Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Thermostable xylanase from *Thermoascus aurantiacus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anisotropically to small-molecule accuracy

R. Natesh,[a]† K. Manikandan,[a] P. Bhanumoorthy,[a]‡ M. A. Viswamitra[a] and S. Ramakumar[a,b]*

[a]Department of Physics, Indian Institute of Science, Bangalore 560 012, India, and [b]Bioinformatics Centre, Indian Institute of Science, Bangalore 560 012, India

† Present address: Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, England.
‡ Present address: Department of Biological Sciences, Purdue University, West Lafayette, Indiana, IN 47907, USA.

Correspondence e-mail: ramak@physics.iisc.ernet.in

Received 22 March 2002
Accepted 5 November 2002

PDB References: xylanase, 1.11 Å structure, 1i1x, r1i1xsf; 0.89 Å structure, 1i1w, r1i1wsf.

Dedicated to the late Professor M. A. Viswamitra.

Thermoascus aurantiacus xylanase is a thermostable enzyme which hydrolyses xylan, a major hemicellulose component of the biosphere. The crystal structure of this F/10 family xylanase, which has a triosephosphate isomerase (TIM) barrel $(\beta/\alpha)_8$ fold, has been solved to small-molecule accuracy at atomic resolution (1.11 Å) at 293 K (RTUX) and at ultrahigh resolution (0.89 Å) at 100 K (CTUX) using X-ray diffraction data sets collected on a synchrotron light source, resulting in $R/R_{\rm free}$ values of 9.94/12.36 and 9.00/10.61% (for all data), respectively. Both structures were refined with anisotropic atomic displacement parameters. The 0.89 Å structure, with 177 476 observed unique reflections, was refined without any stereochemical restraints during the final stages. The salt bridge between Arg124 and Glu232, which is bidentate in RTUX, is water-mediated in CTUX, suggesting the possibility of plasticity of ion pairs in proteins, with water molecules mediating some of the alternate arrangements. Two buried waters present inside the barrel form hydrogen-bond interactions with residues in strands $\beta2$, $\beta3$, $\beta4$ and $\beta7$ and presumably contribute to structural stability. The availability of accurate structural information at two different temperatures enabled the study of the temperature-dependent deformations of the TIM-barrel fold of the xylanase. Analysis of the deviation of corresponding C^{α} atoms between RTUX and CTUX suggests that the interior β-strands are less susceptible to changes as a function of temperature than are the α-helices, which are on the outside of the barrel. $\beta\alpha$-loops, which are longer and contribute residues to the active-site region, are more flexible than $\alpha\beta$-loops. The 0.89 Å structure represents one of the highest resolution structures of a protein of such size with one monomer molecule in the asymmetric unit and also represents the highest resolution TIM-barrel fold structure to date. It may provide a useful template for theoretical modelling studies of the structure and dynamics of the ubiquitous TIM-barrel fold.

1. Introduction

Xylanases (xylan-degrading enzymes) are classified into two distinct families, F/10 and G/11, of the glycoside hydrolases (Henrissat & Davies, 1997). They hydrolyse xylan, the most abundant hemicellulose in plant cell-wall polysaccharides. The study of xylanases has assumed great importance in view of their industrial applications, which include bio-bleaching in the paper industry, which is economical and eco-friendly, and applications in the pharmaceutical industry and in the clarification of fruit juices. The structures from the F/10 family possess the $(\beta/\alpha)_8$ TIM-barrel fold, first observed in triosephosphate isomerase (Banner *et al.*, 1975), while those from the G/11 family predominantly consist of β-sheets. Despite

their structural differences, remarkably they function in the same manner, with net retention of anomeric configuration (double-displacement mechanism) in the enzymatic hydrolysis (Henrissat & Davies, 1997). The mode of action of these enzymes is governed by exquisite details of their three-dimensional structures rather than their globular fold (Davies & Henrissat, 1995). The crystal structure of the thermostable F/10 family *Thermoascus aurantiacus* xylanase has been studied by us at 1.8 Å resolution (Natesh *et al.*, 1999) and by others at 1.14 Å resolution and as complexes at lower resolution (Lo Leggio *et al.*, 1999, 2001; Teixeira *et al.*, 2001). We have determined the crystal structure of the *T. aurantiacus* xylanase to small-molecule accuracy at atomic resolution (1.11 Å; room temperature, 293 K; RTUX) and at ultrahigh resolution (0.89 Å; cryotemperature, 100 K; CTUX). Both structures were refined with anisotropic atomic displacement parameters (ADPs). Anisotropic refinement with restraints based on the 1.14 Å resolution X-ray data collected at room temperature for *T. aurantiacus* xylanase (PDB code 1k6a) has been reported previously by Teixeira *et al.* (2001). The root-mean-square (r.m.s.) deviation of corresponding C^{α} atoms between 1k6a and RTUX is 0.18 Å. However, we have chosen to compare the CTUX structure with the RTUX structure determined in our laboratory since both are from the same strain of the organism isolated from the local Indian soil. This helps to bring out more clearly the temperature-dependent changes in the structure, without having to be concerned about the sequence changes which are a consequence of the protein being isolated from different strains of the organism (the amino-acid sequence of 1k6a has seven differences to that for RTUX or CTUX). The 0.89 Å structure to our knowledge is the highest resolution TIM-barrel fold structure determined to date, which is important considering the fact that roughly 10% (revised to ∼12.2%) of all enzymes of known structure assume the TIM-barrel fold (Pujadas & Palau, 1999; Farber, 1993).

The overall advantages and benefits of atomic resolution (∼1.2 Å) or even beyond to ultrahigh resolution (better than 0.95 Å) X-ray crystallography have been reviewed (Longhi *et al.*, 1998; Dauter, Lamzin *et al.*, 1997; Bott & Boelens, 1999). It has been amply demonstrated that increased resolution along with improved refinement methods results in models with higher accuracy, which can serve as accurate input for theoretical studies (Dauter, Lamzin *et al.*, 1997). They can provide a wealth of information about the structurally and functionally relevant plasticity of macromolecules. Some of the advantages of atomic and ultrahigh-resolution refinement are that anisotropic ADPs may be used, many H atoms can be seen in electron-density maps, multiple (dual) conformations can be identified for the residues and water structure emerges from the solvent continuum.

We have discussed all these features utilizing atomic and ultrahigh-resolution structures of the study enzyme, *T. aurantiacus* xylanase. The structures have enabled us to characterize geometric hidden strain (Karplus, 1996) which may be present in some regions of the protein. Salt bridges between oppositely charged groups are important contributors to protein structure stability (Anderson *et al.*, 1990) and in particular to stability at high temperature (Elcock, 1998). The flexibility of ion pairs in proteins has recently been analysed (Kumar & Nussinov, 2001). A comparison of RTUX and CTUX has provided crystallographic evidence for plasticity as well as water-mediated rearrangements which might occur in ion pairs present in proteins. Our work suggests that the water structure around the protein is not rigid and the flexibility of amino-acid side chains can result in alternate modes of interaction. Based on the availability of accurate structural information at two different temperatures, inferences have been drawn about the temperature-dependent deformations of the ubiquitous TIM-barrel fold from the study of the *T. aurantiacus* xylanase structure.

Table 1
Crystallographic data-collection details.

Values in parentheses are for the last resolution shell.

	CTUX (100 K)	RTUX (293 K)
Space group	$P2_1$	$P2_1$
Unit-cell parameters (Å, °)	a = 41.05, b = 66.99, c = 50.76, β = 113.50	a = 41.52, b = 68.09, c = 51.44, β = 113.56
Wavelength (Å)	0.98	1.009
Unit-cell volume (Å^3)	128419.97	133791.79
Total No. of measured reflections	858643	294946
Multiplicity of data set	4.80	3.25
Total No. of unique reflections	177476	90855
Resolution range (Å)	35.00–0.89 (0.92–0.89)	25.00–1.11 (1.15–1.11)
Completeness (%)	92.0 (83.0)	87.5 (81.3)
$R_{\rm merge}$ (%)	4.7 (18.0)	5.2 (36.0)
$I/\sigma(I)$	32.65 (7.82)	20.60 (2.98)

2. Materials and methods

2.1. Data collection

Xylanase was obtained from a strain of the thermophilic fungus *T. aurantiacus* isolated from local Indian soil by ion-exchange and gel-permeation chromatography (Khandke, 1986; Khandke *et al.*, 1989). Crystals of *T. aurantiacus* xylanase were grown as described previously (Viswamitra *et al.*, 1993).

Both room-temperature (293 K) 1.11 Å and cryotemperature (100 K) 0.89 Å data were collected at the National Synchrotron Light Source, Brookhaven National Laboratory, beamline X9B, using the Quantum 4 CCD detector (ADSC). For cryotemperature data collection, a single crystal of *T. aurantiacus* xylanase was scooped out in a fibre loop and immersed in a stock solution of 10% $(NH_4)_2SO_4$ pH 7.4, 50 m*M* Tris–HCl containing 15% glycerol for few seconds and then placed in a stream of cold nitrogen gas at 100 K. Room-temperature data were collected with a crystal inside a sealed capillary tube. To cover the whole range of intensities, three sets of images were recorded, differing in exposure times and resolution limits, for the cryodata and two sets of images were recorded for the room-temperature data.

The diffraction data were processed with the *HKL*2000 suite (Otwinowski & Minor, 1997). Table 1 shows a summary of the data collection. The overall temperature factor estimated from the Wilson plot (Wilson, 1942) based on data from the frozen crystal is 4.2 Å^2 and from the room-temperature crystal is 7.9 Å^2.

