The EMBO Journal vol.13 no.11 pp.2493–2501, 1994

Three-dimensional structure of endo-1,4-β-xylanase II from *Trichoderma reesei*: two conformational states in the active site

Anneli Törrönen, Anu Harkki[1] and Juha Rouvinen[2]

Department of Chemistry, University of Joensuu, PO Box 111, FIN-80101 Joensuu, Finland and [1]Cultor Ltd Technology Center, FIN-02460 Kantvik, Finland
[2]Corresponding author

Communicated by J.Janin

The three-dimensional structure of endo-1,4-β-xylanase II (XYNII) from *Trichoderma reesei* has been determined by X-ray diffraction techniques and refined to a conventional *R*-factor of 18.3% at 1.8 Å resolution. The 190 amino acid length protein was found to exist as a single domain where the main chain folds to form two mostly antiparallel β-sheets, which are packed against each other in parallel. The β-sheet structure is twisted, forming a large cleft on one side of the molecule. The structure of XYNII resembles that of *Bacillus* 1,3–1,4-β-glucanase. The cleft is an obvious suggestion for an active site, which has putative binding sites for at least four xylose residues. The catalytic residues are apparently the two glutamic acid residues (Glu86 and Glu177) in the middle of the cleft. One structure was determined at pH 5.0, corresponding to the pH optimum of XYNII. The second structure was determined at pH 6.5, where enzyme activity is reduced considerably. A clear structural change was observed, especially in the position of the side chain of Glu177. The observed conformational change is probably important for the mechanism of catalysis in XYNII.

Key words: conformational change/crystal structure/*Trichoderma*/xylanase

Introduction

Xylan is the most abundant hemicellulose in plants, accounting for as much as 35% of the dry weight of higher plants. Xylans are heteropolysaccharides, the backbone of which is composed of ~150–200 β-1,4-linked xylopyranose units. The degree of substitution of xylan varies, depending on its botanical origin. The substituents, like arabinose, 4-*O*-Me-D-glucuronic acid and acetic acid, are bound covalently to the xylopyranose units. The action of several hemicellulolytic enzymes is needed to hydrolyze xylan completely. During the last few years, the enzymatic hydrolysis of xylan has raised significant interest because of its applications, for example, in biobleaching and the food and feed industry (Biely, 1985; Poutanen *et al.*, 1987; Viikari *et al.*, 1990; Wong and Saddler, 1992; Gilbert and Hazlewood, 1993).

The enzymes acting on the backbone of xylans are exo-1,4-β-xylosidases (EC 3.2.1.37) and endo-1,4-β-xylanases (EC 3.2.1.8). Endoxylanases are capable of hydrolyzing the internal 1,4-β-bonds of the xylan backbone and thereby produce several xylo-oligomers of varying length. About 20 amino acid sequences of different low-molecular-weight xylanases of bacterial and fungal origin have been published. The comparison of available β-glycanase sequences by hydrophobic cluster analysis (HCA) has revealed eight families (A–H) (Gilkes *et al.*, 1991; Henrissat, 1991; Törrönen *et al.*, 1993a). So far, X-ray structures of five glycosidic enzymes have been reported: the catalytic domain of cellobiohydrolase II from *Trichoderma reesei* (family B) (Rouvinen *et al.*, 1990), the catalytic domain of endoglucanase E2 from *Thermomonospora fusca* (family B), endoglucanase CelD from *Clostridium thermocellum* (family E) (Juy *et al.*, 1992), endoglucanase V from *Humicola insolens* (Davies *et al.*, 1993) and *Bacillus* H(A16-M) hybrid 1,3–1,4-glucanase (Keitel *et al.*, 1993).

Low-molecular-weight xylanases belong to the β-glycanase family G, and their amino acid sequences are highly conserved (Törrönen *et al.*, 1993a). The crystallization of seven different low-molecular-weight xylanases has been reported (Moriyama *et al.*, 1987; Rose *et al.*, 1987; Golubev *et al.*, 1993; Pickersgill *et al.*, 1993; Törrönen *et al.*, 1993b; Viswamitra *et al.*, 1993). However, the only structural information published so far concerns the folding pattern of IPO endo-1,4-β-xylanase from *Bacillus pumilus* (Arase *et al.*, 1993).

The filamentous fungus *T.reesei* is known to be one of the most effective producers of polysaccharide hydrolyzing enzymes (Kubicek *et al.*, 1993). The two major inducible homologous endo-β-1,4-xylanases (EC 3.2.1.8) secreted by *T.reesei* are XYNI and XYNII (Törrönen *et al.*, 1992). They are both small protein molecules with molecular masses of 19 and 21 kDa, and consisting of 178 and 190 amino acids, respectively. The isoelectric point is 5.2 for XYNI and 9.0 for XYNII. The pH optimum and product distribution for XYNI and XYNII are different. XYNII alone represents >50% of the xylanolytic activity of the culture filtrate of *T.reesei* grown on xylan. We have recently crystallized both enzymes (Törrönen *et al.*, 1993b) and now report the three-dimensional structure of XYNII.

Results

Crystallography

The three-dimensional structure of XYNII has been solved using the multiple isomorphous replacement method augmented with anomalous scattering information. The final model contains two xylanase molecules (called A and B) in an asymmetric unit. All the 190 amino acid residues and 1480 atoms for both enzyme molecules are included in the model. The first residue in both molecules is pyrrolidone carboxylic acid. The number of water molecules in the final model is 344 (0.9/residue). The final model contains a total of 3304 atoms. The r.m.s. fit between the two molecules,

© Oxford University Press

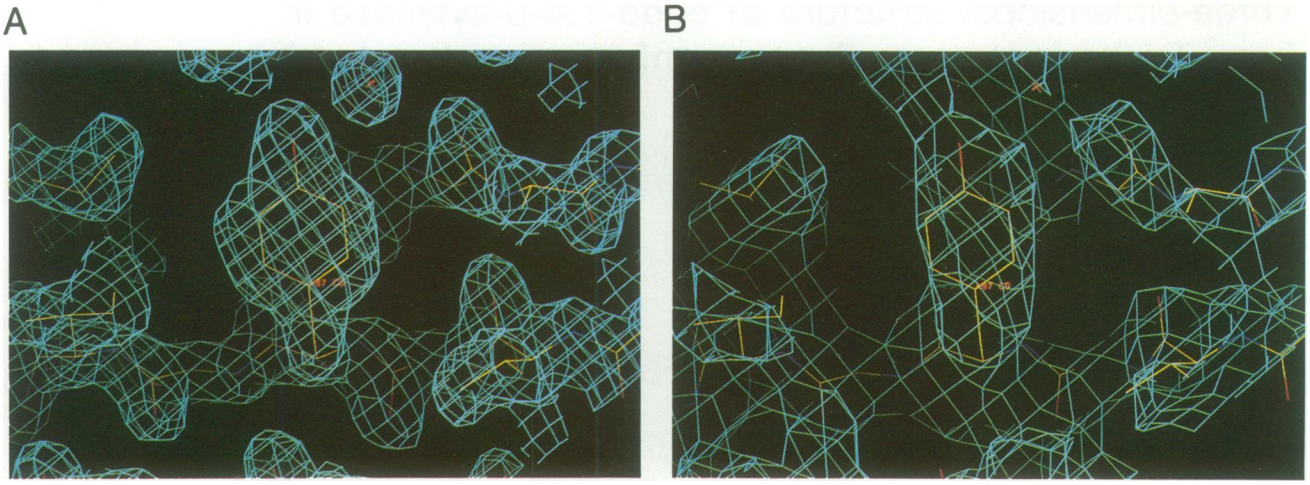

Fig. 1. (**A**) A portion of the final model with electron density (coefficients $2F_{obs}-F_{calc}$, α_{calc}, contoured at 1σ level) around Tyr87. (**B**) The same region of the original (unaveraged) 2.8 Å MIR map with the refined model.

