Functional analysis of the aglycone-binding site of the maize β-glucosidase Zm-p60.1

Radka Dopitová, Pavel Mazura, Lubomír Janda, Radka Chaloupkova, Petr Jeřábek, Jiří Damborský, Tomáš Filipi, Nagavalli S. Kiran and Betislav Brzobohaty

1 Institute of Biophysics AS CR, v.v.i., Brno, Czech Republic
2 Department of Functional Genomics and Proteomics, Masaryk University, Brno, Czech Republic
3 Department of Molecular Biology and Radiobiology, Mendel University of Agriculture and Forestry, Brno, Czech Republic
4 Loschmidt Laboratories, Institute of Experimental Biology and National Centre for Biomolecular Research, Masaryk University, Brno, Czech Republic

Glycoside hydrolases (GH; EC 3.2.1) catalyze the selective hydrolysis of glycosidic bonds within polysaccharides and polycarbohydrates or between carbohydrates and non-carbohydrate moieties. Based on amino acid sequence similarities, GHs are currently classified into 112 families, as described in the CAZy database (http://www.cazy.org) [1]. β-Glucosidases are found in families GH1, GH3 and GH9.

β-Glucosidases such as Zm-p60.1 (Zea mays) and Bgl4:1 (Brassica napus) have implicated roles in regulating plant development by releasing biologically active cytokinins from O-glucosides. A key determinant of substrate specificity in Zm-p60.1 is the F193–F200–W373–F461 cluster. However, despite sharing the same substrates, amino acids in the active sites of Zm-p60.1 and Bgl4:1 differ dramatically. In members of the Brassicaceae we found a group of β-glucosidases sharing both high similarity to Bgl4:1 and a consensus motif A-K-K-L corresponding to the F193–F200–W373–F461 cluster. To study the mechanism of substrate specificity further, we generated and analyzed four single (F193A, F200K, W373K and F461L) and one quadruple (F193A–F200K–W373K–F461L) mutants of Zm-p60.1. The F193A mutant showed a specific increase in affinity for a small polar aglycone, and a deep decrease in $k_{cat}$ compared with the wild-type. Formation of a cavity with decreased hydrophobicity, and significant consequent alterations in ratios of reactive and non-reactive complexes, revealed by computer modeling, may explain the observed changes in kinetic parameters of the F193 mutant. The large decrease in $k_{cat}$ for the W373K mutant was unexpected, but the findings are consistent with the F193–aglycone–W373 interaction playing a dual role in the enzyme’s catalytic action; influencing both substrate specificity, and the catalytic rate by fixing the glucosidic bond in a favorable orientation for attack by the catalytic pair. Investigation of the combined effects of all of the mutations in the quadruple mutant of Zm-p60.1 was precluded by extensive alterations in its structure and almost complete abolition of its enzymatic activity.
glucosides with only carbohydrate moieties. There is considerable interest in plant β-glucosidases, because they are involved in diverse biological processes, ranging from developmental regulation, for example, activation of the plant hormones cytokinins [3] and abscisic acid [4], through cell wall degradation in the endosperm during germination [5], to pathogen defense reactions [6].

Three-dimensional structures of GH1 β-glucosidases from 19 species have been reported, seven of which are plant β-glucosidases (http://www.cazy.org). Although levels of sequence identity vary between 17% and 45% in the GH1 β-glucosidases, their structures have proved to be highly similar. The overall fold of the enzymes is a single domain (β/α)6 barrel which classifies them as members of clan GH-A of related GH families [7]. GH1 β-glucosidases are retaining in that the anomeric configuration of the glucose is the same in the product (β-D-glucose) as it is in the substrate (a β-D-glucosides). Substrate hydrolysis requires the participation of two glutamic acid residues (designated the catalytic pair) within highly conserved TXNEX and ITENG motifs, which reside in the loop regions at the C-terminal ends of β-strands 4 and 7, respectively [8].

Given the tremendous diversity of aglycone moieties in natural glucosides (which reflects their numerous biological functions) the fine-tuning of diverse biological processes in plants must depend (inter alia) on a number of β-glucosidases having high degrees of specificity towards their respective substrate aglycones. However, despite the substantial progress that has been made towards elucidating the mechanism of glucosidic bond cleavage and the roles of the catalytic pair, our knowledge of the molecular determinants of aglycone specificity in β-glucosidases remains limited. Elucidation of the aglycone specificity of β-glucosidases is a key prerequisite for understanding their precise role in biological processes in which glucosylation and de-glucosylslation steps are regulatory elements. In addition, the ability to modulate the specificity of β-glucosidases that would follow its elucidation could have valuable biotechnological applications.

A maize β-glucosidase, Zm-p60.1, a member of the GH1 family, has been shown to release active cytokinins from their O- and N3-glucosides, and thus has implicated roles in the regulation of maize seedling development [3]. The enzyme has been located in plastids [9], and its accumulation in chloroplasts and plastids of transgenic tobacco has been shown to perturb the cytokinin metabolic network [10]. In addition, an allozyme of Zm-p60.1, Zm-Glu1, has been shown to hydrolyze 4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA)-β-D-glucopyranoside (DIM-BOA-β-D-Glc) [11] in a manner similar to a β-glucosidase purified from maize seedlings [12], and has been implicated in defense against pathogens by releasing the toxic aglycone (DIMBOA) from its storage form, DIMBOA-β-D-Glc. However, no direct experimental evidence confirming that Zm-Glu1 is involved in defense responses in planta has been published.

