

The Tertiary Structure at 1.59 Å Resolution and the Proposed Amino Acid Sequence of a Family-11 Xylanase from the Thermophilic Fungus *Paecilomyces varioti* Bainier

P. Rajesh Kumar¹, S. Eswaramoorthy¹, Paul J. Vithayathil¹ and M. A. Viswamitra^{1,2*}

¹Department of Physics, Indian Institute of Science, Bangalore 560 012, India

²Jawaharlal Nehru Centre for Advanced Scientific Research Indian Institute of Science Bangalore, 560 012, India

We report the crystal structure at 1.59 Å and the proposed amino acid sequence of an *endo*-1,4-β-xylanase (PVX) from the thermophilic fungus *Paecilomyces varioti* Bainier (PvB), stable up to 75 °C. This fungus is attracting clinical attention as a pathogen causing post-surgical infections. Its xylanase, known as a skin-contact allergen, is the first protein from this fungus whose three-dimensional structure has been elucidated. The crystals of PVX conform to the space group $P2_12_12_1$ with $a = 38.76$ Å, $b = 54.06$ Å and $c = 90.06$ Å. The structure was solved by molecular replacement techniques using polyaniline coordinates of the *Thermomyces lanuginosus* xylanase (PDB code 1YNA) and a careful model building based on the amino acid sequence known for two trypsin-digested peptide fragments (17 residues), the sequence and structural alignment of family-11 xylanases and electron density maps. The final refined model has 194 amino acid residues and 128 water molecules, with a crystallographic *R*-factor of 19.07% and a free *R*-factor of 21.94%. The structure belongs to an all-β fold, with two curved β-sheets, forming the cylindrical active-site cleft, and a lone α-helix, as present in other family-11 xylanases. We have carried out a quantitative comparison of the structure and sequence of the present thermophilic xylanase (PVX) with other available native structures of mesophiles and thermophiles, the first such detailed analysis to be carried out on family-11 xylanases. The analysis provides a basis for the rationalisation of the idea that the “hinge” region is made more compact in thermophiles by the addition of a disulphide bridge between Cys110 and Cys154 and a N-H...O hydrogen bond between Trp159 near the extremity of the lone α-helix and Trp138 on β-strand B8. This work brings out explicitly the presence of the C-H...O and the C-H...π type interactions in these enzymes. A complete description of structural stability of these enzymes needs to take account of these weaker interactions.

© 2000 Academic Press

Keywords: family-11 xylanase; thermostability; crystal structure; amino acid sequence; weaker interactions

*Corresponding author

Introduction

Xylanases (EC 3.2.1.8) have become a focus of interest in the paper industry (Viikari *et al.*, 1994) because they decrease the demand for chlorine-based chemicals in the wood pulp delignification process, thereby reducing the toxic industrial effluents, which are the by-products of organochlorine (Amant, 1993). The other applications are in the food and beverage industry (Biely, 1985). Structur-

Present address: S. Eswaramoorthy, Building 463, Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA.

Abbreviations used: PVX, *endo*-1,4-β-xylanase; PvB, *Paecilomyces varioti* Bainier; 1YNA, *Thermomyces lanuginosus* xylanase.

E-mail address of the corresponding author: mav@physics.iisc.ernet.in

al and sequence classifications have demonstrated the presence of more than 60 families of glycosyl hydrolases. Two such hydrolases are xylanases, which hydrolyse the polysaccharide xylan, and have been categorised based on sequence similarities, namely, family-10 usually possessing an $(\alpha/\beta)_8$ fold and family-11, having predominantly β -strands with one α -helix (Henrissat & Davies, 1997). Family-11 xylanases have been well studied because of their direct usage in bio-bleaching in the paper industry (Gilkes *et al.*, 1991; Henrissat & Bairoch, 1993; Wakarchuk *et al.*, 1994a), especially the mesophilic xylanases expressed by *Bacillus circulans* (PDB code 1XNB) and *Trichoderma reesei* (PDB codes 1ENX and 1XYN). These enzymes are gaining importance as model enzymes for investigating protein folding, thermal stability, and modes and specificities of enzymatic action on their substrates (Wakarchuk *et al.*, 1994b; Törrönen *et al.*, 1994; Törrönen & Rouvinen, 1995, 1997).

Cellulases, xylanases and various other hydrolytic enzymes used in many industries and households have been studied intensely with a clinical perspective of locating enzyme-allergens (Tarvainen *et al.*, 1991). Some of these enzymes, including xylanases and cellulases, have been found to cause occupational and non-occupational allergies, such as respiratory and irritant contact dermatitis. Based on the clinical examinations of several industrial workers dealing with xylanases and cellulases in their daily chores, there have been suggestions to include these enzymes as "airborne skin-contact allergens". *Paecilomyces varioti* Bainier (PvB) was believed to be rarely a human pathogen. But there are now many clinical reports (Castro *et al.*, 1990; Byrd *et al.*, 1992; Dhindsa *et al.*, 1995; Shing *et al.*, 1996), where PvB fungal growth has been identified as the source of infection in several post-surgical complications. These results open up a new direction of research on xylanases with a clinical perspective, and it becomes necessary to understand the nature of pathogenicity of this fungus at the molecular structure level through some of the major enzymes it expresses. Xylanase is one such enzyme that is expressed extracellularly by PvB, and its structure described here represents the first such study of a protein from this organism.

Because of the tremendous applications in the industries, which require enzymes functioning efficiently at elevated temperatures, a detailed structural and sequence comparison of a thermophilic enzyme system with the members in the mesophilic regime becomes necessary. The improved thermal stability is a concerted effort of several local adjustments on the sequence and the structure of the enzyme. Several such factors contributing to thermal stability have been recorded in the literature (Vogt & Argos, 1997; Vogt *et al.*, 1997; Querol *et al.*, 1996). Three-dimensional native structures of two thermophilic (including *endo*-1,4- β -xylanase (PVX)) and four mesophilic family-11 xylanases have been determined so far. In this work, we

have made a detailed structural comparative analysis, and have derived the differences in the sequences and structural properties among them. PVX shares more sequence and structural similarity with *Thermomyces lanuginosus* xylanase (PDB code 1YNA) than with mesophilic enzymes (Table 1), the key feature being the presence of a disulphide linkage, which is absent in the mesophilic forms. We have designated this region of the molecule as the hinge region, as the stability of the enzyme seems to hinge very much on the S-S bridge and related interactions seen in its neighbourhood. An examination of the PVX structure suggests that there may be a significant link between the functions of this part and the active-site cleft on the opposite side of the molecule, as detailed in the section Engineering possibilities, below. Among other factors that could contribute to better thermal stability is the presence of specific ion pairs and buried water molecules, and a lower packing value as found in a few other families of thermophilic enzymes (Facchiano *et al.*, 1998).

We have reported the thermal stability of the crystals of PVX mounted in a glass capillary tube, containing ammonium sulphate solution at 30% saturation. The optical behaviour of the crystal was studied by heating it in increments of 5 deg. C. The crystal completely lost birefringence at around 65°C and exhibited no X-ray diffraction (Eswaramoorthy *et al.*, 1994). Thermal stability studies using biochemical techniques revealed that the enzyme demonstrated 55% activity for ten minutes at 75°C. On further heating to 80°C for ten minutes, the enzyme displayed only 35% activity and completely lost its activity after an hour (Krishnamurthy, 1989).

At the time of structure determination, the complete amino acid sequence information for PVX was not available. Knowledge of only 17 amino acid residues was available due to the trypsin digestion of PVX. Two such trypsin-digested fragments were sequenced using Edman degradation. The process of determining the structure of PVX using high-resolution X-ray diffraction data (1.59 Å) was augmented by the information derived from sequence and structural alignment of family-11 xylanases and partial amino acid sequence elucidated using biochemical means. The resulting high-quality electron density map (see Figure 3) allowed us to trace the complete primary structure of PVX, which was not possible with our earlier crystals that gave an X-ray intensity data set only up to 2.7 Å resolution (Eswaramoorthy *et al.*, 1994).

