Crystal Structures of a Poplar Xyloglucan Endotransglycosylase Reveal Details of Transglycosylation Acceptor Binding

Patrik Johansson, Harry Brumer III, Martin J. Baumann, Åsa M. Kallas, Hongbin Henriksson, Stuart E. Denman, Tuula T. Teeri, and T. Alwyn Jones

INTRODUCTION

Plant cell walls are composite structures of cellulose, hemicelluloses, lignin, and structural proteins. The insoluble cellulose microfibrils constitute the main load-bearing component, whereas the hemicelluloses are needed for flexibility. The shape, size, and sometimes even the function of a plant cell are defined by the structure and properties of its cell wall. Cell wall reconstruction is an important prerequisite of many central processes of plant life, such as germination, growth, fruit ripening, organ abscission, vascular differentiation, and responses to pathogens (Carpita and McCann, 2000). Structural changes in the cell wall components are brought about by enzymatic modification, and the wall-modifying enzymes play important roles during cell wall morphogenesis. One of the key interactions in the primary cell walls of dicotyledons is the intimate association of cellulose and xyloglucan. Xyloglucan is a soluble hemicellulose with a backbone composed of $\beta$-(1→4)-linked glucose residues similar to cellulose. The shape and properties of the primary cell walls (Chanliaud et al., 2002) are defined by the structure and properties of its cell wall. Cell wall reconstruction is an important prerequisite of many central processes of plant life, such as germination, growth, fruit ripening, organ abscission, vascular differentiation, and responses to pathogens (Carpita and McCann, 2000). Structural changes in the cell wall components are brought about by enzymatic modification, and the wall-modifying enzymes play important roles during cell wall morphogenesis. One of the key interactions in the primary cell walls of dicotyledons is the intimate association of cellulose and xyloglucan. Xyloglucan is a soluble hemicellulose with a backbone composed of $\beta$-(1→4)-linked glucose residues similar to cellulose. However, the glucan backbone of xyloglucan is abundantly substituted with $\alpha$-(1→6)-linked xylopyranose branches that in turn may be further derivatized by $\beta$-(1→2)-linked galactopyranosyl residues. In some cases, the galactose residues are fucosylated. The nature and extent of the substitution on the glucan chain are both species- and tissue-dependent (Vincken et al., 1997; Vierhuis et al., 2001). Xyloglucan binds noncovalently to cellulose, thereby coating and cross-linking the adjacent cellulose microfibrils (Hayashi, 1989; McCann et al., 1990). In vitro studies of model systems suggest that these xyloglucan/cellulose networks are the main contributors of the extensibility and strength required of the primary cell walls (Chanliaud et al., 2002). Therefore, xyloglucan-metabolizing enzymes play a central role in practically all processes requiring remodeling of the cell walls (Fry, 1989).

During processes such as fruit ripening, hydrolytic enzymes can be produced to achieve efficient cell wall degradation. However, upon cell wall expansion and elongation, cell wall loosening is a temporary requirement that must be followed by rapid reinforcement of the wall structure. For this purpose, plants have evolved unique transglycosylating enzymes, the xyloglucan endotransglycosylases (XETs). XET catalyzes, in a first step, an endolytic cleavage of a cross-linking xyloglucan polymer that permits cellulose microfibrils to separate and the cell to expand. In the second reaction step, XET transfers the newly generated end to another sugar polymer, thereby restoring stable cell wall structure (Smith and Fry, 1991; Xu et al., 1995; Purugganan et al., 1997; Campbell and Braam, 1999). This general mechanism has been demonstrated by adding XET to xyloglucan in vitro (Lorences and Fry, 1993; Sulová et al., 1998; Sulová and Farkaš, 1998; Baran et al., 2000; Steele et al., 2001). XET activity also has been directly demonstrated in plant tissues undergoing wall expansion (Vissenberg et al., 2000, 2001) and restructuring (Thompson et al., 1997; Ito and Nishitani, 1999; Thompson and Fry, 2001). Recent evidence also suggests that XET may play...
a role during the early phases of secondary cell wall deposition, possibly to reinforce the connections between the primary and secondary wall layers (Bourquin et al., 2002). The XETs are typically encoded by multigene families. For example, >30 XET genes have been identified in the Arabidopsis (*Arabidopsis thaliana*) genome. Based on sequence similarities, three to four different subgroups have been recognized (reviewed in Rose et al., 2002). Comprehensive expression analyses show that the XET gene family members are individually regulated, which suggests that the corresponding enzymes operate in different tissues or during different developmental processes. In the absence of comprehensive biochemical analyses, it is not yet known whether the tissue-specific expression or the division in subgroups reflect differences in the substrate specificity or the mode of action of the different XET isoenzymes.