Three-dimensional structures have been obtained for Zm-p60.1 [13], Zm-Glu1 and its complex with the non-hydrolyzable inhibitor p-nitrophenyl β-D-thiogluco-pyranoside [14], and co-crystals of an inactive mutant of Zm-Glu1 and DIMBOA-β-D-Glc [15]. Analysis of these structures has provided indications that the enzymes’ specificity toward substrates with aryl aglycones is conferred by the aromatic aglycone system stacking with W373, and van der Waals interactions with edges of F193, F200, and F461 located opposite W373 in a slot-like aglycone-binding site [13,15]. In addition, kinetic analysis and computer simulations of F193L/Y/W mutants have demonstrated that F193-aglycone-W373 interactions not only contribute to aglycone interactions, but also codetermine the catalytic rate by fixing the glucosidic bond in an orientation favorable for attack by the catalytic pair [13].

A distinctly different member of the GH1 family – a β-glucosidase hydrolyzing a cytokinin-O-glucoside – has been found in Brassica napus and designated Bgl4:1 [16]. Bgl4:1 and Zm-p60.1 display 44% identity at the amino acid sequence level. However, when we inspected the Bgl4:1 sequence, we found no hydrophobic cluster corresponding to the F193–F200–W373–F461 cluster of Zm-p60.1. Analysis of these two distinct β-glucosidases, which appear to have very similar tertiary structures and substrate specificity, but differ dramatically in the architecture of their aglycone-binding sites, offers exciting prospects for identifying molecular determinants of substrate specificity in β-glucosidases. Structurally, the aglycone-binding sites of Zm-p60.1 from Zea mays and Bgl4:1 from B. napus represent two extreme cases in their protein family.

Here, we report a consensus motif found in Bgl4:1 and evolutionarily closely related β-glucosidases of the GH1 family of the Brassicaceae that corresponds to the F193–F200–W373–F461 cluster of Zm-p60.1. We also report the construction of four single mutants and one quadruple mutant introducing features of the consensus motif into the Zm-p60.1 scaffold, an analysis of structural and catalytic properties of the mutants, and simulations of the substrate–enzyme interactions of the wild-type and one of the mutants. The results provide indications of the native enzymes’ catalytic action and determinants of
Results

Design and construction of the mutant \(\beta\)-glucosidases

Findings that cytokinin-\(O\)-glucosides are natural substrates for both of the two \(\beta\)-glucosidases, Zm-p60.1 and Bgl4:1, but the architecture of their sites that recognize the aglycone moieties of these substances differs distinctly, prompted us to initiate a bioinformatic analysis of plant \(\beta\)-glucosidases to obtain insights into the evolution of the molecular sites involved in the two modes of aglycone binding.

The amino acid sequence of Zm-p60.1 was compared with the sequences of 22 other members of the GH1 family from 13 plant genera. The resulting alignment was manually adjusted (Fig. S1) and a phylogenetic tree was inferred (Fig. 1). Interestingly, we found four \(\beta\)-glucosidases closely related to Bgl4:1, all of which belong to the Brassicaceae, forming a separate five-member group. Furthermore, using castp software, we identified 37 amino acid residues forming an active site cavity including the residues that make contact with glucose, an aglycone or both during interactions with their substrates, based on data obtained from the Protein Ligand database (Table S1).

Information obtained using the two approaches allowed us to determine the relative level of variability in amino acid composition at the selected positions corresponding to the amino acid residues forming the active site (Fig. 2). In accordance with previous studies, a higher degree of conservation was found among amino acid residues that contact a sugar, including the fully conserved amino acid residues Q33, H137, N185, Y328, W452, E459, W460, and the catalytic pair E186 and E401. By contrast, a high degree of variability was found in amino acid residues that contact an aglycone; only 5 of 17 such amino acid residues were fully conserved. In accordance with their proposed role in aglycone specificity, F193 and F461 are among the most variable amino acid residues of the active center, and both F200 and W373 are also quite variable (showing almost half as much variability as F193 and F461). In the Brassicaceae group related to Bgl4:1, a consensus motif A-K-K-L was identified, corresponding to the F193–F200–W373–F461 cluster involved in enzyme specificity towards aglycones in Zm-p60.1. Interestingly, both lysine residues and the leucine residue are conserved in all five enzymes of the group, and the alanine residue is found in all but one of the enzymes, namely Bgl4:1, where the same position is occupied by a serine residue (Fig. S1). The results define a novel architecture involved in the molecular recognition of aromatic aglycones in the Brassicaceae group of \(\beta\)-glucosidases. To allow more instructive structural comparisons, amino acid residues of the A-K-K-L consensus motif were modeled into the corresponding positions of the F193–F200–W373–F461 cluster in the Zm-p60.1 aglycone-binding site (Fig. 3).
positions were calculated using the scoring function in 
SWISS-PDBVIEWER v. 3.7, and the results were visualized 
with PYMOL v. 0.97 [17,18].

To initiate a functional comparison of the two distinct 
architectures of the aglycone-binding site, site-
directed mutagenesis was employed to generate four 
single (F193A, F200K, W373K and F461L) mutants 
and one quadruple (F193A–F200K–W373K–F461L) 
mutant introducing features of the A-K-K-L consensus 
into the Zm-p60.1 scaffold.

