PROTEINS: Structure, Function, and Genetics 29:77–86 (1997)

Structural Basis of the Properties of an Industrially Relevant Thermophilic Xylanase

Gillian W. Harris,[1]* Richard W. Pickersgill,[1] Ian Connerton,[1] Philippe Debeire,[2] Jean-Pierre Touzel,[2] Christelle Breton,[3] and Serge Pérez[3]
[1]*Molecular Biology Section, Food Macromolecular Science Department, Institute of Food Research, Reading Laboratory, Earley Gate, Reading, United Kingdom*
[2]*Unite de Physicochimie et Biotechnologie de Polymères, Institut National de la Recherche Agronomique, Villeneuve d'Ascq Cédex, France*
[3]*Ingénierie Moléculaire, Institut National de la Recherche Agronomique, Nantes, France*

ABSTRACT **A thermophilic xylanase from *Bacillus* strain D3 suitable for use as a bleach booster in the paper pulping industry has been identified and characterized. The enzyme is suited to the high temperature and alkaline conditions needed for using xylanases in the pulp industry. The xylanase is stable at 60°C and relatively stable at high temperatures, with a temperature optimum of 75°C. The pH optimum is 6, but the enzyme is active over a broad pH range. The xylanase has been cloned and sequenced, and the crystal structure has been determined. The structure of *Bacillus* D3 xylanase reveals an unusual feature of surface aromatic residues, which form clusters or "sticky patches" between pairs of molecules. These "sticky patches" on the surface of the enzyme are responsible for the tendency of the protein to aggregate at high concentrations in the absence of reagents such as ethylene glycol. The formation of dimers and higher order polymers via these hydrophobic contacts may also contribute to the thermostability of this xylanase. Proteins 29:77–86, 1997.**

Key words: xylanase; family G; thermostability; hydrophobic cluster; sequence; structure

INTRODUCTION

Xylanase pretreatment of paper pulp to remove xylan while preserving cellulose content ideally requires a xylanase that is thermostable, active at alkaline pH, and lacks cellulase activity. The thermophilic xylanase from *Bacillus* strain D3 (Patent FR-9101191[1]) is suitable for industrial use as it fulfils these requirements, is active over a broad pH range, has a high temperature optimum and has higher specific activity than most other xylanases. The unique characteristics of this xylanase which give it industrial desirability are shown to be conferred on the enzyme by unusual structural features.

Xylanases randomly hydrolyze the β-1,4-glycosidic bonds of β-1,4-xylan, the major plant cell wall polysaccharide component of hemicellulose, to produce several xylooligomers of different lengths. On the basis of sequence similarity, xylanases have been classified into two distinct families, F and G[2], also known as glycosyl hydrolase families 10 and 11, respectively[3]. The three-dimensional structures of a number of low-molecular weight (20 kDa) xylanases have been reported (*Bacillus pumilus*[4,5]; *Bacillus circulans*[6,7]; *Trichoderma harzianum*[6]; *Trichoderma reesei*[8,9]). These homologous enzymes belong to family G and are single-domain β-sheet proteins.

Xylanases are used as bleach boosters in the pulp industry. In paper production, the pulp must be treated to remove the lignin component which gives the unwanted colour. The standard method of paper whitening uses environmentally unfriendly chlorine bleaches. Xylanase pretreatment reduces the quantity of chloride needed. As industrial pulping is conducted under conditions of high temperatures and high pH (pH 10–11), the ideal xylanase for pulp bleaching would be thermostable and active at alkaline pH.

The potential biotechnological applications of thermophilic and alkaphilic xylanases has stimulated a search for strains overproducing the enzymes possessing the desired features. In the course of such a quest, a thermophilic *Bacillus* xylanase was isolated[10,11]. We report here the characterization, cloning, sequencing, and three-dimensional X-ray structure of a thermophilic xylanase from *Bacillus* strain D3.

Dr. Breton and Dr. Pérez are currently at CERMAV-CNRS, BP 53, 38041 Grenoble Cédex 9, France.

*Correspondence to: Gillian W. Harris, Molecular Biology Section, Food Macromolecular Science Department, Institute of Food Research, Reading Laboratory, Earley Gate, Whiteknights Road, Reading, RG6 6BZ, UK.
E-mail: gillian.harris@bbsrc.ac.uk

Received 31 January 1997; Accepted 20 February 1997

MATERIALS AND METHODS

Biochemical Characterization

Strain and purification of the enzyme

The bacterial strain XE, selected on the basis of its high xylanolytic activity, was isolated from compost samples.[11] The *Bacillus* strain D3 was obtained by mutagenisis with ethyl methanesulfonate,[1] and the xylanase was purified as described previously.[10,12]

Specific activity, pH optimum, and stability

Xylanase activity was determined by measuring the reducing sugars liberated from xylan by the method of Kidby and coworkers.[13] The specific activity of the purified xylanase was determined at 60°C with soluble birchwood xylan as substrate following previously established protocol.[14] The residual activity is relative to activity when measured at pH 5.8. The protein concentration was determined by gravimetry or from amino acid analysis.

The activity of the xylanase over pH range 2 to 12 was measured with birchwood xylan. This enzyme showed the same enzymic activity in sodium-acetate buffer at pH 6 or in distilled water. The stability of the xylanase was studied at basic pH with different incubation times. Then 150 U of xylanase was put in a solution of birchwood xylan (0.5%, universal buffer) at different pH at room temperature. Samples were withdrawn periodically and assayed for enzymic activites at 60°C.

Adsorption onto Avicel

The incubation mixture of 40 ml of Avicel (2% in water) and the xylanase at 14 U/ml concentration was stirred at 4°C. Samples were periodically withdrawn (from 30 minutes to 6 hours), centrifuged at 13,000 rpm for 5 minutes, and the supernatants were checked for xylanase activity; all the xylanase activity was detected in the supernatants.

Cloning and Sequencing of Xylanase

Strains, vectors and culture conditions

The thermophilic *Bacillus* D3 was grown at 55°C in a liquid medium as outlined previously.[10] The SURE, XL1-Blue, and XLOLR *Escherichia coli* strains, the vectors λ ZAP Express and pBluescript, and the filamentous helper phage ExAssist™ were all purchased from Statagene Cloning Systems.

DNA techniques

Bacterial genomic DNA was extracted from *Bacillus* according to the method of Yang.[15]

PCR was used to amplify a region of the chromosomal DNA coding for the xylanase. The nucleotide sequence of the forward primer (P1) was AAYACNTAYTGGCARTAYTGGACNGAYGG (derived from the sequence NTYWQYWTDG in the N terminus end of the xylanase); that of the reverse primer (P2) was YTGWCKNACRCTCCARTAYTG (corresponding to the sequence QYWSVRQ, a conserved region near the C terminus of other xylanases from different *Bacillus* species). PCR was performed with chromosomal DNA as a template and the primers P1 and P2 on a thermocycle (Perkin-Elmer, France) with the following temperature profile: 1 min 94°C—1 min 50°C—2 min 72°C for 35 cycles. The PCR product was purified and cloned into pBluescript. The chimaeric plasmid (pBX20) was used to transform SURE cells.