2.2. Structure refinement

Initial refinements of the RTUX and CTUX structures were carried out using *CNS* v.0.4 (Brünger *et al.*, 1998). The free *R* set for both the 1.11 and 0.89 Å data sets were set up using *CNS*. A randomly selected total of 1745 reflections (2%) and 1766 reflections (1%) from the 1.11 and 0.89 Å data sets, respectively, were set aside for $R_{\rm free}$ calculations (Brünger, 1992; Kleywegt & Jones, 1995). The CTUX unit-cell parameters were significantly different (see §4.1) from the RTUX unit-cell parameters and hence maintaining the same $R_{\rm free}$ set as in the 1.11 Å data is not meaningful. The free *R* sets were maintained throughout the refinement. All the *B* factors were kept isotropic and no alternate conformations were modelled during the *CNS* refinement. After convergence in *CNS*, the models were subsequently refined using *SHELXL*97 (Sheldrick & Schneider, 1997; Sheldrick, 1997) with the CGLS (conjugate-gradient least squares) option and all the ADPs were kept isotropic. Temperature factors were subsequently made anisotropic, alternate conformations modelled and iterative cycles of model building and refinement led to the final convergence of the 1.11 and 0.89 Å restrained structures. Model building was carried out using *O* (Jones *et al.*, 1991). Any new atoms added to the protein were isotropically refined for at least two cycles before they were made anisotropic. The multiple-conformation site-occupation factors were refined constraining their sum to be unity. Engh & Huber (1991) restraints on the bond distances and angles were used in restrained refinements. The data-to-parameter ratio of 7.02 for the 0.89 Å data enabled us to carry out unrestrained full-matrix least-squares refinement on the CTUX model using *SHELXL*97. The pyrrolidonecarboxylic acid (Pca; also known as pyroglutamic acid) was not restrained in either the 1.11 Å RTUX or the 0.89 Å CTUX structure.

2.3. Details of the restrained and unrestrained refinement of the *T. aurantiacus* xylanase structure at 1.11 and 0.89 Å resolution

2.3.1. Atomic resolution (1.11 Å) structure. Refinement of the RTUX model was carried out first since the room-temperature data became available prior to the cryodata. The starting model was a water-free 1tux (Natesh *et al.*, 1999) model in the RTUX cell, having an initial *R* and $R_{\rm free}$ of 28.42 and 30.79%, respectively, for all data in the resolution range 25–1.11 Å. Some of the few ambiguities that existed in the 1.8 Å model (Natesh *et al.*, 1999) sequence could be easily resolved by simply examining the atomic resolution map (see below). The solvent molecules were added as described previously in the case of the 1.8 Å structure (Natesh *et al.*, 1999). A few cycles of minimization and isotropic individual *B*-factor refinement led to a model with an *R* and $R_{\rm free}$ of 18.83 and 19.99%, respectively. An extra residue Gln302 was modelled (Fig. 1). However, residue 303 was not visible even in the atomic resolution map. Refinement of the 0.89 Å CTUX structure was started at this stage. After completion of the 0.89 Å CTUX structure, the additional information observed from the CTUX final model was used to update the RTUX model in regions of ambiguity. This led to a model having an *R* and $R_{\rm free}$ of 18.73 and 19.91, respectively. Refinement was then continued in *SHELXL*97 (CGLS) which gave an RTUX model with an *R* and $R_{\rm free}$ of 17.66 and 19.54%, respectively, for all data in the resolution range 10–1.11 Å (low-resolution limit defaulted to *SHELXL* value). ADPs of all atoms of the model were converted to anisotropic (instruction ANIS; Sheldrick & Schneider, 1997), leading to a dramatic drop in *R* and $R_{\rm free}$ to values of 12.63 and 14.89%, respectively. At this stage, most of the possible alternate conformations in the RTUX were modelled using the improved Fourier map. A total of 267 waters were picked, H atoms were included in calculated positions in all cases (except for multiple-conformation atoms and heteroatoms other than pyrrolidonecarboxylic acid) and all non-H atoms were refined anisotropically including water molecules (except for protein multiple-conformation atoms). Ten cycles of CGLS refinement reduced *R* and $R_{\rm free}$ to 10.17 and 12.42%, respectively. Nearly a 1.75% drop in $R_{\rm free}$ was witnessed when H atoms were added, indicating the significance of hydrogen contributions in atomic resolution refinement of proteins. After incorporating the Asn, Asp, Gln and Glu information in the 1.11 Å structure (borrowed from the 0.89 Å structure), the

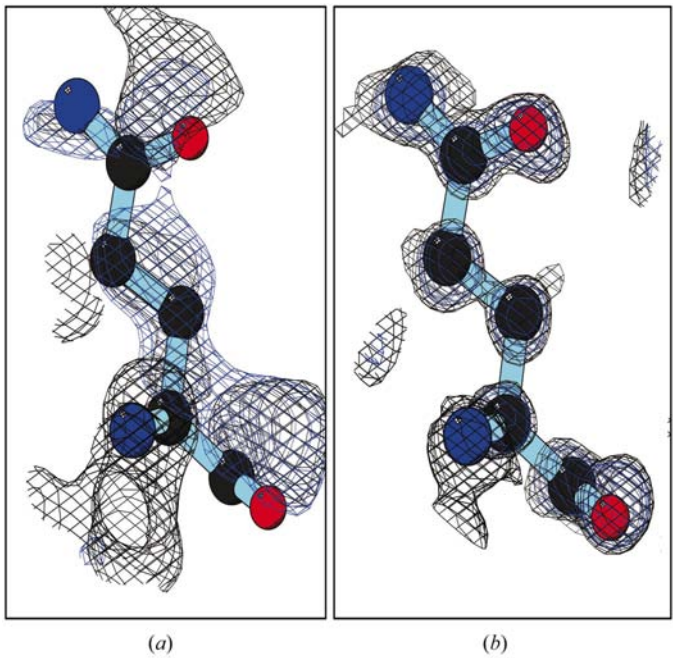

Figure 1
Electron density for residue 302 at (*a*) 1.11 Å and (*b*) 0.89 Å resolution, at a stage when the residue was not included in the refinement. ($2F_o - F_c$) map (black) contoured at 1σ and ($F_o - F_c$) map (blue) contoured at 3σ. As the figure shows, the 0.89 Å map is superior to the 1.11 Å map. Figures were produced using *BOBSCRIPT* (Esnouf, 1997).

Table 2
Refinement statistics of 1.11 and 0.89 Å structures and other details.

	1.11 Å structure (RTUX)	0.89 Å structure (CTUX)
Program	*CNS*/*SHELXL*97	*CNS*/*SHELXL*97
Restraints	Engh & Huber (1991). None for Pca†	Initially Engh & Huber (1991) except Pca†. Finally none (unrestrained)
Restraints/parameters	28345/22784	None/24867
Resolution range (Å)	10.0–1.11	10.0–0.89
Final R factor, $F_o > 4\sigma(F_o)$‡ (%)	9.16 (76475)	8.85 (168479)
Final $R_{\rm free}$, $F_o > 4\sigma(F_o)$‡ (%)	11.67 (1571)	10.41 (1707)
Final R factor, all data‡ (%)	9.94 (85628)	9.00 (174690)
Final $R_{\rm free}$, all data‡ (%)	12.36 (1745)	10.61 (1766)
wR^2§ (%)	23.44	24.95
Goodness of fit (GooF¶)	1.770	2.283

† Pca, pyrrolidonecarboxylic acid. ‡ The number of reflections is given in parentheses. § $wR^2 = \{\sum[w(F_o^2 - F_c^2)^2]/\sum[w(F_o^2)^2]\}^{1/2}$. ¶ GooF = $\{\sum[w(F_o^2 - F_c^2)^2]/(n-p)\}^{1/2}$.

final refinement converged to the RTUX structure with R and $R_{\rm free}$ values of 9.94 and 12.36%, respectively, for all data in the resolution range 10.00–1.11 Å and R and $R_{\rm free}$ of 9.16 and 11.67%, respectively, for $F_o > 4\sigma(F_o)$, with $wR^2 = 0.2344$ and GooF (goodness of fit) = 1.770 (Table 2).

2.3.2. Ultrahigh-resolution (0.89 Å) structure. The intermediate refined room-temperature 1.11 Å structure with all waters removed placed in the CTUX cell was used as a starting model for 0.89 Å refinement. The starting R and $R_{\rm free}$ for the initial model were surprisingly high at 51.0 and 52.3%, respectively, for all data in the resolution range 35.0–0.89 Å and did not refine, presumably because of the use of the room-temperature model, which had a 4% larger unit-cell volume (Table 1). Refinement with stepwise addition of data from 3 to 0.89 Å resolution resulted in a structure with an R and $R_{\rm free}$ of 28.01 and 29.03, respectively. Automated water picking was carried out using *ARP*/*wARP* (Lamzin & Wilson, 1993) in combination with *REFMAC* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The waters thus picked were manually inspected in *O* with a map generated from the model before the *ARP*/*wARP* run. Waters were retained wherever they appeared both in the $(2F_o - F_c)$ and $(F_o - F_c)$ maps with reasonable geometry. This led to a model which converged to an R and $R_{\rm free}$ of 18.07 and 18.51%, respectively. Further *SHELXL*97 (CGLS) refinement with all ADPs being isotropic led to a model with an R and $R_{\rm free}$ of 15.11 and 16.41%, respectively, for all data in the resolution range 10.0–0.89 Å. When the ADPs were refined anisotropically, R and $R_{\rm free}$ rapidly fell to 12.07 and 13.87%, respectively. Multiple conformations, H atoms and water molecules were incorporated and ten cycles of CGLS gave a model with an R and $R_{\rm free}$ of 10.47 and 11.87%, respectively. Six ordered glycerol molecules from the cryoprotectant were modelled in the electron-density map.