Results and Discussion

PVX structure determination

The structure of PVX was determined by applying molecular replacement techniques using the program AMoRe (Navaza, 1994). As the complete amino acid sequence of PVX and the family to

which it belongs were not available, polyalanine models of both family-10 and family-11 xylanases were used as starting models. The rotation and translation solutions confirmed that PVX belongs to the family-11 xylanase, yielding a high correlation coefficient and a low *R*-factor for the model from 1YNA (Table 1). The availability of the sequence of the two trypsin-digested peptide fragments by Edman degradation, GWNPGLNAR - (50-58) and YNAPSIDG - (123-130) provided the starting point to anchor these residues in the electron density map. On the basis of the positions of these two sequence signatures, the other conserved residues were located. Further iterative rounds of model building and refinement resulted in the complete primary and tertiary structure. Information from the classified PROSITE database containing sequence signatures present in the family-11 xylanases was also used to confirm the authenticity of the residues flanking the active-site glutamate residues (Glu86 and Glu178) (Hofmann *et al.*, 1999). Ambiguities concerning the identity of Asn/Asp and Gln/Glu were resolved by maximising the hydrogen bonding capability and from the probability of the occurrence of that particular residue from sequence alignment (McDonald & Thornton, 1994).

Quality of the refined model

The final refined model of PVX consists of 194 amino acid residues and 128 water molecules, with the possibility that the N-terminal Gly could be acylated, as determined biochemically (Krishnamurthy, 1989). PVX contains the same

number of amino acid residues as 1YNA, and the tertiary structure is very similar. The possibility of occurrence of extra residues at the N and C termini was ruled out, based on the stereochemical constraints imposed by the symmetry-related molecules proximal to the terminal residues. The backbone electron density is continuous and hence the possibility of any insertion or deletion with respect to the 1YNA model is ruled out. Of the non-glycine residues, 90 % are present in the most favoured regions and the rest in the additionally allowed regions of the Ramachandran plot as calculated by PROCHECK (Laskowski *et al.*, 1993). There is no non-glycine residue in the generously allowed or disallowed region of the plot.

The structure of PVX

The overall structure of PVX conforms to the fold of other family-11 xylanases, as shown in Figure 1. It is isostructural to 1YNA, yet crystallises in the space group $P2_12_12_1$, whereas 1YNA crystallises in the space group $P2_1$ (Gruber *et al.*, 1998). The final model of PVX contains one α -helix and 14 β -strands forming two curved β -sheets, A and B (Figures 1 and 2). The curved β -sheet B, the cord and the thumb of PVX form the cylindrical active-site cleft. PVX belongs to the class of retaining enzymes, as do other family-11 xylanases. This has been verified experimentally for the family-11 xylanases (Gebler *et al.*, 1992), where the enzymic hydrolysis of a sugar with an equatorial bond results in a product with a retained anomeric configuration. The catalytic mechanism is similar to that found in the hen eggwhite lysozyme (HEWL)

Table 1. Sequence and structural properties of family-11 xylanases

Property	PVX	1YNA	1XNB	1XND	1XYN	1ENX
No. of residues	194	194	190	190	178	190
No. of non-hydrogen atoms	1493	1504	1448	1473	1348	1485
No. of β -strands	14	14	13	14	13	15
Hydrophobic residues (%)						
(A,V,L,I,P,M,F,W)	27.3	32.4	30.2	30.0	33.1	30.0
Surface area ($\text{\AA}^2$)	8130	8068	7820	7805	7472	7983
Normalised ^a surface area ($\text{\AA}^2$)	41.9	41.6	41.2	41.1	42.0	42.0
Volume ($\text{\AA}^3$)	19,070	19,320	18,590	18,820	17,330	18,970
Normalised volume ($\text{\AA}^3$)	98.3	99.6	97.8	99.0	97.4	99.8
Active site volume ($\text{\AA}^3$)	1841.1	1600.5	1732.1	1908.2	1240.1	1829.3
rmsd ($\text{\AA}^2$)	-	0.4	1.4	0.9	1.3	0.8
Sequence identity (with PVX) (%)	-	86	50	60	45	61
Mean packing	0.375	0.386	0.389	0.393	0.384	0.391/0.392 ^b
Buried water	7	6	8	6	8	-
No. of ion pairs 3-8 $\text{\AA}$	17	23	9	15	13	16
No. of aromatic pairs (3-8) $\text{\AA}$	27	30	21	29	16	27
No. of hydrogen bonds	197	197	180	188	157	180/189 ^b
No. of C-H...O hydrogen bonds	44	52	45	53	46	56

^a The values for surface area and volume are normalised over the total number of amino acid residues.

^b The two values refer to the two chains A and B present in 1ENX.

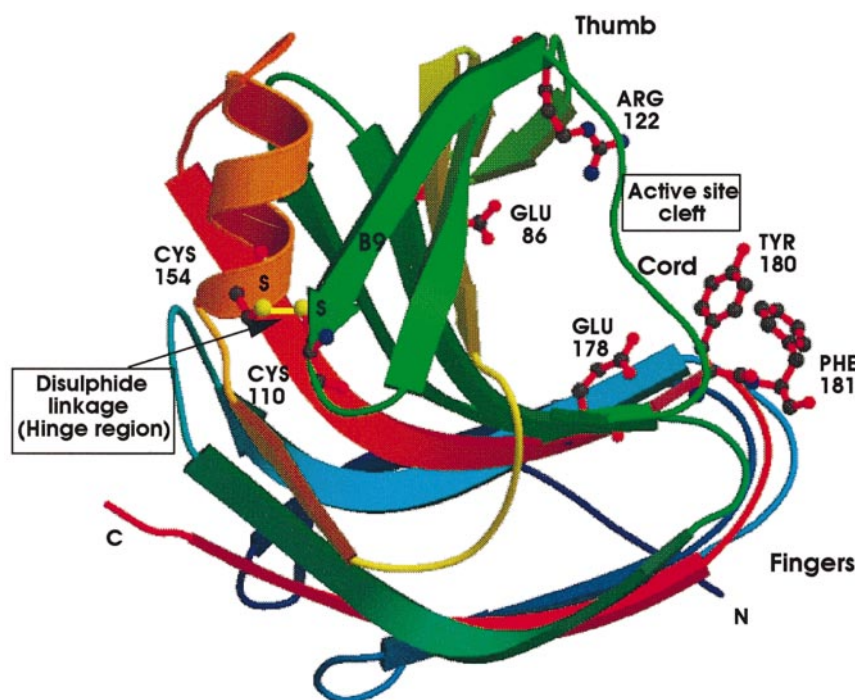


Figure 1. A cartoon representation of the PVX enzyme. The N and C termini are marked and the active-site cleft of the enzyme is shown on the right. The charged residues in the cylindrical active-site cleft, Glu86, Glu178 and Arg122, the two cysteine residues, Cys110 and Cys154, and the aromatic pair between Tyr180 and Phe181 are represented as ball-and-sticks. The single α -helix in the enzyme is on the left-hand side of the Figure. The hinge region referred to in the text is at the position of the disulphide linkage. The strand B9 connecting the thumb region to the hinge region is also shown. The colouring follows a rainbow with blue at the N terminus and red at the C terminus. The Figure was prepared using MOLSCRIPT (Kraulis, 1991).

(Inoue *et al.*, 1992; Gebler *et al.*, 1992). As in the other family-11 xylanases, the catalytic Glu86 (nucleophile) and Glu178 (acid/base catalyst) of PVX are located in the active-site cylindrical cavity of this enzyme where xylan enters. In addition, Glu178, the proposed acid/base catalyst (Lawson *et al.*, 1997) exists in the down conformation where the distance between Glu86 and Glu178 is more than 10 Å (Figure 1). This is in agreement with the structures of 1YNA and 1ENX, where Glu178 adopts the down conformation. This feature can be attributed to the pH of our crystallisation, which was around 4.5, as suggested by other workers (Törrönen *et al.*, 1994; Gruber *et al.*, 1998). This residue is liable to conformational change with respect to change in pH. Inference from the 1ENX structure suggests that Glu178 might prefer the up conformation at its pH optimum of 6.5. This has direct significance for the activity of the enzyme, as the polysaccharide xylan would not be cleaved at pH values less than 6.0, when the distance between the catalytic glutamate residues would be large. The enzyme possesses a disulphide linkage between Cys110 and Cys154, with the most favourable geometry. Arg187 exists in two different conformations, with conformation A forming an ion pair with Glu36 on the surface of the protein.

The deduced primary structure compares well with the sequence and structural patterns of family-11 xylanases as illustrated in Figure 2. Mass spectra of PVX crystals show a major peak corresponding to a molecular mass of 21,374 Da. The molecular mass as derived from the proposed amino acid sequence corresponds to 21,365 Da. This very close agreement between the two masses imparts a high degree of confidence in the assign-

ment of the amino acid sequence to the PVX enzyme from its electron density map.