Construction of a Bacillus D3 genomic library in λ ZAP Express

Chromosomal DNA was partially digested with *Sau*3AI and the resulting DNA fragments in the size range 1.5–8 kb were purified and ligated into *BamH* I-digested λ ZAP Express. The library was constructed using XL1-Blue cells as indicated by the manufacturer.

Screening of the genomic library

pBX20 was digested with *BamH* I and *Hind* III, and the DNA insert was purified and labeled with digoxigenin (Boehringer Mannheim) following the instructions of the manufacturer. The labeled DNA was used to screen the genomic library. After the third screening, positive lambda plaques were isolated, and the recombinant plasmid pBK-CMV inserted in the vector λ ZAP Express was excised by using the filamentous phage ExAssist and then recovered by infecting the XLOLR cells in the presence of kanamycin (10 μg/ml).

DNA and protein sequence analysis

Plasmid preparations for sequence determination were perfomed by using Qiagen tip 100 (Diagen, Coger, France). Double-stranded DNA sequencing was done by the dideoxy chain termination method of Sanger,[16] using Sequenase™ 2.0 DNA sequencing kit from United States Biochemical. Both universal and specific primers were used to sequence the sense and antisense strands of inserts in the plasmids.

The Sequence Analysis Software Package by Genetics Computer, Inc. (the GCG Package) was used throughout this work.

Crystallography

Crystallization and structure solution

Bacillus D3 xylanase was crystallized as previously described[12] in the triclinic spacegroup P1 (a = 49.4 Å, b = 51.6 Å, c = 72.9 Å, $\alpha = 90.5°$, $\beta = 95.3°$, $\delta = 92.4°$) with four molecules in the asymmetric unit. Data were collected to 3.2 Å resolution by using a Siemens area detector (Cu $K\alpha$ radiation), and processed by using the XENGEN package[17] [R_{merge} of 6.6% for 9992 unique reflections; data 81% complete to 3.4 Å resolution and 73% complete to 3.2 Å resolution].

Sequence alignment has shown *Bacillus* D3 xylanase to be most similar to the 20 kDa family G xylanase from *Bacillus circulans*. The coordinates of this xylanase structure were obtained from Dr. Robert Campbell, National Research Council of Canada. These coordinates were used to solve the structure of *Bacillus* D3 xylanase by molecular replacement using the program AMoRe.[18]

The rotation function, calculated at 3.2 Å, gave four strong peaks (peak heights 26.9, 26.2, 25.9, 24.0, relative to the next highest peak of 15.4) corresponding to the four molecules in the asymmetric unit cell. The first molecule defines the origin in the space group P1. The translation function was run to search for the second molecule with the first molecule fixed (*R* value: 48.5%), then to search for the third molecule with the first and second molecules fixed (*R* value: 44.5%), and, finally, to search for the fourth molecule with the first, second, and third molecules fixed (*R* value 40.4%). Rigid body refinement using AMoRe reduced the *R* value for the four solutions to 38.4%.

Crystallograhic refinement

Further rigid body refinement of the molecular replacement model at 4.0 Å using the program X-PLOR[19] gave an *R* value of 36.9%. This was followed by an X-PLOR simulated annealing run with noncrystallographic symmetry restraints between the four molecules, which gave an *R* value of 0.242 at 3.2 Å resolution (2-sigma cutoff, 9119 reflections used). A 2Fo-Fc map was calculated at 3.2 Å resolution (CCP4 program suit[20]) and displayed on the Evans & Sutherland PS390 picture system by using the graphics program O.[21]

The sequence of *Bacillus* D3 xylanase was aligned by hand with that of *Bacillus circulans* xylanase (see Fig. 2). The sequence of the model structure was mutated in O to match that of *Bacillus* D3 xylanase and the model was rebuilt to fit the density. The agreement between the electron density and the expected *Bacillus* D3 xylanase sequence was good. The rebuilt model was put into an X-PLOR simulated annealing run with noncrystallographic symmetry (NCS) restraints applied. This gave an *R* value of 0.231 at 3.2 Å resolution.

The map was averaged using the RAVE program suite.[22] A mask was generated from the model coordinates using the program MAMA. The noncrystallographic symmetry matrices relating the four molecules were obtained by superimposing the molecules in O. These matrices were improved using the program IMP. The 3.2 Å resolution map was averaged using the program AVE.

The model was rebuilt to the averaged map and refined by using the program X-PLOR. Positional refinement, with NCS restraints, was followed by grouped B factor refinement (two B factors per residue, i.e., one for main-chain atoms, and one for side-chain atoms), to give an *R* value of 0.187 at 3.2 Å resolution. Further rounds of manual rebuilding and X-PLOR refinement gave an *R* value of 0.174 for 9119 reflections (2-sigma cutoff) in the resolution range 10.0 to 3.2 Å, with a final model of 1476 protein atoms and 40 well-ordered water molecules per protein molecule.

The free-*R* method of refinement was not used in the X-PLOR refinement. As all reflections have been included in the refinement, it would not be valid to calculate a free-*R* value for this refinement. We elected not to use the free-*R* refinement method because of incomplete data with spacegroup P1 and low resolution coupled with four molecules in the asymmetric unit, we did not consider the exclusion of 10% of the data would be justified. As the model was derived from molecular replacement by using a highly homologous structure, we felt the loss of the free-*R* value as an indicator of correctness was not too severe in this case.

A Ramachandran plot was calculated by using the program PROCHECK[23]; all except one (Thr2) of the 151 nonglycine and nonproline residues per molecule have conformations in the most favored (81.8%) or additionally allowed (17.5%) regions. Thr2 is in a "disallowed" region; the formation of a tight hydrogen bond (2.8Å) between the carbonyl of Asn1 and that of Gly39 could give rise to this distortion. Further, OG1 of Thr2 forms a hydrogen bond (3.5 Å) to the carbonyl of Gly21. These interactions anchor the N terminus in position. The corresponding N-ternimal residues adopt a different conformation in the *Bacillus circulans* structure where there are two additional residues. Pro73 is in the cis conformation (The residue corresponding to Pro73 in the 1.49 Å resolution *Bacillus circulans* structure, Pro75, is also in the cis conformation). The overall G factor is 0.07.

