the FEBS Journal

Carbohydrate binding sites in *Candida albicans* exo-β-1,3-glucanase and the role of the Phe-Phe 'clamp' at the active site entrance

Wayne M. Patrick[1,2,*], Yoshio Nakatani[1,*], Susan M. Cutfield[1], Miriam L. Sharpe[1], Rochelle J. Ramsay[3] and John F. Cutfield[1]

1 Department of Biochemistry, University of Otago, Dunedin, New Zealand
2 Institute of Natural Sciences, Massey University, Auckland, New Zealand
3 Institute of Molecular Biosciences, Massey University, Palmerston North, New Zealand

Keywords
aromatic entranceway/clamp; exoglucanase; glycoside hydrolase; ordered water molecules; protein–carbohydrate interaction; site-directed mutagenesis

Correspondence
J. F. Cutfield, Department of Biochemistry, University of Otago, PO Box 56, Dunedin 9054, New Zealand
Fax: +64 3 479 7866
Tel: +64 3 479 7836
E-mail: john.cutfield@otago.ac.nz

*These authors contributed equally to this work

Database
Structural data for Exg mutants F258I, F144Y/F258Y and E292Q, as well as F229A/E292S, are available in the Protein Data Bank under the accession numbers 2PF0, 3O6A, 2PC8 and 3N9K, respectively

(Received 17 August 2010, revised 1 September 2010, accepted 3 September 2010)

doi:10.1111/j.1742-4658.2010.07869.x

Candida albicans exo-β-1,3-glucanase (Exg; EC 3.2.1.58) is implicated in cell wall β-D-glucan remodelling through its glucosyl hydrolase and/or transglucosylase activities. A pair of antiparallel phenylalanyl residues (F144 and F258) flank the entrance to the active site pocket. Various Exg mutants were studied using steady-state kinetics and crystallography aiming to understand the roles played by these residues in positioning the β-1,3-D-glucan substrate. Mutations at the Phe-Phe entranceway demonstrated the requirement for double-sided CH/π interactions at the +1 subsite, and the necessity for phenylalanine rather than tyrosine or tryptophan. The Tyr-Tyr double mutations introduced ordered water molecules into the entranceway. A third Phe residue (F229) nearby was evaluated as a possible +2 subsite. The inactive double mutant E292S/F229A complexed with laminaritriose has provided the first picture of substrate binding to Exg and demonstrated how the Phe-Phe arrangement acts as a clamp at the +1 subsite. The terminal sugar at the −1 site showed displacement from the position of a monosaccharide analogue with interchange of water molecules and sugar hydroxyls. An unexpected additional glucose binding site, well removed from the active site, was revealed. This site may enable Exg to associate with the branched glucan structure of the *C. albicans* cell wall.

Introduction

The large and diverse catalogue of glycoside hydrolases, together with knowledge of their specificities, mechanisms and structures, provides a logical platform for engineering novel enzyme functions [1]. The general mechanistic features of retaining β-glycoside hydrolases are now reasonably well established, despite the wide variation in their tertiary structures. A double displacement reaction involving the formation of a glycosyl-enzyme intermediate and subsequent hydrolysis (or transglycosylation) most likely proceeds through

Abbreviations
DFG, 2-deoxy-2-fluoro-glucopyranoside; Exg, *Candida albicans* exo-β-1,3-glucanase; GH5, glycoside hydrolase family 5.

oxocarbenium ion-like transition states [2–4]. The catalytic nucleophile and acid/base groups are both usually carboxylates, most frequently glutamate, as shown by labelling, crystallographic and mutagenesis studies [5–7]. There is also good evidence to suggest that these enzymes induce a distorted boat conformation in the sugar ring at the −1 position [4,8,9]. However, even within a particular family of β-glycoside hydrolases, it is still not obvious how the substrate is initially recognized by the particular enzyme and then drawn into the active site. Clearly, a detailed understanding of the protein–carbohydrate interactions that determine specificity and modulate activity is required to guide engineering efforts.

The complexity of substrate recognition by glycoside hydrolases is exemplified in the structurally well-characterized glycoside hydrolase family 5 (GH5) [10], which mainly includes endo-β-1,4- and exo-β-1,3-glucanases, and which has been further divided into various sub-families [11,12]. All members of GH5 have structures that are variations on the $(\beta\alpha)_8$ barrel fold, with their active sites containing the same eight similarly disposed residues. Of these, five are able to interact with the −1 site of the substrate, whereas the other three assist in orienting the two catalytic glutamates [13]. This is achieved through extensive intramolecular and intermolecular hydrogen bonding involving planar amino acid side chains and sugar hydroxyl groups. In addition to these conserved interactions at the −1 site, aromatic side chains further from the catalytic centre interact with individual glucopyranoside units through stacking interactions, as seen in the structures of several cellulases complexed with oligosaccharides [13–15]. For example, five aromatic platforms (including three tryptophans) have been identified in the subfamily 1 enzymes and most of these have topological, although not residue-specific, equivalents in the endoglucanases of other subfamilies [15]. The stacking of aromatic residues against the hydrophobic faces of sugar rings is a recurring feature of protein–carbohydrate interactions [16] and is a critical element of substrate recognition and specificity. It involves a CH/π interaction in which the partial positive charges from the glycoside hydrogens interact with the electron-rich π cloud on the face of an aromatic side chain [17]. Although generally one-sided, such interactions can involve both sides of the sugar ring, as seen for example in the D-galactose-binding protein from *Escherichia coli* [18]. In this case, the pyranose is sandwiched between a Trp and a Phe, as well as being tethered by multiple hydrogen bonds.

The pathogenic yeast *Candida albicans* possesses a cell wall-associated exo-β-1,3-glucanase (Exg; EC 3.2.1.58), which is implicated in cell wall remodelling [19]. We solved the crystal structure of this GH5 enzyme in the presence of two different mechanism-based inhibitors, thereby revealing the close network of interactions that hold the terminal glucose of the β-glucan substrate in the −1 subsite at the bottom of the active site pocket [20]. The entrance to the active site pocket of Exg is flanked by a pair of phenylalanyl residues, Phe144 and Phe258 (both highly conserved in subfamily 9), which are disposed in an antiparallel manner with a ring separation of ~ 8.5 Å. In the crystal structure of Exg complexed with the inhibitor 2-deoxy-2-fluoro-glucopyranoside (DFG), a second isolated glucopyranoside moiety was found sandwiched between these aromatic rings, with no additional stabilization from hydrogen bonding to the protein, suggesting that this aromatic gateway may act as a clamp to control both the entry of β-1,3-glucan substrate and the exit of free glucose product or, alternatively, glucosyl transfer to an acceptor [21]. The Phe-Phe clamp corresponds to the +1 sugar binding subsite on the enzyme. A similarly disposed aromatic clamp, but involving a Trp-Trp pair, is found in the GH3 exo-1,3/1,4-β-glucanase (ExoI) from barley, and it has been proposed that the wider fused ring structure is responsible for the broader range of β-linkages recognized by this enzyme [22].

To elucidate the roles played by the paired phenylalanyl residues and a third Phe (F229) at a putative +2 subsite, we analyzed the effects of specific mutations at these sites on enzyme activity and, where possible, on tertiary structure. We have also examined the structures of two catalytically disabled mutants of Exg, E292Q and E292S/F229A, in the presence of oligosaccharides, in an attempt to visualize an Exg-substrate complex, which thus far has not been seen.

Results

Subsite binding energies

Previous structural work suggested that the +1 sugar binding site of Exg, as defined by the Phe-Phe clamp, was particularly critical for substrate recognition. However, an earlier study suggested the +2 subsite makes the greatest contribution to binding energy [21]. To resolve this discrepancy, we reappraised the kinetic data derived by Stubbs *et al.* [21], using the formulae of Hiromi [23], and showed that carbohydrate-protein interactions at the +1 subsite of Exg do indeed provide the main contribution to the transition state interaction energy (Fig. 1A). There are smaller contributions from sugar binding at the +2 and +3 subsites.