Parallel with the traditional enzyme classification based on substrate specificity and the reaction catalyzed (Webb, 1992), carbohydrate-active enzymes have been classified in families based on sequence similarity (Henrissat and Davies, 1997). In the glycoside hydrolases, the overall fold in the different families exhibits clearly identifiable similarities, and these families have been further clustered to form clans. It also has been generally observed that the enzymes within a given family apparently share the same, either retaining or inverting (Koshland, 1953), reaction

<table>
<thead>
<tr>
<th>Table 1. Data Collection and Refinement Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
</tr>
<tr>
<td>Beamline</td>
</tr>
<tr>
<td>Detector</td>
</tr>
<tr>
<td>Temperature (K)</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
</tr>
<tr>
<td>Number of images</td>
</tr>
<tr>
<td>Oscillation angle per frame/mosaicity (°)</td>
</tr>
<tr>
<td>Space group</td>
</tr>
<tr>
<td>Unit-cell parameters (Å, °)</td>
</tr>
<tr>
<td>Number of unique reflections</td>
</tr>
<tr>
<td>Redundancy</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
</tr>
<tr>
<td>Completeness (%)</td>
</tr>
<tr>
<td>&lt;I/Io&gt;</td>
</tr>
<tr>
<td>R_sym (%)</td>
</tr>
<tr>
<td>Refinement</td>
</tr>
<tr>
<td>R_cryst (%)</td>
</tr>
<tr>
<td>R_free (%)</td>
</tr>
<tr>
<td>RMSD from ideal values</td>
</tr>
<tr>
<td>Bond distance (Å)</td>
</tr>
<tr>
<td>Bond angle (°)</td>
</tr>
<tr>
<td>Ramachandran Plot</td>
</tr>
<tr>
<td>Residues, core regions (%)</td>
</tr>
<tr>
<td>Number of Atoms [Average B Values (Å²)]</td>
</tr>
<tr>
<td>Protein</td>
</tr>
<tr>
<td>Water</td>
</tr>
<tr>
<td>Glycosylation sugars</td>
</tr>
<tr>
<td>XLLG sugars</td>
</tr>
<tr>
<td>Average XLLG B Values (Å²)</td>
</tr>
<tr>
<td>Molecule</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
</tbody>
</table>

---

*a R_sym = \( \sum |I_i - \langle I \rangle| / \sum |I_i| \).

*b R_cryst = \( \sum |F_{calc} - |F_{obs}| | / \sum |F_{calc}| \).

c R_free, calculated from a randomly chosen 5% of all unique reflections.

d Ideal values from Engh and Huber (1991).

* Ramachandran core regions defined according to Kleywegt and Jones (1996).

ESRF, European Synchrotron Radiation Facility; ADSC, Area Detector Corporate Systems.

Values in parentheses refer to the highest resolution shell.
mechanism (Carbohydrate-Active Enzymes server, http://afmb.cnrs-mrs.fr/CAZY/GH_intro.html). All presently known XET and XET-like enzymes belong to glycoside hydrolase family GH16, a member of clan B. In addition to other family 16 enzymes, such as 1,3-1,4 β-glucanases (lichenases), keratan-sulfate-endogalactosidas, κ-carrageenases, β-agarases, and 1,3-β-glucanases (laminarinases), clan B contains cellobiohydrolases and endoglucanases from family GH7. Crystallographic studies of the catalytic modules of enzymes in family GH16 (Keitel et al., 1993; Michel et al., 2001) and family GH7 (Divne et al., 1994) reveal an identical folding topology consisting of two antiparallel β-sheets that stack to form a β-sandwich consisting of one concave and one convex face. The concave face contains the sugar substrate binding sites, with the catalytic side chains located on a single strand that is offset from the center of the sheet (Ståhlberg et al., 1996). Clan B glycoside hydrolases act via a double-displacement reaction that results in the net retention of the configuration of the anomeric carbon. In the first step of the general retaining mechanism, a side chain carboxylate acts as a nucleophile that attacks the anomeric carbon of the sugar ring. A proton from a second side chain carboxylate facilitates fission of the glycosidic bond by general acid catalysis. The glycosyl-enzyme intermediate thus formed (Vocadlo et al., 2001) is decomposed in a second step, wherein a water molecule is