**Secondary structure and dimer assembly of the 
mutant enzymes**

The wild-type and mutant enzymes were expressed in 
*Escherichia coli* BL21(DE3)pLysS and purified close to 
homogeneity as follows. The first step was metal chelate 
affinity chromatography, following a previously 
described protocol [19]. This purified the wild-type and 
single mutants to levels exceeding 85% according to 
densitometric analysis of Coomassie Brilliant Blue R250-stained SDS/PAGE gels (not shown), but failed 
to yield the quadruple (F193A–F200K–W373K–F461L) mutant, designated P2, in > 30% purity, indicating that the accessibility of the His tag is signifi-
cantly altered in P2. Subsequent ammonium sulfate 
precipitation followed by hydrophobic chromatogra-
phy resulted in preparations of P2, as well as the wild-
type and single mutants, with > 94% purity (Fig. S2).

CD spectroscopy was used to assess the relative proportions of secondary structural elements in the 
wild-type and mutant enzymes (using DICROPRT v. 1.0, 
see Fig. 4) and the thermal stability of the mutant 
enzymes. The predictions obtained for the wild-type 
enzyme coincided well with estimates obtained from a 
crystal structure, indicating that they were highly 
reliable [13] (Fig. 4). The relative proportions of α heli-
ces and β sheets in F193A and W373K appear to be 
identical to those in the wild-type, whereas the propor-
tions of α helices appear to be lower in F461L, F200K 
and P2. Furthermore, the F193A and W373K muta-
tions do not result in any change in the thermostability 
of the enzyme (Table S2), and thermal unfolding of the 
wild-type and both the F193A and W373K mutants 
was found to be irreversible (Fig. S3).

The propensity of the wild-type and each of the 
single mutant enzymes to form dimers was analyzed by 
size-exclusion chromatography. The enzymes were 
purified by metal chelate affinity chromatography and 
subjected to size-exclusion chromatography using a 
HighLoad 16/60 Superdex 200 column. The enzymes 
eluted in two peaks, d and m, corresponding to appar-
ent molecular masses of ~110 and ~43 kDa, respec-
tively, (Fig. 5A,B and Table S3). The apparent 
molecular mass of ~110 kDa is in good agreement 
with the 118 kDa calculated for the dimeric forms of 
the enzymes based on their amino acid composition. 
Furthermore, wild-type Zm-p60.1 was found in dimeric 
form in its crystal structure [13]. The E401D mutant of 
Zm-p60.1, which is defective in dimer assembly, [13] 
was used to show that the peak m corresponds to the 
monomeric forms of the enzymes each of which has a 
calculated molecular mass of 59 kDa (based on amino 
acid composition) – consistent with the 60 kDa deter-
mined from the SDS/PAGE analysis (Fig. 5A,B and 
Table S3). Low molecular mass polypeptides found in 
peak m in Coomassie Brilliant Blue-stained SDS/ 
PAGE gels (Fig. 5B) were not detected by either anti-
(Zm-p60) or anti-(His-tag) serum in western blots (not 
shown), suggesting that they represent contaminants of 
the monomer fraction by low molecular mass proteins. 
Based on the same criteria, a ~66 kDa polypeptide
found in peak d represents a minor contaminant of the dimeric form of the enzymes. Whereas the wild-type, F193A and F461L mutant enzymes were found almost exclusively in the form of dimers, the F200K and W373 mutations apparently hindered dimer assembly.

Dimeric and monomeric forms of the enzymes were resolved by native PAGE, and enzymatic activity was found to be associated exclusively with the dimeric forms by in-gel activity staining (Fig. 5C,D), as previously found for the wild-type and a number of mutant enzymes [13,20].

Kinetics of the mutant enzymes

Two general β-glucosidase substrates differing in polarity and the size of their aromatic aglycones, pNPGlc and 4-methylumbelliferyl β-D-glucopyranoside (4MUGlc), were used to evaluate the effects of the mutations on the enzymes’ kinetics (Table 1). F461L increased the enzyme’s relative catalytic efficiency, defined as $(k_{cat}/K_m)_{mutant}/(k_{cat}/K_m)_{WT}$ by 20% compared with the wild-type for both substrates, by increasing $k_{cat}$. By contrast, the F193A, F200K and W373K single mutations had dramatic negative effects on catalytic efficiency. The F193A substitution reduced the enzyme’s efficiency via 195- and 42-fold reductions in $k_{cat}$ values for pNPGlc and 4MUGlc, respectively. Interestingly, this substitution also highly increased the enzyme’s affinity for pNPGlc; reducing the $K_m$ for this substrate > 15-fold and the $K_m$ for 4MUGlc by only ~20%. The F200K mutation resulted in 5- and 10-fold increases in $K_m$, with 18- and 29-fold reductions in $k_{cat}$ for pNPGlc and 4MUGlc, respectively. The W373K mutation caused similar reductions in affinity for the substrates; 3- and 12-fold increases in
Km for pNPGlc and 4MUGlc, respectively. However, these changes were accompanied by 68- and 243-fold reductions in kcat for pNPGlc and 4MUGlc, respectively, indicating that substrate turnover was hampered to a much higher extent by the W373 mutation. In general, reductions in the relative efficiency of F193A, F200K and W373K mutants were more pronounced with 4MUGlc as the substrate, and the W373K mutant showed the lowest efficiency with both substrates.