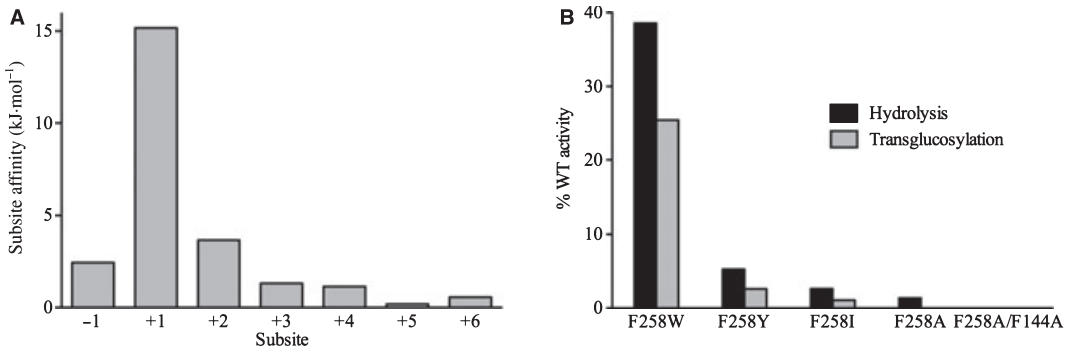

Fig. 1. Key properties of the +1 subsite of Exg. (A) Subsite binding energies based on published kinetic data for Exg-catalysed hydrolysis of β-1,3-linked glucans [21]. Glycosidic bond cleavage occurs between subsites −1 and +1. (B) Relative specific activities of site-directed mutants involving F258 compared to wild-type Exg. Hydrolytic activities are indicated by the black bars and transglucosylation activities by grey bars. Note that the double mutant was inactive in both assays.

Production of enzyme variants

Recombinant native Exg and the E292Q mutant were produced using a previously established *Saccharomyces cerevisiae* expression system in which the orthologous Exg gene had been deleted, whereas the F144A, F258A/I/Y/W and F229A mutants, and the E292S/F229A, F144A/F258A and F144Y/F258Y double mutants, were expressed in the *Pichia pastoris* system when it was apparent that this system provided significantly better yields. In each case, the protein was secreted into the medium, enabling a one-step purification by hydrophobic interaction chromatography, which resulted in $\sim$ 70% recovery of total enzyme. Final yields of pure protein were in the order of 0.5 mg·L^{-1} of culture from the *S. cerevisiae* system and up to 50 mg·L^{-1} from *P. pastoris*. Mutant proteins exhibited the same mobility as wild-type Exg on denaturing gels and possessed the same N-terminal sequence, indicating correct processing by the yeast host. Crystals suitable for X-ray analysis were obtained for the F258I, E292Q, E292S/F229A and F144Y/F258Y variants of Exg.

Mutations involving F258, F144 and F229

Figure 1B shows the results of the activity measurements for hydrolase and transglycosylase assays. A definite trend is seen whereby activity of the F258 mutants falls off steadily in the order Phe > Trp > Tyr > Ile > Ala. This fall-off is more pronounced for the transglycosylase reaction than for the hydrolysis reaction. Although the single mutation variants retained some activity, the double mutant F144A/F258A was essentially inactive in both assays. The more conservative double mutant F144Y/F258Y was considerably less active than native Exg (Table 1). It is also clear from the other kinetic analyses completed (Table 1) that the K_M values are significantly higher when aliphatic substitutions are made, whereas the k_{cat} values for these mutants were reduced by two- to 20-fold. This confirms the important role of the Phe-Phe gateway in substrate recognition and binding for catalysis. The F229A mutation, designed to test the importance of a third Phe residue close to the F258-F144 clamp, resulted in overall loss of hydrolytic efficiency of 17-fold, which is some five times higher than for the F144A mutation.

Examination of the crystal structure of the F258I mutant showed it to be isomorphous with native Exg other than a small adjustment in the backbone conformation of loop 256–263 to accommodate the mutated residue, and a larger movement in the neighbouring loop 312–324. Ile258 adopts a less favoured rotamer, directed away from where the native Phe side chain is disposed and further away from the active site pocket (Fig. 2A,C). The mutation causes a change in local water structure with two new well-ordered water molecules introduced that in turn interact with and shift the external loop 312–324. It is pertinent to note that the most favoured conformer for Ile258 would have

Table 1. Kinetic constants for the hydrolysis reaction with laminarin.

	k_{cat} (min^{-1})	K_M (mg·mL^{-1})	k_{cat}/K_M (min^{-1}·mL·mg^{-1})
Wild-type	9400 ± 450	4.8 ± 0.5	1960
F258W	5300 ± 350	10.6 ± 1.2	500
F258I	540 ± 30	20.1 ± 0.8	27
F144A	460 ± 40	21.1 ± 2.0	22
F229A	1500 ± 100	13.2 ± 1.6	114
F144Y/F258Y	480 ± 20	8.3 ± 0.7	58

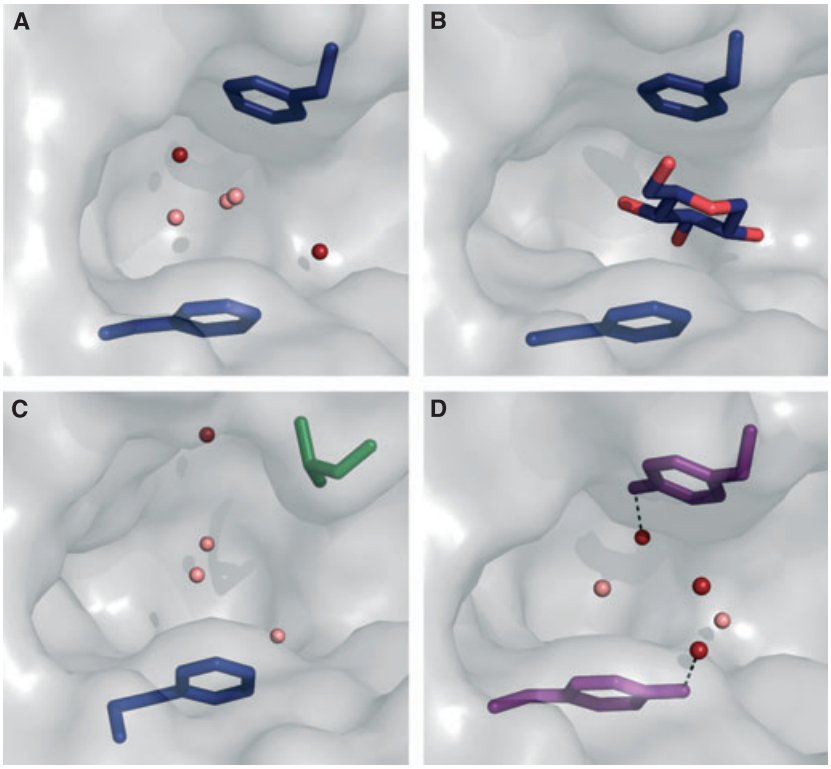

Fig. 2. Structural consequences of mutations at the Phe-Phe clamp of the +1 subsite. Disposition of F144 (lower) and F258 (upper) side chains at the entrance to the active site in native Exg are shown in (A); with a glucose moiety bound between the rings (as in Fig. 4) in (B); the F258I mutation (C); and the double mutation F144Y/F258Y (D). Ordered water molecules are shown as red spheres, less ordered waters are shown in pink and hydrogen bonds are shown as dashed lines.

forced side chain atoms CG2 and CD1 into the aqueous channel that lies between the Phe-Phe gateway. By contrast, the crystal structure of the F144Y/F258Y double mutant is very similar to native Exg with both tyrosyl hydroxyl groups being accommodated without disturbing neighbouring residues (Fig. 2D). Each hydroxyl hydrogen bonds to a water molecule not present in the native structure, so there is a notable difference in local water organization in the clamp region.