The final restrained blocked least-squares refinement, with a round equivalent to two complete cycles of blocked full-matrix refinement with restraints, was performed with each block containing about 50 residues. The successive blocks overlapped by five residues. The final blocked least-squares refinement of CTUX with DELU, SIMU and ISOR restraints (Sheldrick & Schneider, 1997) gave a model with an R and $R_{\rm free}$ of 9.08 and 10.51%, respectively, for all data and 8.93 and 10.32%, respectively, for $F_o > 4\sigma(F_o)$. The GooF was 2.319 and the restrained GooF was 2.181. A parallel unrestrained blocked full-matrix refinement was carried out. This allowed rigorous estimation of the errors associated with each independent refined parameter from the inverse least-squares matrix. A similar protocol was used in the refinement of ferredoxin at 0.94 Å (Dauter, Wilson *et al.*, 1997). This unrestrained refinement converged to a model with an R and $R_{\rm free}$ of 9.00 and 10.61%, respectively, for all data and an R and $R_{\rm free}$ of 8.85 and 10.41%, respectively for $F_o > 4\sigma(F_o)$, $wR^2 = 0.2495$ and GooF = 2.283 (Table 2). The small 0.1 increase in $R_{\rm free}$ between the restrained and unrestrained structures may be attributed to the presence of a very minute proportion of disordered components which were not restrained. It took nearly 5 days to run unrestrained blocked full-matrix refinement on an SGI Octane system at our institution.

Before carrying out the final unrestrained blocked full-matrix refinement, a similar round of refinement was carried out. The coordinates obtained thus were analysed for 'CG—OD1/OD2, OD1/ND2' and 'CD—OE1/OE2, OE1/NE2' bond distances in Asp, Asn and Glu, Gln, respectively. The bond distances (written by BOND instruction) for each of these bonds were tabulated, extracting the values from the '.lst' file of the *SHELXL*97 (Sheldrick & Schneider, 1997) output (details not shown). A clear trend was seen for all these residues with low B factors (<15 Å^2). The 'CG—ND2, CD—NE2' and 'CD—OE1/OE2, CG—OD1/OD2' bond lengths were observed to tend to ideal values (Engh & Huber, 1991), clearly helping to overcome the ambiguity between Asn and Asp and between Gln and Glu in about five cases, in the absence of the actual correct sequence. In fact, the interchange of Asn OD1↔ND2 atoms could be also recognized. Some side chains that had sketchy or no density in room-temperature (1.11 Å) RTUX electron-density maps could be clearly seen in the cryotemperature CTUX map. For example, Gln302 as shown in Fig. 1 and residue 303, which could not at all be seen in the the RTUX electron-density map, were clearly seen in the CTUX electron-density map.

The sequence of the enzyme *T. aurantiacus* xylanase has been determined by us independently from the high-resolution (1.8 Å) structure as described in Natesh *et al.* (1999). The sequence was further improved by analysing the ultrahigh-resolution CTUX structure.

3. Results and discussions

3.1. Quality and features of the 1.11 Å RTUX and 0.89 Å CTUX final refined models

The *T. aurantiacus* xylanase structure has been solved at 1.11 and 0.89 Å resolution to small-molecule accuracy with $R < 10\%$ for both models (Table 2). The difference electron-

Table 3
Statistics from the final 1.11 Å RTUX and 0.89 Å CTUX structures.

Values in parentheses refer to the number of unique atoms, excluding alternate conformers.

Parameter	RTUX	CTUX
Protein atoms	2380 (2302)	2419 (2311)
Solvent atoms	267 (265)	426 (414)
Others	8 (8)	72 (67)
R.m.s. deviations		
Bond distance (Å)	0.016	0.036
Bond angle (°)	2.1	2.54
Dihedral (°)	24.3	23.86
Improper (°)	1.73	2.83
Luzzati coordinate error (e.s.d.) (Å)		
Working set	0.14	0.13
Test set	0.15	0.13
Ramachandran plot statistics (%)		
Residues in most favoured regions	92.1	91.0
Residues in additional allowed regions	7.5	8.6
Residues in generously allowed regions	0.4	0.4
Residues in disallowed regions	0.0	0.0
Average *B* factors (Å^2)		
Main chain	10.39	5.26
Side chain	14.98	6.77
Solvent molecules	30.65	19.2
Others	19.54	14.4
B factors (Å^2)		
Maximum, main chain	33.91	21.15
Minimum, main chain	5.68	2.82
Maximum, side chain	50.00	33.88
Minimum, side chain	2.33	1.72
Maximum, solvent (water)	49.58	38.96
Minimum, solvent (water)	8.56	4.38
Maximum, others	28.26	35.87
Minimum, others	15.97	4.61

density maps ($F_o - F_c$) contoured between 1.5 and 2.0σ for both the 1.11 and 0.89 Å resolution data revealed H atoms even before their contributions were introduced into the structure-factor calculation.

3.1.1. 1.11 Å model (RTUX) quality and features. The final refined protein model consisted of 2380 non-H protein atoms from 302 amino-acid residues and 267 water molecules (Table 3). R and $R_{\rm free}$ for all data are given in Table 2. The r.m.s. deviation values of electron density in the ($F_o - F_c$) maps is 0.05 e Å^{-3}. Other statistics such as the variation of the R factor as a function of resolution from the Luzzati plot (Luzzati, 1952), quality of the model from the Ramachandran (φ, ψ) plot (Ramachandran & Sasisekharan, 1968) as assessed by *PROCHECK* (Laskowski *et al.*, 1993) *etc.* for the final protein model are shown in Table 3.

3.1.2. 0.89 Å model (CTUX) quality and features. The final refined protein model consisted of 2419 non-H protein atoms from 303 amino-acid residues and 426 water molecules (Table 3). Residue 1 was clearly identified as Pca instead of the α-*N*-acetyl group deduced by earlier workers using biochemical methods (Roy, 1989; Srinivasa *et al.*, 1991). An additional residue 303 could be modelled as Gln from the cryoelectron-density map which was not at all visible in the 1.8 and 1.11 Å room-temperature electron-density maps. Many alternate conformations could be seen. An example for the modelling of dual conformations of amino-acid side chains is the clear electron density seen in CTUX for Trp275 and Arg276, which are located in the vicinity of the active site (Fig. 2). An example of H atoms seen in the 0.89 Å map before their contributions were included in the structure-factor calculation is shown in Fig. 3. To quantify their occurrence, a difference Fourier map was calculated at the end of refinement with all H atoms omitted (Esposito *et al.*, 2000) and analysed for peaks within 0.31 Å of calculated H-atom positions. Using this criteria, a total of 79% (1681 of 2125) of the H atoms were localized in one molecule of the asymmetric unit at an electron-density level >1.5σ (0.16 e Å^{-3}). The percentage

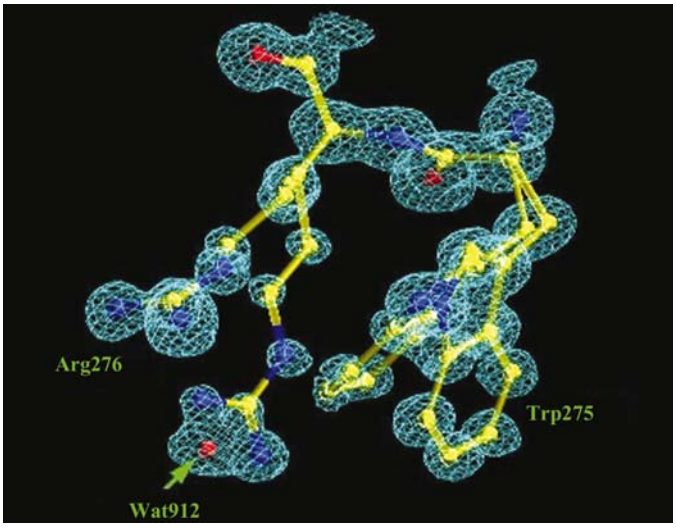

Figure 2
Concerted movement of Trp275 and Arg276 in the vicinity of the active site seen in the ultrahigh-resolution CTUX structure. Figure produced using *O*.

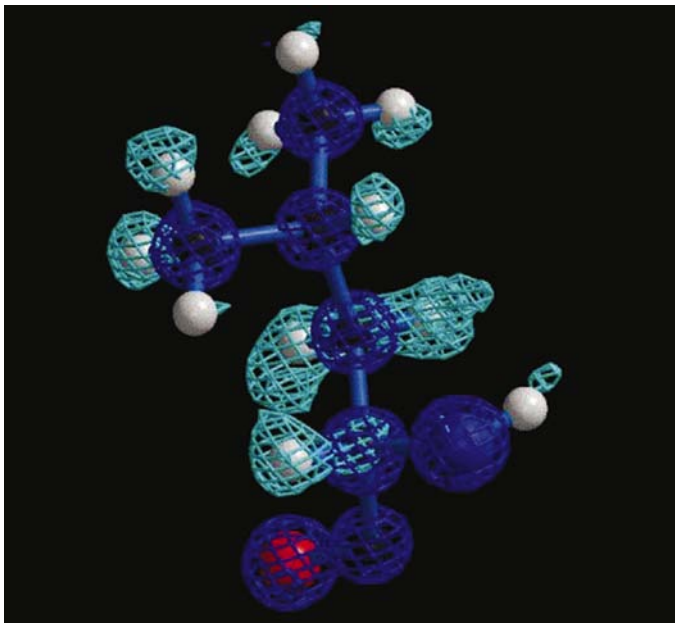

Figure 3
($F_o - F_c$) map (cyan) contoured at ∼1.5σ and ($2F_o - F_c$) map (blue) contoured at 4σ for the residue Leu283. Many H atoms can be clearly seen in the ($F_o - F_c$) map (cyan). The electron-density maps (0.89 Å) were calculated at a stage where no H atoms had yet been added to the model.