A and B, is 0.231 Å (calculated from the positions of Cα atoms). The *R*-factor is 18.3% for 24 911 reflections with $|F| \geq 1\sigma\,|F|$ between 8.0 and 1.8 Å (75% complete). The overall geometry is very good (0.008 Å r.m.s. distances, 2.5° r.m.s. angles and 0.9° r.m.s. for fixed dihedrals). The real-space correlation coefficient (Jones *et al.*, 1991), calculated for the final model using all atoms as a function of residue, showed good values. The coefficient was smallest for residues 164–165 in molecule A (0.64). In molecule B, these residues had a much stronger electron density. Asp140 was the only residue which fell outside the allowed regions on the Ramachandran map. A portion of the final electron density map together with the corresponding MIR map is shown in Figure 1.

The second model for XYNII was calculated using the data from crystal soaked in a solution containing xylose at pH 6.5. The final model contains 3304 atoms. The *R*-factor is 18.4% for 17 684 reflections with $|F| \geq 1\sigma\,|F|$ between 8.0 and 1.8 Å (53.7% complete, 84.1% complete at 2.5 Å resolution). The average deviation from the ideal in bond lengths is 0.009 Å, in bond angles 2.7° and in fixed dihedral angles 0.9°. The r.m.s. difference between the two XYNII models is 0.148 Å and was calculated from the positions of all Cα atoms in the molecule A.

Molecular structure

The schematic drawing and naming of the secondary structure elements are presented in Figure 2A. The overall shape of this globular protein is ellipsoidal, having approximated dimensions of 32 × 34 × 42 Å. XYNII has 190 amino acids; it is a single-domain polypeptide containing two β-sheets (A and B) and one α-helix (Figure 3). The structure contains a total of 15 β-strands. All β-strands have antiparallel hydrogen bonding, except B6b and B7, which have mixed bonding. Both β-sheets A and B twist and form a cleft on one side of the protein. The hydrophobic faces of these β-sheets are packed against each other in parallel in a sandwich-like manner, forming the hydrophobic core of the protein. The hydrophilic face of β-sheet A forms a flat surface with many serine and threonine residues, and is accessible to solvent. The hydrophilic face of β-sheet B makes a surface for the cleft.

β-Sheet A is composed of five antiparallel β-strands. The

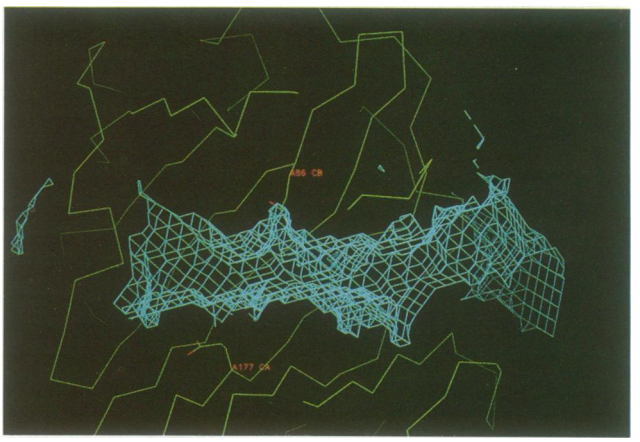

Fig. 4. Solvent-accessible surface of the cleft in XYNII structure. Two catalytic residues Glu86 and Glu177 are shown.

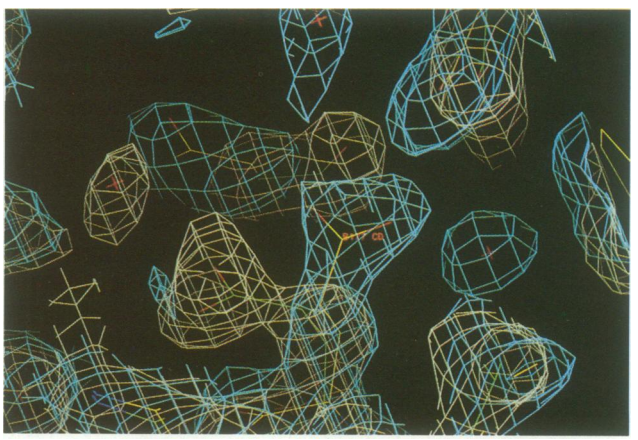

Fig. 5. The observed change in Glu177 position. The electron density (coefficients $2F_{obs}-F_{calc}$, α_{calc}, contoured at 1σ level) around Glu177 for a 'pH 5.0' model (brown) and for a 'pH 6.5' model (cyan).

putative β-strand A1 is missing in XYNII, but may exist in some other homologous xylanases. Strands A2, A3 and A6 are shorter (four or five residues) than the middle strands A4 and A5 (nine and 11 residues). The short strands are quite planar, whereas the long strands are twisted. Four

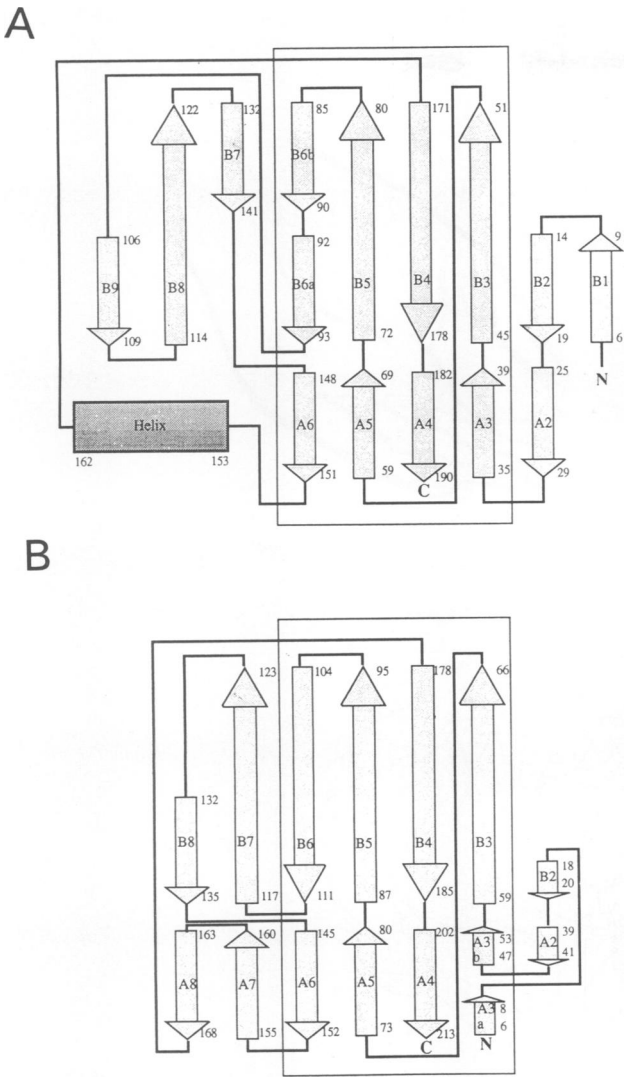

Fig. 2. The schematic drawing of the secondary structure elements (**A**) XYNII from *T.reesei* and (**B**) *Bacillus* 1,3−1,4-β-glucanase. The arrows mark β-strands and cylinder α-helix. The structure elements of 1,3−1,4-β-glucanase are adopted from the work of Keitel *et al.* (1993), from which the small helix between strands B4 and A4 is omitted for clarity. The corresponding elements between two proteins are numbered in a similar way. The common central folding pattern in both proteins is boxed.

strands are connected at the same end of β-sheet A to β-sheet B. These loops have different lengths (2−5 residues). The connection between A6 and B6 is interrupted by a long insertion.