Structure of the active site mutant E292Q-Exg with β-1,3-oligosaccharides revealed an unexpected carbohydrate binding site

Previous studies had shown that mutating the catalytic nucleophile E292 to glutamine resulted in an inactive enzyme [24,25]. The crystal structure analysis of E292Q-Exg alone revealed that the only significant difference observed from native Exg was a rearrangement of water structure near the amide side chain of residue Gln292, which adopted the same conformation as the glutamyl side chain. Various β-1,3-glucan oligomers (laminaritriose, -tetraose, -pentaose) were soaked into the crystals of E292Q at pH 6.2 and difference electron density maps were examined. In each case, weak density was apparent in the active site pocket (−1 subsite) extending out to the Phe-Phe gateway (+1 subsite), equivalent in length to two to three linked glucose residues, although this was difficult to model as an oligosaccharide. The structural analysis of E292Q soaked with laminaripentaose showed a weakly bound sugar residue (BGC2) positioned in the Phe-Phe gateway. Significantly, electron density corresponding to a clearly defined sugar residue (BGC1) was seen in a surface indentation well removed (25–30 Å) from the active site pocket and adjacent to Trp287 (Fig. 3), in a side-by-side stacking arrangement. Multiple hydrogen bonds between the 2-, 3- and 4-sugar hydroxyls and the protein backbone stabilized the interaction. The orientation of this glucosyl residue indicated that it corresponded to the nonreducing terminus of the oligosaccharide. The other connected residues were presumed to be mobile and directed into the solvent space between molecules in the crystal lattice.

Structure of the double mutant E292S/F229A with laminaritriose

Various other catalytically disabled mutants of Exg were prepared and several of these could be crystallized. However, crystal soaking experiments with oligosaccharides did not lead to electron density maps showing ordered carbohydrate in the main sugar-binding sites,

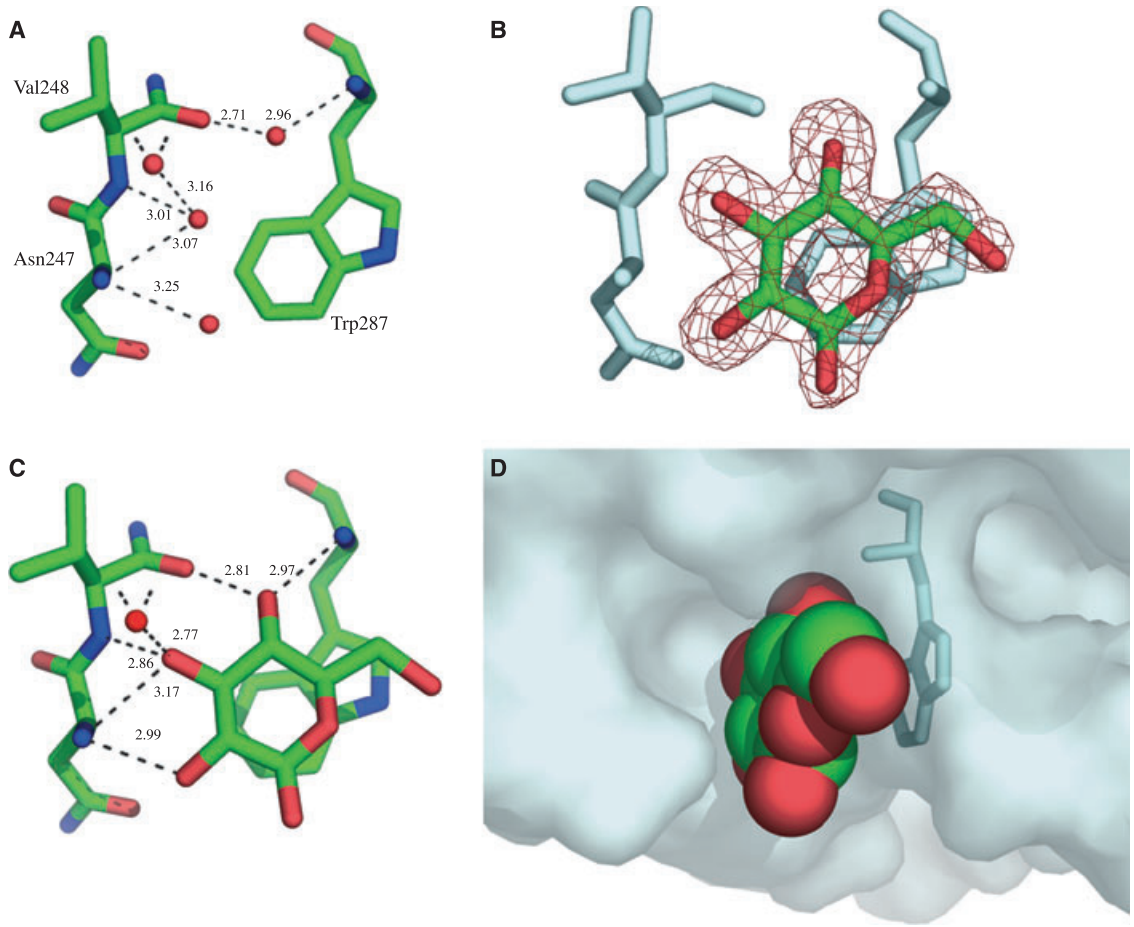

Fig. 3. Binding site for a glucopyranoside remote from the active site. (A) Native exoglucanase structure in the region of Trp287 showing four structural water molecules. (B) Difference electron density contoured at 3σ observed in crystals of inactive mutant E292Q soaked with laminaripentose (L5). (C) Hydrogen bonds formed between the glucose moiety and the protein backbone. A bridging water molecule makes two additional hydrogen bonds to the backbone. (D) Space filling model of the bound glucose stacking against Trp287 in the surface depression.

with the exception of the E292S/F229A double mutant. The crystal structure of this double mutant complexed to laminaritriose revealed the positions of all three glucopyranoside units, with the sugar at the nonreducing end bound in the active site pocket (Fig. 4A,B). This sugar makes multiple hydrogen bonds to amino acid side chains around the pocket or to bridging water molecules. The second sugar is bound in the Phe-Phe clamp (+1 site) held only by CH/π interactions, whereas the third sugar is directed away from the surface of the molecule and makes only the one hydrogen bond to E262. A LIGPLOT diagram (Fig. 5) shows the noncovalent interactions between protein, carbohydrate and water molecules. The same external carbohydrate binding site discovered with the E292Q mutant was also identifiable but, in this case, a biose was seen to bind, with the third sugar presumably being disordered. Figure 6 shows how Exg binds two molecules of laminariotriose with five sugar subsites identified (three in the active site and two on the outside of the protein).