Table 4
Deviations from the target stereochemistry for backbone.

Parameter	E.s.d. (contributor)	Target†
N—C^{α} (Å)	0.009 (297)	0.02
C^{α}—C (Å)	0.010 (297)	0.02
C═O (Å)	0.009 (297)	0.02
N—C^{-} (Å)	0.009 (297)	0.02
C^{α}—C^{β} (Å)	0.011 (276)	0.02
C^{-}—N—C^{α} (°)	0.589 (297)	‡
N—C^{α}—C (°)	0.553 (297)	‡
C^{α}—C—N^{+} (°)	0.609 (297)	‡
Chiral volume (N—C^{β}—C) (Å^3)	0.044 (263)	0.1
ω (C^{α}—C—N^{+}—$C^{\alpha+}$) (°)	0.585 (296)	§

† Target value set for standard uncertainties during restrained refinement. ‡ Angles were restrained as 1,3-distances with uncertainty 0.04. The standard uncertainty against which *PROCHECK* compares is ∼2.0°. § The peptide bond was restrained to planarity (using combination of FLAT and CHIV with uncertainty target 0.1 Å^3). The standard uncertainty against which *PROCHECK* compares is 5.8°.

increases to 86% for main-chain H atoms alone.

R and $R_{\rm free}$ for all data are given in Table 2. The r.m.s. deviation value of electron density in the ($F_o - F_c$) map is 0.07 e Å^{-3}. All the statistics as described above for RTUX are shown for the final refined CTUX model in Table 3. The estimated standard uncertainties (e.s.d.s), as obtained from the final unrestrained blocked full-matrix refinement (from '.lst' file) for the backbone atoms, are given in Table 4.

3.2. Analysis of anisotropic atomic displacement parameters (ADPs) from the RTUX and CTUX structures

Anisotropy is defined as the ratio of the minimum and maximum eigenvalues of the 3 × 3 matrix of ADPs. The ADPs of RTUX and CTUX have been analysed using the web-based program *PARVATI* (Merritt, 1999).

3.2.1. Analysis of anisotropy for the RTUX structure. The distribution using 20 bins for the anisotropy range 1.0–0.0 shows a Gaussian distribution with a mean and a σ value of 0.47 and 0.12, respectively, for 2204 ANISOU records (protein atoms only). The atoms tend to be increasingly non-spherical as the surface of the protein is approached from the centre of mass. The thermal ellipsoid representation for the Glu232 residue is shown in Fig. 4. With all the ISOR, SIMU, DELU restraints on the U_{ij}s (anisotropic ADPs), there are no 'non-positive-definite' atomic descriptors (Sheldrick & Schneider, 1997) present in this structure.

3.2.2. Analysis of anisotropy for the CTUX structure. The ADPs were refined in the final least-squares refinement without any restraints (except for disordered residues, whose ADPs were refined isotropically). It is probable that analysis of these unrestrained refined anisotropic ADPs would be more accurate and closer to reality. The anisotropy ranges from 1.0 to 0.0 and is consistent with a Gaussian distribution with a mean and a σ value of 0.55 and 0.15, respectively, for the 2203 ANISOU records (protein atoms only). These values are comparable with those based on analysis of structures in the Protein Data Bank (PDB; Westbrook *et al.*, 2002), with a mean of 0.45 and a standard deviation of 0.15 (Merritt, 1999). The thermal ellipsoid representation for the Glu232 residue is shown in Fig. 4. As can be seen from Fig. 4, the cryo-temperature thermal ellipsoid clearly has less vibration in all directions (Fig. 4*a*) than the room-temperature thermal ellipsoid (Fig. 4*b*). The mean values of anisotropy of 0.47 and 0.55 for the RTUX and CTUX structures, respectively, suggest that the ADPs are more isotropic in CTUX. Although no restraints were applied in the final refined structure, there are no 'non-positive-definite' atomic descriptors (Sheldrick & Schneider, 1997) present in the CTUX structure.

3.3. Analysis of N—C^{α}—C (τ) angle in RTUX, CTUX and 15 ultrahigh-resolution structures: possible occurrence of context-dependent hidden strain in proteins

One of the factors contributing to the hidden strain in proteins is the distortion in the N—C^{α}—C angle (τ angle; Karplus, 1996; Esposito *et al.*, 2000). The covalent geometry N—C^{α}—C backbone angle was analysed for its dependence on the conformational state of the residues: whether they are in α-helical, β-strand or loop regions. The accuracy of the CTUX structure in particular, with mean errors (e.s.d.s) of 0.01 Å and 0.55° on backbone distances and N—C^{α}—C angles, respectively (Table 4), allows a more thorough investigation of this issue. The mean values for the α-helix, β-strands and coils computed for 15 ultrahigh-resolution structures (resolution better than 0.95 Å from the PDB), CTUX and RTUX structures are shown in Table 5. In all these analyses, outliers >5σ were excluded from the calculations. The mean values clearly suggest that in general residues in β-strand conformational regions show a lower N—C^{α}—C (τ) angle compared with residues in α-helical regions or coil (loop) regions. This is in agreement with the analyses (Ashida *et al.*, 1987) on small molecules from the Cambridge Structural Database and protein structures refined at 1.75 Å or better (Karplus, 1996). The secondary-structure assignment is based on the Kabsh & Sander (1983) method as implemented in *PROCHECK*. The larger τ angle for the α-helical region is also in accordance with the theoretical prediction of Ramakrishnan & Balasubramanian (1972), based on contact criteria, that the area in

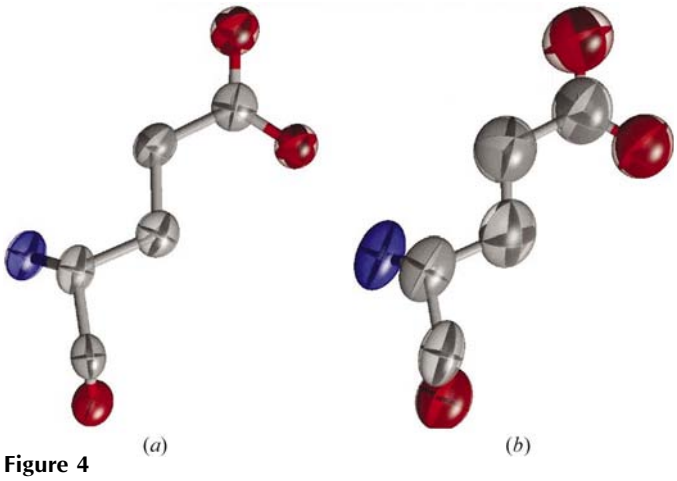

Figure 4
Representation of anisotropic thermal motion of residue Glu232, which is involved in salt-bridge formation (*a*) in CTUX, (*b*) in RTUX. Thermal ellipsoids are drawn with 50% probability using *RASTEP* (Merritt & Murphy, 1994).

Table 5
Analysis of the N—C^{α}—C angle for 16 ultrahigh-resolution structures including CTUX.

The angles in RTUX are also provided for comparison. Values indicated in parentheses are number of residues present in that particular secondary structure and used for the averaging.

PDB code†	Helix	Strand	Coil
1b0y	111.256 (28)	110.441 (12)	111.213 (43)
1bxo	111.675 (59)	109.781 (170)	111.764 (85)
1dy5	111.762 (32)	109.483 (48)	110.542 (39)
1ejg	111.400 (22)	107.978 (5)	112.051 (16)
1g66	111.800 (96)	109.010 (30)	111.540 (81)
1g6x	111.970 (16)	108.640 (14)	111.600 (28)
1gci	111.546 (100)	109.770 (60)	111.580 (106)
1nls	112.203 (18)	109.367 (131)	110.996 (82)
1rb9	112.750 (9)	109.808 (12)	112.034 (30)
1kwf	111.570 (196)	108.710 (8)	111.560 (159)
2fdn	112.077 (11)	109.458 (18)	113.109 (24)
2pvb	111.453 (74)	(0)	111.656 (32)
3lzt	111.541 (64)	111.211 (13)	112.449 (49)
3pyp	111.835 (49)	109.057 (40)	111.305 (34)
7a3h	111.578 (134)	109.449 (63)	111.494 (100)
CTUX	111.455 (157)	109.522 (58)	111.613 (80)
RTUX	111.668 (159)	110.142 (58)	112.001 (80)

† 1b0y, electron-transfer protein HIPIP, 0.93 Å; 1bxo, penicillopepsin, 0.95 Å; 1dy5, ribonuclease A, 0.87 Å; 1ejg, crambin, 0.54 Å; 1g66, esterase, 0.90 Å; 1g6x, BPTI, 0.86 Å; 1gci, subtilisin, 0.78 Å; 1nls, concanavalin A, 0.94 Å; 1rb9, rubredoxin, 0.92 Å; 1kwf, glycosidase, 0.94 Å; 2fdn, ferredoxin, 0.94 Å; 2pvb, parvalbumin, 0.91 Å; 3lzt, lysozyme, 0.92 Å; 3pyp, photoactive yellow protein, 0.85 Å; 7a3h, endoglucanase, 0.95 Å.

the φ, ψ map available for right-handed α-helical conformation increases with increases in the N—C^{α}—C angle.

Since Gly occupies the maximum area in the Ramachandran map (Ramachandran & Sasisekharan, 1968) and is the least bulky of the 20 coded amino acids, the dependence of the τ angle of Gly on φ, ψ angles was analysed. Although the dependence of the τ angle on φ was not so evident, the bond angle in general assumes its maximum value when ψ is close to zero, which corresponds to the bridge region of the Ramachandran map (Fig. 5). Hence, this context-dependent distortion of the τ angle may be a useful criterion for the validation of protein structures.