β-Sheet B contains β-strands B1−B9. Strands B1, B2 and B9 are short (four, six and four residues), and B3, B4, B5 and B8 are longer (seven, eight, nine and nine residues). One residue divides unit B6 into two shorter strands, B6a and B6b, which are twisted ~90°. There is also a twist of ~90° in the middle of the long β-strands B3, B4 and B5. The twisted part of β-sheet B has been described as a separate sheet in xylanase from *B.pumilus* (Arase *et al.*, 1993). The insertion between strands B6 and A6 contains three β-strands: B7, B8 and B9. The only α-helix is located between β-strands A6 and B4, and is packed against the hydrophobic face of β-sheet B.

The overall structure has the shape of a 'right hand' (Figure 3A and B), where the two β-sheets (A and B) form 'fingers' and the twisted part of β-sheet B and the α-helix form a 'palm' (Figure 3A). The long loop of the nine residues between β-strands B7 and B8 makes a 'thumb', which partly closes the cleft. A striking feature is a 12 residue long irregular loop between β-strands B6a and B9. A part of this, residues 96−102 (Tyr−Asn−Pro−Ser−Thr−Gly−Ala), form a 'cord', which partly closes the cleft on one side. The residues in the cord have a clear electron density, indicating that this unit has a well-defined structure. No hydrogen bonds between the cord and the rest of the molecule were observed, but some hydrophobic contacts were obvious.

XYNII contains no cysteine residues, so there are no disulfide bridges. However, there are eight negatively charged (aspartate and glutamate) residues and 13 positively charged (lysine, arginine and histidine) residues in the protein molecule. These form ionic bridges: between Asp170 and Arg81, Glu91 and Arg141/Arg145, Glu107 and Lys104/Arg142, and Asp116 and Lys104/Arg142. None of these residues is completely conserved. Lys49 has an important structural role. It places a positively charged nitrogen in a position inside the protein and makes hydrogen bonds with the main chain carbonyl oxygens of Asn10, Gly30, Gly32 and Gln34.

The active site cleft and conformational change

The shape of the cleft was estimated by using the solvent-accessible surface (Voointholt *et al.*, 1989) (Figure 4). The active site is likely to be located in this cleft, which has a length and depth of ~25 and 9 Å, respectively. The average width of the cleft is ~4 Å, but there are two regions where the width is smaller: one between Trp18 and Pro126 in the middle of the cleft, and the other between Tyr96 and Tyr179 at the end of the cleft.

XYNII was crystallized at pH 5.0, which also corresponds to the pH optimum of the enzyme (Tenkanen *et al.*, 1992; Törrönen *et al.*, 1992). In order to get a complex structure, crystals were soaked in a solution containing xylose at pH 6.5. At this pH, the activity of XYNII is considerably reduced and xylose binding was expected to be better. Clear structural changes for some residues in the cleft were observed in both molecules A and B (Figure 5). Unfortunately, no clear electron density was found for the ligand, probably due to the low binding affinity of xylose. Later, the effect of pH was studied by collecting a new data set at 2.0 Å resolution at pH 6.5, but now without xylose (data not shown). Again, similar conformational change was observed.

The largest change was observed for Glu177, where the torsion angle between Cβ and Cγ atoms was changed ~100° (from −89° to 171°). This places the carboxyl part of the side chain in a different position, the displacement being 2.8 Å calculated from the positions of the Cδ atom. In addition, a change in the conformation of the side chain of Tyr73 and Tyr88 can be detected. The bond between the atoms Cα and Cβ of Tyr73 rotates ~78° (from 55° to −23°), and the bond between the atoms Cβ and Cγ of Tyr73 ~47° (from 82° to 129°). Both rotations cause a change of ~4.0 Å in the position of the Oη atom of Tyr73. The position of the side chain of Tyr88 is less affected. The bond between the atoms Cα and Cβ of Tyr88 rotates ~9° (from −58° to −67°) and the bond between the atoms Cβ and Cγ of Tyr88 ~12° (from 83° to 71°). These rotations lead to an ~0.9