The final coordinates for molecule 1 were structurally aligned with the *Bacillus circulans* structure by using the program MNYFIT 5.0,[24] which gave a root-mean-square (rms) deviation of 0.063 Å for 179 equivalent Cα atoms (omitting the first 2/5 residues and residues 118/120–121, respectively). Overall the structures are very similar.

Previous efforts to solve the structure of *Bacillus* D3 xylanase by molecular replacement using the program X-PLOR[19] with the coordinates of the *Bacillus pumilus* xylanase[4] (kindly supplied by Professor Katsube and Dr. Hata, Osaka University, Japan) were not successful. As a test, these coordinates were also used as the search model, with the molecular replacement program AMoRe, but the correct solution was not found. This shows the critical dependance of molecular replacement on the search model. *Bacillus* D3 xylanase has greater sequence similarity to *Bacillus circulans* xylanase than to *Bacillus*

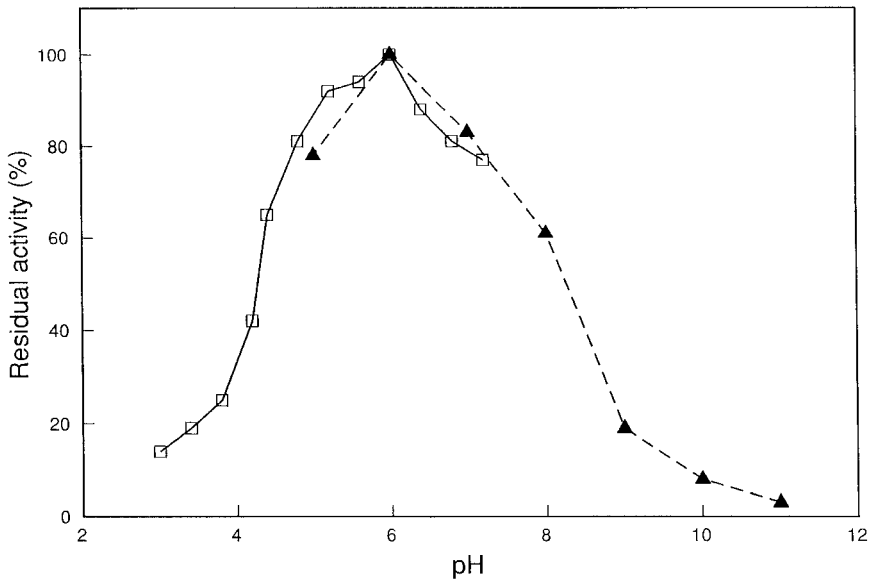

Fig. 1. pH activity profile for *Bacillus* D3 xylanase. The activities were measured with birchwood xylan in citrate buffer (solid line and squares) or Universal buffer (broken line and triangles).

pumilus xylanase (73% compared to 50% sequence identity). Also, the *Bacillus circulans* xylanase structure is refined at 1.49 Å resolution, whereas the resolution of the *Bacillus pumilus* xylanase structure is 2.2 Å.

RESULTS

Characterization of Xylanase From *Bacillus* Strain D3

The molecular mass determined by the relative mobility on SDS/PAGE was 22 kDa, whereas using molecular mass spectrometry or calculation from the sequence gives a value of 20,700 Da. The pI is 7.7.[10] The xylanase does not bind Avicel, thus it has no exoglucanase activity.

Specific activity, pH optimum and stability

The specific activity of the xylanase at 60°C was determined as 2000 U/mg protein. The pH optimum is 6, but the enzyme is active over a broad pH range (Fig. 1). At pH 10, the xylanase hydrolyzed 10% of the soluable xylans, which were hydrolyzed at pH 6. This does not reflect the efficacity of the xylanase toward insoluable substrates.

As the most advantageous working conditions for using xylanases as bleach boosters in the pulp industry are high-temperature and alkaline conditions, the stability of the xylanase was studied at basic pH with different incubation times. At 80°C, no denaturation occurred within 24 hours at pH 9 and 10. For pH 11 the residual activity (at 60°C) was 94%, 85%, and 74% after 4, 6, and 24 hours, respectively, and for pH 12.5 the residual activity was 80%, 58%, and 25% after 4, 6, and 24 hours, respectively.

Temperature optimum and stability

The temperature optimum is 75°C. After 15 minutes incubation time, at 70°C the specific activity is 3000 U/mg of protein, whereas at 75°C the specific activity is 3300 U/mg of protein.

The enzymic stability was studied at 60°C. The thermal stability depends on the enzyme concentration. With enzyme concentrations of 1 to 2 mg/ml (without substrate) in 25% ethylene glycol, the xylanase is stable for 24 hours. For an enzyme concentration of 20 mg/ml, the residual activity is 60–70% after 24 hours of incubation. Addition of ethylene glycol (25% final) did not enhance the thermostability of xylanase at this concentration. For protein concentration of 200 mg/ml, the half-lives at 75°C and 80°C, respectively, are 40 and 25 minutes.

Cloning and Sequencing of *Bacillus* D3 Xylanase

The amino acid sequence of the N-terminal region of the xylanase has been determined as NTYWQYWTDGIGYVNATNGQG. The sequence shows 67% identity with the N terminus of xylanases produced by *Bacillus subtilis* and *Bacillus circulans*. Both these enzymes belong to the G family and share conserved regions along their polypeptide chains, including a region of 7 amino acids near the C terminus. A part of the gene coding for the xylanase has been amplified by PCR by using two degenerate primers, P1 and P2, corresponding to the N-terminal end of the xylanase and to the conserved region near the C terminus, respectively. A 450 bp DNA fragment was obtained