Molecular modeling of enzyme–substrate complexes for wild-type and F193A enzymes

Wild-type and F193A mutant enzyme–substrate complexes were explored by molecular modeling to obtain insights into the molecular interactions underlying the observed changes in the mutants’ enzymatic kinetics. Modeling was only applicable to F193A because interpretation of acquired data requires preservation of the overall tertiary structure in the modeled proteins. W373K also has an indistinguishable structure from the wild-type, according to the CD spectral analysis. However, this mutant could adopt a high number of possible conformations at the W373 position, precluding robust interpretation of any results obtained by molecular modeling with current methods. Furthermore, assembly of W373K mutant homodimers is hindered, indicating that there are alterations in its conformation that are not amenable to CD spectroscopy.

Table 1. Steady-state kinetic parameters for hydrolysis of pNPGlc and 4MUGlc by mutant and wild-type Zm-p60.1 β-glucosidases. Assays were performed using substrates at a minimum of seven concentrations and the parameters were calculated using ORIGIN PRO 7.5 software. Relative efficiency: (kcat/Km)mutant/(kcat/Km)WT × 100.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Km</th>
<th>kcat</th>
<th>kcat/Km</th>
<th>Relative efficiency</th>
<th>Km</th>
<th>kcat</th>
<th>kcat/Km</th>
<th>Relative efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>0.68 ± 0.03</td>
<td>42.80 ± 0.56</td>
<td>62.94 ± 2.89</td>
<td>100.00</td>
<td>0.148 ± 0.013</td>
<td>53.60 ± 1.09</td>
<td>362.16 ± 32.59</td>
<td>100.00</td>
</tr>
<tr>
<td>F193A</td>
<td>0.045 ± 0.0035</td>
<td>0.22 ± 0.003</td>
<td>4.89 ± 0.39</td>
<td>7.77</td>
<td>0.120 ± 0.012</td>
<td>1.29 ± 0.04</td>
<td>10.75 ± 1.13</td>
<td>2.97</td>
</tr>
<tr>
<td>F200K</td>
<td>3.50 ± 0.22</td>
<td>2.43 ± 0.05</td>
<td>0.69 ± 0.046</td>
<td>1.09</td>
<td>1.510 ± 0.101</td>
<td>1.87 ± 0.05</td>
<td>1.24 ± 0.089</td>
<td>0.34</td>
</tr>
<tr>
<td>W373K</td>
<td>2.10 ± 0.21</td>
<td>0.63 ± 0.02</td>
<td>0.30 ± 0.032</td>
<td>0.48</td>
<td>1.736 ± 0.125</td>
<td>0.22 ± 0.01</td>
<td>0.13 ± 0.011</td>
<td>0.04</td>
</tr>
<tr>
<td>F461L</td>
<td>0.65 ± 0.05</td>
<td>49.27 ± 1.15</td>
<td>75.80 ± 6.09</td>
<td>120.43</td>
<td>0.164 ± 0.019</td>
<td>70.88 ± 2.16</td>
<td>432.19 ± 51.75</td>
<td>119.34</td>
</tr>
</tbody>
</table>

Km for pNPGlc and 4MUGlc, respectively. However, these changes were accompanied by 68- and 243-fold reductions in kcat for pNPGlc and 4MUGlc, respectively, indicating that substrate turnover was hampered to a much higher extent by the W373 mutation. In general, reductions in the relative efficiency of F193A, F200K and W373K mutants were more pronounced with 4MUGlc as the substrate, and the W373K mutant showed the lowest efficiency with both substrates.

Fig. 5. Quaternary structure of wild-type and mutant Zm-p60.1 β-glucosidases. (A) Elution profiles of wild-type and mutant Zm-p60.1 β-glucosidases from the HighLoad 16/60 Superdex 200 column. A sample (1.5 mL) of each enzyme purified by metal chelate affinity chromatography was applied to the column and eluted with elution buffer (50 mM Tris/HCl, 500 mM NaCl; pH 7.00). Fractions corresponding to peaks d and m were collected and analyzed by (B) Coomassie Brilliant Blue-stained SDS/PAGE, (C) Coomassie Brilliant Blue-stained native-PAGE and (D) in-gel activity staining of native-PAGE gels. Peaks 1, 2, 3, 4 and 5 correspond to Blue Dextran 2000, ferritin (Mr 440 kDa), aldolase (Mr 158 kDa), BSA (Mr 67 kDa) and ovalbumin (Mr 43 kDa), respectively, used as standards. Arrow marks positions of the wild-type and mutant Zm-p60.1 polypeptides in SDS/PAGE.