Figure 1. Features of the Poplar XET Molecular Structure.
(A) PttXET16A consists of two large β-sheets arranged in a sandwich-like manner. The active site residues E85, D87, and E89 are included in the figure (light blue).
(B) XET topology. N- to C-terminals are color-coded red to blue. The nucleophile and the general acid/base reside on β7, which also holds the N-linked glycosylation site.
(C) Glycosylation of PttXET16A. Two N-acetylglucosamine sugars and a mannose residue were observed attached to N93 at the end of strand β7. 2Fobs−Fcalc electron density within 1 Å of the oligosaccharide. Contouring at 0.4 e/Å³ (1.0σ level).
(D) The three catalytic residues of GH16 enzymes are structurally well conserved despite the overall low sequence identity in the family. Populus XET16A in yellow-green with light blue sidechains, Bacillus 1,3-1,4 β-glucanase in red, and Pseudoalteromonas κ-carrageenase in dark blue. The coloring scheme is in all figures.
activated, leading to a nucleophilic attack on the anomic center by general base catalysis from the deprotonated form of the second carboxylate (Sinnott, 1990). This second carboxylate is commonly referred to as the general acid/base residue because of its dual role in catalysis. Most of the family GH16 enzymes share a general active-site motif of ExDxE, whereas an insert among the κ-carrageenases and the family GH7 enzymes gives a conserved motif of ExDxxE (Michel et al., 2001). The first and last glutamyl residues in these motifs have been unambiguously identified as the catalytic nucleophile and the general acid/base, respectively (Keitel et al., 1993; Hahn et al., 1995; Ståhlberg et al., 1996). In both families, the middle Asp forms a tight hydrogen bond with the nucleophile.

XET acts similarly to the glycoside hydrolases in the first reaction step but prefers a carbohydrate acceptor over a water molecule in the second reaction step. This strong preference for transglycosylation instead of hydrolysis allows XET to perform its natural function (i.e., religation of the nascent donor end of one xyloglucan molecule to the nonreducing, acceptor, end of another xyloglucan molecule). Many glycoside hydrolases can catalyze transglycosylation reactions, although non-natural activated substrates or special low water activity conditions are typically required (Vocadlo and Withers, 2000). Curiously, some isoenzymes of XET show detectable rates of hydrolysis, whereas other isoenzymes are clearly pure transglycosylases (see Henriksson et al., 2003 and references therein). To fully understand the role of XET in the different processes of plant development, detailed enzymological studies of the individual enzymes are needed. Key questions include the mechanistic foundation of the transglycosylation versus hydrolytic activity of XET and the molecular interactions in the enzyme active site that allow efficient binding of the heavily branched xyloglucan polymers. Detailed studies of the amino acid residues involved in sugar recognition and catalysis require a high-resolution crystal structure of XET. We have reported earlier the crystallization and preliminary x-ray analysis of the enzyme XET16A from Populus tremula x tremuloides (Johansson et al., 2003). Here, we describe the three-dimensional structure of this enzyme at 1.8-Å resolution followed by a detailed structural comparison of XET with other clan B enzymes. We also present the structure of XET in complex with a xyloglucan nonasaccharide bound in the acceptor site of the enzyme.

RESULTS

Overall Structure

The native three-dimensional structure of P. tremula x tremuloides XET16A (PtxET16A, nomenclature of Henrissat et al., 1998) was determined by single isomorphous replacement with anomalous scattering using a three-site gold derivative. The model was refined to a resolution of 2.1 Å. Data collection and refinement statistics are provided in Table 1. This particular crystal form contained two molecules, A and B, in the asymmetric unit that could be superimposed with a root-mean-square deviation (RMSD) of 0.45 Å and 0.64 Å for Cα atoms and all atoms, respectively.

PtxXET16A exhibits the β-jellyroll-type structure typical of other family GH16 enzymes. However, a notable structural feature arises because of an insertion of 68 residues at the C terminus of XET (Figure 1). In other family GH16 enzymes, the C terminus is located after the final β-strand on the lower β-sheet. In PtxXET16A, however, the C-terminal extension crosses this sheet, forming the only α-helix in the molecule and an extra β-strand at the edge of the upper sheet. Two disulfide bonds stabilize this linker region. The C207-C216 linkage connects strand β14 and the extension, whereas C253-C266 connects strand β15 and the C terminus. These Cys bridges are well-conserved among known XET gene family members. Nearly all of the XET genes studied to date encode a conserved potential N-linked glycosylation site, situated 5 to 15 residues after the conserved active site residues. In PtxXET16A, the equivalent residue N93 was found to be glycosylated, and two N-acetylgalactosamine rings and one mannose residue were clearly visible in the electron density. This well-defined structure is stabilized by several hydrogen bonds to the protein (Figure 1C).