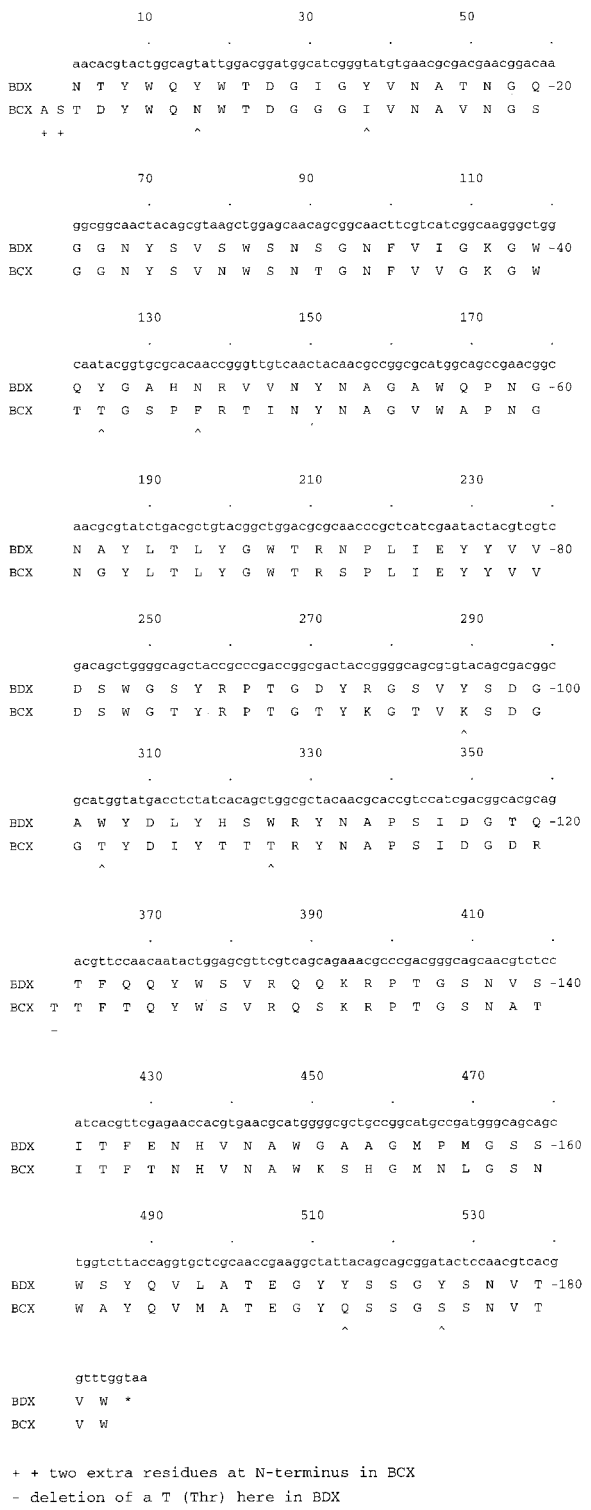

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the mature protein *Bacillus* D3 xylanase (BDX, **top**) and alignment with the amino acid sequence of *Bacillus circulans* xylanase (BCX, **bottom**). The nonconserved aromatic residues are indicated with a caret.

TABLE I. Sequence Homology of *Bacillus* D3 Xylanase With Family G Xylanases, Derived From a FASTA Search of Genbank (Release 94), Swiss-Prot (Release 32), and PIR (Release 48), for the 10 Best Scores of Identity

Acc. no.*	Source	%id†	No. aa‡
G:U15985	*Bacillus stearothermophilus*	76	211
P:S48126	*Bacillus* sp. YA-14 (xynS)	74	213
G:D32065	*Aeromonas caviae* (xylI)	74	211
S:P18429	*Bacillus subtilis* (xynA)	73	213
S:P09850	*Bacillus circulans*	73	213
G:U01242	*Thermomonospora fusca* (TFXA)	61§	338
S:P26220	*Streptomyces lividans* (xylC)	60	240
G:Z49892	*Aspergillus nidulans*	60	225
P:S47512	*Streptomyces* sp. *EC3*	58	240
S:P26515	*Streptomyces lividans* (xynB)	58§	333

*Accession number in G: GenBank, S: Swiss-Prot, P: PIR.
†Percentage of identity with *Bacillus* D3 xylanase protein sequence.
‡Number of amino acids of the precursor protein.
§N-terminal.

and cloned into pBluescript. The sequence of the resultant plasmid pBX20 can be attributed with certainty to *Bacillus* D3 xylanase. To obtain the complete gene of the xylanase, a genomic library of *Bacillus* D3 was prepared in *E. coli* XL1-blue by using the phage vector λ ZAP Express. This library was screened with the insert of the plasmid pBX20. One positive plaque, designated pBX52α2, was shown to contain the complete gene for *Bacillus* D3 xylanase.

The nucleotide sequence, together with the deduced protein sequence of the mature enzyme, is shown in Figure 2. The N-terminal sequence deduced from the nucleic acid sequence is identical to the result of the protein sequence determination, thereby proving that the gene coding for the desired enzyme had been cloned successfully. The xylanase is composed of 182 amino acids, corresponding to a calculated M_r of 20 683 Da, consistent with the value of 20,700 Da determined by mass spectroscopy for the purified enzyme.[12] The calculated pI value of 7.7 is in agreement with that determined by isoelectric focusing.[10]

The results of a FASTA search of the databases Genbank, PIR, and Swiss-Prot yielded more than 40 xylanase sequences. The xylanase shares sequence homology with other xylanases of the G family (see Table I). The protein sequence shows highest identity (more than 70% identity) as expected with the xylanases from *Bacillus subtilis* and *Bacillus circulans* and also with the recently published xylanase sequences from *Bacillus stearothermophilus*,[25] *Aeromonas caviae* (GenBank/EMBL Data Bank accession number D32056) and *Bacillus* sp. YA-14 (PIR accession number S48126). Unambiguously *Bacillus* D3 xylanase is a new protein that possesses a unique amino acid sequence.