Discussion

Examination of the previously determined crystal structures of native and inhibited forms of Exg from *C. albicans* revealed that this enzyme recruits the same set of eight active site residues as does a group of cellulases, yet shows quite different specificity [20]. It was suggested that the Phe144:Phe258 pairing that lines opposite sides of the entrance to the pocket was particularly relevant because it largely defined the +1 subsite. Indeed, the importance of the +1 subsite is seen

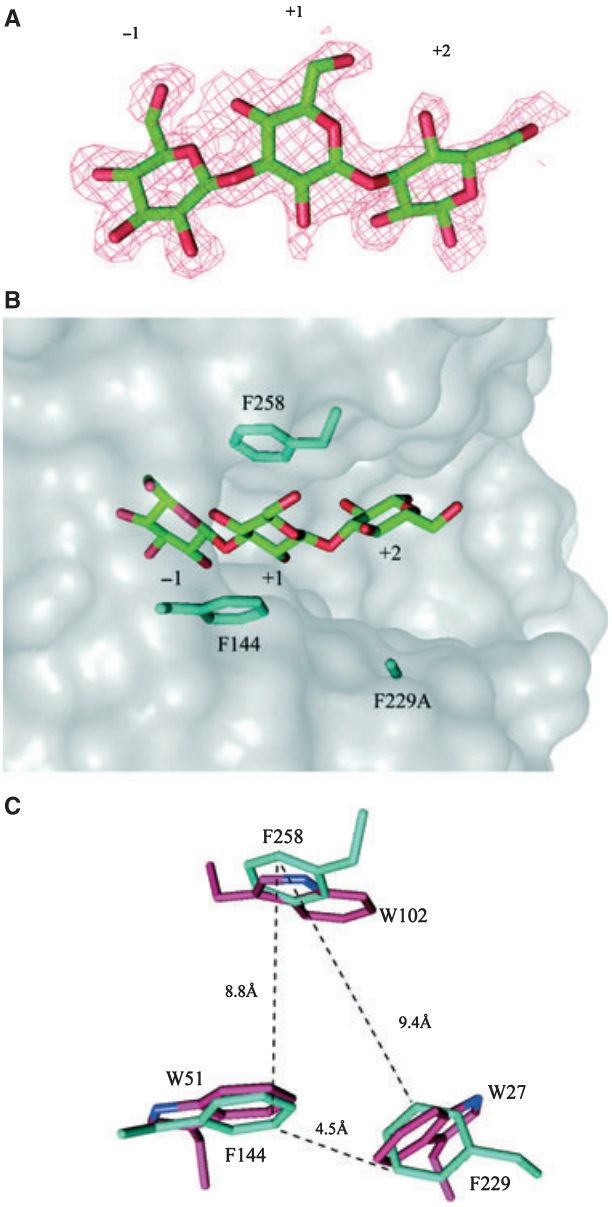

Fig. 4. Binding of laminaritriose (L3) to Exg double mutant E292S/F229A. (A) Relevant electron density from the 2Fo-Fc map is contoured at 0.5σ to show the third sugar residue (labelled +2). (B) Cut-away view of L3 binding in the active site pocket showing the Phe-Phe clamp and the F229A mutation. Carbohydrate binding subsites are labelled −1, +1 and +2. Glycosidic bond cleavage occurs between −1 and +1. (C) Comparison of the aromatic triads of Exg and a carbohydrate binding module CBM 4-2 from *T. maritima* (PDB: 1gui) is shown with F144, F258 and F229 from Exg (cyan), and W51, W102 and W27 from CBM 4-2 (magenta).

from the associated binding energy derived from kinetic studies (Fig. 1A). Site-directed mutagenesis of the two phenylalanines, which are well conserved amongst fungal exo-β-1,3-glucanases, was implemented to help explain their importance to the specificity and catalytic efficiency of the enzyme.

The role of the Phe-Phe entranceway

The fortuitous discovery of an unreacted molecule of the mechanism-based inactivator 2′,4′-dinitrophenyl-2-deoxy-2-fluoro-β-D-glucopyranoside bound in the crystal structure of the Exg:DFG covalent complex highlighted the special nature of the entrance to the active site pocket [20]. With the sugar moiety in $^4C_1$ chair conformation lying parallel between residues Phe144 and Phe258, it could be seen that it was held to the protein purely through stacking interactions, as if clamped. In the native Exg structure, the space between the phenylalanyl rings was occupied by a thin trail of weak electron density indicative of largely disordered water molecules. We aimed to determine whether other aromatic residues could be substituted here, given the known greater propensity for Trp and Tyr to be found in sugar binding sites of proteins [26]. Modelling showed that such substitutions should be able to be accommodated. We also aimed to test the requirement for a two-sided aromatic clamp with respect to GH5 exoglucanase function.

The results obtained showed that, although aromaticity was a vital property of both sides of the clamp, the preference for Phe was also clear, with both K_M and k_{cat} affected by mutation. As an aliphatic substitution results in the loss of one set of CH/π interactions, the resulting decrease in catalytic efficiency was not unexpected. However, the significantly reduced activity of the Tyr-Tyr mutant is harder to explain. Rearrangement of the water molecules in and around the clamp region as seen in the crystal structure may be relevant.

The relative activities of the mutants tested showed the same trend for both the hydrolysis and transglycosylation assays. The observation that the transglycosylation assay appeared to be somewhat more sensitive to mutation than the hydrolysis assay is interesting. This could suggest that, for the mutant enzymes, there is different partitioning of the covalent glucosyl intermediate between the competing acceptors of oligosaccharide and water. Turnover of the glucosyl intermediate via hydrolysis is kinetically less favoured than transglycosylation by at least a factor of ten for native Exg [21,25] and requires activation of a water molecule by the catalytic base E192. Protection of the covalent intermediate from the water nucleophile via occlusion of the catalytic base by bound product or incoming acceptor has been proposed as the likely reason for glucan transglycosylase (Gas2) from *S. cerevisiae* to favour transglycosylation and thereby allow