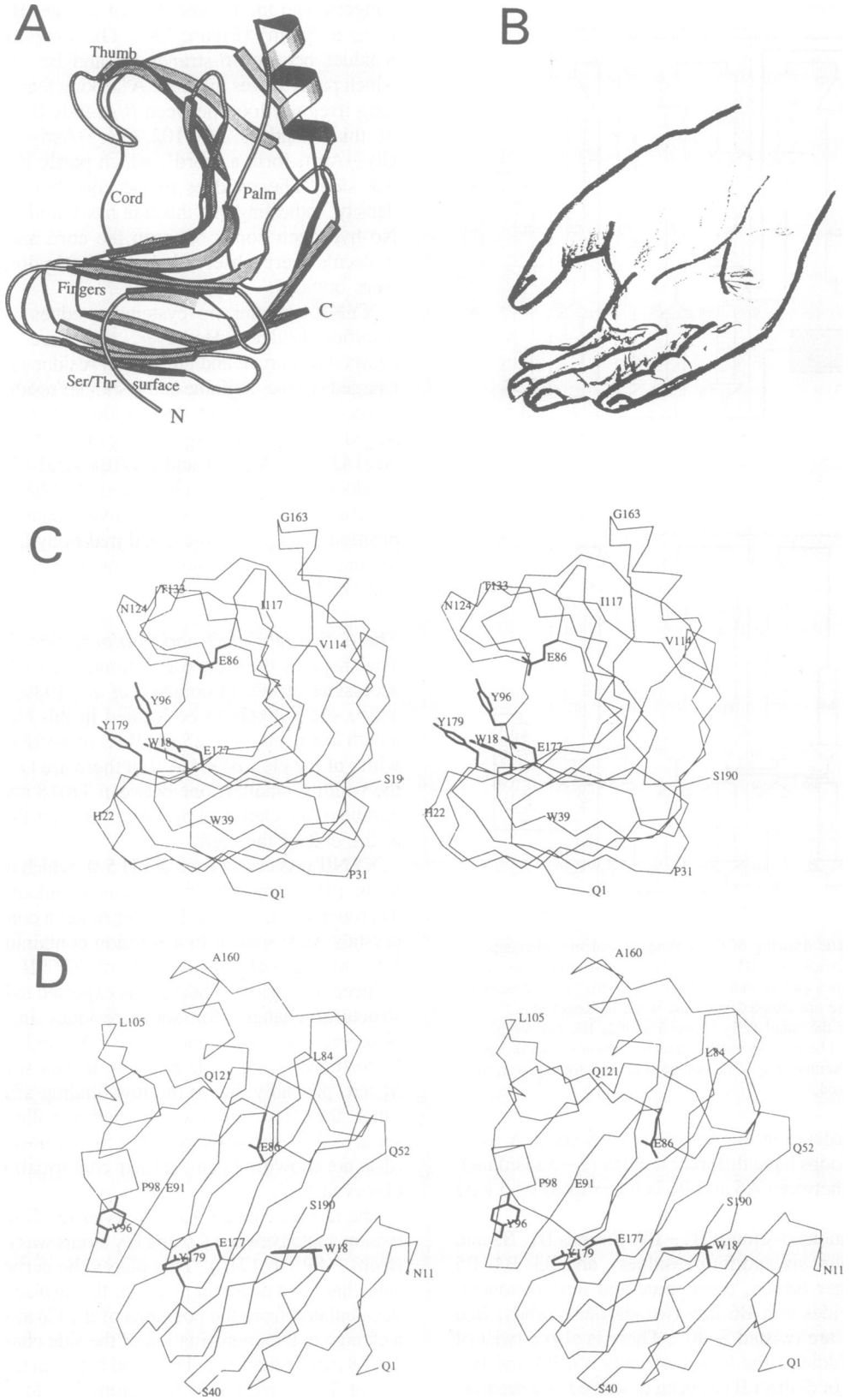

Fig. 3. (**A**) Ribbon representation of the XYNII molecule showing an α-helix and β-strands. The structure is reminiscent of the shape of a 'right hand' (**B**) and this analogy is shown. The picture was generated with the MolScript program (Kraulis, 1991). (**C**) The Cα skeleton of XYNII in stereo. The putative catalytic residues (Glu86 and Glu177) and aromatic binding residues (Tyr96, Tyr179 and Trp18) are shown. (**D**) As in (C), but rotated along the *y*-axis ~90°.

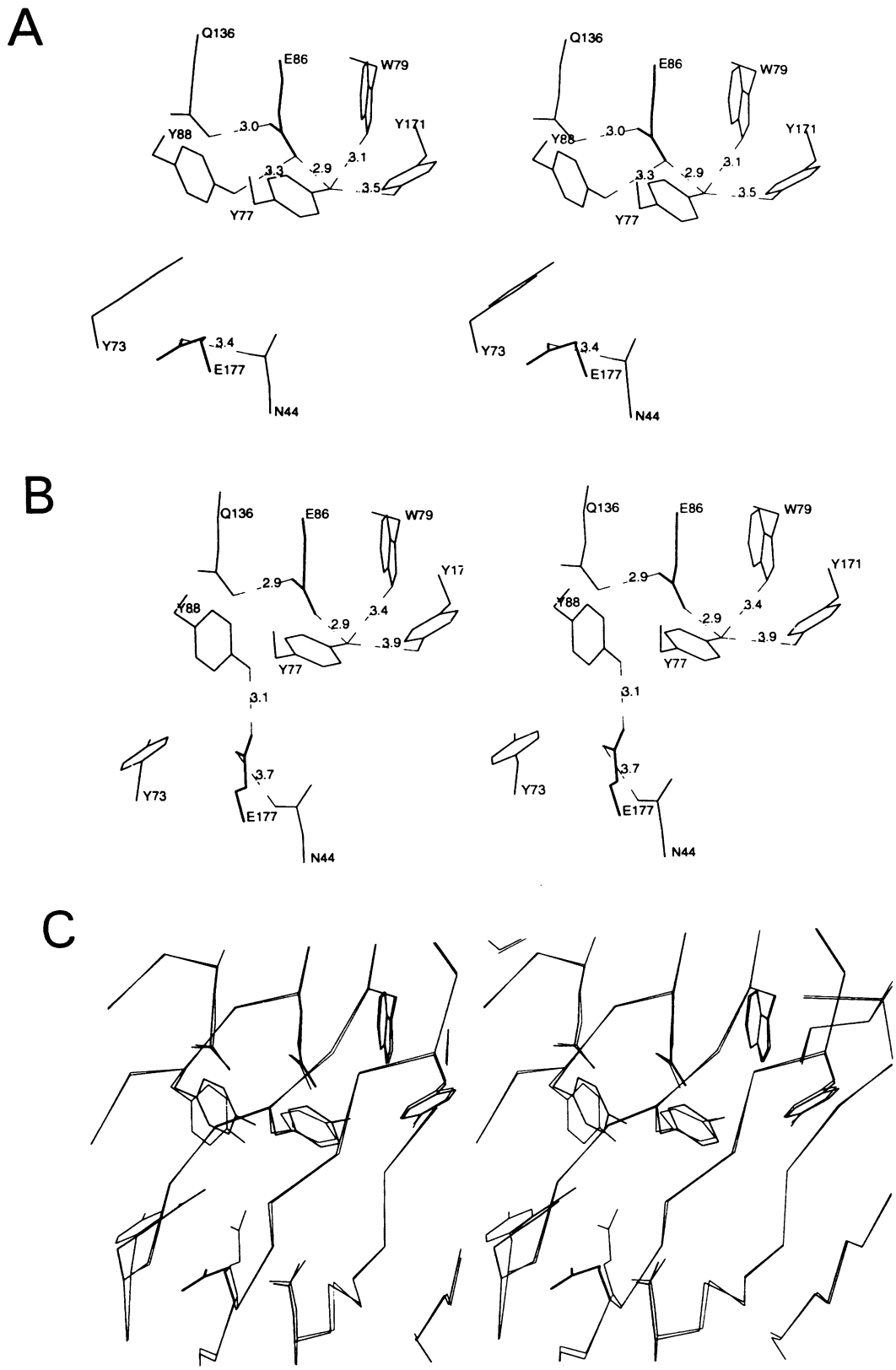

Fig. 6. Hydrogen bond network around Glu86 and Glu177 at (**A**) pH 5.0, (**B**) pH 6.5 and (**C**) both together with the main Cα chain.

Å change in the position of Oη of Tyr88. In spite of the ~0.6 Å change in the position of the 'cord', the differences between the two structures elsewhere in the protein molecule were small, close to the experimental errors.