3.4. An example of plasticity of ion pairs in proteins: Arg124–Glu232 salt bridge in the RTUX and CTUX structures

An interesting feature has been observed when comparing the RTUX and CTUX structures regarding the Arg124–Glu232 salt bridge. Occupancy refinement of Arg124 in the room-temperature RTUX structure shows that the 'bidentate' conformation (*A* conformation) of Arg124 is 60% occupied and a monodentate conformation (*B* conformation) is weakly occupied (Fig. 6*a*). The salt bridge in conformation *B* is mediated through a water molecule (Wat710). The occupancies of Wat710 and the *B* conformer of Arg124 were coupled during *SHELXL* refinement. In the cryotemperature CTUX structure, the bidentate *A* conformer of the Arg124–Glu232 salt bridge was very weakly observed. No clear electron density was observed (Fig. 6*b*) and it refined to 15% occupancy, with conformer *B* having 85% occupancy. In this case, a water molecule (Wat563) mediates the salt bridge. To confirm the authenticity of this observation in CTUX, the occupancy of Wat563 was not coupled with the occupancy of the *B* conformer of Arg124 in *SHELXL* and it refined to 87% occupancy, close to the 85% occupancy of Arg124 conformer *B* (Fig. 6*b*). In cryotemperature CTUX the bidentate species is weakly occupied and a water molecule compensates for the loss to some extent, but in room-temperature RTUX the bidentate species pushes out the water molecule. The bidentate species of the salt bridge survives better at room temperature than the water-mediated salt bridge seen at cryotemperature.

Another example of invasion by water of a salt bridge at lower temperature is seen in the plasticity of the interaction between Glu7 and Lys11. In the salt bridge Glu7–Lys11, Lys Nz is at almost equal distances from Glu OD1 and OD2 (2.93 and 2.86 Å, respectively) in the room-temperature RTUX structure. However, in the cryotemperature CTUX structure, Lys11 Nz forms a hydrogen bond with Wat577 (not seen in RTUX) and has moved so as to be at a shorter distance (2.62 Å) from Glu7 OD1. The alternate interactions seen for the two rearranged salt bridges are associated with different rotamers of amino-acid side chains Arg124 and Lys11.

Recently, Kumar & Nussinov (2001), based on the NMR structures of proteins, investigated the fluctuations of ion pairs and their stability in proteins. Their major conclusion was that salt bridges observed in the crystal structure may break and new salt bridges may be formed. The disruption of protein-surface salt bridges has been recently found to be important in protein–DNA interactions (Saecker & Record, 2002). Crystallographic evidence in the present analysis for the rearrangement of the salt bridge between the RTUX and CTUX structures is a further example of plasticity. Moreover, a network has been made with a water molecule in CTUX.

Water molecules can compete for charged atoms which have the potential to form salt bridges. It appears that the competition of water molecules for interacting partners

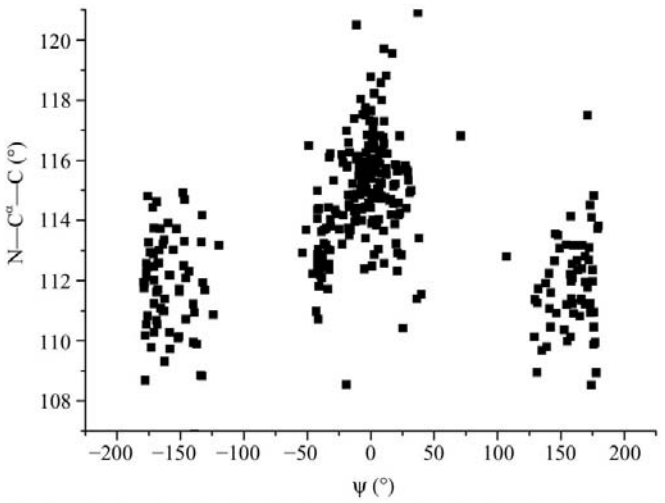

Figure 5
The conformational dependence of N—C^{α}—C (τ) angles in 16 ultrahigh-resolution structures including CTUX on ψ angle for Gly residues. τ assumes larger values when $\psi \simeq 0^{\circ}$, corresponding to the bridge region of the Ramachandran map.

becomes more pronounced at lower temperature than at higher temperatures, presumably because of entropy considerations. A greater number of water molecules can be located in electron-density maps at lower temperatures than at higher temperatures (267 in RTUX and 426 in CTUX), suggesting that water molecules are more localized at lower temperature.

Since the dielectric constant of water decreases with increasing temperature, salt-bridge formation may be energetically more favourable owing to its electrostatic component (Pace, 2000) at high temperatures than at lower temperatures.

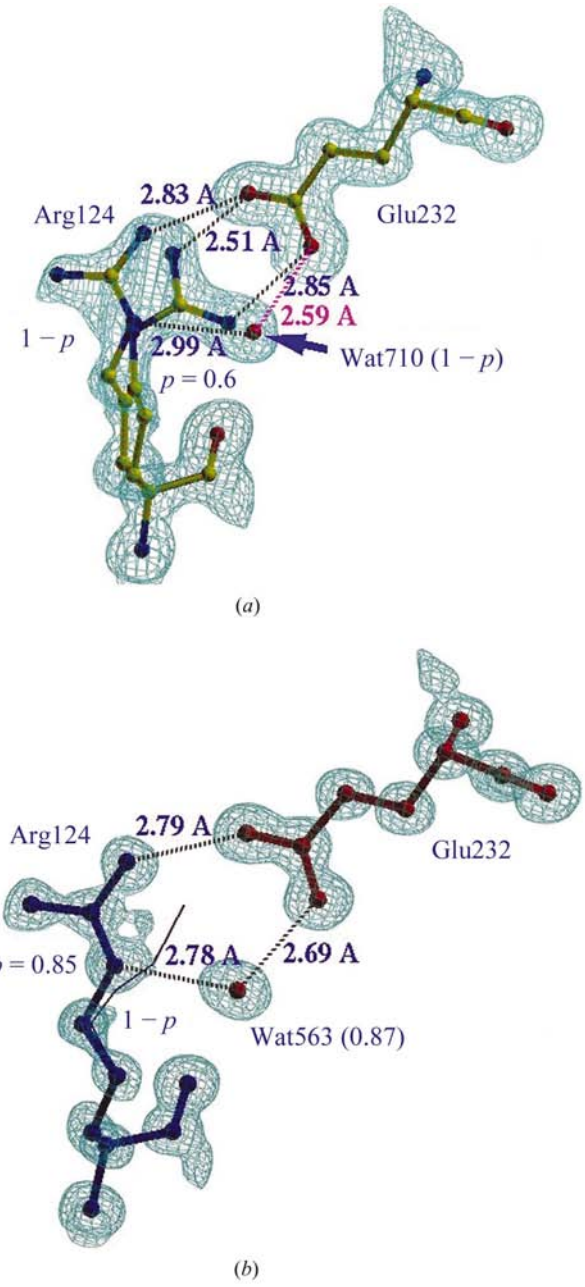

Figure 6
Temperature-dependent plasticity of the salt bridge Arg124–Glu232. (*a*) RTUX structure, $2F_o - F_c$ map, 1.6σ contour; (*b*) CTUX structure, $2Fo - F_c$ map, 1.5σ contour. Note that the salt bridge, which is to a large extent bidendate in the RTUX structure, becomes a water-mediated interaction in the CTUX structure. The solid black line in panel (*b*) represents the lower occupancy (0.15) rotamer of Arg124.

Our study points to the possible dependence of salt-bridge formation upon temperature and the role of water molecules in disrupting salt bridges. Accurate crystal structure analysis of a thermophilic protein at higher temperature may be useful in this context, since salt bridges are implicated as one of the determinants of the thermostability of proteins (Yip *et al.*, 1995). It appears that it will be meaningful to compare salt bridges in different proteins for crystal structures determined at the same temperature rather than at different temperatures. Also, our analysis suggests that caution has to be exercised before extrapolating the results from cryocrystallography to the situation at room temperature or even higher temperatures (Sandalova *et al.*, 1999).

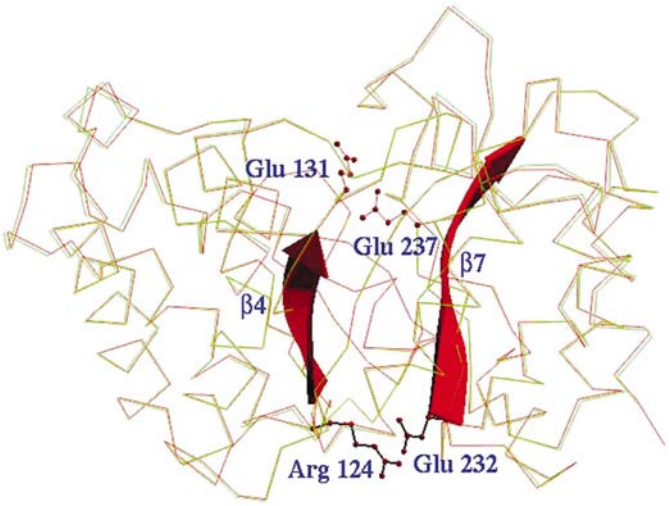

Figure 7
C^{α} trace of the RTUX (green) and CTUX (red) structures superimposed using *O*. The β-strands β4 and β7 are also shown for the CTUX structure along with active-site residues Glu131 and Glu237 and the salt bridge Arg124–Glu232. Figure produced using *O*.