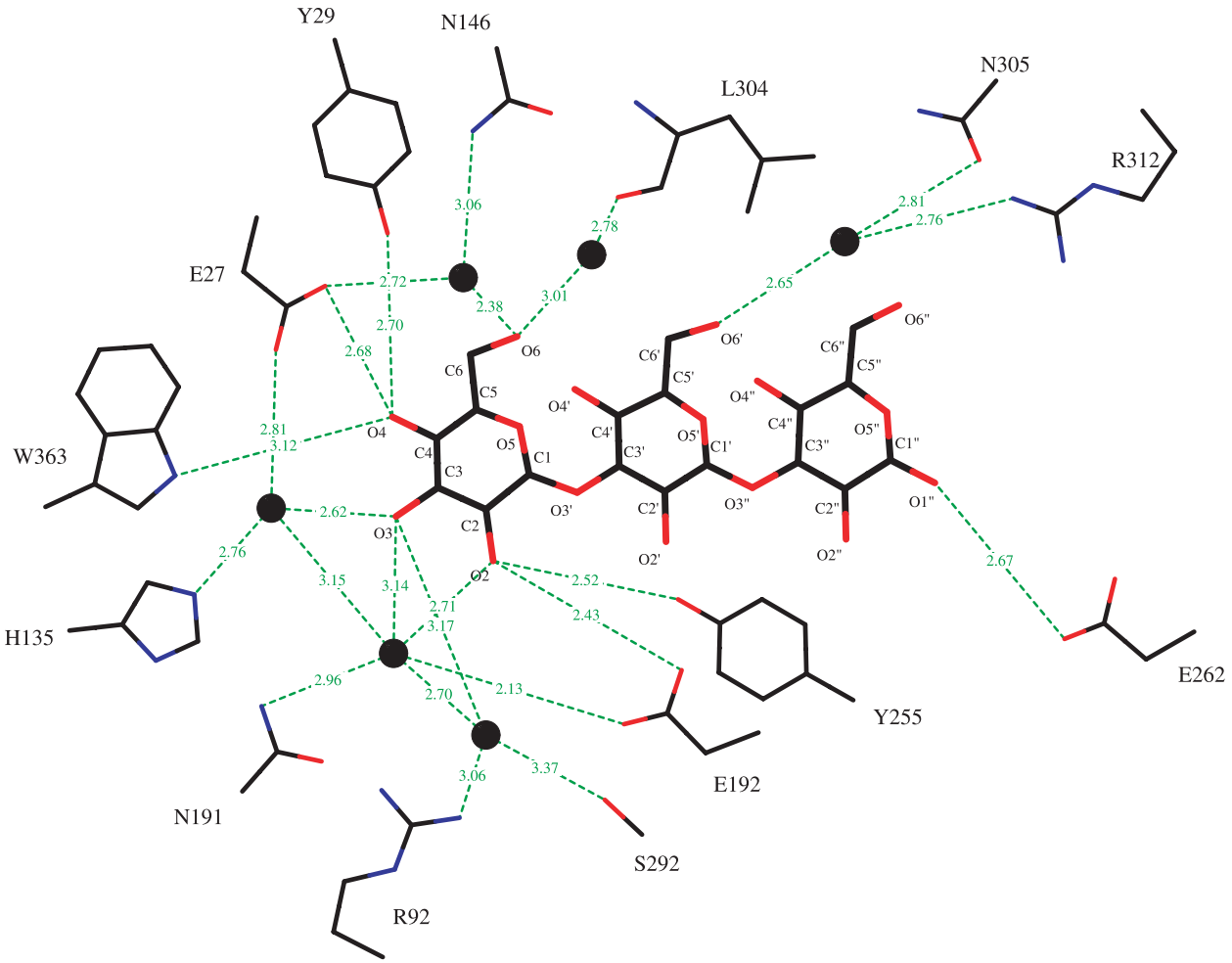

Fig. 5. A LIGPLOT diagram showing hydrogen-bonding interactions between laminaritriose, the Exg double mutant F229A/E292S and connecting water molecules (black spheres).

glucan extension during yeast cell wall remodelling [27]. In the structure of the DFG:Exg complex, a water molecule situated 2.8 Å from E192 is well placed for nucleophilic attack on the anomeric carbon of the terminal sugar. This water molecule is not seen in the laminaritriose (L3):mutant Exg structure, being displaced by the atoms of the glycosidic linkage between the first two sugar residues. Such base occlusion is also likely to be important for glucan remodelling in the *Candida* cell wall.

It is possible that the hydrogen-bonding potential of the Trp or Tyr side chains in contrast to Phe would impose restrictions on glucan interactions with Exg, either slowing entry into and exit from the active site pocket, or possibly altering specificity. The latter idea was discussed by Hrmova *et al.* [22], who studied the structure and specificity of a barley β-D-glucan glucohydrolase, a member of GH3, which possesses a similar aromatic clamp to Exg but which is made up of a Trp-Trp pair. They suggested that this wider-sided clamp allowed a broader substrate specificity than the smaller Phe-Phe clamp of Exg, such that not only β-1,3-glucan linkages, but also β-1,2-,β-1,4- and β-1,6-linkages could be hydrolyzed. A Trp-Trp clamp was also observed in the structure of 4-α-glucanotransferase from *Thermotoga maritima*, a member of GH13, suggesting that such aromatic clamps can be associated with even more diverse specificity; in this case, the transfer of maltosyl and longer dextrinyl residues [28]. Another GH13 enzyme, cyclodextrin glycosyltransferase, has a pair of Phe residues that interact with D-glucose bound at the +2 site, although they are situated further apart and are more angled than the nearly antiparallel Phe-Phe pair seen in Exg [29]. A pair of Phe residues in trehalulose synthase offers a different kind of 'aromatic clamp', being disposed at right angles to

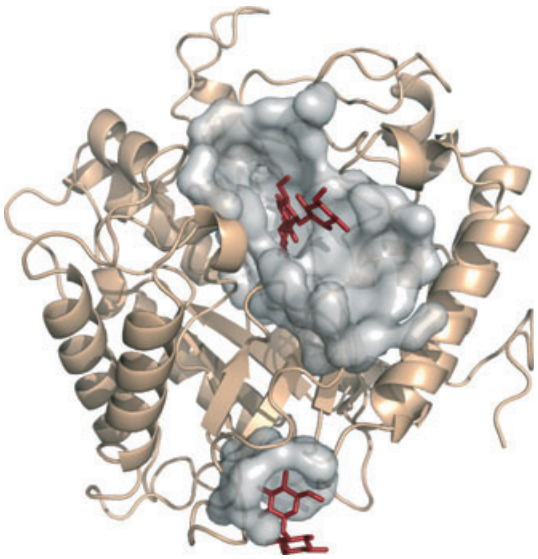

Fig. 6. Surface representation of F229A/E292S:Exg showing one molecule of laminaritriose bound to the active site (subsites −1, +1 and +2) and another to the remote site at which two ordered glucose residues were seen.

each other and only 6 Å apart, but also involved in substrate specificity [30]. If the role of the Phe-Phe clamp in Exg is to influence specificity, then we would expect conservation within subfamily 9. Indeed, a BLAST search using Exg as the query sequence suggests that the two Phe residues are conserved in 23 of the top 25 hits, representing 20 different fungal and yeast species. The two exceptions were putative glucanases from *Schizosaccharomyces pombe*, with F258Y, and *Cryptococcus neoformans*, with F144Y. Interestingly, there is no sequence evidence for a possible Tyr-Tyr clamp. Although these substitutions can be sterically accommodated in the *C. albicans* Exg structure, the accompanying 50-fold reduction in catalytic efficiency would appear to mitigate against such a variant.

We have referred to the two-sided aromatic entrance way as a clamp, although this term does not necessarily provide the best analogy. In the Exg double mutant structure with laminaritriose (L3), described in the present study, the second sugar is seen to be held through CH/π interactions between F144 and F258 in a manner similar to (but much clearer than) the second DFG moiety identified in the previously determined structure of native Exg [20]. However, the first sugar in L3 has swivelled away from the position adopted by DFG in the −1 site, as indeed it must do to maintain the β-1,3-linkage between the two sugars. Although the C4-hydroxy group makes exactly the same three hydrogen bonds to the protein as does DFG, the C3-OH group no longer contacts the protein and instead binds to three water molecules, whereas the C2-OH forms two new hydrogen bonds with the protein. Two of these water molecules coincide with the C2-OH and C3-OH groups in native Exg, whereas the third occupies the space created by the E292S mutation. This situation provides the stabilization energy that enables the first sugar to be held close to its catalytic binding position, assisted by the CH/π interactions involving the second sugar. The F144-F258 pair has to act as a releasable clamp to allow both docking and release of β-glucan. Productive binding for catalysis would require the terminal sugar to displace solvent at the −1 site, form compensating hydrogen bonds to the protein and stack against W363 at the base of the pocket [20].

Does Phe229 contribute to the +2 sugar binding site?

The precise disposition and relative orientation of aromatic platforms, particularly those involved in sandwich interactions with sugar residues, is clearly a major determinant of specificity in glycoside hydrolases. In Exg, a third phenylalanine (Phe229) is situated close to Phe144 apparently in a position to interact with the β-1,3-glucan substrate. Strikingly, the carbohydrate-binding module CBM 4-2 of a bacterial laminarinase, which recognizes the same substrate (laminarin) but which is structurally unrelated to Exg [31], contains a cluster of tryptophans positioned similarly to the three phenylalanines in Exg (Fig. 5C). Initially, this appears to be an example of convergent evolution towards aromatic triads that can accommodate the twists specifically associated with β-1,3-glucan polymers. The nature of the aromatic residues in the two triads might then be reflecting the different functional constraints for the two proteins: Exg requiring precise positioning of laminarin substrate for efficient exoglucanase action and CBM 4-2 for tight binding of laminarin. However, the structure of CBM 4-2 complexed to laminarihexose (PDB:1gui) shows that the oligosaccharide is threaded through the Trp triad in a different orientation to that seen for the Phe triad in the structure of the mutant Exg:L3 complex. Neither of the two sugar residues interacting with the Trp triad overlaps with the sugar seen in the Phe-Phe clamp of Exg.

