrazole molecule in the active site of ZmGlu1-E191D. The Fo - Fc Fourier difference maps before refinement are shown, calculated using only the enzyme model phases without substrate, contoured at 2.5σ above the mean density. D, superimposition of the active sites of myrosinase (blue) and ZmGlu1 (yellow), both in complex with the glucotetrazole inhibitor molecule.

**Figure 5.** Superimposition of the glucose moieties of DIMBOA-Glc in ZmGlu1 and dhurrin in SbDhr1. The glucose ring of DIMBOA-Glc has rotated by ~60° with respect to that of dhurrin. Consequently, the residues binding the sugar groups O-2, O-3, and O-4 in dhurrin bind O-3, O-4, and O-6 in DIMBOA-Glc. See also Table III. In each box, the top residue occurs in ZmGlu1, and the bottom residue occurs in SbDhr1. Of the 2 glutamates shown by stick representation, Glu464 occurs in ZmGlu1, whereas Glu460 occurs in SbDhr1.
the glutotetrazole inhibitor that mimics the coplanar reaction transition state.

Analysis of the interactions of the substrate molecules with specific amino acids in the active-site cleft showed that it is the interactions involving the aglycone moiety that influence the position of the glycosyl moiety. In ZmGlu1, aromatic stacking interactions are dominant in the aglycone-binding pocket, and they are capable of accommodating different aromatic aglycones. Therefore, this enzyme binds and hydrolyzes a number of β-linked glucosides, having in common an aromatic aglycone like those present in 4-methylumbelliferol-β-D-glucoside (K_m = 370 μM and k_cat = 30.3 s⁻¹) (17) and p-nitrophenyl-β-D-glucoside (K_m = 380 μM and k_cat = 24.2 s⁻¹), o-nitrophenyl-β-D-glucoside (K_m = 1300 μM and k_cat = 19.8 s⁻¹), and DIMBOA-β-D-glucoside (K_m = 98 μM and k_cat = 29.2 s⁻¹) (19). In SbDhr1, polar interactions involving two functional groups of the dhurrin aglycone are dominating, and the precise relative positions of these groups on the aglycone are crucial for substrate binding and catalysis to take place. As a consequence, the enzyme displays high selectivity for its natural substrate, for which the aglycone-binding pocket is tailored (K_m = 61 ± 10 μM and k_cat = 14.3 ± 1.7 s⁻¹) (17). No relevant activity for other substrates is detectable (19).

Another aspect of the catalytic mechanism is revealed by the complex of ZmGlu1 with the transition state-mimicking inhibitor glutotetrazole. As in myosin (33), it binds in the pure 

4H2 half-chair conformation, further supporting the S3 → 4H2 → 4C1 pathway for the glycosylation step in family 1 β-glucosidases. In the complex described here, the glutamate specifically binding to O-6 (Glu462 in ZmGlu1), a characteristic group of gluco-configured substrates, has undergone a conformational change compared with the native “empty” enzyme and with the complex of ZmGlu1 with its natural substrate. For SbDhr1 in complex with dhurrin, the equivalent glutamate is in the same conformation as in ZmGlu1-glucotetrazole. Interestingly, in the double mutant ZmGlu1-E191D/F198V, the conformation of Glu462 is the same in the complexed or uncomplexed enzyme (17), thus displaying a unique conformation, as in SbDhr1. This glutamate, displaying two conformational states, might play the role of a trigger in ZmGlu1, guiding the movement of the glycoal moiety into the correct distorted position required for nucleophilic attack to occur. This observation is in agreement with the earlier postulated mechanism for this enzyme, in which both ends of the substrate are bound tightly, leaving the necessary flexibility to the atoms surrounding the glycosidic linkage and therefore allowing the stabilization of the different reaction intermediates. The substrate-binding pocket is adapted to stabilize several different conformers of substrate molecules, also providing flexibility to accommodate different aglycone groups, making this enzyme active on a broad spectrum of substrates. In contrast, the substrate-binding pocket of SbDhr1 is tailored only to bind its natural substrate in an activated distorted conformation, leading to a highly specific enzyme.