Elucidation of probable thermostability parameters in family-11 xylanases

Various sequence and structural parameters have been suggested as possible reasons for thermostability in proteins (e.g. see Vogt *et al.*, 1997; Querol *et al.*, 1996). It is also known from experiment that the energy barrier between mesophilic and thermophilic forms of an enzyme is quite small and can be overcome by the addition of a few hydrogen bonds or the presence of a disulphide linkage (Fersht & Serrano, 1993). Increase in the aromatic-aromatic interactions on the surface has been suggested to be the reason for the increased thermostability of the xylanase from the *Bacillus* D3 strain (Harris & Perez, 1997), and the presence of a disulphide linkage and increased number of ion pairs have been found to be the reason for thermostability in *T. lanuginosus* xylanase (1YNA) (Gruber *et al.*, 1998). In this work, we have analysed the following parameters as suggested in the literature (Vogt *et al.*, 1997; Querol *et al.*, 1996) among the family-11 xylanases; environment of the disulphide linkage, electrostatic interactions, weak and strong hydrogen bonds, hydrophobic packing, aromatic-aromatic interactions, role of buried water molecules, residue substitutions based on sequence alignment, C-H...O and C-H... π interactions and several other parameters related to molecular surface area and volume of the enzyme and the volumes of the active-site cleft. All the above-mentioned structural parameters have been calculated for four

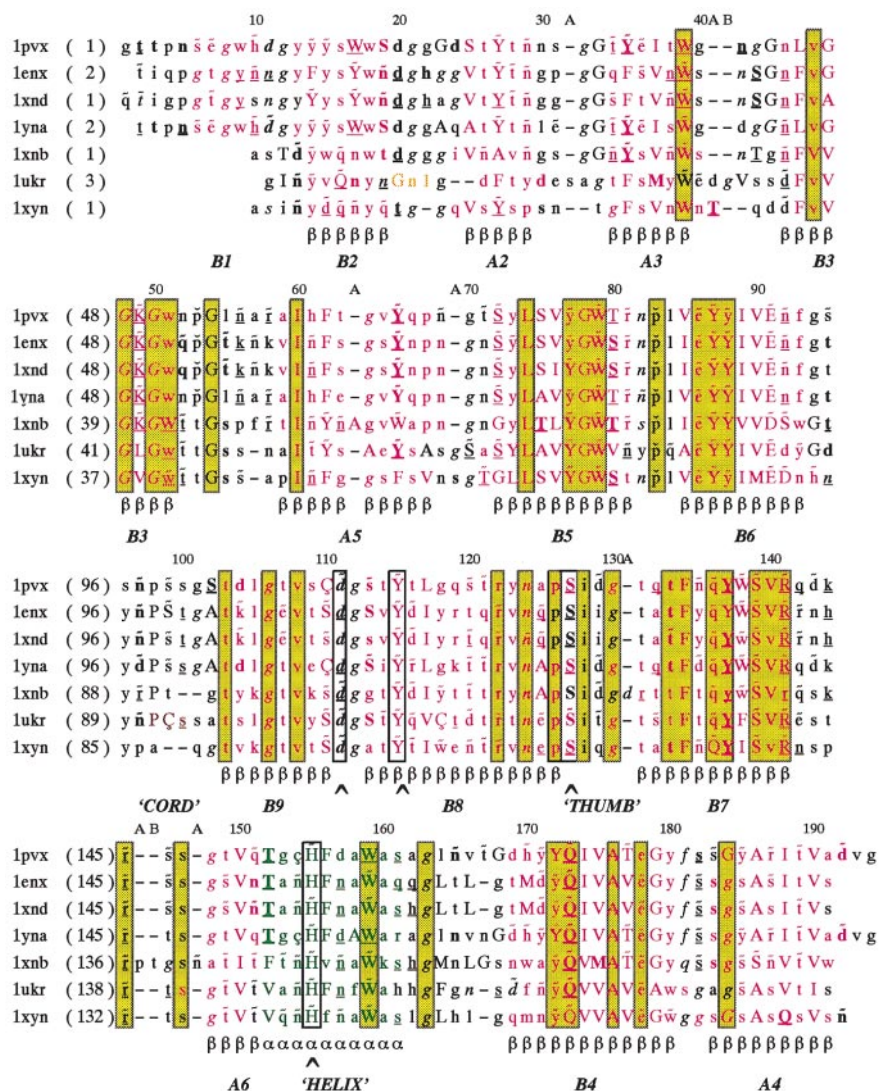


Figure 2. Sequence and structural alignment of PVX with other family-11 xylanases. 1YNA, xylanase from *Thermomyces lanuginosus*; 1XNB, *Bacillus circulans*; 1XND, *Trichoderma harzianum*; 1XYN and 1ENX, xylanases I and II from *Trichoderma reesei*; 1UKR, *Aspergillus niger*. Residues in green represent the α -helix, magenta represents residues in the β -strands, and orange represents residues in the 3_{10} -helix. This representation was made using the programs COMPARER (Sali & Blundell, 1990; Zhu *et al.*, 1992) and JOY (Mizuguchi *et al.*, 1998). The secondary structural elements (sheets A and B, helix) and other regions (cord, thumb) are denoted below the alignment. The yellow patches on the alignment refer to the highly conserved residues in family-11 xylanases. Also shown is the conserved motif His155-Tyr115-Asp111-Ser139 using the caret (^) symbol. The other conventions used in the Figure are: upper case, solvent-inaccessible; lower case, solvent-accessible; italics, positive ϕ ; breve, *cis*-petide; tilde, H-bond to other side-chain; bold, H-bond to main-chain amide group; underline, H-bond to main-chain carbonyl group; cedilla, disulphide bond.

mesophilic and two thermophilic family-11 xylanases, and their respective amino acid sequences were analysed for differences. This is the first such detailed analysis to be carried out on this family. The idea was to look for any clear emerging trend towards the differences between the thermophilic and mesophilic xylanases contributing to an improved thermal stability.

The disulphide linkage in thermophilic xylanases

Sequence comparison studies reveal that, while mesophilic xylanases have Ser and Asn at position

110 and 154, respectively, the thermophilic xylanases have Cys at both these positions (Figure 2). Structural comparisons demonstrate the replacement of a hydrogen bond between Ser and Asn in the mesophilic xylanases, by a much stronger disulphide bridge between Cys110 and Cys154 in both PVX and 1YNA. As a consequence, the distance between the two C α atoms of Cys110 and Cys154 is shorter by approximately 0.5 Å in thermophiles than in their mesophilic counterparts. This disulphide bond is least accessible to the solvent region and is in the immediate vicinity of the structural motif (His-Tyr-Asp-Ser-Wat) in the hinge

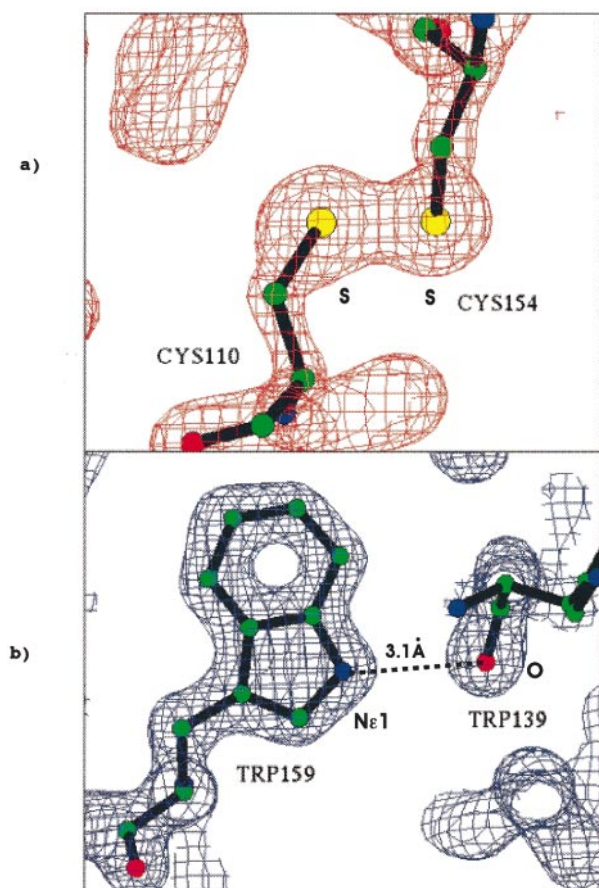


Figure 3. (a) Figure shows the $(2F_{\text{obs}} - F_{\text{calc}})$ electron density map contoured at 1.0σ , after the final round of refinement, for the disulphide region comprising residues Cys110 and Cys154. (b) Electron density around Trp139 and Trp159. The hydrogen bond between N ϵ^1 of Trp159 and the carbonyl oxygen atom of Trp139 is also indicated. The Figure was prepared by BOBSCRIPT (Esnouf, 1997).