Structural analysis of the E292S/F229A complex with laminaritriose does indeed show a quite different orientation of the triose from that seen in CBM 4-2. The question arises as to whether mutating Phe229 to the non-aromatic alanine has influenced this reorientation? Other than making a van der Waals' contact with the delta carbon of E192 it serves no other structural

role, yet its mutation to alanine results in a significant loss of catalytic efficiency. On the other hand, this loss of activity may be because E192, the catalytic acid-base, now has greater mobility. It is possible that the loss of the F229 aromatic platform has resulted in the third sugar being redirected and so F229 could still be involved in the true +2 site. Regardless, the fact that we were unable to observe stable binding of 1,3-β-glucan oligosaccharides in crystals of E292 single mutants does suggest that F229 interacts with the substrate.

An unexpected additional glucose binding site

It was hoped that soaking the catalytically disabled mutant E292Q-Exg with substrate might reveal productive binding extending from subsite −1 to at least subsite +2 (i.e. beyond the aromatic clamp); however, the density was not sufficiently clear to model, irrespective of the various oligosaccharides tried, with laminaripentaose appearing the best of these. Unexpectedly, a well-defined glucosyl residue from the nonreducing end was observed bound to the protein in a depression on the outside of the molecule, for each of the three β-1,3-oligosaccharides soaked into the crystal. This sugar moiety stacks against Trp287 and forms four hydrogen bonds to the protein through its 2-OH, 3-OH and 4-OH groups (Fig. 3). Such interactions are typical of functional carbohydrate binding sites and suggest that we are not merely observing an artefact as a result of the crystal being soaked with relatively high concentrations of oligosaccharide. Typical carbohydrate binding sites are said to be preformed in that only small protein conformational changes occur upon binding, whereas water molecules are positioned to mimic the sugar hydroxyl groups in the unbound form of the protein [32]. Both of these characteristics were observed upon comparing various native and mutant Exg structures. The external site was revealed again in the structure of the double mutant complexed with laminaritriose where now two (of the three) sugar residues could be seen. The second sugar is not bound to the enzyme but is stabilized by interactions with a neighbouring molecule in the crystal.

Trp287 is highly conserved amongst GH5 sub-family 9 members and lies in a depression that may be designed to tether the enzyme to a nonreducing end of the branched β-glucan homopolymer that constitutes a large part of the *C. albicans* cell wall. Although this is not part of a discrete carbohydrate-binding module [33], it may help prevent Exg, a secreted enzyme, from diffusing too quickly away from the cell wall. Surface-based carbohydrate binding sites of weak affinity are well known (e.g. in lectins and haemagglutinin) where shallow indentations contrast with the deeper clefts found in the active sites of carbohydrate processing enzymes [32]. If Trp287 does have a functional role then it would represent an unusual feature for a $(\beta\alpha)_8$-barrel enzyme, in that it is located at the N-terminal end of the β7 strand facing away from the active site [34]. Interestingly, it is only five residues removed from the catalytic nucleophile, E292, at the C-terminal end of β7. A recent modelling study of glucan interacting with Exg has predicted the involvement of two loop regions in the enzyme to accommodate the glucan chain [35] but our structural studies have yet to confirm this.

In conclusion, the results obtained from mutagenesis of the Phe-Phe gateway have emphasized not only the need to preserve the aromatic nature of this entranceway for efficient catalytic turnover, but also the necessity for the completely nonpolar and less bulky F144/F258 pairing to position the substrate for glycosidic bond cleavage at the nonreducing end. The successful visualization of bound β-1,3-glucan trisaccharide to inactivated enzyme represents the first complexed structure of Exg involving an oligosaccharide and provides an insight into the enzyme's mechanism of action in the *C. albicans* cell wall. Finally, the unexpected discovery of an isolated binding site remote from the active site poses the intriguing possibility of a novel evolutionary solution to the problem of maintaining association with the *C. albicans* cell wall.

Experimental procedures

Substrates and reagents

All reagents and buffer chemicals were obtained from Sigma Chemical Co. (St Louis, MO, USA) unless otherwise noted. Restriction enzymes were obtained from Roche (Basel, Switzerland) and New England Biolabs (Beverly, MA, USA). Growth media components were obtained from Difco (Franklin Lakes, NJ, USA) and Gibco BRL (Gaithersburg, MD, USA). Geneticin was obtained from Boehringer Mannheim (Mannheim, Germany). Laminaritriose, -tetraose and pentose (fine grade) were obtained from Seikagaku Kogyo (Tokyo, Japan).

PCR mutagenesis of *EXG*

The pFOX vectors, containing fragments of the *EXG* gene, were constructed previously [25]. Site-directed mutations of F144A, F229A, F258A, F258I, F258Y, F258W and E292S were generated using overlap extension PCR [36] and E292Q was already available. The F144A/F258A, F144Y/F258Y and F229A/E292S double mutants were

made by sequential mutation. The oligonucleotides used to generate the mutations were (positions of mismatches underlined): F144A: 5′-CAA AAT GGG GCT GAC AAC TCC-3′; F144Y: 5′-CAA AAT GGG TAT GAC AAC TCC-3′; F229A: 5′-CAC GAT GCT GCC CAA GTC TTT-3′; F258A: 5′-TAC CAA GTG GCT TCC GGT GGT-3′; F258I: 5′-TAC CAA GTG ATT TCC GGT GGT-3′; F258Y: 5′-TAC CAA GTG TAT TCC GGT GGT-3′; F258W: 5′-TAC CAA GTG TGG TCC GGT GGT-3′; E292S: 5′-GG AAC GTC GCT GGT TCA TGG TCT GCT GCT TTG-3′. Outside primers (also used for DNA sequencing) were T3: 5′-ATT AAC CCT CAC TAA AG-3′ and T7: 5′-AAT ACG ACT CAC TAT AG-3′. Expand high fidelity DNA polymerase (Roche) was used in all PCR reactions. Products were subcloned back into the appropriate pFOX vector, and mutations were confirmed by DNA sequencing.

Expression of exoglucanase mutants

Wild-type Exg was produced in *S. cerevisiae* strain AWY-1 as described previously [24,25]. The F144, F229, F258 and E292 mutant Exg species were produced in *P. pastoris* strain KM71 (Invitrogen, Carlsbad, CA, USA). To facilitate cloning into the integrative expression plasmid pPIC9K [37], it was first necessary to subclone the mutation-containing pFOX fragments into vector pGSB1, a pUC19 derivative containing the complete *EXG* ORF. The entire ORFs containing each mutation were then cloned into the *Sna*BI site of pPIC9K. After linearization of each resulting plasmid with *Sal*I, *P. pastoris* was transformed using an electroporation method adapted from that described for *S. cerevisiae* [38]. Briefly, a 200-mL culture of *P. pastoris* KM71 was grown at 27 °C in YPD medium [1% (w/v) yeast extract, 2% (w/v) casein hydrolysate, 2% (w/v) glucose] until A_{600} of 1.3–1.6 was reached. Cells were harvested and washed in ice-cold water and then in ice-cold 1.0 M sorbitol, before being resuspended in 0.6 mL of ice-cold 1.0 M sorbitol. An 80-μL aliquot of cells was added to 5–10 μg of DNA and the sample was electroporated at 1500 V, 186 Ω and 50 μF in an Electro Cell Manipulator 600 (BTX Inc., San Diego, CA, USA). Ice-cold 1.0 M sorbitol (1 mL) was added to the cells immediately following electroporation. Transformants were screened for histidine prototrophy and then for multiple integration of the plasmid by plating on increasing concentrations of geneticin (0–4 mg·mL^{-1}), as described previously [37]. Expression of the mutant Exg proteins was induced in 50-mL cultures of minimal medium [39] containing 1% (w/v) casamino acids and buffered to pH 6.0 with 100 mM potassium phosphate. Cultures were inoculated to high optical densities (A_{600} = 5–25) and grown in shake flasks at 27 °C with an agitation rate of 250 r.p.m. for 72–102 h. Fresh methanol (1%, v/v) was added every 24 h, to ensure continued induction of expression.