Specifically, the data suggest that Asn259, Phe261, and Ser462, located in the aglycone-binding site of SbDhr1, are crucial for aglycone recognition and binding in SbDhr1. The first residue unexpectedly plays a role in dhurrin hydrolysis since the corresponding substitution in ZmGlu1 (D261N) is unable to transform this protein to hydrolyze dhurrin, whereas the two other individual substitutions (M263F and F466S) allow little dhurrin hydrolysis by ZmGlu1 (K_m = 66 ± 21 and 65 ± 10 μM and k_cat = 0.21 ± 0.08 and 0.55 ± 0.09 s⁻¹, respectively) (17). The contribution of the latter residue (Ser462) has also been shown in chimera 39, based on SbDhr1, containing residues 462–488 of ZmGlu1 (19). This chimeric enzyme still has a strict specificity for dhurrin, but the k_cat is reduced by 87%, which probably is due mainly to the S462F substitution. When one compares the residues at position 462 throughout the whole family 1 β-glucosidases of Arabidopsis, it is striking to note a high divergence of residues found at this position (Supplemental Fig. 1). At this position, there are 6 alanines, 3 asparagines, 9 leucines, 1 valine, 3 glycines, 3 isoleucines, 8 glutamines, 1 lysine, 1 phenylalanine, 1 serine, 7 glutamic acids, 1 aspartic acid, and 3 cysteines. This variability may reflect the diversity of substrates hydrolyzed by these enzymes. This is also true for the residue at position 463, equally located in the aglycone-binding pocket, which is as variable as position 462 (3 alanines, 9 serines, 1 glutamine, 1 tyrosine, 2 phenylalanines, 1 valine, 5 glycines, 2 lysines, 2 threonines, 1 methionine, 1 glutamic acid, 6 asparagines, 8 aspartic acids, and 5 histidines) and therefore most likely highly important for the substrate specificity of these enzymes.

Plants contain more genes that encode carbohydrate-active enzymes than any other organism whose genome has been sequenced to date (42). For example, the human genome con-

| Site | Distance | Site | Distance | Site | Distance | Site | Distance |
|------|----------|------|----------|------|----------|------|----------|
| O-2  | Asp169 O-82 2.29 | Asp191 O-82 2.70 | Asp191 O-ε1 3.13 | Glu165 O-ε1 3.41 |
| Glu401 O-ε1 2.67 | Glu 401 O-ε1 3.19 | Glu406 O-ε1 2.85 | Glu406 O-ε1 2.87 |
| Glu404 O-ε2 3.07 | Asn186 O-52 3.17 | Glu 406 O-ε2 3.24 | Glu406 O-ε2 3.00 |
| Asn186 O-52 3.38 | His142 O-ε2 3.52 | Asn186 O-52 2.94 | Asn186 O-52 2.94 |
| His143 O-ε2 3.64 | His143 O-ε2 3.38 | Asn186 O-52 2.94 | Asn186 O-52 2.94 |
| O-3  | Gln39 O-ε1 2.60 | Glu 406 O-ε1 3.68 | Gln38 O-ε1 2.63 | Glu 406 O-ε2 3.75 |
| Gln39 N-ε2 3.50 | Glu 406 O-ε1 3.12 | Gln38 N-ε2 3.57 | Glu 406 O-ε2 3.75 |
| His143 N-ε2 2.90 | Asn186 O-52 3.23 | His142 N-ε2 3.54 | His141 N-ε2 2.98 |
| Trp461 N-ε1 2.93 | His143 N-ε2 3.28 | Trp465 N-ε1 3.17 | Trp465 N-ε1 3.17 |
| Gln460 O-ε1 2.63 | Glu 406 N-ε2 3.35 | Gln38 N-ε2 3.26 | Gln38 N-ε2 3.26 |
| Trp465 N-ε1 3.18 | His143 N-ε2 3.10 | Glu464 O-ε1 2.00 | Glu464 O-ε1 2.61 |
| O-6  | Glu460 O-ε2 2.74 | Glu464 O-ε2 2.56 | Glu464 O-ε2 2.60 | Glu464 O-ε2 2.56 |

**Table III**

**Substrate Specificity in Family 1 β-Glucosidases**

"Site" indicates the glycone-binding site (subsite −1).