Table 1. Steady-state kinetic parameters for hydrolysis of pNPGlc and 4MUGlc by mutant and wild-type Zm-p60.1 β-glucosidases. Assays were performed using substrates at a minimum of seven concentrations and the parameters were calculated using ORIGIN PRO 7.5 software. Relative efficiency: (kcat/Km)mutant/(kcat/Km)WT × 100.
The structures of enzyme–substrate complexes were obtained for both the wild-type and F193A enzymes by docking the substrate molecules 4MU\(\text{Glc}\) and \(\rho\text{NPGlc}\) into their active sites. The structures obtained from the docking were divided into reactive and non-reactive complexes, depending on the orientation of the sugar moiety (Fig. 6A,B), and the results from 50 dockings for each complex are summarized in Tables S4 and S5. In each case the most highly populated binding mode was a reactive complex. However, the number of non-reactive clusters and the proportion of lightly populated reactive clusters were higher for F193A than for the wild-type enzyme, and non-reactive binding generally appears to be energetically preferred in the F193A mutant. The most highly populated binding modes from the docking were selected for further optimization, but this did not result in significant repositioning of the substrate molecule inside the enzyme active site. Reactive enzyme–substrate complexes of the wild-type enzyme and F193A mutant are geometrically similar (Fig. 6C–F), showing no significant differences in the distances of reacting atoms. The only noted difference was in the orientation of the aromatic ring of \(\rho\text{NPGlc}\) in the F193A mutant (Fig. 6D), owing to lost van der Waals contact with the side-chain of the substituted phenylalanine residue. However, the overall orientation of the aglycone moiety remains the same for both proteins because of the strong stacking interaction with W373.

**Discussion**

We identified a group of \(\beta\)-glucosidases in members of the Brassicaceae that are closely related evolutionarily to Bgl4:1, a \(\beta\)-glucosidase of *B. napus* that cleaves cytokinin-\(O\)-glucosides, thus sharing natural substrates with maize \(\beta\)-glucosidase Zm-p60.1. Despite also having the same overall fold, a \((\beta/\alpha)_8\) barrel, and levels of amino acid sequence similarity ranging from 45% to 53%, the architecture of the aglycone-binding site of Zm-p60.1 differs distinctly from that of Bgl4:1 and its homologs. These findings offer exciting prospects for comparative analysis of the molecular determinants of substrate specificity in the GH1 family of \(\beta\)-glucosidases. Sequence
comparisons of the Brassicaceae group identified a consensus motif, A-K-K-L, corresponding to the F193–F200–W373–F461 cluster of Zm-p60.1 that is involved in its interactions with aglycones. Therefore, we constructed four single (F193A, F200K, F461L and W373K) mutants and one quadruple (F193A–F200K–W373K–F461L) mutant introducing features of the consensus motif into the Zm-p60.1 scaffold, then subjected the mutant and wild-type enzymes to structural, kinetic and molecular modeling analyses to seek insights into the catalytic action of the β-glucosidases.

Kinetic analysis of the F193A mutant indicated that its $K_m$ for $p$NPGlc was greatly reduced (15-fold), whereas its $K_m$ for 4MU Glc was practically unaltered compared with the wild-type, and thus that the mutation caused a substantial selective increase in its affinity for $p$NP Glc (Table 1). Its $k_{cat}$ values decreased for both substrates, but the decrease was more pronounced for 4MU Glc (Table 1). The apparently unaltered structure of the F193A mutant compared with the wild-type, according to CD spectral analysis (Fig. 4), allowed us to interpret the kinetic parameters using molecular modeling of enzyme–substrate complexes. Molecular docking did not indicate any significant differences in the geometry of the most highly populated energetically favorable reactive enzyme–substrate complexes of the wild-type and F193A enzymes that could be responsible for the determined differences in their kinetic parameters. However, the proportions of non-reactive clusters and lightly populated reactive clusters were significantly higher for the F193A mutant than for the wild-type. Such changes are expected to lead to reductions in $k_{cat}$ because of mis-positioning of the glucosidic bond in higher fractions of lightly populated reactive enzyme–substrate complexes and increases in enzyme occupation in non-reactive enzyme–substrate conformations. The decrease in the F193 mutant’s $K_m$ for $p$NP Glc, compared with the wild-type, is likely to reflect the higher frequency of energetically preferred, non-reactive complexes it apparently forms. Furthermore, the F193A substitution widens the slot between amino acid residues at positions 193 and 373, and reduces its hydrophobicity, which may allow substrates with small polar aromatic aglycones, for example, $p$NP Glc, to enter the active site without removal of a water hydration shell, saving energy otherwise needed for its dehydration, and thus preferentially increasing the enzyme’s affinity for these substrates. The data are consistent with our previous results indicating that F193–aglycone–W373 interactions not only contribute to aglycone recognition, but also codetermine catalytic rates by fixing the glucosidic bond in a favorable orientation for attack by the catalytic pair [13]. A dramatic reduction in enzyme activity was observed in the F193V mutant, but this was likely because of an unexpected rearrangement in three other amino acid residues that are also involved in the substrate binding site according to previous structural analysis [21].

The W373K mutant exhibited the most pronounced reductions in relative efficiency for both substrates analyzed. Unexpectedly, the dramatic decrease in $W373K$’s specificity constant is caused mainly by a decrease in its $k_{cat}$. Based on enzyme structure analysis and molecular docking, W373 stacking interactions with the aglycone aromatic system and van der Waals interactions with the edges of the phenyl rings provided by F193, F200 and F466 appear to be the major determinants of aglycone recognition and specificity in Zm-p60.1 [13–15]. Thus, the dramatic reductions in $k_{cat}$ conferred by the W373K mutation indicate a previously unrecognized function of W373 in the determination of the catalytic rate of the enzyme, albeit one that is consistent with the involvement of F193–aglycone–W373 interactions in both substrate affinity and determination of the catalytic rate inferred from previous analyses of the F193I mutant [13].