Comparison with Known Family GH16 Structures

A superposition of the three known structures of family GH16 enzymes revealed strong structural similarity despite a striking lack of sequence identity within the family. With the exception of the C-terminal linker, XET shares the common β-sandwich fold with both 1,3-1,4 β-glucanases and κ-carrageenases in family GH16 as well as the endoglucohydrolases and xylanases in family GH7. The RMSD values for the Cα atoms within the conserved portions of the structures vary between 1.3 and 1.9 Å (Table 2). The similarity to family 16 enzymes is more extensive than to family 7 as indicated by the number of Cα atoms that could be aligned in the comparison. Other glycoside hydrolases in families GH11 and GH12 also exhibit similar β-sandwich structures, but they are topologically different than PtxXET16A.

The loops connecting the strands in the conserved structural core of the clan B enzymes govern the shape and the properties of their substrate binding sites. Consistent with their action on very different substrates, the size and conformation of these loops are significantly different between the structures. The donor binding sites (negative subsite identifiers, using the nomenclature of Davies et al., 1997) of XET are much more open than those of the β-glucanases and κ-carrageenases (Figure 2). By contrast, the acceptor binding (positive) subsites are

---

### Table 2. Structural Similarities between Core Cα Atoms of PtxXET16A and Some Other Clan B Enzymes

<table>
<thead>
<tr>
<th>Structure</th>
<th>1,3-1,4 β-glucanase, 1AYH</th>
<th>κ-Carrageenase, 1DYP</th>
<th>Endoglucanase I, 1OVW</th>
<th>Cellobiohydrolase I, 1CEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSD (Å)</td>
<td>1.33</td>
<td>1.81</td>
<td>1.67</td>
<td>1.61</td>
</tr>
<tr>
<td>No. Cα atoms</td>
<td>115</td>
<td>115</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>

---

Xyloglucan Endotransglycosylase Structure 877
Figure 2. Comparison of the Known GH16 Structures.

(A) Both the lichenases (red) and the κ-carrageenases (dark blue) feature a narrower cleft than the endotransglycosylase (gray surface) with entrance loops that cover the –2 subsite of XETs and family 7 enzymes. The bound XLLG sugar (gold) has been included for reference.

(B) Superposition of PttXET16A (light blue) and the Bacillus lichenase (red). Several of the aromatics in the substrate binding cleft of XETs are conserved within family 16. Residue numbering refers to PttXET16A.

(C) Structure-based sequence alignment of PttXET16A, the Bacillus 1,3-1,4 β-glucanase 1AYH, the Pseudoalteromonas κ-carrageenase 1DYP, and three representative sequences from the different XET subgroups.
generally more open in XET than in the β-glucanases but more closed than in the κ-carrageenases.

Like many other sugar binding enzymes, the active cleft of XET is lined with several aromatic residues that can establish van der Waals interactions with potential oligosaccharide or polysaccharide substrates (Figure 2B). Some of these, such as the residue Y75 in subsite –1, which is responsible for coordinating the sugar ring undergoing nucleophilic substitution, are conserved throughout clan B (Michel et al., 2001). Even the more distantly related GH7 enzymes share a few conserved aromatics with XET. In particular, a Trp residue equivalent to W179 in the subsite +1 of XET is found in the family GH7 enzymes as well as among the lichenases. κ-Carrageenases, by contrast, feature a Trp and a Phe residue at approximately the same position in the sequence, but the chain folds back in a long loop stretching over the catalytic cleft to interact with the sugar residue bound in the +1 subsite. However, some aromatic residues in the active site cleft (e.g., Y170 and Y250) are unique to XET and might thus be specifically involved in xyloglucan recognition.

The potential –2 subsite, defined by residue W174, of PtxXET16A adopts a different side chain conformation in the two independent molecules A and B. In molecule A, the side
chain adopts a conformation leading to a packing interaction with a residue from a symmetry-related molecule. The side chain conformation observed in molecule B is the same as that observed for the structurally equivalent Trp in GH7 enzymes, where it plays a key role in the formation of the −2 site (Divne et al., 1998). Although the corresponding residue is conserved in 1,3-1,4 β-glucanases (e.g., W184 in Bacillus 1,3-1,4 β-glucanase), it is not likely to be directly involved in substrate binding because of the positioning of the long A21-R35 loop of lichenases (Figure 2A).

Both 1,3-1,4 β-glucanases and κ-carrageenases feature a calcium binding site on the convex side of the molecule, which has been suggested to enhance thermostability (Keitel et al., 1994). However, this binding site, previously predicted to be conserved in family 16 enzymes (Michel et al., 2001), has no counterpart in XET.