region, conserved in both mesophiles and thermophiles, which includes the buried water Wat278, the neutral buried His155 and the residues Tyr115, Asp111 and Ser139. Interestingly, in the thermophiles, Ser139 adopts a different rotamer compared to the mesophiles, making a markedly more stable interaction between His155 N δ^1 and Ser139 O γ . The introduction of the disulphide bridge and the above-mentioned hydrogen bond suggests that the hinge region in thermophilic xylanases is inherently more stable than in the mesophiles. Mutational studies carried out on 1XNB showed that the introduction of a disulphide linkage at this position confers a slight increase in thermal stability (Wakarchuk *et al.*, 1994a). A comparison of the sequences of family-11 xylanases reveals this as a major difference, and suggests it as a main factor in the hierarchy of factors contributing to the increased thermostability. This feature was pointed

out earlier in the case of the thermophilic xylanase 1YNA (Gruber *et al.*, 1998).

Ion pairs

The presence of additional ion pairs in a thermophilic protein can be expected to increase its thermal stability compared to its mesophilic counterparts (Vogt *et al.*, 1997; Querol *et al.*, 1996). We have calculated the ion pairs in all the family-11 xylanases (Table 1) (Barlow & Thornton, 1983). The distances were calculated between N δ^1 /C ϵ^2 of His or N ϵ of Lys or C ϵ of Arg and C δ of Glu or C γ of Asp. Interestingly, the cluster Glu91-Arg141-Arg145-Asp111-His155 is almost completely buried and forms a part of the conserved, rigid, His-Tyr-Asp-Ser-Wat motif (Figure 4). Although we could detect no obvious correlation between the number of ion pairs and thermal stability, we find that there are regions where there are a greater number of ion pairs. In particular, the ion pairs between residues His10 and Asp11, and Glu36 and Arg187, which feature only in the thermophiles can be thought of as factors that bring greater cohesion to this region than in the mesophiles. Interestingly, four other ionic residues His61, Asp104, His170 and Asp192 are found only in the thermophiles, but they make ion pairs with residues that are present also in the mesophiles. There are many other ion pairs in the structure, but the advantage that they may confer either as specific to thermophiles or as common to the family-11 xylanases could not be clearly seen.

Buried water molecules

We have made a detailed comparison of family-11 xylanases for the possibility of thermophilic xylanases possessing more buried water molecules (Vogt *et al.*, 1997) using the program PRO_ACT (Williams *et al.*, 1994). The buried water molecule in the histidine motif, Wat278, is highly conserved in terms of its environment. The environment of Wat238 is partially conserved. We also find in PVX several frameworks of three to five amino acid residues stabilised by buried water molecules not found in the mesophiles; namely, Wat204, 207, 216 and 297. One would expect that thermostability would require that water molecules that are stabilising the structure should stay attached to the enzyme even at higher temperatures. These water molecules are bound by three or more hydrogen bonds with neighbouring donor or acceptor atoms, as shown in Figure 5(a).

Packing

Better packing of the protein interior would favour more van der Waals interactions between interacting groups and thus contributing to the increase in the stability of the protein (Vogt & Argos, 1997). We have made a detailed calculation of the mean packing value for each of the family-

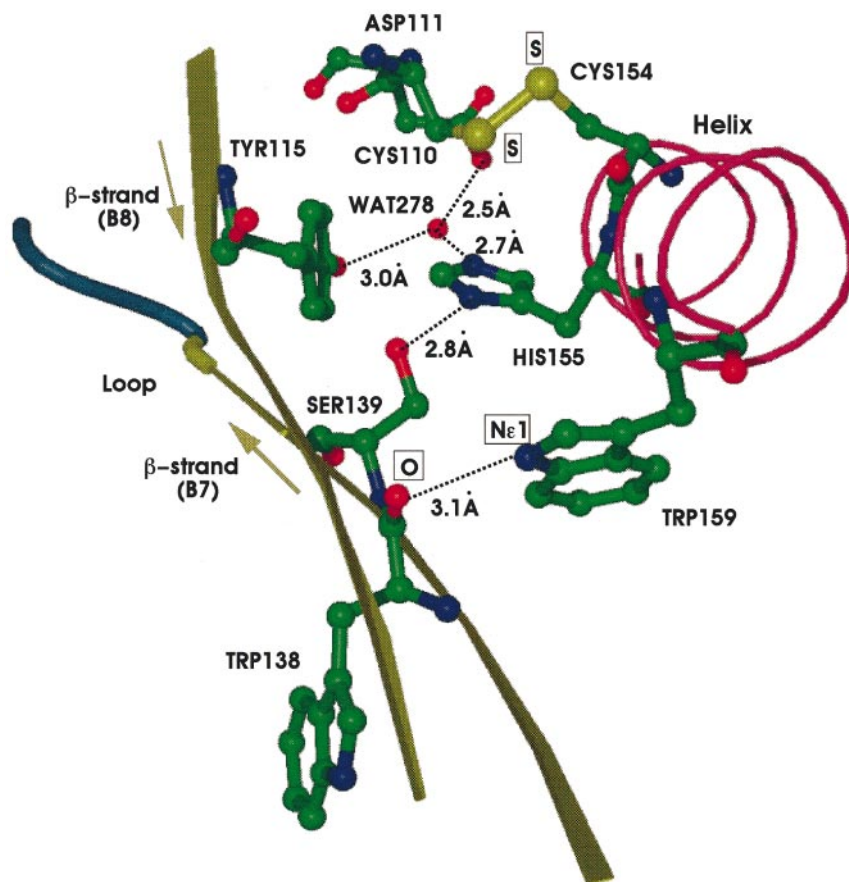


Figure 4. The highly conserved His-Tyr-Asp-Ser-Wat motif present in family-11 xylanases. The participating residues in this motif are represented along with the secondary structures they belong to. The disulphide linkage is formed between Cys110 and Cys154 in PVX. Ser and Asn occupy the corresponding positions in the mesophiles, respectively. The hydrogen bond between N ϵ 1 of Trp159 and the carbonyl oxygen atom of Trp138 is found only in the thermophiles. The Figure was prepared using VMD (Humphrey *et al.*, 1996) and rendered using Raster3D (Merrit & Bacon, 1997).

11 xylanases (Table 1). This calculation was performed on the basis of the occluded surface of each atom in a given protein (Pattabiraman *et al.*, 1995). We find that PVX possesses the lowest packing value, followed by one of the mesophiles, 1XYN and the thermophile 1YNA. All the other mesophiles have higher values. The observed decrease is admittedly only a small percentage ($\sim 3\%$), but is in line with the recent recognition that such fractional changes in intramolecular packing are an important factor of thermophilicity (DeDecker *et al.*, 1996) and that thermophilic proteins in certain families tend to possess the lowest packing value compared to the mesophilic counterparts (Karshikoff & Ladenstein, 1998).

C-H...O interactions

That the C-H group in biological systems can act as another conventional hydrogen bond donating group like N-H and O-H has been highlighted by several workers (Derewenda *et al.*, 1994, 1995; Fabiola *et al.*, 1997). Using locally available programs, we have now explored the possibility of such hydrogen bonds being present in PVX and the related structures. In the process of determining the C-H...O bonds, the hydrogen atoms were fixed on the basis of stereochemistry using the software BIOSYM (INSIGHTII, 1993). Every possible C-H...O interaction was visualised using a

graphics system, with care taken to select only the best, with distances between C and O atoms within 2.5 to 3.5 Å and the hydrogen bond angle in the range 140° to 180° . Even using this stringent hydrogen bonding criterion, we find a large number of amino acid residues favourably exposed to form the C-H...O type interaction (Table 1) and a strong C-H...O type interaction is shown in Figure 5(c).