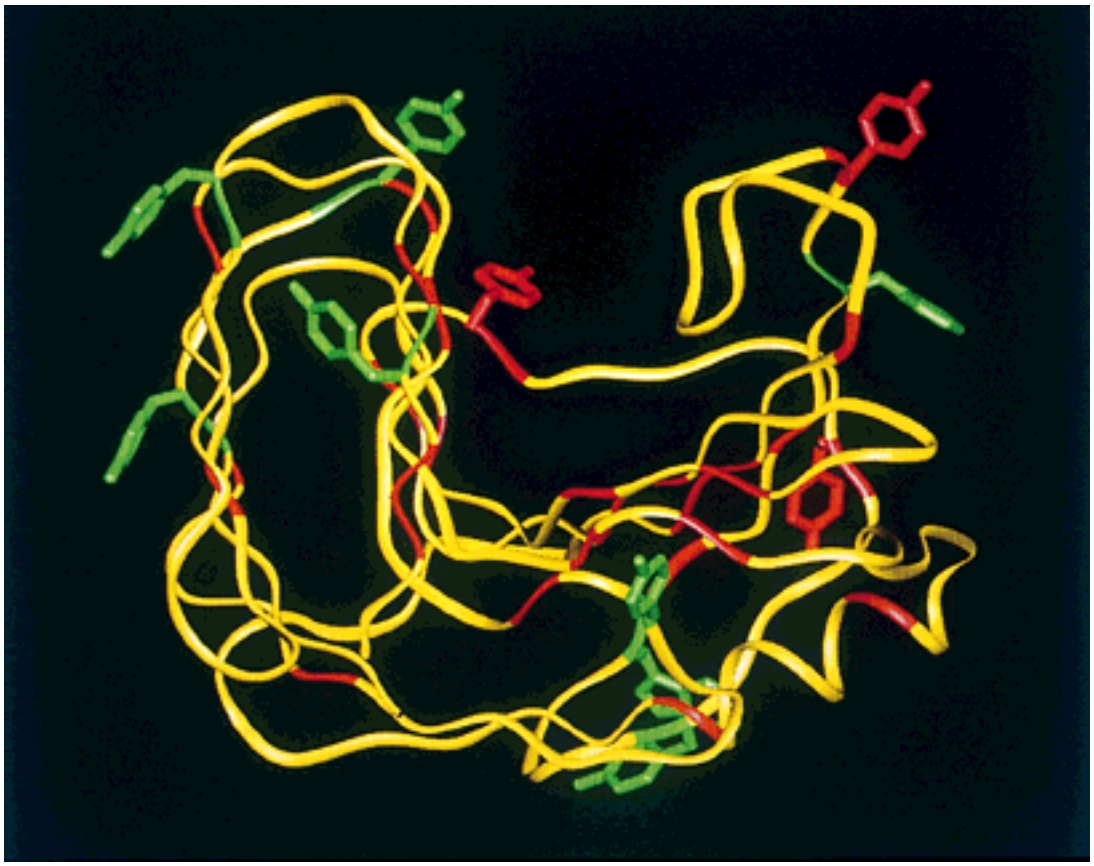

Fig. 3. Ribbon representation of the structure of *Bacillus* D3 xylanase (prepared by using the graphics program *Quanta*[42]). The three surface aromatic residues (Tyr86, Tyr 92, and Tyr 111) that are common to this structure and that of *Bacillus circulans* xylanase[7] are shown in red; the eight surface aromatic residues that are found only in *Bacillus* D3 xylanase are shown in green. The postions of all the aromatic residues (Trp, Tyr, Phe) that are common to both structure are indicated by the red stripes on the protein ribbon backbone.

Crystal Structure of Xylanase

Bacillus D3 xylanase crystallizes in the triclinic spacegroup P1 with four molecules in the asymmetric unit cell. There is apparently no simple noncrystallographic relationship between the four molecules. Molecules 1 and 4 are related by an approximate twofold axis down *X*, and molecules 2 and 3 are also related by the same approximate twofold axis down *X*. Viewed down *Y*, molecules 1 and 2, and 3 and 4, appear to be related by approximate twofold screw axes; these axes are in the same direction but not coincident. Similarly, viewed down *Z*, molecules 1 and 3, and 2 and 4, appear to be related by approximate twofold screw axes; again these axes are in the same direction but not coincident.

The xylanase structure consists of a single domain of regular antiparallel β-pleated sheet and one portion of α helix (Fig. 3). The overall shape of the molecule resembles a half-open hand with β sheets 1 and 2 packing together to form the thumb, and β sheet 3, which is essentially a continuation of β sheet 2 but perpendicular to it, forming the palm. The substrate binding site is formed by a deep cleft between the palm and the thumb. The single α helix lies across the upper palm of the hand on the opposite side to the cleft. The overall fold is similar to that of other family G xylanases[4–9] and most closely resembles that of *Bacillus circulans* with which *Bacillus* D3 xylanse has a high sequence homolgy (Fig. 2).

DISCUSSION

Industrial paper pulping ideally requires a xylanase that is thermostable, active at alkaline pH, and lacks cellulase activity. The thermophilic xylanase from *Bacillus* strain D3 has been shown to fulfil these stringent requirements and the production of the xylanase on industrial scale in optimal conditions is able to produce 1000units/ml.

The xylanase has one of the highest specific activities reported for xylanases (with temperature optima of 50–65°C).[26–35] Further, the xylanase has no detectable cellulase activity. The temperature opti-

mum is 75°C, though at this temperature some denaturation occurs. The enzyme is stable at 60°C and relatively stable at high temperatures.

As the most advantageous working conditions for using xylanases as bleach boosters in the pulp industry are high-temperature and alkaline conditions, the stability of the xylanase was studied at basic pH. The pH optimum is 6, but the enzyme is active over a broad pH range (see Fig. 1). The xylanase is relatively stable in extreme pH conditions encountered in industrial processes.

Highly thermophilic xylanases from the hyperthermophilic bacterium *Thermotoga maritima*[36–37] have temperature optima of 90–105°C, but the pH optima are 5.4–5.5. Even though these xylanases are highly thermophilic, their lack of stability at alkaline pH would render them of little use in the paper pulping process.

The thermophilic xylanase from *Bacillus* strain D3 (M_r 20 700 Da) can be assigned, on the basis of sequence similarity, to family G according to Henrissat's classification of cellulases and xylanases.[2] The xylanase has been shown by X-ray structural studies to be a single domain β-sheet protein (Fig. 3) consisting predominately of regular antiparallel β-pleated sheet and one portion of α helix. The fold is unique to the family G xylanases. The overall shape of the molecule resembles a half-open hand with a deep cleft between the palm and the thumb. Although the overall fold is similar to that of other family G xylanases,[4–9] the xylanase has distinctive structural features resulting from its specific sequence and these are responsible for the unique properties of this thermostable enzyme.

Multiple sequence alignment of *Bacillus* D3 xylanase with other family G xylanases confirms that the residues conserved throughout the family G xylanases, including two glutamic residues previously noted by Wakarchuk *et al*[7], are also conserved in the *Bacillus* D3 xylanase (data not shown). The three dimensional structure of the xylanase confirms these two glutamates, Glu76 and Glu169, situated at the bottom of the cleft, as the catalytic residues. The two active site glutamates are separated by a distance of 4.7Å, consistent with a retention of conformation reaction mechanism with Glu76 acting as the nucleophile and Glu169 as the acid-base catalyst. The catalytic glutamate residues are connected via hydrogen bonding to a water molecule (O184). This water molecule is well defined in the electron density map, as are six further water molecules (O202, O201, O208, O183, O185 and O211) lying in the active site cleft. The seven water molecules (three on either side of O184) form a chain 25 Å in length, which stretches along the substrate binding cleft. These water molecules would be displaced when the substrate binds. Aromatic residues (Trp126, Tyr78, Tyr67, Tyr163) line the base of the cleft and further aromatic residues are to be found along the sides of the cleft. These can interact with the substrate on binding.

Despite the high similarity of *Bacillus* D3 xylanase with *Bacillus circulans* xylanase both in terms of primary structure (73% sequence identity) and tertiary structure (rms deviation of 0.063 Å for Cα atoms), *Bacillus circulans* xylanase is not stable at temperatures above 60°C,[38] whereas *Bacillus* D3 xylanase is stable at 75°C. It is possible to speculate on structural reasons for the higher thermostability of the *Bacillus* D3 xylanase.