Active Site Structure

In spite of significant variation in their immediate surroundings, the three active site acidic residues are conserved with remarkably similar conformations in the family GH16 structures (Figure 3A). Because the function of these residues has been experimentally verified in the related enzymes, we infer that in PttXET16A, the catalytic nucleophile is E85, and the general acid/base residue is E89, both of which are found on strand β7. In the case of the putative nucleophile, the corresponding residue in Arabidopsis XET TCH4 has been mutated to a Gln, resulting in the reduction of enzyme activity by two orders of magnitude (Campbell and Braam, 1998). The OE2 of the PttXET16A putative nucleophile E85 forms a short hydrogen bonding interaction with OD1 of D87 (distance 2.5 Å). The predicted acid/base residue E89 has no hydrogen bonds with protein atoms but forms two hydrogen bonds with the hydration shell. The nucleophile in GH16 enzymes usually forms a hydrogen bond with the side chain of a residue two residues toward the N terminus. In PttXET16A, this corresponds to a hydrogen bond from NE2 of H83 to OE1 of E85 (distance 2.8 Å), which, based on sequence alignments, may be a recurring feature of the XET family of enzymes.

Complex Structures with Xyloglucan Nonasaccharide

To provide more insight into how XET interacts with its natural substrate, the structure of PttXET16A in complex with a xyloglucan nonasaccharide, XLLG (Figure 4A, nomenclature according to Fry et al., 1993), was solved and refined to a resolution of 1.8 Å (Table 1). The sugar-enzyme complex could be superimposed onto the apo-structure of PttXET16A with pairwise RMSDs of ~0.3 Å per monomer on equivalent Cα atoms. Portions of the loop connecting strands β13 and β14 slid by up to ~1 Å as compared with the apo-enzyme, causing a constriction of the binding site. Electron density maps displayed the XLLG ligand bound in the crevice between this loop and the strands β8 and β9 (Figure 3B). Three β(1→4)-linked glycosyl units (Glc1, Glc2, and Glc3) were located in the +1, +2, and +3 binding sites. A xylose residue (Xyl1) is connected via a α(1→6) linkage to Glc1, but no electron density was observed for another sugar linked from Xyl1. The second glucose, Glc2, also carries an α(1→6)-linked xylose, Xyl2, which in turn bears a β(1→2)-linked galactose (Gal1). Glc3 is the most exposed sugar and has no clear electron density for other attached sugars. The average temperature factors for the sugar rings were in the range 27 to 55 Å² when refined with full occupancy (Table 1). Even though the quality of the electron density was sufficient to identify the different sugar moieties and their linkages (Figure 3C), only six out of the possible nine sugar rings in XLLG were observed. No density for binding in the negative (donor) subsites was observed in either molecule in the asymmetric unit nor were we successful in soaking smaller oligosaccharides into these sites.

Figure 4. The Two Xyloglucan-Derived Oligosaccharides Used.

(A) XLLG. The oligosaccharide found in the crystal structure is assumed to be identical to the six sugars within the dashed line.
(B) XLLG-CNP.
The core glycosyl residues of the ligand adopt an extended conformation that results in the formation of a hydrogen bond between the ring oxygen and the 3-hydroxyl group of the preceding sugar. Glc1 is the best-determined sugar and is involved in a network of hydrogen bonds with several protein side chains (Figure 3B). E89, the predicted general acid/base residue, interacts with both the 3- and 4-OH groups of Glc1. The indole ring of W179 plays a role in the formation of three sugar binding sites. As well as forming stacking interactions with glucosyl units in the +1 and +2 sites, it has a conformation that allows it to donate a hydrogen bond to the first xylose (Xy1), which in turn also forms a hydrogen bond with the protein side chain of D178. The linkages from Glc2 to Gal1 are such that Glc2 is sandwiched between Gal1 and W179 and forms no hydrogen bonds to the protein. Glc3 has few interactions with the protein: one direct hydrogen bond to the carbonyl of G183 and one mediated by a water molecule to R116. The three hydroxyl groups of Xy1 are each involved in a single hydrogen bond; two of them interacts with protein side chains, whereas the third interacts with the 3-OH of Glc2. The linkage to Glc1 is such that only one face of the sugar is exposed to solvent. Xyl2 lacks direct hydrogen bonds to the enzyme but is firmly held in place, slotting into a volume defined by the interacting side chains of R116, Y250, R258, E114, and Q102. Gal1 forms two protein hydrogen bonds and has one buried and one solvent-accessible face, resulting in a net loss of $\sim150 \AA^2$ of solvent-accessible surface area, compared with a fully extended conformation. The bound ligand, tightly wedged into the $-1$ to $-3$ subsites, results in a total accessible surface loss of $\sim410 \AA^2$ for both the oligosaccharide and the protein.