Recent crystal structure determination and subsequent homology modeling revealed that hydrophobic interactions are the major contributors to the binding of aglycone moieties to a human cytosolic β-glucosidase (hCBG) [22]. Structural superimposition showed that W345 of hCBG has a similar conformation to W373 of Zm-p60.1, lining the aglycone-binding site in a way that enables stacking interactions with an aromatic aglycone. Dramatic reductions in the specificity constants for a number of glycosides were found in kinetic analyses of W345 mutants. Similar to our results, these reductions in specificity constants were because of reductions in $k_{cat}$, whereas $K_m$ values increased much less, and even decreased for several β-glucosides, including three of five natural substrates tested. Investigation of hCBG’s 3D structure showed that the amine group of the W345 indole ring is located close (~3.9 Å) to the O6 of the sugar. This finding led to a proposal that W345 may be a key residue ensuring that the glucosidic bond is positioned in a favorable orientation for attack by the catalytic pair by a combination of aromatic stacking with the aromatic aglycone and hydrogen bonding to the sugar moiety of the substrate [22]. However, our inspection of the structures of ZM-Glu1 and its catalytically inactive mutant in co-crystals with the non-hydrolysable substrate $p$-nitrophenyl β-d-thioglucoside, the competitive inhibitor dhurrin and the substrate DIMBOA-β-d-Glc indicated that the corresponding distances are ~5.3, 4.8 and
Accumulation of the four mutations in a single molecule of the quadruple P2 mutant resulted in the polypeptide chain folding into a distinct structure characterized by an inversed ratio of $\alpha$ helices and $\beta$ strands compared with the wild-type (Fig. 4). In addition, the electrophoretic mobility of the P2 mutant in native PAGE is slower than the wild-type, and it forms dimers to a low, albeit detectable, extent (not shown). Furthermore, its enzymatic activity decreased dramatically, precluding determination of kinetic parameters. This indicates that, in future work, sequence analysis should be focused on other parts of the sequences (outside the four-residue signature) in order to explain the eventual effects of the mutations.

Conclusion

In conclusion, this study corroborates and extends previous knowledge of the dual role of F193–aglycone–W373 interactions in the catalytic action of the Zm-p60.1 $\beta$-glucosidase; contributing both to the enzyme’s affinity for substrates with aromatic aglycones and codetermination of the catalytic rate by fixing the glucosidic bond in a favorable orientation for attack by the catalytic pair. Furthermore, our computer modeling of the wild-type and F193A enzymes’ interactions with two substrates provides indications of the mechanisms involved in these roles, inter alia that the F193A mutation leads to the formation of a cavity with decreased hydrophobicity, and significant consequent alterations in ratios of reactive and non-reactive complexes. Wider exploration by computer modeling was precluded by unexpected structural alterations. These are mirrored in the most extreme case of the quadruple mutant in almost complete abolishment of enzyme activity, which also excluded investigation of the effects of accumulation of the mutations in a single protein molecule.

Experimental procedures

Structural analysis

The structural analysis of Zm-p60.1 was based on X-ray data presented previously [13,24]. Its active site was determined using the CASTp server [25], and the amino acid residues within the frame shaping the active site making calculated contacts with the tested ligands were identified using data in the Ligand Protein Contacts database [26].

Sequence analysis and phylogenetics

Protein sequences were selected for alignment that met several criteria, notably apparently robust characterization

7.8 Å, respectively [14,15,23]; clearly too long to allow formation of a hydrogen bond, for which a distance of $\sim 3$ Å is required. Taken together, the results obtained regarding $\beta$-glucosidases from organisms as distantly related as maize and humans performing distinct functions clearly indicate that the role of the tryptophan residue in the position equivalent to W373 in the enzyme’s catalytic action is more complex than anticipated in previous studies [13–15] in that it appears to influence the catalytic rate more than substrate binding parameters.

The F200K substitution resulted in the second most severe reductions in specificity constants of all the single-point mutations analyzed (Table 1). Interpretation of these reductions in kinetic parameters in molecular terms is precluded by a significant structural alteration deduced from the results of CD spectroscopy (Fig. 4). The high degree of structural alteration might indicate an involvement of F200 in folding of Zm-p60.1. Interestingly, an F200L mutation was shown to cause an increase in the specificity constant for $p$NPGlc, although it remained practically unaltered for o-nitrophenyl $\beta$-D-glucoside and 4MUGlc. However, the structure of this mutant was not investigated [21].

The specificity constants of the F461L mutant were increased by $\sim 20\%$ for both substrates compared with the wild-type (Table 1). As for the F200K mutant, the F461L mutation also resulted in altered proportions of secondary structural elements, precluding interpretation of the changes in molecular terms, although the core of its ($\beta/\alpha)_8$ barrel might have remained unaltered because the changes were because of a reduction in its content of $\alpha$ helices, whereas its $\beta$-sheet content remained unchanged (Fig. 4). A positive effect of a F461S mutation on specificity constants for all investigated artificial substrates has been previously reported [21], but the effect of this mutation on enzyme structure was not determined in the cited study. Interestingly, however, the increases were mainly because of increases in turnover number, although the affinity for $p$NPGlc and 4MUGlc decreased about twofold. Furthermore, the F461S mutant gained low but detectable enzymatic activity towards dhurrin, a natural substrate of a related $\beta$-glucosidase (SbDhr1) and a competitive inhibitor of Zm-p60.1. These findings indicate that variations in the amino acid residue at position 461 may have stronger effects on $k_{cat}$ than on $K_m$, and thus significant effects on the enzyme’s specificity towards natural substrates.