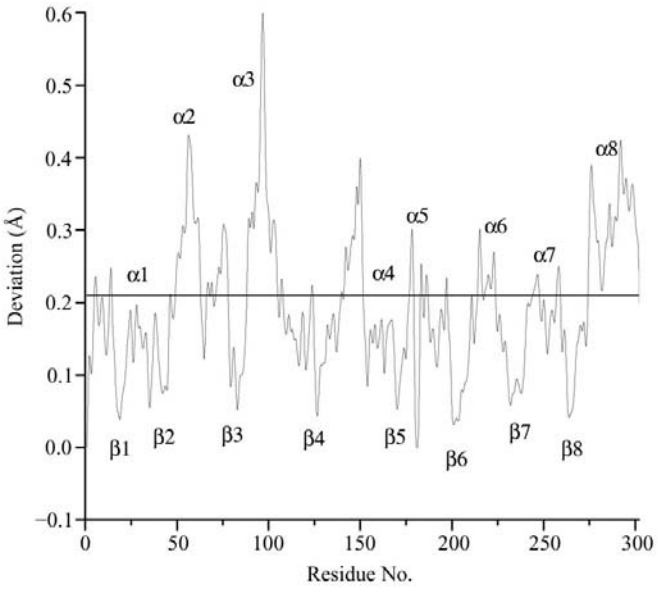

Figure 8
Plot of deviation between corresponding C^{α} atoms (between RTUX and CTUX) plotted against residue number.

4. Temperature-dependent deformations of *T. aurantiacus* xylanase with $(\beta/\alpha)_8$ TIM-barrel fold

4.1. Comparison of unit-cell parameters, unit-cell volumes and overall structure

RTUX and CTUX structures were solved using crystals belonging to space group $P2_1$, with unit-cell parameters as shown in Table 1. The unit-cell volume (Table 1) has shrunk by about 4% in CTUX compared with RTUX, a significant change. Fig. 7 compares the molecules in RTUX and CTUX. From Fig. 7 it can be seen that the CTUX structure (red) bends towards the central axis of the barrel compared with the RTUX structure (green). Also shown is a cartoon representation of the strands β4 and β7, towards the C-terminus of which the active-site glutamates Glu131 and Glu237 reside and towards the N-terminus of which the salt-bridge-forming residues Arg124 and Glu232 reside.

4.2. Comparison of radii of gyration

The radius of gyration provides information about the molecular shape and dimensions. We have calculated and compared the radius of gyration of the protein molecule in the RTUX (16.70, 15.31 and 12.61 Å) and CTUX (16.60, 15.24 and 12.53 Å) structures. It may be seen that the percentage increase of the length of the principal axis from CTUX to RTUX is larger for the major and minor axes (0.65 and 0.60) than for the intermediate axis (0.46), indicating that the molecule undergoes anisotropic expansion with the increase in temperature. Interestingly, the intermediate axis about which the percentage change is the least coincides with the barrel axis of the $(\beta/\alpha)_8$ TIM barrel (Natesh *et al.*, 1999). It has been suggested that in enzymes the active-site residues are close to one of the principal axes, a feature found also to be valid in the case of the study enzyme (Foote & Raman, 2000).

4.3. Least-squares superposition of the corresponding C^α atoms between RTUX and CTUX

Structural comparison of the RTUX and CTUX structures has been carried out. The corresponding C^α atoms of both structures were superposed using the program *ALIGN* (Cohen, 1997). The r.m.s. deviation is found to be 0.21 Å. It indicates that the deviation between the cryo- and room-temperature structure is small, even though the temperature difference between them is nearly 193 K.

Compared with α-helices (0.21 Å), the β-strands (0.08 Å), which are in the interior of the protein, have a lower mean deviation (Fig. 8). Similarly, it is found that $\alpha\beta$-loops have less positional flexibility (0.15 Å) than $\beta\alpha$-loops (0.25 Å). The positional flexibility in $\beta\alpha$-loops, which are longer than $\alpha\beta$-loops, might facilitate substrate binding, enzymatic activity and product release. Of the 26 alternate-conformation residues present in the CTUX structure, 12 are in $\beta\alpha$-loops. Interestingly, helix 8 (α8) (Natesh *et al.*, 1999) at the C-terminus of the protein has a high mean C^α deviation (0.35 Å) compared with other secondary-structural elements present in the enzyme (Fig. 8).

4.4. Analysis of temperature factors

The extent of motion of the atoms about their mean position in macromolecules is characterized by the *B* factor (temperature factor); *B*-factor analysis might reveal insights into the stability, dynamics and the most flexible part of the proteins (Halle, 2002). To compare the structure at different temperatures, we have calculated the *B* factor (equivalent isotropic temperature factor) ratio of the corresponding main-chain atoms (N, C^α, C and O) of the cryo- and room-temperature structures. The overall protein mean ratio is found to be 2.0. The mean ratio for the β-strands (1.8) is less than the overall mean ratio of the protein, which may be a consequence of the fact that the β-strands are highly packed in the interior of the protein. The mean *B*-factor ratio is the same for the $\alpha\beta$-loops and $\beta\alpha$-loops (2.1 and 2.0, respectively), suggesting that the vibrational flexibility of the loop regions is smaller than the positional flexibility. However, the mean ratio for the helix α8 (2.3) (Natesh *et al.*, 1999) is higher than the overall protein mean ratio.

Results of molecular-dynamic studies have suggested that protein unfolding might be initiated at sites that are prone to large thermal fluctuations (Daggett & Levitt, 1992; Lazaridis *et al.*, 1997). It may be speculated from the above analyses that helix 8 (α8) (Natesh *et al.*, 1999) at the C-terminus of the study enzyme, the last secondary-structural element in the sequence, may be the one which might contribute to the initial stages of the unfolding of the $(\beta/\alpha)_8$ TIM-barrel structure.

5. Water structures in RTUX and CTUX

Water is a natural medium for protein molecules and has a great influence on their dynamics, stability and function (Rupley & Careri, 1991; Levitt & Park, 1993). In particular, water rearrangement plays an important role in carbohydrate–protein complexation (Christopher *et al.*, 2001). The role of water in stabilizing protein tertiary structure has been reported (Sadasivan *et al.*, 1998). Theoretical studies have modeled water molecules around proteins as probability densities allowing a given water molecule to occupy multiple sites around a protein molecule and hence predicting the possible hydration patterns (Phillips & Pettitt, 1995). Because of the high resolution, accurate crystal structure data and the presence of more localized waters at lower temperature, we could locate more waters in the CTUX structure (426), one and half times the number located in the RTUX structure (267). The CTUX structure has 12 dual-conformation waters. In both RTUX and CTUX, nearly half of the water molecules are found in pairs or networked with other waters (cutoff distance 3.5 Å). We have compared the hydration shell (Esposito *et al.*, 2000) around the protein and examined the role of buried waters in the RTUX and CTUX structures.

5.1. Waters common to RTUX and CTUX

The protein C^α superposition matrix (§4.3) was applied to RTUX water coordinates for comparison with the water structure in CTUX. Based on a distance cutoff of 1.00 Å, the

Table 6
Waters common to RTUX and CTUX and hydrogen-bond interaction statistics.

(*a*) Waters common to RTUX and CTUX.

Water distances (Å) between RTUX and CTUX	No. of waters	Coordination No.† RTUX	Coordination No.† CTUX
0.00–0.25	61	3.05	3.61
0.25–0.50	92	2.37	3.00
0.50–0.75	43	1.93	2.72
0.75–1.00	17	2.06	2.41

(*b*) Hydrogen-bond interaction statistics. On the whole, there are more protein–water interactions present in CTUX as it has more waters than RTUX.

	No. of hydrogen bonds (protein alone)			No. of hydrogen bonds (protein–heteroatoms)		Total	
	MM	MS/SM	SS	MH/HM	HS/SH	P	PH
RTUX	190	26/39	43	49/144	119/105	298	715
CTUX	190	27/42	41	54/169	144/140	300	807

† Average number of hydrogen bonds.

distribution of common waters in the defined deviation range 0.00–1.00 Å is given in Table 6(*a*). Of the 267 waters in RTUX, 83% are present in CTUX within this deviation range, implying that upon temperature changes, even though many of the waters are relatively in the same position, the water structure around the protein shows some flexibility.

5.2. Comparison of hydrogen-bond interactions in RTUX and CTUX

Hydrogen-bonding interactions stabilize the folded structure of macromolecules and it is of interest to study the effect of temperature on them. Therefore, we have compared the intramolecular and intermolecular hydrogen bonds in RTUX and CTUX using the program *HBPLUS* (Mcdonald & Thornton, 1994) to calculate hydrogen bonds (Table 6*b*). As may be seen from Table 6(*b*), the intramolecular main-chain-to-main-chain (MM) hydrogen bonds do not change with temperature, presumably because they stabilize the secondary-structural elements in proteins. There are marginal changes in the main-chain-to-side-chain (MS) and side-chain-to-side-chain (SS) interactions. However, the largest differences are in the substantially larger number of hydrogen-bonding interactions between protein side-chain atoms and heteroatoms, most of which are water molecules. This may be seen in the context of the increase in the number of waters in the first hydration shell in CTUX compared with RTUX (Table 7).

5.3. Comparison of hydration shells at different temperatures

Understanding the general features of hydration around the protein structures is important for drug-design studies, the principles of protein folding and thermodynamics and dynamics of protein systems (Jiang & Brünger, 1994; Shih *et al.*, 1995). Table 7 shows the distribution of waters in various

Table 7
Hydration shells in RTUX and CTUX.