Aromatic pairs

Aromatic pairs are supposed to contribute 1–2 kcal mol $^{-1}$ (Fersht & Serrano, 1993) and are an important factor in intramolecular packing in proteins. Figure 5(d) shows the angular geometry of a selected aromatic pair (Tyr88-Trp138) observed in PVX. As performed in an earlier analysis (Burley & Petsko, 1985), the distances between the centroids of the aromatic residues Trp, Phe and Tyr in the family-11 xylanases were calculated. The interaction was considered to be of the type aromatic-aromatic only if the distances were between 3.5 Å and 8 Å. On this basis, the strong aromatic-aromatic interactions found are Y88-W138, W79-F134, Y87-F156, Y180-F181 and Y14-W9 in PVX. In some of them, the dihedral angle between the two aromatic rings is almost 90° and the H-atom of a C-H group in a phenyl ring is directed towards the middle of its aromatic neighbour making a T type

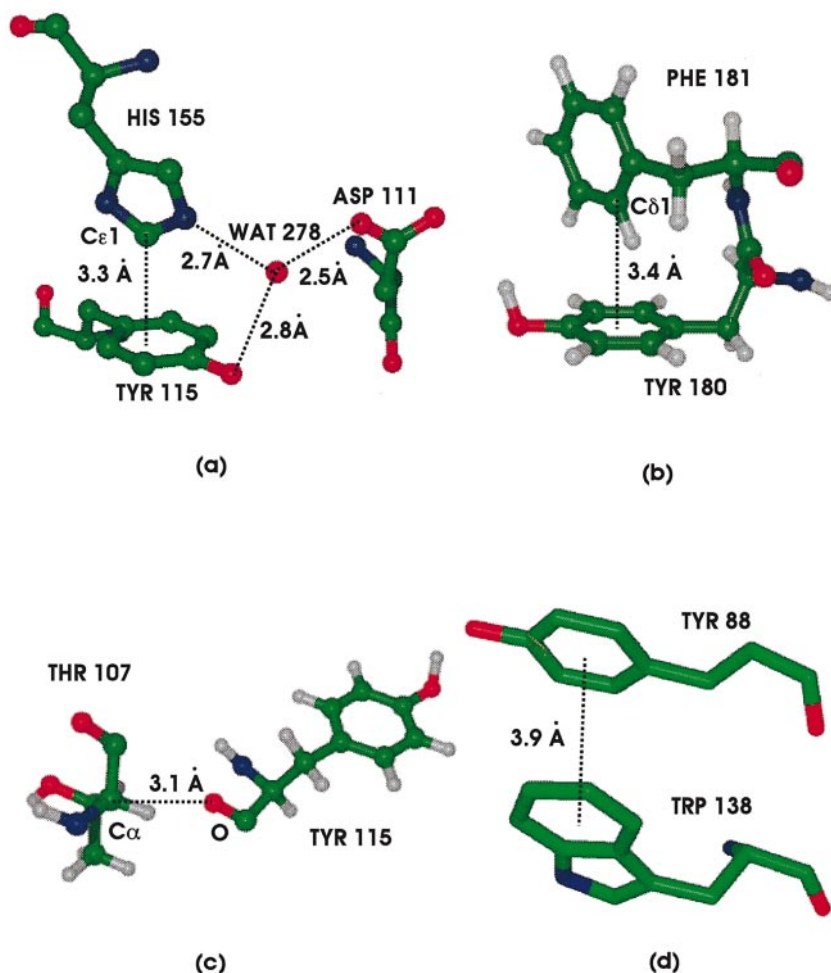


Figure 5. Representative interactions found in the PVX structure. (a) Buried water 278, forming three hydrogen bonds with the neighbouring donor and acceptor atoms of His155, Tyr115 and Asp111. This region is in the hinge portion of the molecule. The Figure shows the C-H... π interaction from His155 to Tyr115. (b) A C-H... π interaction between C δ^1 of Phe181 and Tyr180. (c) A C-H...O interaction between Thr107 and Tyr115. (d) The aromatic-aromatic pair Tyr88 and Trp138. The Figure was prepared using VMD (Humphrey *et al.*, 1996) and rendered using Raster3D (Meritt & Bacon, 1997).

of interaction (Figure 5(b)). There is also the occurrence of the type where the C-H group makes a close approach to only two or three carbon atoms of an aromatic neighbour.

C-H... π interactions

Several aromatic pairs have geometries characteristic of C-H... π type interactions between far-off aromatic residues in the polypeptide chain. X-ray structures of small molecules have shown a great deal of evidence for the presence of such interactions. For a good C-H... π interaction, only those distances between the interacting C atom and the centroid of the phenyl group lying between 2.5 and 3.5 Å were chosen. One such weaker interaction conserved among all the family-11 xylanases is between C ϵ^1 of His155 and the π -electron cloud of Tyr115 (Figure 5(a)), in the conserved motif His-Tyr-Asp-Ser-Wat containing the buried Wat278 (Connelly & McIntosh, 1998). The distances are consistent with the metrics expected for a C-H... π interaction. This interacting motif occurs in close proximity to the disulphide linkage as shown in Figure 4. It has been noted (Fersht & Serrano, 1993) that a histidine-phenyl pair could stabilise an α -helix and does contribute ~ 1 kcal

mol $^{-1}$ to protein stability. In the case of the thermophilic family-11 xylanases, the histidine ring His155 is found closer to the conserved Tyr115 by ~ 0.3 Å than in the mesophiles. This difference can be attributed to the presence of the S-S bridge surrounding this conserved motif. The residues of this conserved motif are thus more tightly packed in thermophiles with the buried Wat278 having perhaps less tendency to be disturbed in comparison with mesophiles, because of the surrounding S-S bridge. Another notable C-H... π interaction is from the side-chain C β belonging to Glu86, one of the catalytic residues, to the centre of Trp79.

Molecular surface area and volume

Deviations from globularity of proteins have been correlated to their thermal stability (Vogt *et al.*, 1997). The molecular surface area and volume of the mesophilic as well as the thermophilic xylanases were calculated using the softwares NACCESS (S.J. Hubbard & J.M. Thornton, Department of Biochemistry and Molecular Biology, University College, London, UK) and VOIDOO (Kleywegt & Jones, 1994). They are listed in Table 1, along with values normalised with respect to the total number of amino acid residues in each protein. They do

Table 2. Data collection and refinement statistics of PVX

A. Data collection	
Resolution range (Å)	20.0-1.59
Observed reflections	124,152
Unique reflections	24,200
R_{merge} (%) ^a	9.30/35.84 ^b
Completeness (%)	92.5/78.0 ^b
$I/\sigma I$	25.7/3.2 ^c
B. Refinement	
Resolution range (Å)	20.0-1.59
R_{cryst} (%)	19.07
R_{free} (%) ^d	21.94
No. of protein atoms	1493
No. of water molecules	128
rmsd bond lengths (Å)	0.015
rmsd angles (deg.)	1.3
Overall B -factor (Å ²)	23.56
Mean B -factor, protein main chain (Å ²)	21.21
Mean B -factor, protein side-chain (Å ²)	23.10
Mean B -factor, solvent (Å ²)	40.4

^a $R_{\text{merge}} = \sum |I_i - \langle I \rangle| / \sum I_i$, where I_i is the intensity of an individual reflection and $\langle I \rangle$ is the mean intensity of that reflection.

^b The value after the slash refers to the highest resolution shell, 1.7-1.59 Å.

^c The first value is the average and the second value refers to the highest resolution shell, 1.7-1.59 Å.

^d R_{free} is a cross-validation agreement factor computed for a test set of reflections (10% of the total) that were not used in the refinement process.

not exhibit any clear trend that, could be related to thermostability of xylanases. The active-site cleft volumes of family-11 xylanases calculated using the software SURFNET (Laskowski, 1995) (Table 1) also do not show any distinct differences to suggest that a characteristic change in this volume is an essential feature of thermophiles. The active-site cleft volume is shown in Figure 6, superimposed on the secondary structure representation of PVX.