Interestingly, all the mutations except F461L had more severe effects on the enzyme’s interactions with 4MUGlc than with $p$NPGlc, thus apparently shifting its specificity slightly towards substrates with small, polar, aromatic aglycones.
Expression, purification and size-exclusion chromatography of the wild-type and mutant enzymes

To express wild-type and mutant enzymes in E. coli strain BL21(DE3)pLysS (Novagen, Darmstadt, Germany), a previously described procedure [19] was followed. Cells were cultured in Luria–Bertani medium supplemented with ampicillin (100 μg·mL⁻¹), chloramphenicol (50 μg·mL⁻¹), 0.1% glucose and 5 mM Na₂HPO₄ pH 7 at 37 °C to an A₆₀₀ of 0.5–0.6. Recombinant protein expression was then induced by adding 0.1 mM isopropyl-1-thio-β-D-galactoside and 3 mM cellobiose. Three hours after induction at 22 °C, cells were harvested by centrifugation at 3500 g for 10 min at 4 °C. The cell pellets obtained from 500 mL portions of culture were each resuspended in 6 mL of extraction buffer containing 20 mM phosphate buffer (pH 7.9), 0.5 M NaCl, 0.1% Triton X-100 and stored at −20 °C. After thawing, the cells were broken by sonication using a Sonoplus GM7035 W (Bandelin, Berni, Germany) with 3 × 60 s pulses, on ice. The cell lysate was then centrifuged at 47 446 g for 30 min at 4 °C to remove insoluble cell debris. The protein-containing supernatant was applied to an Ni Sepharose high performance column (GE Healthcare, Chalfont St Giles, UK) equilibrated with buffer A (20 mM Na₂HPO₄ pH 7.9, 0.5 M NaCl). The ballast proteins were washed out from the column with 15 column volumes of buffer B (50 mM Na₂HPO₄ pH 7.9, 1 M NaCl, 20 mM imidazole) and 15 column volumes of buffer C (50 mM Na₂HPO₄ pH 7.9, 1 M NaCl, 50 mM imidazole). (His)₉Zmp-60.r was eluted in buffer D (20 mM Na₂HPO₄ pH 7.9, 1 M NaCl, 20% glycerol, 100 mM EDTA). Ammonium sulfate (pH 7) was added to eluted fractions to a final concentration of 1.0 M and the resulting solutions were centrifuged at 16 500 g for 15 min. The supernatants were applied to a HiTrap Phenyl-HP column (GE Healthcare) and the proteins were purified using a linear gradient of 0.8–0.2 M (NH₄)₂SO₄, pH 7.0. Flow-through fractions were pooled, desalted and concentrated using an Amicon Ultra-4 ultrafiltration cell with 10 kDa cut-off (Millipore, Bedford, MA, USA). The purity of the wild-type and mutant enzymes was determined by SDS/PAGE followed by Coomassie Brilliant Blue staining and densitometry using a GS800 densitometer and Quantity One 1-D software (Bio-Rad, Hercules, CA, USA).

To determine the degree of dimer assembly in the wild-type and mutant enzymes, the enzyme preparations obtained from the metal chelate affinity chromatography were concentrated using the Amicon Ultra-15 ultrafiltration cell with 30kDa cut-off (Millipore), and each retentate (1.5 mL) was applied to a HighLoad 16/60 Superdex 200 prep grade column (GE Healthcare Biophysics, Uppsala, Sweden) then eluted with elution buffer (50 mM Tris/HCl, 500 mM NaCl; pH 7.00) using ÄKTA FPLC system (GE Healthcare Biophysics). Ferritin (Mr 440 kDa), aldolase (Mr 158 kDa), bovine serum albumin (Mr 67 kDa) and ovalbumin (Mr 43 kDa) were used as molecular mass standards, and the void volume was determined using Blue Dextran 2000 (GE Healthcare Biophysics). Apparent molecular masses of eluting proteins were determined from a log Mr versus V/V₀ plot, where Vₑ represents an elution volume and V₀ a void volume. The content and purity of the enzymes in individual fractions were determined from Coomassie Brilliant Blue-stained SDS/PAGE gels (see above). Migration of the enzymes to positions corresponding to an apparent molecular mass of 60 kDa was confirmed by western blot and immunostaining. Proteins separated by SDS/PAGE were transferred to a poly(vinylidine difluoride) membrane (Immobilon P; Millipore, Bedford, MA, USA) by semidry western blotting [31]. Positions of (His)₉Zmp-60.r were then visualized by an alkaline phosphatase-mediated immunostaining procedure [32], using: (a) polyclonal anti-(Zmp-p.60) serum raised in rabbits.
against recombinant (His)_6Zm-p60.1 produced in E. coli and anti-rabbit IgG conjugated to alkaline phosphatase, supplied by Sigma (Deisenhofen, Germany); and (b) anti-polyhistidine mAbs raised in mouse (Sigma) against the polyhistidine [(His)_6] domain and goat anti-mouse IgG conjugated to alkaline phosphatase (Sigma).