The hydrogen bonds of Glu86 and Glu177

In previous structural studies it has been suggested that Glu93 and Glu182 are the catalytic residues in *B.pumilus* IPO xylanase (corresponding to Glu86 and Glu177 in XYNII)

```
Ak                                           SAGINYVQNYNGNLADFTYDES.A    23
TrI                                           ASINYDQNYQT.GGQVSYSPSNT    22
Cla  ATNLNTTESTFSKEVLSTQKTYSAFNTQAAPKTITSNEIGVNGGYDYELWKD.YGNTSMTLKNG    63
Bc                                       ASTDY.....WQNWTDGGGIVNAVNGSG    23
Bp                                  RTITNNEMGNHSGYDYELWKD.YGNTSMTLNNG    32
Sl                              ATTITTNQTGTDGM....YYSFWTDGGGSVSMTLNGG    33
TrII                                 QTIQPGTGYNNGYFYSYWNDGHGGVTYTNGPG    32
                                          ===>    =====>     ====>
                                           B1        B2         A2

Ak   GTFSMYWEDGVSSDFVVGLGW..TTGSSN......AISYSAEYSASGSSSYLAVYGWVNYPQA     78
TrI  G.FSVNWNTQ..DDFVVGVGW..TTGSSA......PINFGGSFSVNSGTGLLSVYGWSTNPLV     74
Cla  GAFSCQWSN..IGNALFRKGKKFNDTQTYKQLGNISVNYDCNYQPY.GNSYLCVYGWTSSPLV    123
Bc   GNYSVNWSNT..GNFVVGKGW..TTGSPFR.....TINYNAGVWAPNGNGYLTLYGWTRSPLI     77
Bp   GAFSAGWNN..IGNALFRKGKKFDSTRTHHQLGNISINYNASFNPG.GNSYLCVYGWTQSPLA     92
Sl   GSYSTQWTN..CGNFVAGKGW..STGDG.......NVRYNGYFNPV.GNGYGCLYGWTSNPLV     84
TrII GQFSVNWSN..SGNFVGGKGW..QPGTKNK.....VINFSGSYNPN.GNSYLSVYGWSRNPLI     85
       ====>       ======>              ==========>   ========>    =
         A3          B3                      A5           B5

     *
Ak   EYYIVED.YGDYNPCSSATS.LGTVYSDGSTYQVCTDTRTNEPSITGT.STFTQYFSVRES    136
TrI  EYYIMEDNH.NY.PA.QGTVK.GTVTSDGATYTIWENTRVNEPSIQGT.ATFNQYISVRNS    130
Cla  EYYIVD.SWGSWRPP.GGTSK.GTITVDGGIYDIYETTRINQPSIQGN.TTFKQYWSVRRT    180
Bc   EYYVVD.SWGTYRP..TGTYK.GTVKSDGGTYDIYTTTRYNAPSIDGDRTTFTQYWSVRQS    134
Bp   EYYIVD.SWGTYRP..TGAYK.GSFYADGGTYDIYETTRVNQPSIIGI.ATFKQYWSVRQT    148
Sl   EYYIVD.NWGSYRP..TGTYK.GTVSSDGGTYDIYQTTRYNAPSVEGT.KTFQQYWSVRQS    140
TrII EYYIVE.NFGTYNP.STGATKLGEVTSDGSVYDIYRTQRVNQPSIIGT.ATFYQYWSVRRN    143
     ====>  =>             ===>    ========>          ==========>
     B6b    B6a  'cord'     B9         B8      'thumb'      B7

                                          *
Ak   TRTSGTVT...VANHFNFWAQHGFGNSDFN.YQVMAVEAWSGAGSASVTISS             184
TrI  PRTSGTVT...VQNHFNAWASLGLHLGQMN.YQVVAVEGWGGSGSASQSVSN             178
Cla  KRTSGTIS...VSKHFAAWESKGMPLGKM.HETAFNIEGYQSSGKADVNSMSINIGK        233
Bc   KRPTGSNATITFTNHVNAWKSHGMNLGSNWAYQVMATEGYQSSGSSNVTVW              185
Bp   KRTSGTVS...VSAHFRKWESLGMPMGKM.YETAFTVEGYQSSGSANVMTNQLFIGN        201
Sl   KVTSGSGT.ITTGNHFDAWARAGMNMGQFRYYMIMATEGYQSSGSSNITVSG             191
TrII HRSSGSVN...TANHFNAWAQQGLTLGTMD.YQIVAVEGYFSSGSASITVS              190
         ===>    ==========         =======>   ========>
          A6          Helix            B4          A4
```

Fig. 7. Multiple sequence alignment among endo-1,4-β-xylanases (family G). The totally conserved amino acids are in bold-face type and the stars indicate the putative catalytic residues. The secondary structure assignment of XYNII (*TrII*) is indicated by arrows (see Figure 2). The positions of the 'thumb' and 'cord' are marked. *Ak, Aspergillus kawachii* (Ito *et al.*, 1992); *TrI, Trichoderma reesei* XYNI (Törrönen *et al.*, 1992); *Cla, Clostridium acetobutylicum* (Zappe *et al.*, 1990); *Bc, Bacillus circulans* (Yang *et al.*, 1988); *Bp, Bacillus pumilus* (Fukusaki *et al.*, 1984); *Sl, Streptomyces lividans* (Shareck *et al.*, 1991); *TrII, Trichoderma reesei* XYNII (Törrönen *et al.*, 1992).

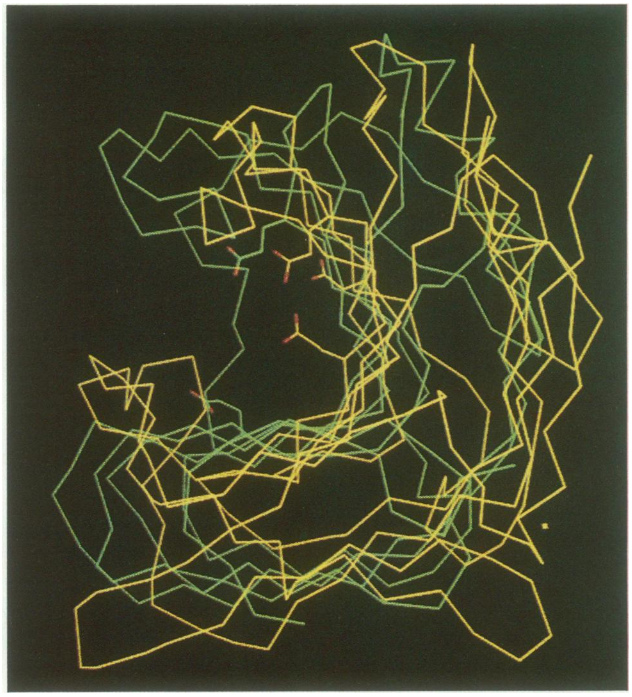

Fig. 8. The superimposition of XYNII from *T.reesei* (green) and 1,3−1,4-β-glucanase from *Bacillus* (yellow).

(Ko *et al.*, 1992). The conformational change in XYNII observed in this study alters the hydrogen bonding pattern of both residues (Figure 6). In the pH 5.0 structure, the Oϵ1 of Glu177 forms a hydrogen bond to the Nδ2 of Asn44, but the Oϵ2 does not form any hydrogen bonds. Thus, it is possible that Oϵ2 has a proton and Glu177 is neutral in this conformation. In the other (pH 6.5) structure, the position of the carboxyl group shifted. There Oϵ1 still forms a weak hydrogen bond to Nδ2 of Asn44, whereas strong hydrogen bondings to Oη of Tyr88 and to one water molecule were observed. Glu177 is likely to be in a charged form in this second observed conformation (pH 6.5), because both its Oϵ1 and Oϵ2 accept hydrogen bonds.

No clear conformational differences were detected in the position of Glu86 in these two structures. However, some changes were observed in the hydrogen bonding of Glu86. In the native structure, Oϵ1 of Glu86 makes a strong hydrogen bond to Nϵ2 of Gln136. Oϵ2 forms one hydrogen bond to Oη of Tyr 77 and one to Oη of Tyr88. The hydrogen bond network continues from Oη of Tyr77 to Nϵ1 of Trp79 and to Oη of Tyr171. In the other (pH 6.5) structure, the distance between Oϵ2 of Glu86 and Oη of Tyr88 increases from 3.3 to 4.1 Å, because of the position shift of 0.9 Å of Oη of Tyr88, and the hydrogen bonding disappears. To compensate for this loss, the hydrogen bond between Oϵ2 of Glu86 and Nϵ2 of Gln136 shortens from 3.0 to 2.9 Å. The link between the putative catalytic residues Glu86 and Glu177 is a hydroxyl group of Tyr88. The proton attached to Oη of Tyr88 points towards Glu86 in 'pH 5.0' conformation and towards Glu177 in 'pH 6.5' conformation.