The contribution of the latter residue (Ser462) has also been shown in chimera 39, based on SbDhr1, containing residues 462–488 of ZmGlu1 (19). This chimeric enzyme still has a strict specificity for dhurrin, but the k_cat is reduced by 87%, which probably is due mainly to the S462F substitution. When one compares the residues at position 462 throughout the whole family 1 β-glucosidases of Arabidopsis, it is striking to note a high divergence of residues found at this position (Supplemental Fig. 1). At this position, there are 6 alanines, 3 asparagines, 9 leucines, 1 valine, 3 glycines, 3 isoleucines, 8 glutamines, 1 lysine, 1 phenylalanine, 1 serine, 7 glutamic acids, 1 aspartic acid, and 3 cysteines. This variability may reflect the diversity of substrates hydrolyzed by these enzymes. This is also true for the residue at position 463, equally located in the aglycone-binding pocket, which is as variable as position 462 (3 alanines, 9 serines, 1 glutamine, 1 tyrosine, 2 phenylalanines, 1 valine, 5 glycines, 2 lysines, 2 threonines, 1 methionine, 1 glutamic acid, 6 asparagines, 8 aspartic acids, and 5 histidines) and therefore most likely highly important for the substrate specificity of these enzymes.

Plants contain more genes that encode carbohydrate-active enzymes than any other organism whose genome has been sequenced to date (42). For example, the human genome con-
Substrate Specificity in Family 1 β-Glucosidases

...tains only a few β-glucosidases, yet up to 48 putative β-glucosidases are encoded by the Arabidopsis genome. This large multiplicity of enzymes in plants is a reflection of an even larger diversity of substrates since hundreds of different β-glucosides have been isolated and characterized from plants, and an enzyme isoform may be hydrolyzing two or more substrates sharing similar aglycone moieties. Assigning a biological function to each of these 40 plus β-glucosidase isoenzymes is a challenge for plant biologists. But, using homology modeling in conjunction with the improved knowledge of the mechanism of β-glucosidases, we expect to be able to predict in silico the nature of a substrate for a given β-glucosidase in the near future.

Acknowledgments—We thank the European Synchrotron Radiation Facility staff for provision of beam time and the team at beamlines ID14-EH1 and ID14-EH2 for technical assistance during data collection.