Electrophoresis and in-gel activity staining

Wild-type and mutant proteins were separated from other proteins in their respective preparations by native PAGE using 10% (w/v) gels [33]. They were then subjected to in-gel activity staining (zymography) by incubating the gels for 30 min at 37 °C with 5-bromo-4-chloro-3-indolyl-β-D-glucopyranoside (Biosynth International Inc, Itasca, IL, USA) dissolved in N,N'-dimethylformamide and diluted to the final working concentration of 0.6 mM in McIlvaine citrate-phosphate buffer (pH 5.5, 50 mM), a procedure developed by Mazura and Filipi (unpublished results). Proteins were visualized by Coomassie Brilliant Blue staining.

CD spectra

CD spectra were recorded at room temperature using a Jasco J-810 spectrometer (Jasco, Tokyo, Japan), collecting data from 185 to 260 nm, at 100 nm min^{-1} with a 1 s response time and 2 nm bandwidth using a 0.1 cm quartz cuvette containing the wild-type and mutant enzymes. Each spectrum shown is the average of 10 individual scans corrected for absorbance by the buffer. Collected CD data were expressed in terms of mean residue ellipticity (θ_{MRE}) using the equation:

\[ \theta_{MRE} = \frac{\theta_{obs} M \times 100}{nc} \]

where \( \theta_{obs} \) is the observed ellipticity in degrees, \( M \) is the protein’s molecular mass, \( n \) is the number of residues, \( c \) is the cell path length, \( I \) is the protein concentration and the factor 100 converts the resulting value to deg dmol^{-1} cm^{-1}. The proteins’ contents of secondary structural elements were calculated from the spectra using Self Consistent [34], K2D [35] and CONTIN [36] methods implemented in the program DICOProt (http://dicoprot-phil.ibcp.fr).

Enzyme and protein assays

The enzymatic activities of the wild-type and mutant proteins were assayed using 4MUGlC and pNPGlc as fluorogenic and chromogenic substrates, respectively [12,20], and the kinetic constants were calculated using ORIGIN PRO 7.5 software (OriginLab Corp., Northampton, MA, USA). The concentrations of the proteins in the preparations were determined using the DC Protein Assay (Bio-Rad) with BSA as a calibration standard.

Molecular modeling

The structures of the substrate molecules pNPGlc and 4MUGlC were built in INSIGHTII v. 95 (Biosym/MSI, San Diego, CA, USA) and energy-minimized by the AM1 semi-empirical quantum mechanics method, using the keyword PRECISE for optimization. A model of the F193A mutant was constructed using the experimental structure of the β-glucosidase Zm-p60.1 obtained from the Protein Data Bank (PDB ID 1HXJ). The substitution was introduced to the structure using the program AUTODOCK v. 3.05 [37]. The grid maps (81 × 81 × 81 points with 0.25 Å grid spacing) were calculated using AUTOGRAF v. 3.06. Fifty dockings were performed for each substrate using a Lamarckian genetic algorithm [37] with a population size of 50 individuals, a maximum of 1.5 × 10^6 energy evaluations and 27 000 generations, an elitism value of 1, and mutation and cross-over rates of 0.02 and 0.5, respectively. Local searches were based on a pseudo Solis and Wets algorithm [38] with a maximum of 300 iterations per search. Final orientations from every docking were clustered with a clustering tolerance for the root-mean-square positional deviation of 0.5 Å. The most highly populated complexes obtained in the molecular dockings were further optimized using the quantum mechanic program MOPAC2002 (Fujitsu, Kawasaki, Japan). All protein residues were fixed during optimization except E186, F193, F200, W373, E401 and F461. Heavy atoms of the backbone were fixed in all residues to keep the overall geometry of the protein active site intact. MOPAC calculations were carried out using the AM1 Hamiltonian and the BFGS geometry optimization algorithm. Results from the calculations were analyzed using the program TRITON v. 3.0 (Masaryk University, Czech Republic).

Acknowledgements

This project was supported by grants GACR203/02/0865 from the Grant Agency of the Czech Republic, LC06034, 1M06030, LC06010, MSM0021622415 and MSM0021622412 from the Ministry of Education, Youth and Sports of the Czech Republic, and AV0Z50040507 and AV0Z50040702 from the Academy of Sciences of the Czech Republic.

References


Supporting information

The following supplementary material is available:

**Fig. S1.** Multiple sequence alignment of 23 β-glucosidases from 13 plant species.

**Fig. S2.** Purity of the wild-type and mutant Zm-p60.1 β-glucosidases used for CD spectroscopy and enzyme kinetics analysis.

**Fig. S3.** Thermostability of F193A and W373K mutant and wild-type Zm-p60.1 β-glucosidases.

**Table S1.** Active site amino acids of Zm-p60.1 β-glucosidases forming contacts with glycone (G), aglycone (A) or no part (-) of substrate molecule as selected from the Ligand-Protein Contacts database (http://bioportal.weizmann.ac.il/oca-bin/lpccsu).

**Table S2.** Melting temperatures ($T_m$) of the mutant and wild-type Zm-p60.1 β-glucosidases derived from thermal denaturation curves.

**Table S3.** Apparent molecular masses of dimeric and monomeric forms of the wild-type and mutant Zm-p60.1 β-glucosidases determined by size-exclusion chromatography.

**Table S4.** Populations of 4MUGlc docked to the active site of mutant and wild-type Zm-p60.1 β-glucosidases.

**Table S5.** Populations of pNPGlc docked to the active site of mutant and wild-type Zm-p60.1 β-glucosidases.

This supplementary material can be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.