Hydration shell	RTUX No. of waters	RTUX Coordination No.†	CTUX No. of waters	CTUX Coordination No.†
Zeroth‡	8	3.12	11	3.45
First	220	2.50	280	2.87
Second	32	1.56	96	2.25
Third	2	1.00	15	1.80
Fourth	0	0.00	2	1.00
Isolated	3	0.00	10	0.00
Total	265	2.39	414	2.69

† Coordination No. refers to the average number of hydrogen bonds. ‡ Isolated buried waters.

hydration shells as seen in the room and cryotemperature structures. As may be expected, more water molecules are present in each hydration shell in the CTUX. The percentages of water molecules present in the first hydration shell of RTUX and CTUX are 83 and 68%, respectively. In comparison, 82% of 199 waters located in the room-temperature structure 1k6a (Teixeira *et al.*, 2001; see §1) belong to the first hydration shell. It clearly indicates that the extent of hydration around the protein molecule is more limited to the first hydration shell at higher temperatures. Waters in the higher order hydration shells can be localized at lower temperature, as seen in CTUX. The coordination number (Table 7) decreases from the zeroth hydration shell (isolated buried waters) to higher order hydration shells for a given structure. Also, within a hydration shell the coordination number increases from higher temperature to lower temperature. In general, the water molecules are better ordered and interact better at lower temperatures.

5.4. Analysis of buried waters

Buried waters contributing to protein thermal stability are known and their roles in a number of proteins have been discussed (Shih *et al.*, 1995). We have used the program *PRO_ACT* (Williams *et al.*, 1994) to identify 20 buried waters which are present in both the RTUX and CTUX structures (the CTUX structure has 11 more buried waters). Interestingly, they are located in the 0.00–0.25 Å distance-deviation range (Table 6*a*) and the hydrogen-bonding pattern is maintained. This small deviation in their location might suggest that they are important for protein stability. The buried waters have been pooled into two groups: isolated buried waters (8) and buried waters networked with other waters (12). The water molecules belonging to the latter group are mostly found as dimers (four pairs, with coordination number 3.37); both are buried and hydrogen bonded with each other, which might mean that a dimer of buried waters may be an important interaction motif in the protein. Supplementary Table 1[1] shows the interaction of secondary-structural elements and the number of waters involved and indicates that whether the

[1] Supplementary data have been deposited in the IUCr electronic archive (Reference: ad0183). Details for accessing these data are given at the back of the journal.

buried water molecules are networked or isolated, they mostly interact with residues belonging to the $\beta\alpha/\alpha\beta$ loops (16 out of 20), in particular with $\beta\alpha$ loops (13 out of 20). Fig. 9(*a*) shows hydrogen-bond interactions of one of the isolated buried water [Wat525 (W3); *B* factor is 4.98 Å^2] and in all the four interactions only main-chain atoms of protein molecule are involved (Supplementary Table 2*a*).

Isolated buried waters have more interactions with main-chain atoms as they are well inside the protein. Of a total of 25 hydrogen bonds, 15 are formed with atoms from the secondary-structural elements such as β-strands/α-helices and hence stabilize the structure by cross-linking them. However, the buried waters which are part of water networks have more interactions with side chain than with main chains (Supplementary Table 2*b*).

Of the 20 buried waters, six waters interacting with $\beta\alpha$ loops and β-strands (Supplementary Table 1) are present in the active-site cleft region and may be important for substrate binding. They are part of water networks and two make hydrogen bonds with the active-site residues, presumably mimicking the O atoms in the xylooligomers, the substrate of the enzyme.

It is found that the average number of hydrogen bonds (coordination number) is 3.50 for all buried waters (Supplementary Table 2). The average *B* factors for the isolated buried waters in RTUX and CTUX are 10.96 and 5.29 Å^2, respectively.

Recently, Pace *et al.* (2001) have shown that hydrogen bonds involving buried polar amino acids increase the protein stability. Of 58 water–protein hydrogen-bond interactions, 26 interactions are between the buried waters and atoms of the buried polar amino acids of the protein molecule. In particular, amino acids having hydroxyl (—OH) groups (Ser, Tyr and Thr) are involved in 19 interactions of the 26 (Supplementary Table 2). This might be important since the study protein is a thermostable enzyme. Our analysis shows that many interactions between buried waters and buried polar atoms are present and suggests that these interactions are likely to contribute to the structural stability of proteins.

A special mention may be made of the water dimer formed by Wat533 (W1) and Wat511 (W2), which is present inside the β-barrel and bridges the β-strands β2, β3, β4 and β7. The residues of the protein molecule involved in hydrogen bonding to water W1, including the active-site residue nucleophile Glu237, are fully conserved and for the water W2 the Thr236 is fully conserved and Thr44 and Arg81 are partially conserved across the F/10 family xylanases whose three-dimensional structures have been solved (Fig. 9*b*) (Supplementary Table 2). The *B* factor for this water dimer at the active site is low (6.84 and 4.58 Å^2 for W1 and W2, respectively, in CTUX). This water dimer is also found in the 1k6a (Teixeira *et al.*, 2001) and *Penicillium simplicissium* xylanase (a thermostable enzyme; PDB code 1bg4) structures and these dimers have same hydrogen bonds to the protein atoms as in the study enzyme. The conservation of the location and hydrogen-bonding scheme of the water dimer suggests that it may play a structural and functional role in F/10 family xylanases.

From Supplementary Table 2, it may be seen that buried water molecules have more interactions with main-chain

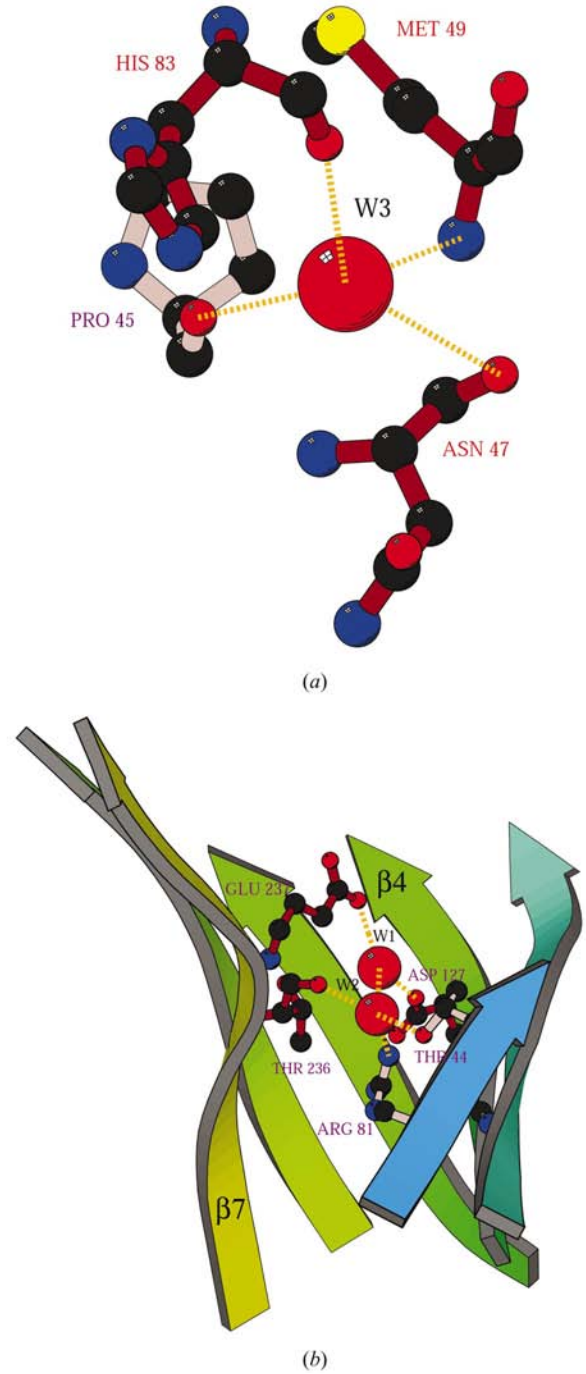

Figure 9
Examples of buried water interactions with protein atoms. (*a*) Interactions of one of the isolated buried waters, Wat525 (W3). Conserved residues are labelled in red. Pro45 and His83 are residues belonging to β2 and β3, respectively. Asn47 and Met49 are part of the $\beta\alpha$-loop connecting β2 and α2. Interestingly, all the water hydrogen bonds are with main-chain atoms. (*b*) A water dimer formed by Wat511 (W1) and Wat533 (W2) and its interactions. Conserved residues are labelled in red. Interactions involving water molecules appear to contribute to the stability of residues in the active-site region. β-Strands 1 and 8 are not shown.

carbonyl O atoms which carry a partial negative charge. The carbonyl O atoms in many cases are buried in the protein. Sometimes, the same water molecule bridges two carbonyl atoms from adjacent/nearby residues through hydrogen bonds, perhaps compensating for the unfavourable electrostatic interactions between carbonyl O atoms. In this context, it is of interest to recall the earlier suggestions that water molecules may facilitate binding of proteins to other molecules such as DNA by screening unfavourable electrostatics (Reddy *et al.*, 2001).

6. Conclusions

The structures of *T. aurantiacus* xylanase at 0.89 Å (cryo-temperature, 100 K) and 1.11 Å (room temperature, 293 K) have been solved to small-molecule accuracy. To our knowledge, the 0.89 Å structure is one of the highest resolution protein structures of this size (303 amino acids, ∼32.9 kDa), with a single monomer chain in the asymmetric unit, to have been solved to date.

Experimental crystallographic evidence has been obtained for the plasticity of salt bridges in proteins with water molecules mediating some of the alternate arrangements: the salt bridge between Arg124–Glu232, which is to a large extent bidentate in RTUX, is water-mediated in CTUX. This indicates that the contribution of electrostatic interactions to the stability of proteins through ion pairs may be expected to be more pronounced at higher temperatures. Based on comparative analysis of the RTUX and CTUX structures at atomic and ultrahigh resolution, we have for the first time pointed out the possible dependence of salt bridges upon temperature and the role played by water molecules at different temperatures.