Although the identity of Glc1 is ambiguous because of the lack of clear electron density for three of the nine sugar rings in the ligand, the solvent accessibility of the Glc3 site suggests that Glc1 corresponds to the fourth glucose residue of XLLG (Figure 4A), counting from the reducing end of the molecule in the ligand structure.

### 2-Chloro-4-Nitrophenyl XLLG Does Not Act as a Donor for the XET Reaction

To overcome limitations in currently available assays (Fry et al., 1992; Sulová et al., 1995) and to explore substrate binding to the donor subsites of XET, we synthesized an XLLG bearing a chromogenic aglycone, 2-chloro-4-nitrophenyl XLLG (XLLG-CNP) (Figure 4B). However, no steady state release of the chromophoric 2-chloro-4-nitrophenylate group was observed under conditions in which XLLG-CNP was incubated with high concentrations of PttXET16A (Table 3, Reactions A and B). Furthermore, under these conditions, no burst of the phenylate was observed as would be expected if XLLG-CNP were cleaved to form a covalent glycosyl-enzyme intermediate. However, XLLG-CNP was able to act as a transglycosylation acceptor when high $M_r$ xyloglucan (XG) was used as a donor substrate (Figure 5). Nevertheless, release of the CNP group from these molecules was never observed by visible spectrophotometry during the course of the reaction (Table 3, Reactions C and D). Binding to the acceptor subsites was confirmed by determining the XET/XLLG-CNP complex structure, which was essentially identical to the XET/XLLG complex. Specifically, well-ordered electron density from several sugar residues of XLLG-CNP was observed in the acceptor subsites, whereas saccharide binding in the donor subsites was not observed (data not shown).

### Molecular Modeling of a Potential Xyloglucan Donor Saccharide

In the absence of suitable oligosaccharide binding to the donor subsites of XET, molecular modeling was used instead to provide insight into potential donor binding interactions. Because the $-1$ subsite is relatively well conserved in clan B, it is instructive to compare the XET-XLLG structure with other clan members for which more extensive structural data is available. Sulzenbacher et al. (1996) described the structure of *Fusarium oxysporum* endoglucanase I in complex with a nonhydrolyzable substrate analog bound in sites $-2$ to $+1$ (Protein Databank [PDB] code 1OYV). The structure of the glycosyl unit in the $-1$ site was later used with several substrate complexes from active site mutants of *Trichoderma reesei* Cel7A to create a mosaic model of cellulose spanning the active site tunnel from sites $-7$ to $+2$ (Divne et al., 1998; PDB code 8CEL). Superposition of the catalytic core of the cellubiohydrolase model and the XET-XLLG complex results in a substantial overlap of the glucose residues in the $+1$ subsites. Glc1 of XLLG is shifted away from the nucleophile by $\sim0.8 \AA$ when compared with the cellulose counterpart (Figure 3D). The distorted glucose ring modeled in the $-1$ site of TrCel7A can be easily accommodated in PttXET16A and could form equivalent interactions with the enzyme. Somewhat surprisingly, the cellulose conformation observed in the $-2$ site of TrCel7A also can be fitted readily to the PttXET16A structure to produce interactions with the side

### Table 3. XET Reaction Conditions

<table>
<thead>
<tr>
<th>Reaction</th>
<th>PttXET16A Concentration (g/L)</th>
<th>XLLG-CNP Concentration (mM)</th>
<th>XG Concentration (g/L)</th>
<th>Reaction Time (h)</th>
<th>Transglycosylation Observed</th>
<th>CNP Release Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2.6</td>
<td>2.0</td>
<td>0</td>
<td>15</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>B</td>
<td>2.6</td>
<td>10.0</td>
<td>0</td>
<td>15</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>C</td>
<td>0.58</td>
<td>0.2</td>
<td>2.5</td>
<td>15</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>D</td>
<td>0.58</td>
<td>2.0</td>
<td>2.5</td>
<td>15</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

*Other reaction conditions were as specified in Methods. XG, xyloglucan.*
chains of W174 and S172. Subsite –3, on the other hand, is not as easily predicted, in part because of the decreasing similarity between XET and the family 7 enzymes but also because of the extensive branching of the xyloglucan molecule. However, the modeling does provide an explanation for why XET preferentially cleaves the glycosidic bonds of unsubstituted glucose units in xyloglucan and allows branched sugars in other sites. In the distorted sugar in the –1 subsite, the 6-OH points toward the protein to form hydrogen bonds with the acid/base residue and with W174. However, in the –2 subsite, the 6-OH, where a xylose branch would normally occur in many xyloglucans, is solvent accessible.