Sequence differences

Several other parameters based on the sequence differences between mesophiles and thermophiles have been quoted in the literature (Vogt *et al.*, 1997; Fersht & Serrano, 1993) as possible reasons for thermostability. The differences in terms of the environment of each residue (i.e. hydrogen bonding, solvent accessibility, etc.) were scanned based on the COMPARE alignment (Sali & Blundell, 1990; Zhu *et al.*, 1992). The resulting representation from the program JOY is shown in Figure 2 (Mizuguchi *et al.*, 1998). The most notable change in the helix is the mutation of Asn to Cys at position 154. This contributes to the disulphide linkage with Cys110, as already discussed. The other notable difference between the mesophiles and the thermophiles is the hydrogen bond formed between N^ε1 of Trp159, near the end of the helix, and the carbonyl group of Trp138, from β -strand B7 (Figures 3(b) and 4). Such side-chain/main-chain interactions impart additional stability, making the helix better adapted for retaining the structure (Facchiano *et al.*,

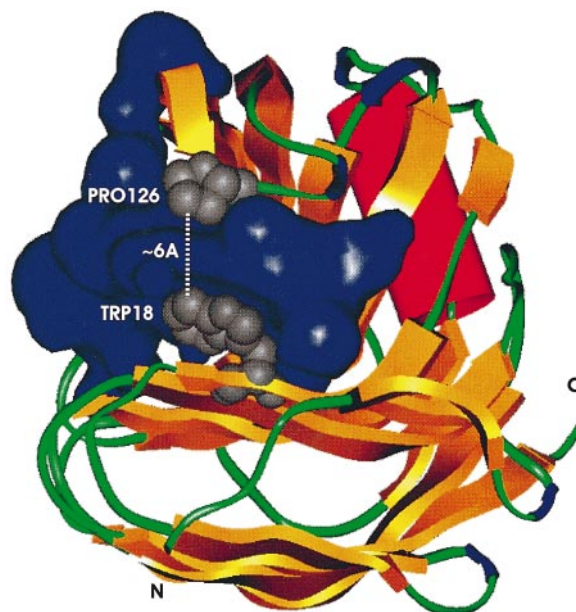


Figure 6. The secondary structure of the PVX enzyme and the active-site cleft volume superimposed on it. The orientation shows the two residues, Trp18 and Pro126, that have the closest contact of ~6 Å in the active-site cleft. In some of the mesophiles, Tyr replaces Trp at position 18. These two residue positions are highly conserved among the family-11 xylanases. The Figure was prepared by SURFNET (Laskowski, 1995) and rendered using BIOSYM (INSIGHTII, 1993).

1998). The capping of a helix with a negatively charged residue on the N terminus and a positively charged residue on the C terminus has been suggested as a factor in the thermal stability of a protein (Querol *et al.*, 1996). We do not find any such mutation altering the capping in the α -helix of the family-11 xylanases. The sequence variations, Lys $\rightarrow$ Arg, Asp $\rightarrow$ Glu and Gly $\rightarrow$ Ala, referred to as Argos replacements (Querol *et al.*, 1996), are also supposed to impart greater stability to protein structures. Our analysis does not find any Argos replacements in the family-11 xylanases.

The sequence comparison of family-11 xylanases, however, shows several other important amino acid changes, notably His61 and the aromatic residue Tyr185, which are replaced by Asn and Ser in the mesophiles. The other notable differences are at the following positions; Trp17, Asn52, Leu55, Asp104, Gln132, Trp159, Gly168, His170 and Tyr172. The basic differences between the thermophiles and the mesophiles must be embedded in such contrasts in amino acid sequences. These changes are perhaps designed in parallel to bring about greater stability at various points of the enzyme, but it is difficult to assess how they might contribute to increased stability; we can, however, talk about the more specific change, the disulphide bridge.

Engineering possibilities

In the literature pertaining to family-11 xylanases, there has been no proposal to link the conserved histidine motif on the helix in the hinge side of the molecule, to the thumb region on the opposite side of the molecule (Figure 1). A careful analysis of the structure reveals that Cys110, which forms a disulphide bridge with Cys154 in the hinge, also forms part of the strand B9 (Figure 1), leading to the thumb, above the active-site cleft. This, together with the fact that the distance between the two C α atoms of 110 and 154 is shorter by approximately 0.5 Å in thermophiles than in their mesophilic counterparts, suggests that there may be a significant link between the functions of these two parts, perhaps in the opening of the cleft. Mutational studies of this region (Figure 4) should hopefully lead us to understand the nature of the role played by this motif in the enzyme and aid in engineering more stable family-11 xylanases. The constitution of xylan is diverse, depending on its botanical origin, and different xylans can have different side-chain substituents attached to the xylose backbone. There will be a great deal of interest if one were to design a xylanase enzyme that has activity against a broad spectrum of xylans. The details of such design must depend in a complicated way on several factors, but one of the reasons why only particular types of xylan are permitted into the active-site cleft may be related to the spatial separation between the residues lining the cleft. The conserved residue Pro126, at the tip of the thumb and Trp/Tyr18, at the end of strand B2, could be of interest in this context. The active-site cleft is closest between these two residues, measuring approximately 6 Å (Figure 6), which may well restrict the entry of the xylan substrate into the active-site cleft to a particular size and type. Any mutation carried out on these two residues will bring about changes in this distance and it would be of interest to see whether this helps engineer family-11 xylanases for broader substrate specificity.

Conclusion

We describe the crystal structure of a family-11 xylanase from the thermophilic organism *Paecilomyces varioti* Bainier. The structure has been determined at high resolution (1.59 Å), permitting the assignment of the complete amino acid sequence from the electron density. We have used the structure as a basis for analysing the factors that contribute to thermostability in this family by comparing it with one other thermophilic xylanase and four mesophilic enzymes from the same family. It provides a basis to the idea that the S-S bridge between the helix and the loop B7, and other close-by interactions yield a more compact structure in the hinge region that can be expected to raise the barrier to thermal instability. Other determinants include the associations involving

buried water molecules, additional ion pairs and C-H...O and C-H... π type interactions dispersed in selected areas, perhaps as a necessary requirement for thermal stability throughout the structure. We do not find any deviations in packing density, molecular volume or surface area that can be attributed to the thermophiles, except that the thermophiles show a slight tendency for a lower packing value.

Materials and Methods

Isolation, purification and crystallisation of PVX

The fungus PvB was grown in 300 ml batches in conical flasks with orbital shaking for 16 days at 60 °C, with crude paper as the major source of xylan in the growth medium at pH 6.5. The culture medium was then filtered and the dark brown protein solution was used for further purification steps. To the filtrate, ammonium sulphate was added to 70% saturation and left at 4 °C overnight for the proteins to precipitate. The precipitated proteins were removed by centrifugation and suspended in distilled water. This was extensively dialysed against double-distilled water. The entire protein fraction was desalted using a G-25 Sephadex column. The major constituent of this fraction was PVX and was directly crystallised to obtain pure xylanase without any further chromatographic purification, which is normally necessary in purifying proteins from culture filtrates. These crystals were collected and re-crystallised with 20% (w/v) ammonium sulphate in the reservoir using the sitting drop/vapour diffusion method at pH 4.5. The purity of the sample was verified using SDS-PAGE (Laemmli, 1970).

Amino acid sequencing and mass spectra

The re-crystallised protein was subjected to reduction and carboxymethylation. The isolated product was suspended in 0.2 M ammonium bicarbonate solution and treated with trypsin in the ratio 1:50 (w/w), and the reaction proceeded for 20 hours at 30 °C. After this time, the solution was brought to pH 1.0 by the addition of trifluoroacetic acid. The digested peptide fragments were separated on a reverse-phase chromatography column using a gradient of 0%–40% (v/v) binary mixture of water and acetonitrile in 0.1% trifluoroacetic acid. The N-terminal sequencing of the protein yielded no N-terminal amino acid, possibly due to the blocking of the amino group. Two selected peptides, which were obtained in a highly pure state by rechromatographing on the same column, were sequenced using Edman degradation with a Shimadzu gas-phase protein sequencer. The sequencing was carried out using standard manufacturer's protocols.

Mass spectra of a PVX crystal were recorded using electron spray ionisation mass spectrometry (ESIMS) on an HP mass spectrometer (model HP-1100) interfaced with an HPLC instrument, at the Molecular Biophysics Unit of this Institute. The spectra were recorded in positive ionisation mode, with the sample in water, adjusted to pH 3.0 maintaining the flow rate at 0.1 ml/minute.

Data collection and processing

The complete three-dimensional X-ray diffraction data were collected using a MAR imaging plate detector coupled with a Rigaku RU200 X-ray generator. Low and high-resolution data were collected in two separate sets to maximise the completeness and data quality. The details of the completeness and the quality of the data collected are given in Table 2. The crystals diffracted well beyond 1.5 Å, but the best quality data could be derived only up to 1.59 Å resolution. The intensity data were integrated using the MARXDS suite of programs, and the different data sets were scaled and merged using the program MARSCALE (Kabsch, 1988, 1993). The crystal belongs to the orthorhombic space group $P2_12_12_1$, with one PVX molecule in the asymmetric unit. The unit cell dimensions are $a = 38.76$ Å, $b = 54.06$ Å and $c = 90.06$ Å. Assuming one molecule of PVX in the asymmetric unit, the calculated value of Matthews's coefficient, $V_M = 2.19$ cm³/g, for a molecular mass of 21.5 kDa, which is within the limits reported (Matthews, 1968).