Enzyme purification

Wild-type Exg was purified as described previously [21]. For the mutant Exg proteins, *P. pastoris* culture medium was harvested by centrifugation (13 800 ***g*** for 10–15 min), concentrated and buffer-exchanged with 50 mM potassium phosphate buffer (pH 7.0) containing 1.0 M $(NH_4)_2SO_4$ using Vivaspin 20-mL concentrators (5 kDa cut-off; Vivascience Ltd., Stonehouse, UK). The recombinant enzymes were then purified in a single step by application of the respective concentrates to a phenyl superose HR 5/5 column (Pharmacia LKB Biotechnology AB, Uppsala, Sweden), which had previously been equilibrated in 50 mM potassium phosphate buffer (pH 7.0) containing 0.6 M $(NH_4)_2SO_4$. A linear reverse salt gradient of 0.6–0.0 M $(NH_4)_2SO_4$, in 50 mM potassium phosphate buffer (pH 7.0) was applied over 35 min at a flow rate of 0.5 mL·min^{-1}. Fractions (1 mL) containing the enzyme were pooled and concentrated to final volumes of 0.5–1.0 mL using Centricon-10 and Microcon-10 microconcentrators (Amicon Corp., Danvers, MA, USA) and stored at 4 °C. It should be noted that, before chromatography, the column was stringently washed and the eluate tested for any residual enzyme activity. Native Exg was not purified on the same column as the mutants. Yields were estimated by a modified Lowry method [40] using BSA as standard and by Nanodrop spectroscopy (NanoDrop, Wilmington, DE, USA).

Enzyme activity analysis

Glycoside hydrolase activity of the Exg mutants was determined with laminarin, a β-1,3-linked polymer of glucose with an average degree of polymerization of 28. Assays were carried out in 80 mM sodium acetate buffer (pH 5.6) using 7.8 mg·mL^{-1} laminarin. This concentration corresponds to approximately twice the K_M value for the wild-type enzyme; at higher concentrations of laminarin, the competing transglycosylase reaction becomes significant. Assays (125 μL total volume) were incubated at 37 °C for 30–120 min, and the reactions were stopped by heating at 100 °C for 10 min. Glucose formation was measured using the glucose oxidase method [41]. The kinetic constants, k_{cat} and K_M, of each Exg mutant tested were determined by assaying with five to seven different concentrations of laminarin, in the range 0.56–28 mg·mL^{-1}. Kinetic data were analysed Prism5 (GraphPad Software Inc., San Diego, CA, USA) using nonlinear regression analysis or double-reciprocal plots. Transglucosylation activity of the recombinant proteins was estimated using 40 mg·mL^{-1} laminaritriose (Seikagaku Kogyo), as described previously [21].

Crystallography

Crystallization conditions for Exg mutants were similar to those previously established for wild-type recombinant Exg

Table 2. Data processing and refinement statistics for exoglucanase mutants.

	F258I	F144Y/F258Y	E292Q:L5	F229A/E292S:L3
Resolution (outer shell) (Å)	32.65–1.90 (1.95–1.90)	38.23–2.00 (2.11–2.00)	19.85–1.80 (1.85–1.80)	38.19–1.70 (1.79–1.70)
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters	*a* = 60.33	*a* = 58.57	*a* = 59.95	*a* = 58.72
	b = 65.39	*b* = 64.63	*b* = 65.32	*B* = 64.40
	c = 96.49	*c* = 94.48	*c* = 96.54	*C* = 94.87
	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°	α = β = γ = 90°
R_{merge} (outer shell)	0.052 (0.21)	0.069 (0.43)	0.039 (0.16)	0.074 (0.63)
Mean $I/\sigma I$ (outer shell)	13.8 (3.3)	17.4 (3.7)	14.2 (4.7)	14.1 (2.2)
Completeness (outer shell)	98.9 (95.3)	100.0 (100.0)	99.3 (94.1)	99.8 (99.9)
Multiplicity (outer shell)	3.3 (2.6)	5.4 (5.3)	3.2 (2.4)	4.2 (4.1)
Number of unique reflections	30 432	25 035	35 563	40 247
R_{cryst}	0.135	0.154	0.136	0.152
R_{free}	0.168	0.200	0.163	0.197
r.m.s.d. for bonds (Å)	0.014	0.020	0.012	0.023
r.m.s.d. for angles (deg)	1.35	1.46	1.25	1.94
r.m.s.d. for chiral volume (Å^3)	0.09	0.11	0.08	0.14
Number of protein atoms	3211	3233	3214	3254
Number of ligand atoms			24	58
Number of water atoms	334	308	289	405
Average main chain *B*-factor (Å^2)	21	21	17	15
Average side chain *B*-factor (Å^2)	24	22	19	17
Average ligand *B*-factor (Å^2) BGC1[a]			24	
Average ligand *B*-factor (Å^2) BGC2[b]			75	
Average ligand *B*-factor (Å^2) L3-1[c]				38
Average ligand *B*-factor (Å^2) L3-2[d]				24
Average ligand *B*-factor (Å^2) Ca				16
Average water *B*-factor (Å^2)	39	30	33	29
Protein Data Bank entry	2PF0	3O6A	2PC8	3N9K

[a] BGC1 is the glucose residue from the nonreducing end of laminaripentose (L5) bound to the exterior of the protein. [b] BGC2 is the weakly bound glucose residue of L5 bound in the Phe-Phe clamp. [c] L3-1 is laminaritriose bound in the active site. [d] L3-2 is a second molecule of laminaritriose bound to the exterior site where the third residue is disordered.

[20]. Briefly, crystals were grown at 17 °C in hanging drops containing 75 mM Hepes-KOH (pH 7.3), 150 mM $CaCl_2$ and 14–17% PEG 8000 at a protein concentration of $\sim$ 7 mg·mL^{-1}. Crystals of E292Q were adjusted to pH 6.2 and then soaked with various β-1,3-glucan oligomers (n = 3, 4 and 5) before data collection. Co-crystallization experiments with the same group of oligosaccharides were also conducted but without success. Crystals of the double mutant F229A/E292S were soaked at pH 7.3 with laminaritriose before data collection. X-ray diffraction data for F258I, E292Q and native Exg were collected at room temperature on a Rigaku R-AXIS II system (Rigaku, The Woodlands, TX, USA) (at the School of Biological Sciences, University of Auckland, Auckland, New Zealand), whereas an in-house Rigaku R-AXIS IV++ system was used for the double mutants F144Y/F258Y and F229A/E292S at 213 K. Data were processed and refined using DENZO/SCALEPACK [42] and the CCP4 crystallographic software suite [43], with COOT [44] employed for model building. Data collection and refinement statistics for the four solved structures are listed in Table 2. Atomic coordinates were deposited with the Protein Data Bank. Structural diagrams employed the software LIGPLOT [45], PYMOL [46] and CASTP [47].

Acknowledgements

This work was supported in part by a University of Otago Research Grant (KLH B02). Professor Pat Sullivan provided access to HPLC for sugar analysis and Bronwyn Carlisle assisted with the illustrations.

References

1 Hancock SM, Vaughan MD & Withers SG (2006) Engineering of glycosidases and glycosyltransferases. *Curr Opin Chem Biol* **10**, 509–519.
2 Sinnott ML (1990) Catalytic mechanisms of enzymic glycosyl transfer. *Chem Rev* **90**, 1171–1202.