REFERENCES

1. Henrissat, B., and Davies, G. (1997) Curr. Opin. Struct. Biol. 7, 637–644
2. Nielsen, H. M. (1988) Phytochemistry 27, 3249–3358
3. Poulton, J. E. (1990) Plant Physiol. 94, 401–405
4. Brazeby, B., Moore, I., Kristofferson, P., Bako, L., Campos, N., Schell, J., and Palme, K. (1993) Science 262, 1051–1054
5. Falk, A., and Rask, L. (1995) Plant Physiol. 108, 1369–1377
6. Dharmawardhana, D. P., Ellis, B. E., and Carlson, J. E. (1995) Plant Physiol. 107, 331–339
7. Leah, R., Kigel, J., Svendsen, I., and Mundy, J. (1995) J. Biol. Chem. 270, 15789–15797
8. Withers, G. S., Warren, R. A. J., Street, I. P., Rupitz, K., Kempton, J. B., and Aebersold, R. (1990) J. Am. Chem. Soc. 112, 5887–5890
9. Wang, Q., Trimbru, D., Graham, R., Warren, R. A., and Withers, S. G. (1995) Biochim. Biophys. Acta 34, 14554–14562
10. Sanz-Aparicio, J., Hermoso, J. A., Martinez-Ripoll, M., Lequerica, J. L., and Polaina, J. (1998) Mol. Biol. 275, 491–502
11. Davies, G., and Henrissat, B. (1996) Structure (Lond.) 3, 853–859
12. Rye, C. S., and Withers, S. G. (2000) Curr. Opin. Chem. Biol. 4, 573–580
13. Davies, G. J., Mackenzie, L., Varrot, A., Dauter, M., Brzozowski, A. M., Schulein, M., and Withers, S. G. (1998) Biochemistry 37, 15280–15287
14. Sulzenbacher, G., Schulein, M., and Davies, G. J. (1997) Biochemistry 36, 5902–5911
15. Davies, G. J., Ducros, V. M., Varrot, A., and Zechel, D. L. (2003) Biochem. Soc. Trans. 31, 525–527
16. Cizek, M., Cizek, M., Zamboni, V., Bevan, D. R., Henrissat, B., and Eser, A. (2003) Proc. Natl. Acad. Sci. U. S. A. 97, 13555–13560
17. Verdoucq, L., Cizek, M., Morinieres, J., Bevan, D. R., and Eser, A. (2003) J. Biol. Chem. 278, 25505–25506
18. Zouhar, J., Vevodova, J., Marek, J., Damborsky, J., Su, X. D., and Brzozobathy, B. (2001) Plant Physiol. 127, 973–985
19. Cizek, M., Blanchard, D. J., Bevan, D. R., and Eser, A. (2000) J. Biol. Chem. 275, 20082–20081
20. Cizek, M., and Eser, A. (1998) Plant Physiol. 116, 1469–1478
21. Emert, P., Vaseela, A., Weber, M., Rupitz, K., and Withers, S. G. (1993) Carbohydr. Res. 256, 113–128
22. Cizek, M., Cizek, M., Zamboni, V., Burmeister, W. P., Bevan, D. R., Henrissat, B., and Eser, A. (2001) Biochem. J. 354, 37–46
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
24. Ho, S. N., Hunt, H. D., Herton, R. M., Pullen, J. K., and Pease, L. R. (1989) Gene (Amst.) 77, 51–59
25. Owinozski, Z., and Minor, W. (1997) in Macromolecular Crystallography (Carter, C. W., and Sweet, R. M., eds) Part A, Vol. 276, pp. 307–326, Academic Press, Inc., New York
26. CCP4 (1994) Acta Crystallogr. Sect. D 50, 760–763
27. Navazza, J. (1994) Acta Crystallogr. Sect. A 50, 157–163
28. Nicholls, A., Sharp, K. A., and Hunig, B. (1991) Proteins 11, 281–296
29. Hoel, W., Tober, I., Eklund, S. H., and Conn, E. E. (1987) Arch. Biochem. Biophys. 252, 152–162
30. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Biochem. J. 321, 557–559
31. Tews, I., Pernakos, A., Oppenheim, A., Dauter, Z., Wilson, K. S., and Vorgias, C. E. (1996) Nat. Struct. Biol. 3, 638–648
32. Vasella, A., Davies, G. J., and Bohn, M. (2002) Curr. Opin. Struct. Biol. 6, 619–629
33. Burmeister, W. P., Cottaz, S., Rollin, P., Vasella, A., and Henrissat, B. (2000) J. Biol. Chem. 275, 39385–39393
34. Nambuchuk, M. N., and Withers, S. G. (1995) Biochemistry 34, 16194–16202
35. Norinbooom, V., Birsan, C., Nitz, M., Rose, D. R., Warren, R. A., and Withers, S. G. (1998) Nat. Struct. Biol. 5, 812–818
36. Varrot, A., and Davies, G. J. (2003) Acta Crystallogr. Sect. D 59, 447–452
37. Ducros, V. M., Zechel, D. L., Murshudov, G. N., Gilbert, H. J., Stabb, L., Stoll, D., Withers, S. G., and Davies, G. J. (2002) Angew. Chem. Int. Ed. Engl. 41, 2824–2827
38. Sulzenbacher, G., Driguez, H., Henrissat, B., Schulein, M., and Davies, G. J. (1996) Biochemistry 35, 15280–15287
39. Vocadlo, D. J., Davies, G. J., Laine, R., and Withers, S. G. (2001) J. Biol. Chem. 276, 2824–2827
40. Zechel, D., and Withers, S. G. (1999) in Comprehensive Natural Products in Chemistry (Barton, D., Nakaniishi, K., and Polutter, C. D., eds) Vol. 5, pp. 279–314, Elsevier Science Publishers B.V., Amsterdam
41. Zechel, D., and Withers, S. G. (2000) Acc. Chem. Res. 33, 11–18
42. Coutinho, P. M., Stam, M., Blane, E., and Henrissat, B. (2003) Trends Plant Sci. 8, 563–565
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

Lionel Verdoucq, Jeanne Morinière, David R. Bevan, Asim Esen, Andrea Vasella, Bernard Henrissat and Mirjam Czjze

J. Biol. Chem. 2004, 279:31796-31803.
doi: 10.1074/jbc.M402918200 originally published online May 17, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M402918200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/05/26/M402918200.DC1

This article cites 39 references, 11 of which can be accessed free at
http://www.jbc.org/content/279/30/31796.full.html#ref-list-1