DISCUSSION

PttXET16A has an overall structure that is typical for GH16 enzymes except for the long C-terminal linker that crosses the convex surface and forms a short additional β-strand on the concave side of the molecule. This extension of the acceptor binding site, together with variations in length and conformation of loops connecting the strands in the β-sandwich, produce an active site that is unique to the XET family of enzymes. In particular, the loop that connects strands β13 and β14 folds back toward the concave surface of the β-sandwich to create a more constricted binding site where a xyloglucan-derived nonasaccharide and its CNP β-glycoside derivative were observed to bind. We see only the central core of these oligosaccharides, which correspond to the sugar rings making the closest contacts with the enzyme. The xyloglucan core takes on a unique conformation whereby the branched rings are packed onto the more extended glucan backbone, causing a significant loss in solvent-accessible surface area. W179 is a key residue in this binding site, interacting with three of the six glycosyl rings by hydrogen bond and stacking interactions. The ligand core, as a whole, forms three internal hydrogen bonds and makes eight more with the protein. According to sequence alignments (Figure 2C), all of these ligand–protein interactions are conserved within the XET family. Most importantly, the side chain of the putative acid/base residue E89 forms a hydrogen bond with the 4-hydroxyl group of the glycosyl unit in the +1 site. A close interaction between these two moieties is an essential requirement for both the formation and the breakdown of the glycosyl-enzyme intermediate. Few sugar structures with such a diverse pattern of linkages have been previously solved by x-ray crystallography. In addition to the extensive set of intrasaccharide and saccharide–protein hydrogen bonds, the complex is noteworthy for the unique packing of Glc2 that is sandwiched between the protein and another sugar ring.

It has been shown previously that the enzymatic removal of the N-linked glycan attached to XETs results in loss of transglycosylase activity to varying degrees (Campbell and Braam, 1998, 1999; Henriksson et al., 2003). By contrast, PttXET16A retains enzymatic activity upon deglycosylation with Streptomyces plicatus endoglycosidase H (EC 3.2.1.96) (Å.M. Kallas, S.E. Denman, K. Piens, H. Henriksson, J. Fält, P. Johansson, T.A. Jones, H. Brumer III, and T.T. Teeri, unpublished data). It is therefore doubtful that N-glycosylation plays a direct role in catalysis, but it may directly influence XET folding and stability as a result of several well-ordered interactions with the protein.

XET-catalyzed transglycosylation most likely occurs through a double-displacement mechanism similar to that outlined in Figure 6. As with all retaining glycosidases, the covalent enzyme intermediate can potentially be decomposed by transfer of the glycosyl unit to either water or another suitable (typically hydroxylated) acceptor molecule. PttXET16A, like many XET enzymes for which comparative biochemical data are available, is a strict transglycosylase and does not hydrolyze xyloglucan to a measurable extent (Å.M. Kallas, S.E. Denman, K. Piens, H. Henriksson, J. Fält, P. Johansson, T.A. Jones, H. Brumer III, and T.T. Teeri, unpublished data). This poses an interesting question: Why is glycosyl transfer to xyloglucan oligosaccharides or polysaccharides the dominant reaction under in vitro assay conditions in which water is present as an alternate acceptor at 55 M? Clearly, the PttXET16A glycosyl-enzyme intermediate is sufficiently long-lived that the departing xyloglucan chain

Figure 5. XET-Catalyzed Incorporation of XLLG-CNP into High M₆ Xyloglucan Monitored by High-Performance Size-Exclusion Chromatography.

(A) Final product analysis of Reaction C, Table 3. RI, refractive index. (B) Final product analysis of Reaction D, Table 3. Solid line, UV detection; dotted line, refractive index (RI) detection; dashed vertical line, approximate column exclusion limit for pullulan standards (approximate M₆ 47,300).
produced in the first reaction step is able to diffuse away and be replaced by a new sugar before enzymic deglycosylation (Figures 6C to 6E). In glycoside hydrolases, this intermediate may typically only be forced to accumulate through the use of artificial substrate analogs (Withers et al., 1987), sometimes in combination with active site mutants (Vocadlo et al., 2001). Because XLLG-CNP does not act as a donor substrate, we infer that binding of sugars in the acceptor site is a prerequisite for the first step, and possibly both steps, of catalysis. The oligosaccharide structure that we observed in the acceptor sites appears to be well placed relative to the catalytic machinery for participation in the glycosyl transfer step (Figures 6E and 6F).