The analysis of the $N-C^{\alpha}-C$ (τ) angle in protein structures determined to ultrahigh resolution, according to which residues in α-helical and coil regions have larger τ angles than those in β-strand conformational regions, suggests the occurrence of hidden strain in proteins. This feature may be useful for the validation of protein crystal structures.

It is known that 10% or more of all enzymes whose structures are known assume the TIM-barrel fold, suggesting that the TIM-barrel fold may be thought of as being catalytically efficient from the point of view of both architecture and dynamics. The functional region of the TIM-barrel fold enzymes varies; the residues in $\beta\alpha$-loops at the C-terminus of the β-strands are particularly variable. The flexibility of the $\beta\alpha$-loops is quite important for the activity of the enzyme. From our analysis of crystal structure at two different temperature (least-squares superposition and *B*-factor ratio analysis), we have noticed that the $\alpha\beta$-loops are relatively less flexible than the $\beta\alpha$-loops. The β-strands are least affected structurally by the increase in temperature. Thus, the TIM-barrel fold in the study enzyme, though having a single domain, may be dissected into parts based on the relative flexibility and described as having a rigid core constituted by the β-barrel and a less rigid exterior formed by the surrounding α-helices composed of stretches which are not contiguous in sequence, forming discontinuous domains. This may be a general feature of all proteins belonging to the TIM-barrel fold since β-strands are located in the interior of the $(\beta/\alpha)_8$-barrel structure.

Buried waters interacting with $\beta\alpha$-loops might play a crucial role in the structure and function of the enzyme since most of them make hydrogen bonds with residues conserved across F/10 family xylanases whose crystal structures are known. Two buried waters present inside the barrel forming hydrogen-bond interactions with residues in strands $\beta2$, $\beta3$, $\beta4$ and $\beta7$ contribute to structural stability, presumably by compensating for unfavourable electrostatic interactions.

Our analysis suggests that water molecules in proteins may be classified into different categories such as being isolated and networked buried waters. Different categories of waters have different patterns of hydrogen-bonding interactions with the protein atoms, playing intriguing roles in the structure, function and stability of protein molecules.

Features determined experimentally and discussed in this report such as temperature-dependent deformations, details of water structure at two different temperatures and salt-bridge plasticity are of general interest and can further our understanding of the ubiquitous TIM-barrel fold assumed by the *T. aurantiacus* xylanase and also serve as useful inputs to theoretical and modelling studies.

The authors thank Dr K. R. Rajashankar for collecting synchrotron X-ray data at the National Synchrotron Light Source, BNL, NY, USA and Dr Z. Dauter for valuable discussions at the stage of refinement of the structures. We thank DBT and DAE–BRNS for financial support. Facilities at the Supercomputer Education and Research Centre, the Interactive Graphics Based Molecular Modeling Facility and The Distributed Information Centre (both funded by DBT) were used in this work. MAV thanks Indian National Science Academy for financial support and KM thanks CSIR (India) for a fellowship. We are extremely grateful to all the referees for a very thorough reading of the manuscript and valuable suggestions.

References

Anderson, D. E., Becktel, W. J. & Dahlquist, F. W. (1990). *Biochemistry*, **29**, 2403–2408.
Ashida, T., Tsunogae, Y., Tanaka, I. & Yamane, T. (1987). *Acta Cryst.* B**43**, 212–218.
Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. (1975). *Nature (London)*, **255**, 609–614.
Bott, R. & Boelens, R. (1999). *Curr. Opin. Biotechnol.* **10**, 391–397.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Christopher, C., Woods, R. J., Gluska, J., Cooper, A., Nutley, M. A. & Boons, G. J. (2001). *J. Am. Chem. Soc.* **123**, 12238–12247.
Cohen, G. E. (1997). *J. Appl. Cryst.* **30**, 1160–1161.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Daggett, V. & Levitt, M. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 5142–5146.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.
Dauter, Z., Wilson, K. S., Sieker, L. C., Meyer, J. & Moulis, J. M. (1997). *Biochemistry*, **36**, 16065–16073.
Davies, G. & Henrissat, B. (1995). *Structure*, **3**, 853–859.
Elcock, A. H. (1998). *J. Mol. Biol.* **284**, 489–502.
Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.
Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.
Esposito, L., Vitagliano, L., Sica, F., Zagari, G. S. A. & Mazzarella, L. (2000). *J. Mol. Biol.* **297**, 713–732.
Farber, G. K. (1993). *Curr. Opin. Struct. Biol.* **3**, 409–412.
Foote, J. & Raman, A. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 978–983.
Halle, B. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 1274–1279.
Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
Jiang, J. S. & Brünger, A. T. (1994). *J. Mol. Biol.* **243**, 100–115.
Jones, T. A., Zou, J. Y., Cowman, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsh, W. & Sander, C. (1983). *Biopolymers*, **22**, 2577–2637.
Karplus, P. A. (1996). *Protein Sci.* **5**, 1406–1420.
Khandke, K. M. (1986). PhD Thesis. Indian Institute of Science, Bangalore, India.
Khandke, K. M., Vithaythil, P. J. & Murthy, S. K. (1989). *Arch. Biochem. Biophys.* **274**, 491–500.
Kleywegt, G. J. & Jones, T. A. (1995). *Structure*, **3**, 535–540.
Kumar, S. & Nussinov, R. (2001). *Proteins Struct. Funct. Genet.* **43**, 433–454.
Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–149.
Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lazaridis, T., Lee, I. & Karplus, M. (1997). *Protein Sci.* **6**, 2589–2605.
Levitt, M. & Park, B. H. (1993). *Structure*, **1**, 223–226.
Lo Leggio, L., Kalogiannis, S., Bhat, M. K. & Pickersgill, R. W. (1999). *Proteins Struct. Funct. Genet.* **36**, 295–306.
Lo Leggio, L., Kalogiannis, S., Eckert, K., Teixeira, S. C. M., Bhat, M. K., Andrei, C., Pickersgill, R. W. & Larsen, S. (2001). *FEBS Lett.* **509**, 303–308.
Longhi, S., Czjzek, M. & Cambillau, C. (1998). *Curr. Opin. Struct. Biol.* **8**, 730–737.
Luzzati, P. V. (1952). *Acta Cryst.* **5**, 802–810.
Mcdonald, I. K. & Thornton, J. M. (1994). *J. Mol. Biol.* **238**, 777–793.
Merritt, E. A. (1999). *Acta Cryst.* D**55**, 1109–1117.
Merritt, E. A. & Murphy, M. E. P. (1994). *Acta Cryst.* D**50**, 869–873.
Natesh, R., Bhanumoorthy, P., Vithaythil, P. J., Sekar, K., Ramakumar, S. & Viswamitra, M. A. (1999). *J. Mol. Biol.* **288**, 999–1012.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pace, C. N. (2000). *Nature Struct. Biol.* **7**, 345–346.
Pace, C. N., Horn, G., Hebert, E. J., Bechert, J., Shaw, K., Urbanikova, L., Scholtz, J. M. & Sevcik, J. (2001). *J. Mol. Biol.* **312**, 393–404.
Phillips, G. N. & Pettitt, B. M. (1995). *Protein Sci.* **4**, 149–158.
Pujadas, G. & Palau, J. (1999). *Biologia (Bratislava)*, **54**, 231–254.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **23**, 283–438.
Ramakrishnan, C. & Balasubramanian, R. (1972). *Int. J. Pept. Protein Res.* **4**, 79–90.
Reddy, C. K., Das, A. & Jayaram, B. (2001). *J. Mol. Biol.* **314**, 619–632.
Roy, R. P. (1989). PhD Thesis. Indian Institute of Science, Bangalore, India.
Rupley, J. A. & Careri, G. (1991). *Adv. Protein Chem.* **41**, 37–172.
Sadasivan, C., Nagendra, H. G. & Vijayan, M. (1998). *Acta Cryst.* D**54**, 1343–1352.
Saecker, R. M. & Record, M. T. Jr (2002). *Curr. Opin. Struct. Biol.* **12**, 311–319.
Sandalova, T., Schneider, G., Käck, H. & Lindqvist, Y. (1999). *Acta Cryst.* D**55**, 610–624.
Sheldrick, G. M. (1997). *The SHELX*97 *Manual*. University of Göttingen, Germany.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.
Shih, P., Holland, D. R. & Kirsch, J. F. (1995). *Protein Sci.* **4**, 2050–2062.
Srinivasa, B. R., Swaminathan, K. R., Ganapathy, C., Roy, R. P., Murthy, S. K. & Vithayathil, P. J. (1991). *Protein Seq. Data Anal.* **4**, 15–20.
Teixeira, S. C. M., Lo Leggio, L., Pickersgill, R. W. & Cardin, C. (2001). *Acta Cryst.* D**57**, 385–392.
Viswamitra, M. A., Bhanumoorthy, P., Ramakumar, S., Manjula, M. V., Vithayathil, P. J., Murthy, S. K. & Naren, A. P. (1993). *J. Mol. Biol.* **232**, 987–988.
Westbrook, J., Feng, Z., Jain, S., Bhat, T. N., Thanki, N., Ravichandran, V., Gilliland, G. L., Bluhm, W. F., Weissig, H., Greer, D. S., Bourne, P. E. & Berman, H. M. (2002). *Nucleic Acids Res.* **30**, 245–248.
Williams, M. A., Goodfellow, J. M. & Thornton, J. M. (1994). *Protein Sci.* **3**, 1224–1235.
Wilson, A. J. C. (1942). *Nature (London)*, **150**, 151–152.
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. (1995). *Structure*, **3**, 1147–1158.