3 McCarter JD & Withers SG (1994) Mechanisms of enzymatic glycoside hydrolysis. *Curr Opin Struct Biol* **4**, 885–892.
4 Zechel DL & Withers SG (2000) Glycosidase mechanisms: anatomy of a finely tuned catalyst. *Acc Chem Res* **33**, 11–18.
5 Davies G & Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 853–859.
6 Withers SG & Aebersold R (1995) Approaches to labeling and identification of active site residues in glycosidases. *Protein Sci* **4**, 361–372.
7 Ly HD & Withers SG (1999) Mutagenesis of glycosidases. *Annu Rev Biochem* **68**, 487–522.
8 White A & Rose DR (1997) Mechanism of catalysis by retaining β-glycosyl hydrolases. *Curr Opin Struct Biol* **7**, 645–651.
9 Vocadlo DJ & Davies GJ (2008) Mechanistic insights into glycosidase chemistry. *Curr Opin Chem Biol* **12**, 539–555.
10 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.
11 Lo LeggioL & Larsen S (2002) The 1.62 Å structure of *Thermoascus aurantiacus* endoglucanase: completing the structural picture of subfamilies in glycoside hydrolase family 5. *FEBS Lett* **523**, 103–108.
12 Larsson AM, Anderson L, Xu B, Muñoz IG, Usón I, Janson JC, Ståhlbrand H & Ståhlberg J (2006) Three-dimensional crystal structure and enzymic characterization of β-mannanase Man5A from blue mussel *Mytilus edulis*. *J Mol Biol* **357**, 1500–1510.
13 Sakon J, Adney WS, Himmel ME, Thomas SR & Karplus PA (1996) Crystal structure of thermostable family 5 endocellulase E1 from *Acidothermus cellulolyticus* in complex with cellotetraose. *Biochemistry* **35**, 10648–10660.
14 Dominguez R, Souchon H, Lascombe M & Alzari PM (1996) The crystal structure of a family 5 endoglucanase mutant in complexed and uncomplexed forms reveals an induced fit activation mechanism. *J Mol Biol* **257**, 1042–1051.
15 Davies GJ, Dauter M, Brzozowski AM, Bjørnvad ME, Andersen KV & Schülein M (1998) Structure of the *Bacillus agaradherans* family 5 endoglucanase at 1.6 Å and its cellobiose complex at 2.0 Å resolution. *Biochemistry* **37**, 1926–1932.
16 Quiocho FA (1993) Probing the atomic interactions between proteins and carbohydrates. *Biochem Soc Trans* **21**, 442–448.
17 Nishio M, Umezawa Y, Hirota M & Takeuchi Y (1995) The CH/π interaction: significance in molecular recognition. *Tetrahedron* **51**, 8665–8701.
18 Vyas NK, Vyas MN & Quiocho FA (1988) Sugar and signal-transducer binding sites of the *Escherichia coli* galactose chemoreceptor protein. *Science* **242**, 1290–1295.
19 Chaffin WL, López-Ribot JL, Casanova M, Gozalbo D & Martínez JP (1998) Cell wall and secreted proteins of *Candida albicans*: identification, function, and expression. *Microbiol Mol Biol Rev* **62**, 130–180.
20 Cutfield SM, Davies GJ, Murshudov G, Anderson BF, Moody PC, Sullivan PA & Cutfield JF (1999) The structure of the exo-β-(1,3)-glucanase from *Candida albicans* in native and bound forms: relationship between a pocket and groove in family 5 glycosyl hydrolases. *J Mol Biol* **294**, 771–783.
21 Stubbs HJ, Brasch DJ, Emerson GW & Sullivan PA (1999) Hydrolase and transferase activities of the β-1,3-exoglucanase of *Candida albicans*. *Eur J Biochem* **263**, 889–895.
22 Hrmova M, De GoriR, Smith BJ, Fairweather JK, Driguez H, Varghese JN & Fincher GB (2002) Structural basis for broad substrate specificity in higher plant β-D-glucan glucohydrolases. *Plant Cell* **14**, 1033–1052.
23 Hiromi K (1970) Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. *Biochem Biophys Res Commun* **40**, 1–6.
24 Chambers RS, Walden AR, Brooke GS, Cutfield JF & Sullivan PA (1993) Identification of a putative active site residue in the exo-β-(1,3)-glucanase of *Candida albicans*. *FEBS Lett* **327**, 366–369.
25 Mackenzie LF, Brooke GS, Cutfield JF, Sullivan PA & Withers SG (1997) Identification of Glu-330 as the catalytic nucleophile of *Candida albicans* exo-β-(1,3)-glucanase. *J Biol Chem* **272**, 3161–3167.
26 Taroni C, Jones S & Thornton JM (2000) Analysis and prediction of carbohydrate binding sites. *Protein Eng* **13**, 89–98.
27 Hurtado-Guerrero R, Schüttelkopf AW, Mouyna I, Ibrahim AF, Shepherd S, Fontaine T, Latgé JP & van Aalten DM (2009) Molecular mechanisms of yeast cell wall glucan remodeling. *J Biol Chem* **284**, 8461–8469.
28 Roujeinikova A, Raasch C, Sedelnikova S, Liebl W & Rice DW (2002) Crystal structure of *Thermotoga maritima* 4-α-glucanotransferase and its acarbose complex: implications for substrate specificity and catalysis. *J Mol Biol* **321**, 149–162.
29 Haga K, Kanai R, Sakamoto O, Aoyagi M, Harata K & Yamane K (2003) Effects of essential carbohydrate/aromatic stacking interaction with Tyr100 and Phe259 on substrate binding of cyclodextrin glycosyltransferase from alkalophilic *Bacillus* sp. 1011. *J Biochem (Tokyo)* **134**, 881–891.
30 Ravaud S, Robert X, Watzlawick H, Haser R, Mattes R & Aghajari N (2007) Trehalulose synthase native and carbohydrate complexed structures provide

insights into sucrose isomerization. *J Biol Chem* **282**, 28126–28136.
31 Boraston AB, Nurizzo D, Notenboom V, Ducros V, Rose DR, Kilburn DG & Davies GJ (2002) Differential oligosaccharide recognition by evolutionarily-related β-1,4 and β-1,3 glucan-binding modules. *J Mol Biol* **319**, 1143–1156.
32 Weis WI & Drickamer K (1996) Structural basis of lectin-carbohydrate recognition. *Annu Rev Biochem* **65**, 441–473.
33 Boraston AB, Bolam DN, Gilbert HJ & Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382**, 769–781.
34 Höcker B, Jürgens C, Wilmanns M & Sterner R (2001) Stability, catalytic versatility and evolution of the $(\beta\alpha)_8$-barrel fold. *Curr Opin Biotechnol* **12**, 376–381.
35 Moura-Tamames SA, Ramos MJ & Fernandes PA (2009) Modelling β-1,3-exoglucanase-saccharide interactions: structure of the enzyme-substrate complex and enzyme binding to the cell wall. *J Mol Graph Model* **27**, 908–920.
36 Horton RM, Cai ZL, Ho SN & Pease LR (1990) Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *BioTechniques* **8**, 528–535.
37 Scorer CA, Clare JJ, McCombie WR, Romanos MA & Sreekrishna K (1994) Rapid selection using G418 of high copy number transformants of *Pichia pastoris* for high-level foreign gene expression. *Bio/Technology* **12**, 181–184.
38 Becker DM & Guarente L (1991) High-efficiency transformation of yeast by electroporation. *Methods Enzymol* **194**, 182–187.
39 Wickerham LJ (1946) A critical evaluation of the nitrogen assimilation tests commonly used in the classification of yeasts. *J Bacteriol* **52**, 392–401.
40 Peterson GL (1977) A simplification of the protein assay method of Lowry *et al*. which is more generally applicable. *Anal Biochem* **83**, 346–356.
41 Bruss ML & Black AL (1978) Enzymatic microdetermination of glycogen. *Anal Biochem* **84**, 309–312.
42 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307–326.
43 CCP4 (1994) The *CCP4* suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760–763.
44 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
45 Wallace AC, Laskowski RA & Thornton JM (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng* **8**, 127–134.
46 De Lano Scientific LLC, Palo Alto, CA, USA. http://www.pymol.org.
47 Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y & Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* **34**, W116–W118.