The thermostability (though not necessarily the activity) of the xylanase from *Bacillus circulans* can be enhanced by the introduction of disulfide bonds.[39] As the *Bacillus* D3 xylanase has no cysteine residues for the formation of disulfide bonds, the thermostability must result from other factors such as electrostatic or hydrophobic interactions.

Comparing the sequence of *Bacillus* D3 xylanase to that of *Bacillus circulans* xylanase, there are eight aromatic residues that are present in the *Bacillus* D3 xylanase structure, but not in the *Bacillus circulans* xylanase structure (in *Bacillus* D3 xylanase these residues are Tyr6, Tyr13, Tyr42, Tyr97, Trp102, Trp109, Tyr172, and Tyr176; see Fig. 2). All these residues are on the surface of the protein (Figs. 3 and 4). By contrast, there is only one aromatic residue in *Bacillus circulans* xylanase (Phe48) that is not correspondingly an aromatic in *Bacillus* D3 xylanase (see Fig. 2). The 29 aromatic residues that are common to both xylanases are mostly found in the interior of the protein, packing between the β sheets, as is usual for protein structures, or in the active site cleft where they play a role in the binding of substrate. Only three of these 29 aromatics are on the surface of the protein (Tyr86, Tyr92, and Tyr 111; see Fig. 3). In *Bacillus* D3 xylanase, the eight additional surface aromatics form clusters or "sticky patches" (Fig. 4). These patches tend to occur between pairs of molecules (e.g., Tyr97 and Trp102 in molecule 1 form a patch with Trp109 in molecule 2). The three surface tyrosines also present in *Bacillus circulans* xylanase are also involved in these hydrophobic clusters. This tendency to aggregate could explain why the enzyme crystallizes with four molecules in the P1 unit cell (Fig. 4). Aromatic–aromatic interactions have been shown to contribute significantly to protein stability[40] as have surface-exposed hydrophobic pockets.[41] The extra stability conferred on the enzyme by the hydropobic interactions resulting from these "sticky patches" on the surface of the enzyme could account for the high thermostability of this xylanase.

These hydrophobic interactions could also explain the unusual behavior of the *Bacillus* D3 xylanase in solution. The enzyme requires high concentrations of ethylene glycol for solubility.[12] In addition, higher enzyme concentrations require higher concentrations of ethylene glycol to keep the protein in solution.

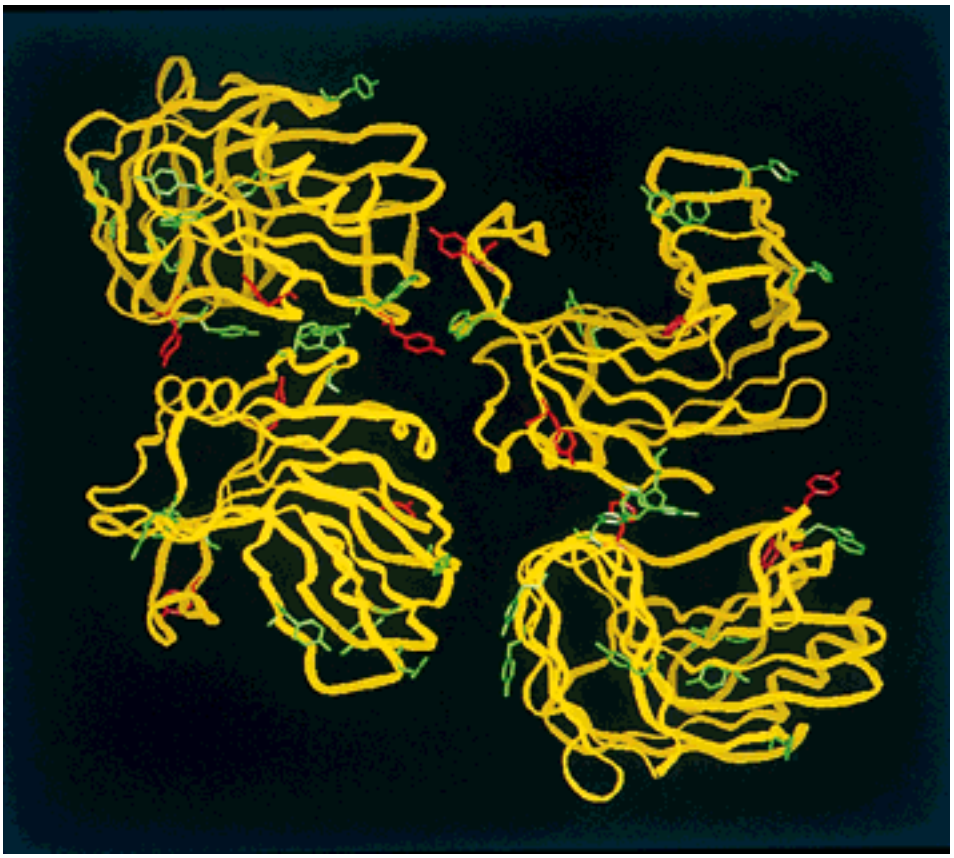

Fig. 4. Crystal packing of the molecules of *Bacillus* D3 xylanase (ribbon representation, prepared by using the graphics program *Quanta*[42]) showing the hydrophobic clusters or "sticky patches" between molecules. The eight surface aromatic residues that are found only in *Bacillus* D3 xylanase are shown in green, and the three surface aromatic residues (Tyr86, Tyr 92, and Tyr 111) that are common to *Bacillus* D3 xylanase and *Bacillus circulans* xylanase[7] are shown in red. The four unique molecules in the P1 asymmetric unit cell are shown.

The thermal stability of *Bacillus* D3 xylanase depends on enzyme concentration, with the stability decreasing with decreasing enzyme concentration. This reduction of thermal stability can be explained in terms of the loss of stabilizing hydrophobic contacts between molecules as the enzyme solution is made more dilute.

There have been several studies reported where the structure of an enzyme from a thermophile has been compared to its mesophilic counterpart and conclusions drawn with regard to the structural basis of thermostability. For triosephosphate isomerase from *Bacillus stearothermophilus,* hydrophobic stabilization of dimer formation by burying the largest amount of hydrophobic surface area was the most important factor[43]. For 3-phosoglycerate kinase from *Bacillus stearothermophilus,* contributing factors found were increased internal hydrophobicity, additional ion pairs and better α-helix stability resulting from removal of helix destabilizing residues and extra helix–dipole/helix–side chain ionic interactions.[44] Helix stabilization, as well as shorter, less flexible loops and tighter packing were considered important in citrate synthase from *Thermoplasma acidophilum*.[45] Ion-pair networks on the surface of the protein buried at domain or subunit interfaces were considered the key factor for glutamate dehydrogenase from *Pyrococcus furiosus*.[46,47] For malate dehydrogenase from *Thermus flavus,* additional ion pairs at interfaces and α-helix stabilization were implicated.[48] Additional salt bridges were also important for hologlyceraldehyde-3-phosphate dehydrogenase from *Themotoga maritima*[49] and cyclodextrin glycosyltranferase from *Thermoaerobacterium thermosulfurigenes* EM1[50] and for D-glyceraldehyde-3-phosphate dehydrogenase from *Thermus aquaticus,* charged-neutral hydrogen bonds as well as salt links were considered important.[51] Although several factors giving enhanced thermostability have been found in the enzyme structures studied, the hydropobic interactions of aromatic residues on the surface of the enzyme, suggested as the structural basis for stability of the *Bacillus* D3

xylanase structure has not been observed in any other enzyme.