Discussion

Family G xylanases

Glycanase family G was first described by Henrissat (1991). Recently, hydrophobic cluster analysis was used to analyze the amino acid sequences of 17 different low-molecular-weight endo-β-1,4-xylanases (Törrönen *et al.*, 1993a). The amino acid homology in this family is high, indicating a similar fold. The alignment of seven amino acid sequences in this family is shown in Figure 7, together with the secondary structure assignment of XYNII. XYNII has a total of 32 residues (17%), which are common in all these proteins. Both the putative catalytic residues Glu86 and Glu177 of XYNII are conserved. There is a clear cluster of conserved residues around Glu86. The direct hydrogen bonds to Glu86 (from Gln136, Tyr77 and Tyr88) are conserved, as is a second 'layer' hydrogen bond from Trp79 to Tyr77. The hydrogen bond from Tyr171 to Tyr77 also exists in other xylanases, although the tyrosine residue is replaced by histidine in some members of the family. This cluster of five residues around Glu86 forms a hydrogen bond network, which is capable of 'fine tuning' the properties of Glu86.

The situation is different around Glu177, where the surrounding is much less conserved. If Glu177 initiates the reaction by donating a hydrogen, the pK_a value of its ionizable side chain directly affects the catalytic properties of the whole enzyme. The pK_a value for 'free' glutamic acid is 4.6−5.0. XYNII has been reported to be most active in a higher pH range: 5.0−5.5 (Tenkanen *et al.*, 1992; Törrönen *et al.*, 1992). In other xylanases, the pH optimum has been estimated to vary between 3.5 and 7.0 (Wong *et al.*, 1992). The variations are probably due to the different amino acid residues in the neighborhood of Glu177.

Many of the conserved residues of endo-1,4-β-xylanases are believed to be structurally important in order to confirm

the correct folding and packing. In addition, there are three residues in which the conservation is interesting. One of them, Pro98 in *trans*-conformation, located in the middle of the 'cord', seems to be structurally important. It also addresses the question of whether or not its conformation is capable of changing from *trans* to *cis*. If this is possible, it could lead to a large structural change in the stucture. Naturally, this hypothesis should be demonstrated experimentally. Two other interesting conserved residues are the Asn124 and Thr133 above the 'thumb' (Figure 3C). The side chains of these two residues have no effect on packing. Both of them point towards solvent without making any important hydrogen bonds to other residues of the protein molecule. The role of this conservation is unclear, although we can speculate that they take part in substrate binding.

The flat Ser/Thr face of β-sheet A is also quite conserved in family G. The function of this face is difficult to understand, but it may take part in the penetration of the enzyme into small pores, which exist in wood components. Its functional role may, to some extent, be similar to separate cellulose-binding domains observed in many cellulases. In addition, it can be noted that the linker region between the cellulase catalytic and substrate-binding domains normally contains numerous serines and threonines, although many of them are believed to be glycosylated (Gilbert and Hazlewood, 1993).

Comparison with Bacillus 1,3−1,4-β-glucanase

The three-dimensional structure of the β-glucan hydrolyzing enzyme, 1,3−1,4-β-glucanase (also called lichenase) from *Bacillus* has recently been reported by Keitel *et al.* (1993). The protein is a hybrid containing residues 1−16 of the mature *Bacillus amyloliquefaciens* enzyme and residues 17−214 from the *Bacillus macerans* enzyme. The general shape of this molecule resembles the XYNII structure, as shown in Figure 8. The r.m.s. fit between the Cα atoms of 83 residues is 2.2 Å, indicating a rather weak structural similarity. The reason for the common general shape, but weak local structural similarity, was found by comparing the topological diagrams of the secondary structure elements (Figure 2). The comparison shows that the central folding pattern (boxed in Figure 2) is similar, indicating the same kind of connection between the central four β-strands. Both structures have an insertion between β-strands B6 and A6. The differences are in the arrangements of the secondary structure elements located outside the central region.

There are three cationic residues, Glu105, Asp107 and Glu109, in the active site cleft of 1,3−1,4-β-glucanase, all having side chains pointing towards the cleft. The study of site-directed mutagenesis has suggested that at least Glu105 is important for catalysis. Interestingly, all three residues are located in β-strand B6, which corresponds to the B6b strand containing Glu86 in the XYNII structure. In comparison, the glucanase structure does not have a corresponding residue for the Glu177 of XYNII.

Implications for catalysis

Without high-resolution complex structures, it is difficult to determine the ligand binding accurately. The soaking experiment with xylose, xylobiose and xylotriose at pH 5.0 and 6.5 did not reveal the binding mode for the ligands. It seems that the binding affinity of xylo-oligomers is too low. For example, Ryazanova *et al.* (1993) have reported that D-xylose binds weakly to different *Aspergillus japonicus* xylanases (K_i = 0.22−1.5 mM).

However, by analyzing the conserved residues in the cleft, some assumptions can be made. In this respect, the most interesting residues are the aromatic residues, which have the hydrophobic face of the side chain on the surface of the cleft. In many carbohydrate-binding proteins, tyrosine, tryptophan or phenylalanine residues participate in the binding of a carbohydrate residue by packing against a flat face of the carbohydrate ring (Vyas, 1991). In the cleft of XYNII, three residues of this kind exist: Trp18, Tyr96 and Tyr179 (Figure 3C and D). Trp18 is located on β-strand B2 at the entrance of the cleft, Tyr179 on β-strand A4 in the middle of the cleft and Tyr96 is located on the 'cord', quite near Tyr179, but on the other side. By using the molecular model of the xylo-oligomer, we can estimate that there is space for at least four xylose units in the cleft (subsites which we describe as A, B, C and D). It is probable that Tyr96 and Tyr179 are part of adjacent subsites (A and B) for xylose residues, whereas Trp18 is part of the fourth subsite (D) further away. Between residues Tyr179 and Trp18 there is space for one xylose residue (subsite C). In addition, we can note that one face of the phenyl ring of Phe180 is on the surface of the cleft, which might interact with the substituents of xylan, although this residue is not conserved.

The site of cleavage is probably located between subsites B and C or between subsites C and D. Tenkanen *et al.* (1992) have studied the distribution of the produced xylo-oligosaccharides of different substituted and unsubstituted xylans in *T.reesei* xylanase-catalyzed reactions. The hydrolysis of unsubstituted xylan was quite slow, and small amounts of xylotriose and longer xylo-oligomers were produced. The hydrolysis of substituted xylan was more efficient, producing larger amounts of xylo-oligomers. The main product was xylotriose. These data would suggest that the site of cleavage is between subsites C and D.