The catalytic residues of GH families 16 and 7 are localized on a single β-strand such that they are displayed on the concave surface of the β-sandwich. Their conformations are highly conserved, but the putative nucleophilic residue in PttXET16A,
The gene encoding PttXET16A from *P. tremula x tremuloides* was cloned into a pPIC9 shuttle vector (Sterky et al., 1998), the gene was cloned into a pPIC9 shuttle vector containing an α-factor secretion signal and the alcohol oxidase promoter. Active recombinant poplar PttXET16A protein was purified from culture filtrates by a sequential combination of strong cation-exchange gel filtration and strong cation-exchange chromatography steps. The full details of the protein expression, purification, and characterization will be published elsewhere. The first trigonal crystal form was obtained at 293K in group P63, with unit cell dimensions a and b = 188.7 Å and c = 46.1 Å and a calculated Matthews coefficient (Matthews, 1968) of 4.0 Å³ D⁻¹, corresponding to a solvent content of 68%.

Expression, Purification, and Crystallization of Native PttXET16A

The crystallographic refinement was initially performed in the CNS package (Brünger et al., 1998), using a simulated annealing-slowcool protocol while maintaining strict twofold NCS constraints. In subsequent refinements, the NCS restraints were released using REFMAC5 (Murshudov et al., 1997). After each cycle of refinement, the α,β-weighted (Read, 1986) 2Fobs – Fcalc and Fobs – Fcalc maps were used for further model rebuilding in O. The first water molecules were added conservatively to peaks >2σ of the 2Fobs – Fcalc electron density map that made at least one hydrogen bond with a protein atom or another water molecule. In the final stages, the α cutoff in CNS Water-Pick (Brünger et al., 1998) was reduced to 1.5σ, and water molecules with a B-factor >60 Å² were removed. Molecular replacement solution of the XET-XLLG complex structure was preformed using the AMoRe package (Navaza, 1994).

Crystallographic refinement was performed in the CNS package (D. Madsen, P. Johansson, and G.J. Kleywegt, unpublished data). Molecular figures were made in biological systems. Molecular figures were made in O and rendered in Molray (Harris and Jones, 2001). MOLMOL (Koradi et al., 1996) was used to make solvent accessible-surface calculations. Coordinates and structure factors of the complex and apo-enzyme have been deposited at the PDB with codes 1UMZ and 1UN1, respectively.

Synthesis of XLLG-CNP

XLLG was produced as described previously (Henriksson et al., 2003). XLLG-CNP was synthesized from the nonasaccharide XLLG by a sequence of per-O-acetylation, C-1 α-bromination, König-Knorr glycosylation, and Zemplén deprotection, the details of which will be published elsewhere.

Enzyme Reactions and Product Analysis

Tamarind xyloglucan (Mn = 202 kDa, 35:45:16:4 Xyl/Glc/Gal/Ara) was obtained from Megazyme (Bray, Ireland). All assays were performed at 22.0°C ± 0.1°C in 12.5 mM NaOAc, pH 5.6. The release of 2-chloro-4-nitrophenolate was monitored at 405 nm (Claeyssens and Aerts, 1992) on a Cary 300 Bio UV-Vis spectrophotometer (Varian, Solna, Sweden). High-performance size-exclusion chromatography was performed in DMSO on a Tosoh Biosep TSK gel G3000 HHR column (7.8 cm × 30 cm; Tosoh, Tokyo, Japan) using detection by refractive index and UV light at 300 nm. Before high-performance size-exclusion chromatography analysis, 100-µL samples were concentrated to dryness in vacuo and
redissolved in 100 µL of high-performance liquid chromatography–grade DMSO. M_w values of xylogluco-oligosaccharides were estimated from a standard line of log(M_w) versus retention time for pullulan standards (M_w range 180 to 112,000). The colorimetric assay of Farkas and coworkers was used for routine XET activity measurements (Sulova et al., 1995).

Sequence data from this article have been deposited with the EMBL/GenBank data libraries under accession number AF515607.

ACKNOWLEDGMENTS

The authors would like to thank Yrgeve Cerenius at beamline I711 (MaxLab) and Raimond Ravelli and Sigrid Kozielecki at beamlines ID14-EH4 and ID14-EH1 (ESRF) for support during data collection. Kathleen Piens and Jerry Ståhlberg are thanked for fruitful discussions. The present work was supported by the Knut and Alice Wallenberg Foundation through the Wallenberg Wood Biotechnology Center.

Received December 16, 2003; accepted January 26, 2004.

REFERENCES


Crystal Structures of a Poplar Xyloglucan Endotransglycosylase Reveal Details of Transglycosylation Acceptor Binding
Patrik Johansson, Harry Brumer III, Martin J. Baumann, Åsa M. Kallas, Hongbin Henriksson, Stuart E. Denman, Tuula T. Teeri and T. Alwyn Jones
Plant Cell 2004;16;874-886; originally published online March 12, 2004; DOI 10.1105/tpc.020065

This information is current as of July 8, 2017