Structure solution

The initial phases were obtained by molecular replacement techniques, using AMoRe (Navaza, 1994). Polyalanine models of the family-10 and family-11 xylanases were used to find the model that best satisfies the observed diffraction data, between 15 and 4 Å. The best results, a higher correlation coefficient and lower *R*-factor, were achieved using the polyalanine model from 1YNA. There was no short contact between symmetry-related molecules.

Model building, electron density sequencing and refinement

The rotation and translation matrices were applied to the 1YNA coordinates to bring it into the orthorhombic cell and were used for calculating the structure factors using the program X-PLOR (ver. 3.851) (Brünger, 1992a). A fraction (10%) of the reflections was kept aside throughout for free *R* cross-validation (Brünger, 1992b; Kleywegt & Brünger, 1996). A few cycles of rigid body refinement followed by positional refinement reduced the *R*-factors to $R_{\text{cryst}} = 42.0\%$ and $R_{\text{free}} = 46.0\%$. A $(2F_o - F_c)$ sigma-weighted map (Read, 1986) computed at this stage had no discontinuity in the backbone electron density. The map also clearly showed no extra density beyond the N and C termini or elsewhere along the protein chain, suggesting that 1YNA and PVX have the same number of amino acid residues. At this stage we could trace the sequence corresponding to the two tryptic peptides unambiguously, using the bulky aromatic residues as markers. Because of a large degree of similarity between 1YNA and PVX, we could use the sequence of 1YNA as the reference and trace the other side-chains. This model was subjected to further refinement by extending the structure factor calculations to a lower resolution of 20 Å and by the application of bulk-solvent correction (Jiang & Brünger, 1994). With the availability of a rotamer library in the model-building program O (Jones *et al.*, 1991), we could fit the side-chains unambiguously, resulting in further reduction in both R_{cryst} and R_{free} . Though some of the residues did not have the exact rotamer for the corresponding electron density, especially Tyr170, the bulkier and longer side-chain residues could be fitted with ease, with the

rotamer library available in O. A few more iterative rounds of model building, sequence interpretation, positional refinement, followed by simulated annealing at 5000 K and individual *B*-factor refinement, further reduced the *R*-factors to $R_{\text{cryst}} = 25.6\%$ and $R_{\text{free}} = 27.9\%$ (Engh & Huber, 1991; Brünger *et al.*, 1987, 1990). At this stage, $(F_o - F_c)$ maps showed a large number of spherical densities, suggesting solvent molecules. Peaks above 4σ and having good hydrogen bonding interactions with potential atoms of the protein were assigned as a water molecule. After successive rounds of water addition, the *R* factors were $R_{\text{cryst}} = 20.1\%$ and $R_{\text{free}} = 23.6\%$. At this stage, simulated annealing omit maps were calculated (starting temperature $T = 4000$ K) to verify the ambiguous regions of the electron density and the model bias in PVX (Brünger *et al.*, 1987; Hodel *et al.*, 1992). Thereafter, the REFMAC program (Murshudov *et al.*, 1997) of the CCP4 suite of programs (Collaborative Computation Project, No. 4, 1994), was used to refine the model anisotropically. This gave an improved model with agreement factors $R_{\text{cryst}} = 19.07\%$ and $R_{\text{free}} = 21.94\%$ with all data in the resolution range 20.0–1.59 Å. At this stage, approximately 90% of the non-glycine residues were in the most favoured regions and the rest in the additionally allowed regions of the Ramachandran plot, as verified using the program PROCHECK (Laskowski *et al.*, 1993). The rms deviations of the bond length and angles from standard values are 0.015 Å and 1.3°, respectively.

Accession numbers

The coordinates and structure factors have been deposited in the PDB (Bernstein *et al.*, 1977) with accession codes 1PVX for the coordinates and r1pvxsf for the structure factors, respectively. The amino acid sequence has been deposited in the SWISS-PROT (Bairoch & Apweiler, 1999) sequence data bank with the accession number P81536.

Acknowledgements

We thank Professor M. Vijayan, Professor P. Balram and Dr K. Suguna for the use of the National Area detector facility and the Mass Spectrograph facility, MBU, IISc, for the 1.59 Å X-ray intensity data collection and mass spectra of PVX, respectively. We thank Professor K. Ravi Acharya, University of Bath, UK, for collecting preliminary lower-resolution X-ray data at their facility. We thank Drs Savitha Krishnamoorthy, R. Maheswari, U. Varshney, A. Navaraj and Swapna Thanedar for their help at various stages of the biochemical work. We thank Dr N. Srinivasan for help with the COMPARER and JOY programs. We thank the Department of Biotechnology and Board of Research in Nuclear Sciences (BRNS) for support. M.A.V. thanks the Indian National Science Academy for the award of a Senior Scientist position. P.R.K. thanks the Council of Scientific and Industrial Research, India, for a Senior Research Fellowship. This work forms part of research undertaken by the Centre for Human Genetics, Bangalore, India.

References

- Amant, I. (1993). The crusade against chlorine. *Science*, **261**, 152–154.