The reaction mechanism of family G xylanases has been studied experimentally by Gebler *et al.* (1992). This study has revealed that this family of enzymes are 'retaining'. By using the description of Sinnott (1990), the mechanism is e → e. This would mean that the leaving group and the reaction product are both in the equatorial position. Thus, the mechanism would belong to the same class as the widely studied and discussed reaction mechanism of the hen egg white (HEW) lysozyme. During the last few years, many variations for the reaction mechanism of HEW lysozyme have been suggested (Kirby, 1987). As in the HEW lysozyme, the catalytic residues of XYNII have carboxylic acid side chains, which are situated on both sides of the substrate. The distance between the carboxyl groups of catalytic residues is 7.2 Å in the HEW lysozyme. The value has been obtained from the lysozyme structure refined in our laboratory (unpublished data). The corresponding distances identified in this study are 10.7 Å for 'pH 5.0' XYNII and 8.1 Å for 'pH 6.5' XYNII. In XYNII, a new and striking feature is the observed conformational change in the catalytic residue Glu177. We could observe this change through a rising pH, which led to a loss of an acidic proton in Glu177. We suggest that this conformational change also occurs in catalysis (at pH 5.0). The conformational change is able to alter the pK_a value of the side chain of Glu177, making it able to accept or donate a proton, depending on

Table I. Data collection and MIR statistics

Crystal and soaking data

Data set	Soaked compound	pH	Soaking conc. (mM)	Soaking time (h)	Unit cell dimensions			
					a (Å)	*b* (Å)	*c* (Å)	*β* (°)
Native		5.0			81.57	60.63	38.25	94.4
PT1	$K_2Pt(CN)_4$	5.0	50	17	81.69	60.82	38.10	94.5
PT2	$K_2Pt(CN)_4$	5.0	50	17	81.51	60.72	38.05	94.4
HG1	$Hg(OAc)_2$ + thioxylose	5.0	20	20	81.71	60.96	38.40	94.4
HG2	$Hg(OAc)_2$ + thioxylose	8.0	20	10	81.35	60.67	38.23	94.4
HG3	MeHgOAc + thioxylose	5.0	20	18	81.47	60.72	38.16	94.4
Native2	xylose	6.5	10	22	81.70	60.96	38.05	94.3

Diffraction data

Data set	dmin (Å)	No. of measurements	No. of unique reflections	Completeness of data (%)	R_{merge} (%)
Native	1.8	84 025	25 322	75	4.95
PT1	2.0	77 296	21 886	86	7.31
PT2	2.5	38 961	12 114	93	7.63
HG1	2.5	43 235	12 336	94	4.91
HG2	2.5	29 763	11 047	85	10.74
HG3	2.0	45 605	19 484	77	8.42
Native2	1.8	44 107	18 071	54	8.77

MIR analysis

	Occupancy	X	Y	Z	R.m.s.*F*h/residual	R_{Cullis}
PT1 + PT2	31.01	0.620	0.195	0.102	1.15	0.858
	23.64	0.144	0.757	0.421		
	10.51	0.592	0.185	0.413		
	9.07	0.089	0.751	0.109		
HG1	18.23	0.197	0.000	0.111	1.53	0.723
	20.93	0.694	0.942	0.454		
	8.97	0.588	0.536	0.272		
	8.05	0.080	0.396	0.262		
HG2	27.59	0.198	0.009	0.113	1.77	0.608
	30.64	0.694	0.948	0.545		
	18.05	0.589	0.541	0.272		
	17.55	0.078	0.404	0.264		
HG3	27.11	0.269	0.509	0.442	0.99	0.808
	23.72	0.237	0.945	0.022		

Mean figure of merit = 0.50
where $R_{merge} = 100 \times \Sigma\Sigma |F^2(i) - <F^2(i)>| / \Sigma\Sigma F^2(i)$, R.m.s.*F*h/residual = $[(\Sigma F_H^2)/(\Sigma F_{DER} - F_{PH})^2]^{1/2}$, where F_H is the heavy atom form factor, F_{DER} and F_{PH} are the derivative structure factor and calculated structure factor. $R_{Cullis} = \Sigma(|F_{DER} - F_{PH}|)/\Sigma(|F_{DER} - F_{NAT}|)$, where F_{NAT} is the native structure factor and the summation is taken over the centric reflections.

its position. This feature could be useful in catalysis, where the glutamic acid residue has a dual role. First, it will initiate the reaction by donating a proton to the substrate; second, it may assist (in a charged form) a water molecule to attack the substrate. XYNII has been reported also to have transxylosidase activity (Tenkanen *et al.*, 1992). The second (pH 6.5) conformation is probably involved in that type of catalysis.

Materials and methods

Crystallization and data collection

Purification and crystallization of native endo-1,4-β-xylanase II have been described by Törrönen *et al.* (1992, 1993b). The crystals are monoclinic, belonging to the space group $P2_1$. Intensity data were collected from XYNII crystals on a Rigaku R-AXIS IIC area detector using CuKa radiation produced by a Rigaku RU200HB rotating anode (50 kV, 180 mA). The area detector data were processed using R-AXIS IIC software. Data collection statistics are given in Table I. We deduced from the unit cell volume that there are two xylanase molecules in the asymmetric unit. The self-rotation function showed a sharp peak in the position of $\Psi = 90°$, $\phi = 86°$ and $\kappa = 180°$, indicating the presence of a non-crystallographic two-fold element.

Structure determination

In total, 52 different heavy-atom experiments were performed in order to find suitable derivatives for MIR phasing. The difference Patterson maps were calculated with the XtalView program. Only $K_2Pt(CN)_4$ was found to produce a clear substitution for heavy atoms. Later, thioxylose was synthesized by replacing a hydroxyl group of C1 carbon with a thiol group. The XYNII crystals were first soaked in thioxylose solution and then transferred to a mercury acetate solution. Difference Patterson maps showed clear peaks for mercury in the Harker section. The co-ordinates of heavy atoms were calculated from the Patterson maps and their positions were refined with the Protein program (Steigemann, 1992). The electron density

maps were skeletonized with the Bones program and the result was studied with the program O (Jones *et al.*, 1991). The first Bones maps were calculated using one platinum data set combined from two different crystals and two separate data sets from different crystals soaked in mercury acetate. Two maps were calculated, corresponding to both hands. One map was better, showing the overall shapes of four molecules in the unit cell. However, it was not possible to trace the chain completely from this map. Later, a new map was calculated with the data set from the crystal which was soaked in thioxylose solution and then in methylmercury acetate. This map showed improved continuity and we were able to determine the fold of the protein and assign the sequence. The atomic model for molecule B was built with the O program. This molecule was then shifted visually by using the Bones map to the approximate position of molecule A. The position of the model was refined by using the real-space refinement feature of the O program, and the co-ordinates for both molecules A and B were obtained. Later, the electron density maps were calculated to all heavy-atom derivative data sets using model phases. The electron density maps showed clearly that His144 was the principal binding site in all mercury soaks. However, no density was found for thioxylose. The reason why mercury binding was obtained only in the presence of thioxylose remains unknown.

Structure refinement

The structure refinement was performed with the X-PLOR program (Brünger *et al.*, 1987) with the aid of a graphical user interface developed in our laboratory. The initial model was refined by simulated annealing. Data in the range of 8.0–2.5 Å were used in the first refinement. The standard slow-cooling protocol was used in the refinement, starting at 3000 K and performing 50 steps of molecular dynamics (time step 0.5 fs) and reducing the temperature by 25 K until 300 K was reached. In the first cycle, an *R*-factor was decreased from 59.1 to 37.7% (*R*-factor = $\Sigma(|F_o - F_c|)/\Sigma(F_o)$. The resulting model was studied with the O program, and manual corrections to the structure were made. The refinement cycles, later only with energy minimization, and manual rebuilding were repeated until the *R*-factor decreased to 18.3%. Water molecules were added to the model when the *R*-factor was 24.4%.