ACKNOWLEDGMENTS

This work was supported by the BBSRC and INRA. We thank Dr. Robert Campbell for sending us the coordinates of the *Bacillus circulans* xylanase structure before they were available from the Brookhaven Protein Database, and Professor Katsube and Dr. Hata, Osaka University, Japan, for allowing us to use the coordinates of the *Bacillus pumilus* xylanase structure.

REFERENCES

1. Samain, E., Debeire, P. Debeire-Gosselin, M., Touzel, J.P. Xylanase, souches de *Bacillus* productrices de xylanase et leurs utilisations. Patent FR-9101191, 1991.
2. Gilkes, N.R., Henrissat, B., Kilburn, D.G., Miller, R.C. Jr.,, Warren, R.A.J. Domains in microbial β-1,4-glycanases: Sequence conservation, function, and enzyme families. Microbiol. Rev. 55:303–315, 1991.
3. Henrissat, B. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280:309–316, 1991.
4. Arase, A., Yomo, T., Urabe, I., Hata, Y., Katsube, Y., Okada, H. Stabilization of xylanase by random mutagenesis. FEBS Lett. 316:123–127, 1993.
5. Dauter, Z., Lamzin, V.S., Dauter, M., Bisgard-Frantzen, H., Halkier, T., Wilson, K.S. X-ray structure of the seed protein Narbonin and of xylanase from *Bacillus pumilus*. Proceedings of the Third AFRC Protein Engineering Conference, Cambridge, U.K., 27–28 September 1993.
6. Campbell, R.L., Rose, D.R., Wakarchuk, W.W., To, R., Sung, W., Yaguchi, M. A comparison of the structures of the 20 kd xylanases from *Trichoderma harzianum* and *Bacillus circulans*. Found. Biotechnol. Ind. Ferment. Res. Publ. 8:63–72, 1993.
7. Wakarchuk, W.W., Campbell, R.L., Sung, W.L., Davoodi, J., Yaguchi, M. Mutational and crystallographic analyses of the active site residues of the *Bacillus circulans* xylanase. Protein Sci. 3:467–475, 1994.
8. Törrönen, A., Harkki, A., Rouvinen, J. Three-dimensional structure of endo-1,4-β-xylanase II from *Trichoderma reesei*: Two conformational states in the active site. EMBO J. 13:2493–2501, 1994.
9. Törrönen, A., Rouvinen, J. Structural comparison of two major endo-1,4-β-xylanases from *Trichoderma reesei*. Biochemistry 34:847–856, 1995.
10. Debeire-Gosselin, M., Loonis, M., Samain, E., Debeire, P. Purification and properties of a 22kDa endoxylanase excreted by a new strain of thermophilic *Bacillus*. In "Xylans and Xylanases." Visser, J., et al. (eds.). Amsterdam: Elsevier, 1992:463–466.
11. Samain, E., Touzel, J.P., Brodel, B., Debeire, P. Isolation of a thermophilic bacterium producing high levels of xylanase. In "Xylans and Xylanases." Visser, J., et al. (eds.). Amsterdam: Elsevier, 1992:467–470.
12. Pickersgill, R.W, Debeire, P., Debeire-Gosselin, M., Jenkins, J.A. Crystallization and preliminary X-ray analysis of a thermophillic *Bacillus* xylanase. J. Mol. Biol. 230:664–666, 1993.
13. Kidby, D.K., Davidson, D.J. A convenient ferricyanide estimation of reducing sugars in the nanomole range. Anal. Biochem. 55:321–325, 1973.
14. Debeire, P., Priem, B., Strecker, G., Vignon, M. Purification and properties of an endo-1,4-xylanase excreted by a hydrolytic thermophilic anaerobe, Clostridium thermolacticum: A proposal for its action mechanism on larchwood 4-*O*-methylglucuronoxylan. Eur. J. Biochem. 187:573–580, 1990.
15. Yang, R.C.A., Mackenzie, R., Bilous, D., Seligy, V.L., Narang, S.A. Molecular cloning and expression of a xylanase gene from *Bacillus polymyxa* in *Escherichia coli*. Appl. Environ. Microbiol. 54:1023–1029, 1988.
16. Sanger, F., Nicklen, S., Coulson, A.R. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. 74:5463–5467, 1977.
17. Howard, A.J. et al. The use of an imaging proportional counter in macromolecular crystallography. J. Appl. Crystallogr. 20:383–387, 1987.
18. Navaza, J. AMoRe:an automated package for molecular replacement. Acta Crystallogr. A50:157–163, 1994.
19. Brünger, A.T. "X-PLOR Version 3.1: A system for X-ray Crystallography and NMR." New Haven: Yale University Press, 1992.
20. Collaborative Computational Project, Number 4, SERC Daresbury Laboratory. The CCP4 suite: Programs for protein crystallography. Acta Crystallogr. D50:760–763, 1994.
21. Jones, T.A., Zou, J.-Y., Cowan, S.W., Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A47:110–119, 1991.
22. Kleywegt, G.J., Jones, T.A. "RAVE: The Manual, Version 1.0." Uppsala, Sweden: University of Uppsala, 1994.
23. Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M. PROCHECK: A program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26:283–291, 1993.
24. Sutcliffe, M.J., Haneef, I., Carney, D., Blundell, T.L. Knowledge based modeling of homologous proteins. 1. Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. Protein Eng. 1:377–384, 1987.
25. Cho, S.G., Choi, Y.J. Nucleotide sequence analysis of an endo-xylanase gene (xynA) from *Bacillus stearothermophilus*. J. Microbiol. Biotechnol. 5:117–124, 1995.
26. Uchino, F., Nakane, T. A thermostable xylanase from a thermophilic acidophilic *Bacillus* sp. Agric. Biol. Chem. 45:1121–1127, 1981.
27. Berenger, J.-F., Frixon, C., Bigliardi, J., Creuzet, N. Production, purification and properties of thermostable xylanase from *Clostridium stercorarium*. Can. J. Microbiol. 31:635–643, 1985.
28. Debeire-Gosselin, M., Touzel, J.P., Debeire, P. Isoxylanases from the thermophile *Clostridium thermolacticum*. In "Xylans and Xylanases." Visser, J., et al. (eds.). Amsterdam: Elsevier, 1992:471–474.
29. Anand, L., Krishnamurthy, S., Vithayathil, P.J. Purification and properties of xylanase from the thermophilic fungus *Humicola lanuginosa* (Griffon and Maublanc) Bunce. Arch. Biochem. Biophys. 276:546–553, 1990.
30. Nanmori, T., Watanabe, T., Shinke, R., Kohno, A., Kawamura, Y. Purification and properties of thermostable xylanase and β-xylosidase produced by a newly isolated *Bacillus stearothermophilus* strain. J. Bacteriol. 172:6669–6672, 1990.
31. Grabski, A.C., Jeffries, T.W. Production, purification and characterization of thermostable xylanase and β-(1,4)-endoxylanase of *Streptomyces roseiscleroticus*. Appl. Environ. Microbiol. 57:987–992, 1991.
32. Monti, R., Terenzi, H.F., Jorge, J.A. Purification and properties of an extracellular xylanase from the thermophilic fungus *Humicola Grisea* var. *thermoidea*. Can. J. Microbiol. 37:675–681, 1991.
33. Gilbert, M., Breuil, C., Yaguchi, M., Saddler, J.N. Purification and characterization of a xylanase from the thermophilic ascomycete *Thielavia terrestris* 255B. Appl. Biochem. Biotechnol. 34–35:247–259, 1992.
34. Khasin, A., Alchanati, I., Shoham, Y. Purification and characterization of a thermostable xylanase from *Bacillus stearothermophilus* T-6. Appl. Environ. Microbiol. 59:1725–1730, 1993.
35. Ritschkoff, A.C., Buchert, J., Viikari, L. Purification and characterization of a thermophilic xylanase from the brown-rot fungus *Gloeophyllum trabeum*. J. Biotechnol. 32:67–74, 1994.