- Bairoch, A. & Apweiler, R. (1999). The SWISS-PROT protein sequence data bank and its supplement TrEMBL in 1999. *Nucl. Acids Res.* **27**, 49-54.
- Barlow, D. J. & Thornton, J. M. (1983). Ion-pairs in proteins. *J. Mol. Biol.* **168**, 867-885.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535-542.
- Biely, P. (1985). Microbial xylanolytic systems. *Trends Biotechnol.* **3**, 286-290.
- Brünger, A. T. (1992a). *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University, New Haven, CT.
- Brünger, A. T. (1992b). Free R-value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature*, **355**, 472-474.
- Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R-factor refinement by molecular dynamics. *Science*, **229**, 23-28.
- Brünger, A. T., Krukowski, A. & Erickson, J. W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallog. sect. A*, **46**, 585-593.
- Burley, S. K. & Petsko, G. A. (1985). Aromatic-aromatic interaction: a mechanism of protein structure stabilisation. *Science*, **229**, 23-28.
- Byrd, R. P., Jr, Roy, T. M., Fields, C. L. & Lynch, J. A. (1992). *Paecilomyces varioti* pneumonia in a patient with diabetes mellitus. *J. Diabetes Complicat.* **6**, 150-153.
- Collaborative Computational Project Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. A*, **47**, 110-119.
- Castro, L. G., Salebian, A. & Sotto, M. N. (1990). Hyalohyphomycosis by *Paecilomyces lilacinus* in a renal transplant patient and a review of human *Paecilomyces* species infections. *J. Med. Vet. Mycol.* **28**, 15-26.
- Connelly, G. P. & McIntosh, L. P. (1998). Characterization of a buried neutral histidine in *Bacillus circulans* xylanase: internal dynamics and interaction with a bound water molecule. *Biochemistry*, **37**, 1810-1818.
- DeDecker, B. S., O'Brien, R., Flemin, P. J., Geiger, J. H., Jackson, S. P. & Sigler, P. B. (1996). The crystal structure of a hyperthermophilic archaeal TATA-box binding protein. *J. Mol. Biol.* **264**, 1072-1084.
- Derewenda, Z. S., Derewenda, U. & Kobos, P. M. (1994). (His)C^α-H...O=C < hydrogen bond in the active-sites of serine hydrolases. *J. Mol. Biol.* **241**, 83-93.
- Derewenda, Z. S., Lee, L. & Derewenda, U. (1995). The occurrence of C-H...O hydrogen bonds in proteins. *J. Mol. Biol.* **252**, 248-262.
- Dhindsa, M. K., Naidu, J., Singh, S. M. & Jain, S. K. (1995). Chronic suppurative otitis media caused by *Paecilomyces varioti*. *J. Med. Vet. Mycol.* **33**, 59-61.
- Engh, R. A. & Huber, R. (1991). Accurate bond angle parameters of X-ray protein structure refinement. *Acta Crystallog. sect. A*, **47**, 392-400.
- Esnouf, R. (1997). BOBSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Mol. Graph.* **15**, 132-134.
- Eswaramoorthy, S., Vithayathil, P. J. & Viswamitra, M. A. (1994). Crystallisation and preliminary X-ray crystallographic studies of thermostable xylanase crystals isolated from *Paecilomyces varioti*. *J. Mol. Biol.* **243**, 806-808.
- Fabiola, G. F., Krishnaswamy, S., Nagarajan, V. & Pattabi, V. (1997). C-H...O hydrogen bonds in β-sheets. *Acta Crystallog. sect. D*, **53**, 316-320.
- Facchiano, A. M., Colonna, G. & Ragonne, R. (1998). Helix stabilizing factors and stabilisation of thermophilic proteins: an X-ray based study. *Protein Eng.* **11**, 753-760.
- Fersht, A. R. & Serrano, L. (1993). Principles of protein stability derived from protein engineering experiments. *Curr. Opin. Struct. Biol.* **3**, 75-83.
- Gebler, J., Gilkes, N. R., Caleysens, M., Wilson, D. B., Benguin, P., Wakaruchuk, W. W., Kilburn, D. G., Miller, R. C. Jr., Warren, R. A. J. & Withers, S. G. (1992). Stereoselective hydrolysis catalyzed by related β-1,4-glucanases and β-1,4-xylanases. *J. Biol. Chem.* **267**, 12559-12561.
- Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr & Warren, R. A. J. (1991). Domains in microbial β-1,4-glycanases: sequence conservation, function and enzyme families. *Microbiol. Rev.* **55**, 303-515.
- Gruber, K., Klintschar, G., Hayn, M., Schlacher, A., Steiner, W. & Kratky, C. (1998). Thermophilic xylanase from *Thermomyces lanuginosus*: high resolution X-ray structure and modelling studies. *Biochemistry*, **37**, 13475-13485.
- Harris, G. W. & Perez, S. (1997). Structural basis of the properties of an industrially relevant thermophilic xylanase. *Proteins: Struct. Funct. Genet.* **29**, 77-86.
- Henrissat, B. & Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **293**, 781-788.
- Henrissat, B. & Davies, G. (1997). Structural and sequence based classification of glycoside hydrolases. *Curr. Opin. Struct. Biol.* **7**, 637-644.
- Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). Model bias in macromolecular crystal structures. *Acta Crystallog. sect. A*, **48**, 851-858.
- Hofmann, K., Bucher, P., Falquet, L. & Bairoch, A. (1999). The PROSITE database, its status in 1999. *Nucl. Acids Res.* **27**, 215-219.
- Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD - Visual molecular dynamics. *J. Mol. Graph.* **14**, 33-38.
- Inoue, M., Yamada, H., Hashimoto, Y., Yasukochi, T., Hamaguchi, K., Miki, T., Horiuchi, T. & Imoto, T. (1992). Stabilization of a protein by removal of unfavourable abnormal pK_a: substitution of undissociable residue for glutamic acid-35 in chicken lysozyme. *Biochemistry*, **31**, 8816-8821.
- INSIGHTII, (1993). *InsightII user guide*, Biosym Technologies, San Diego, CA.
- Jiang, J.-S. & Brünger, A. T. (1994). Protein hydration observed by X-ray diffraction. Solvation properties of pencillopepsin and neuraminidase crystal structures. *J. Mol. Biol.* **243**, 100-115.
- Jones, T. A., Zou, J. Y., Cowman, S. W. & Kjeldgaard, M. (1991). Improved methods for building models in electron-density maps and the location of errors in these models. *Acta Crystallog. sect. A*, **47**, 110-119.
- Kabsch, W. (1988). Evaluation of single-crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Crystallog.* **21**, 916-924.
- Kabsch, W. (1993). Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallog.* **26**, 795-800.
- Karshikoff, A. & Ladenstein, R. (1998). Proteins from thermophilic and mesophilic organisms essentially do not differ in packing. *Protein Eng.* **11**, 867-872.

- Kleywegt, G. J. & Brünger, A. T. (1996). Checking your imagination: applications of the free *R* value. *Structure*, **4**, 897-904.
- Kleywegt, G. J. & Jones, T. A. (1994). Detection, delineation, measurement and display of cavities. *Acta Crystallog. sect. D*, **50**, 178-185.
- Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946-950.
- Krishnamurthy, S. (1989). Purification and properties of xylanase and β -glucosidase elaborated by the thermophilic fungus *Paecilomyces varioti* Bainier. PhD thesis, Indian Institute of Science, Bangalore, India.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.
- Laskowski, R. A. (1995). SURFNET: a programs for visualising molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graph.* **13**, 323-330.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
- Lawson, S. L., Wakarchuk, W. W. & Withers, S. G. (1997). Positioning the acid/base catalyst in a glycosidase: studies with *Bacillus circulans* xylanase. *Biochemistry*, **36**, 2257-2265.
- Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491-497.
- McDonald, I. K. & Thornton, J. M. (1994). The application of hydrogen bonding analysis in X-ray crystallography to help orientate asparagine, glutamine and histidine side-chains. *Protein Eng.* **8**, 217-224.
- Merriett, E. A. & Bacon, D. J. (1997). Raster3D-photorealistic molecular graphics. *Methods Enzymol.* **277**, 505-524.
- Mizuguchi, K., Deane, C. M., Blundell, T. L., Johnson, M. S. & Overington, J. P. (1998). JOY: protein sequence-structure representation and analysis. *Bioinformatics*, **14**, 617-623.
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240-255.
- Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Crystallog. sect. A*, **50**, 157-163.
- Pattabiraman, N., Ward, K. B. & Fleming, P. J. (1995). Occluded molecular surface: analysis of protein packing. *J. Mol. Recogn.* **8**, 334-344.
- Querol, E., Perez-Pons, J. A. & Mozo-Villarias, A. (1996). Analysis of protein conformational characteristics related to thermostability. *Protein Eng.* **9**, 265-271.
- Read, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallog. sect. A*, **42**, 140-149.
- Sali, A. & Blundell, T. L. (1990). Definition of topological equivalence in protein structures - a procedure involving comparison of properties and relationships through simulated annealing and dynamic programming. *J. Mol. Biol.* **212**, 403-428.
- Shing, M. M., Ip, M., Li, C. K., Chik, K. W. & Yuen, P. M. (1996). *Paecilomyces varioti* fungemia in a bone marrow transplant. *Bone Marrow Transplant*, **17**, 281-283.
- Tarvainen, K. & Keskinen, H. (1991). Allergy from cellulase and xylanase enzymes. *Clin. Exp. Allergy*, **21**, 609-615.
- Törrönen, A. & Rouvinen, J. (1995). Structural comparison of two major endo-1,4-xylanases from *Trichoderma reesei*. *Biochemistry*, **34**, 847-856.
- Törrönen, A. & Rouvinen, J. (1997). Structural and functional properties of low molecular weight endo-1,4- β -xylanases. *J. Biotechnol.* **57**, 137-149.
- Törrönen, A., Harkki, A. & Rouvinen, J. (1994). Three-dimensional structure of endo-1,4- β -xylanase II from *Trichoderma reesei*: two conformation states in the active-site. *EMBO J.* **13**, 2493-2501.
- Viikari, L., Kantelinen, A., Sundquist, J. & Linko, M. (1994). Xylanases in bleaching: from an idea to the industry. *FEMS Microbiol. Rev.* **13**, 335-350.
- Vogt, G. & Argos, P. (1997). Protein thermal stability: hydrogen bonds or internal packing? *Fold. Des.* **2**, S40-S46.
- Vogt, G., Woell, S. & Argos, P. (1997). Protein thermal stability, hydrogen bonds and ion pairs. *J. Mol. Biol.* **269**, 631-643.
- Wakarchuk, W. W., Sung, W. L., Campbell, R. L., Cunningham, A., Watson, D. C. & Yaguchi, M. (1994a). Thermostabilization of the *Bacillus circulans* xylanase by the introduction of disulfide bonds. *Protein Eng.* **7**, 1379-1386.
- Wakarchuk, W. W., Campbell, R. L., Sung, W. L., Davoodi, J. & Yaguchi, M. (1994b). Mutational and crystallographic analyses of the active-site residues of the *Bacillus circulans* xylanase. *Protein Sci.* **3**, 467-475.
- Williams, M. A., Goodfellow, J. M. & Thornton, J. M. (1994). Buried waters and internal cavities in monomeric proteins. *Protein Sci.* **3**, 1224-1235.
- Zhu, Z.-Y., Sali, A. & Blundell, T. L. (1992). A variable gap penalty function and feature weights for protein 3-D structure comparisons. *Protein Eng.* **5**, 43-51.

Edited by A. Klug

(Received 2 September 1999; received in revised form 2 November 1999; accepted 5 November 1999)