The second XYNII model was determined by using the data set based on xylose soaked crystal (pH 6.5). The final model of the native XYNII (pH 5.0) was used as an initial model for the refinement. After some corrections in the side chain and water positions, the final model was obtained. The pictures in this paper have been created by the O (Jones *et al.*, 1991), XtalView (McRee, 1992) and MolScript (Kraulis, 1991) programs. The co-ordinates will be deposited in the Brookhaven Protein Data Bank.

Acknowledgements

We thank Reetta Kallio for her skillful technical assistance. Financial support from the Academy of Finland is gratefully acknowledged.

References

Arase,A., Yomo,T., Urabe,I., Hata,Y., Katsube,Y. and Okada,H. (1993) *FEBS Lett.*, **2**, 123–127.

Biely,P. (1985) *Trends Biotechnol.*, **3**, 286–290.

Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) *Science*, **235**, 458–460.

Davies,G.J., Dodson,G.G., Hubbard,R.E., Tolley,S.P., Dauter,Z., Wilson,K.S., Hjort,C., Mikkelsen,J.M., Rasmussen,G. and Schülein,M. (1993) *Nature*, **365**, 362–364.

Fukusaki,E., Panbangred,W., Shinmyo,A. and Okada,H. (1984) *FEBS Lett.*, **171**, 197–201.

Gebler,J., Gilkes,N.R., Claeyssens,M., Wilson,D.B., Béguin,P., Wakarchuk,W.W., Kilburn,D.G., Miller,R.C.,Jr, Warren,R.A.J. and Withers,S.G. (1992) *J. Biol. Chem.*, **267**, 12559–12561.

Gilbert,H.J. and Hazlewood,G.P. (1993) *J. Gen. Microbiol.*, **39**, 187–194.

Gilkes,N.R., Henrissat,B., Kilburn,D.G., Miller,R.C. and Watten,R.A.J. (1991) *Microbiol. Rev.*, **55**, 303–315.

Golubev,A.M., Kilimnik,A.Y., Neustroev,K.N. and Pickersgill,R.W. (1993) *J. Mol. Biol.*, **230**, 661–663.

Henrissat,B. (1991) *Biochem. J.*, **280**, 309–316.

Ito,K., Iwashita,K. and Iwano,K. (1992) *Biosci. Biotech. Biochem.*, **56**, 1338–1340.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.

Juy,M., Amit,A.G., Alzari,P.M., Poljak R.J., Claeyssens,M., Béguin,P. and Aubert,J.-P. (1992) *Nature*, **357**, 89–91.

Keitel,T., Simon,O., Borriss,R. and Heinemann,U. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 5287–5291.

Kirby,A.J. (1987) *CRC Crit. Rev. Biochem.*, **22**, 283–315.

Ko,E.P., Akatsuka,H., Moriyama,H., Shinmyo,A., Hata,Y., Katsube,Y., Urabe,I. and Okada,H. (1992) *Biochem. J.*, **288**, 117–121.

Kraulis,P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.

Kubicek,C.P., Messner,R., Gruber,F., Mach,R.L. and Kubicek-Pranz,E.M. (1993) *Enzyme Microbiol. Technol.*, **15**, 90–99.

McRee,D.E. (1992) *J. Mol. Graph.*, **10**, 44–46.

Moriyama,H., Hata,Y., Yamaguchi,H., Sato,M., Shinmyo,A., Tanaka,N., Okada,H. and Katsube,Y. (1987) *J. Mol. Biol.*, **193**, 237–238.

Pickersgill,R.W., Debeire,P., Debeire-Gosselin,M. and Jenkins,J.A. (1993) *J. Mol. Biol.*, **230**, 664–666.

Poutanen,K., Rättö,M., Puls,J., and Viikari,L. (1987) *J. Biotechnol.*, **6**, 49–60.

Rose,D.R., Birnbaum,G.I., Tan,L.U.L. and Saddler,J.N. (1987) *J. Mol. Biol.*, **194**, 755–756.

Rouvinen,J., Bergfors,T., Teeri,T., Knowles,J.K.C. and Jones,T.A. (1990) *Science*, **249**, 380–386.

Ryazanova,L.P., Uk,K.D., Beletskaya,O.P. and Kulaev,I.S. (1993) *Biochemistry (USSR)*, **58**, 346–352.

Shareck,F., Roy,C., Yaguchi,M., Morosoli,R. and Kluepfel,D. (1991) *Gene*, 107, 75–82.

Sinnott,M.L. (1990) *Chem. Rev.*, 90, 1171–1202.

Steigemann,W. (1992) *Protein, User's Guide*. München.

Tenkanen,M., Puls,J. and Poutanen,K. (1992) *Enzyme Microb. Technol.*, **14**, 566–574.

Törrönen,A., Mach,R.L., Messner,R., Gonzalez,R., Kalkkinen,N., Harkki,A. and Kubicek,C.P. (1992) *Bio/Technology*, **10**, 1461–1465.

Törrönen,A., Kubicek,C.P. and Henrissat,B. (1993a) *FEBS Lett.*, **321**, 135–139.

Törrönen,A., Rouvinen,J., Ahlgren,M., Harkki,A. and Visuri,K. (1993b) *J. Mol. Biol.*, **233**, 313–316.

Viikari,L., Kantelinen,A., Poutanen,K. and Ranua,M. (1990) In Kirk,T.K. and Chang,H.M. (eds), *Biotechnology in Pulp and Paper Manufacture. Applications and Fundamental Investigations*. Butterworth–Heinemann, Boston, MA, pp. 145–151.

Viswamitra,M.A., Bhanumoorthy,P., Ramakumar,S., Manjula,M.V., Vithayathil,P.J., Murtjy,S.K. and Naren,A.P. (1993) *J. Mol. Biol.*, **232**, 987–988.

Voointholt,R., Kosters,M.T., Vegter,G., Vriend,G. and Hol,W.G.J. (1989) *J. Mol. Graph.*, **7**, 243–245.

Vyas,N.K. (1991) *Curr. Opin. Struct. Biol.*, **1**, 732–740.

Wong,K.K.Y. and Saddler.,J.N. (1992) In Visser,J. (ed.), *Xylan and Xylanases*. Elsevier, Amsterdam, pp. 171–187.

Yang,R.C.A., MacKenzie,C.R. and Narang,S.A. (1988) *Nucleic Acids Res.*, **16**, 7187.

Zappe,H., Jones,W.A. and Woods,D.R. (1990) *Nucleic Acids Res.*, 18, 2179.

Received on February 14, 1994; revised on March 22, 1994

Noted added in proof

Recently, Campbell and coworkers have solved the structures of two xylanases from family G, namely xylanases from *Trichoderma harzianum* and *Bacillus circulans* [Campbell,R.L., Rose,D.R., Wakarchuk,W.W., To,R., Sung,W. and Yaguchi,M. (1993) In Suominen,P. and Reinikainen,T. (eds), *Proceedings of the Second TRICEL Symposium on Trichoderma reesei Cellulases and Other Hydrolases*. Foundations for Biotechnical and Industrial Fermentation Research 8, Espoo, pp. 63–72].