36. Simpson, H.D., Haufler, U.R., Daniel, R.M. An extremely thermostable xylanase from the thermophilic eubacterium *Thermotoga*. Biochem. J. 277:413–417, 1991.
37. Winterhalter, C., Liebl, W. Two extremely thermostable xylanases of the hyperthermophilic bacterium *Thermotoga maritima* MSB8. Appl. Environ. Microbiol. 61:1810–1815, 1995.
38. Davoodi, J., Wakarchuk, W.W., Campbell, R.L., Carey, P.R., Surewicz, W.K. Abnormally high pK_a of an active-site glutamic acid residue in *Bacillus circulans* xylanase. Eur. J. Biochem. 232:839–843, 1995.
39. Wakarchuk, W.W., Sung, W.L., Campbell, R.L., Cunningham, A., Watson, D.C., Yaguchi, M. Thermostabilization of the *Bacillus circulans* xylanase by the introduction of disulfide bonds. Protein Eng. 7:1379–1386, 1994.
40. Serrano, L., Bycroft, M., Fersht, A.R. Aromatic-aromatic interactions and protein stability. Investigation by double-mutant cycles. J. Mol. Biol. 218, 465–475, 1991.
41. Frigerio, F., Margarit, I., Nogarotto, R., de Filippis, V., Grandi, G. Cumulative stabilizing effects of hydrophobic interactions on the surface of the neutral protease from *Bacillus subtilis*. Prot. Eng. 9:439–445, 1996.
42. Molecular Simulations Inc. Quanta 4.1. Molecular Simulations Inc., 200 Fifth Avenue, Waltham, MA 02154, USA, 1994.
43. Delboni, L.F., Mande, S.C., Rentier-Delrue, F., Mainfroid, V., Turley, S., Vellieux, F.M.D., Martial, J.A., Hol, W.G.J. Crystal structure of recombinant triosephosphate isomerase from *Bacillus stearothermophilus*: An analysis of potential thermostability factors in 6 isomerases with known 3-dimensional structures points to the importance of hydrophobic interactions. Protein Sci. 4:2594–2604, 1995.
44. Davies, G.J., Gamblin, S.J., Littlechild, J.A., Watson, H.C. The structure of a thermally stable 3-phosoglycerate kinase and a comparison with its mesophilic equivalent. Proteins 15:283–289, 1993.
45. Russell, R.J.M., Hough, D.W., Danson, M.J., Taylor, G.L. The crystal structure of citrate synthase from the thermophilic Archaeon, *Thermoplasma acidophilum*. Structure 2:1157–1167, 1994.
46. Yip, K.S.P., Stillman, T.J., Britton, K.L., Artymiuk, P.J., Baker, P.J., Sedelnikova, S.E., Engel, P.C., Pasquo, A., Chiaraluce, R., Consalvi, V., Scandurra, R., Rice, D.W. The structure of *Pyrococcus furiosus* glutamate dehydrogenase reveals a key role for ion-pair networks in maintaining enzyme stability at extreme temperatures. Structure 3:1147–1158, 1995.
47. Rice, D.W., Yip, K.S.P., Stillman, T.J., Britton, K.L., Fuentes, A., Connerton, I., Pasquo, A., Scandurra, R., Engel, P.C. Insights into the molecular basis of thermostability from the structure determination of *Pyrococcus furiosus* glutamate dehydrogenase. FEMS Microbiol. Rev. 18:105–117, 1996.
48. Kelly, C.A., Nishiyama, M., Ohinishi, Y., Beppu, T., Birktoft, J.J. Determinants of protein thermostability observed in the 1.9 Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. Biochemistry 32:3913–3922, 1993.
49. Korndörfer, I., Steipe, B., Huber, R., Tomschy, A., Jaenicke, R. The crystal structure of holo-glyceraldehyde-3-phosphate dehydrogenase from the hyperthermophilic bacterium *Themotoga maritima* at 2.5 Å resolution. J. Mol. Biol. 246:511–521, 1995.
50. Knegtel, R.M.A., Wind, R.D., Rozeboom, H.J., Kalk, K.H., Buitelaar, R.M., Dijkhuizen, L., Dijkstra, B.W. Crystal structure at 2.3Å resolution and revised nucleotide sequence of the thermostable cyclodextrin glycosyltranferase from *Thermoaerobacterium thermosulfurigenes* EM1. J. Mol. Biol. 256:611–622, 1996.
51. Tanner, J., Hecht, R.M., Krause, K.L. Determinants of enzyme thermostability observed in the molecular structure of *Thermus aquaticus* D-glyceraldehyde-3-phosphate dehydrogenase at 2.5 Å resolution. Biochemistry 35:2597–2609, 1996.