Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of dethiobiotin synthetase at 0.97 Å resolution

Tatyana Sandalova,* Gunter Schneider, Helena Käck and Ylva Lindqvist

Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden

Correspondence e-mail: tanja@alfa.mbb.ki.se

Received 11 August 1998
Accepted 22 October 1998

PDB Reference: dethiobiotin synthetase, 1byi.

The crystal structure of the 224-residue protein dethiobiotin synthetase from *Escherichia coli* has been refined using X-ray diffraction data at 0.97 Å resolution at 100 K. The model, consisting of 4143 protein atoms including 1859 H atoms and 436 solvent sites, was refined to a final *R* factor of 11.6% for all reflections, and has an estimated mean standard uncertainty for the atomic positions of 0.022 Å, derived from inversion of the blocked matrix. The structure was refined with a full anisotropic model for the atomic displacement parameters using *SHELX*97. Stereochemical restraints were applied throughout the refinement. In the last cycles, the planarity of the peptide bonds was not restrained, resulting in a mean ω value of 179.6°. Analysis of the most anisotropic regions of the protein shows that they form four clusters of residues. Alternate conformations for the side chains of 15 residues and for the main-chain atoms of six residues from three loops were included in the model. An analysis of C—H···O hydrogen bonds shows that such interactions occur rather frequently in DTBS; in total, 16 such hydrogen bonds were found. In the central β-sheet, 13 C—H···O bonds between carbonyl O and Cα H atoms were found. Other interactions of this type involve main-chain–side-chain and side-chain–side-chain C—H···O bonds. The model includes 436 water sites, of which 233 molecules form the first hydration shell. Analysis of the protein–solvent interactions shows that about one third of the accessible surface of the enzyme is not covered by ordered solvent. No difference in propensity for ordered solvent close to hydrophilic or hydrophobic surface areas was found. The comparison of the 100 K structure with the structure of the enzyme determined at room temperature shows several regions with different conformation, including areas in the active site, suggesting that structural transitions can occur during flash freezing. This observation questions one of the basic assumptions in the analysis of enzymatic reaction mechanisms using cryocrystallography.

1. Introduction

In spite of the rapidly increasing number of protein structures determined by X-ray crystallography, there is still a surprisingly small number of structures reported at atomic resolution. A generally accepted (albeit somewhat arbitrary) definition of atomic resolution was suggested by Sheldrick (1990), who proposed that such data should extend to higher resolution than 1.2 Å, with more than 50% of the theoretically measurable reflections in the outer shell having $I > 2\sigma(I)$. The advantages of atomic resolution for the understanding of protein structure and function and the present state of affairs have been reviewed recently (Dauter *et al.*, 1995, 1997). Obvious benefits from atomic resolution are, for instance,

more precise estimates of stereochemistry in proteins rather than stereochemical parameters derived from small peptides and the analysis of solvent structures in protein crystals. In functional terms, there has yet to be an enzyme mechanism fully analyzed by structural enzymology at atomic resolution. Such data are of utmost importance if chemical catalysis at the active site of enzymes is to be understood more fully in theoretical terms. Quantitative treatments of enzyme catalysis based on three-dimensional structures suffer from the fact that catalytically important structural changes have not yet been determined reliably owing to limits in the resolution of the X-ray data.

Dethiobiotin synthetase (DTBS; E.C. 6.3.3.3) is the enzyme catalyzing the penultimate step in the synthesis of the vitamin biotin in microorganisms and plants. This reaction consists of the formation of dethiobiotin from diaminopelargonic acid, CO_2 and Mg-ATP (Eisenberg & Krell, 1979; Schneider & Lindqvist, 1997). DTBS is a homodimer composed of 224 amino-acid residues per subunit. Crystallographic studies of dethiobiotin synthetase have revealed that the structure of the enzyme subunit consists of one domain folded into a seven-stranded parallel β-sheet flanked by helices (Huang *et al.*, 1994; Alexeev *et al.*, 1994). Subsequent crystallographic studies of binary, ternary and quaternary complexes of DTBS with various substrates and substrate analogues (Huang *et al.*, 1995; Alexeev *et al.*, 1995) have led to considerable mechanistic insights. In addition, the structure of a phosphorylated reaction intermediate has been determined at 1.6 Å resolution using kinetic crystallography (Käck *et al.*, 1998).

Here, we report the three-dimensional structure of dethiobiotin synthetase at 0.97 Å resolution. With 224 amino acids per subunit, this is one of the largest proteins so far analyzed at this resolution.

2. Methods

2.1. Data collection

Crystals of DTBS were obtained by the hanging-drop vapour-diffusion technique, with polyethylene glycol 8000 as a precipitant in a 0.1 *M* cacodylate buffer at pH 6.5 containing 200 m*M* magnesium acetate, following a previously described protocol (Huang *et al.*, 1994). The crystals grow to a size of up to 0.5 × 0.5 × 0.5 mm.

Data were collected at beamline X11 at the EMBL outstation, DESY, Hamburg at a wavelength of 0.927 Å at 100 K. Before mounting in the nitrogen stream, the crystals were transferred into mother solution containing 20% 2-methyl-2,4-pentanediol for several seconds.

A 30 cm MAR Research image-plate scanner was used as detector. Four data sets were collected on the same crystal at different resolutions and with different exposure times to cover the whole range of intensities. The data were processed with *DENZO* and *SCALEPACK* (Otwinowski, 1993). The crystal belonged to space group *C*2 with unit-cell parameters $a = 72.45 \pm 0.03$, $b = 47.81 \pm 0.02$, $c = 61.20 \pm 0.04$ Å, $\beta = 107.22 \pm 0.02^\circ$. These unit-cell dimensions are slightly smaller than those found at room temperature (73.2, 49.2 and 61.8 Å, respectively), indicating a shrinkage of the unit cell at the cryogenic temperature.

Table 1
Data-processing statistics.

Resolution range (Å)	15.0–0.97
Number of observations	494742
Number of unique reflections	116232
Completeness (%)	
Overall	98.5
Lowest shell (15–3.3 Å)	99.9
Highest shell (0.98–0.97 Å)	94.3
R_{sym} (%)	
Overall	2.5
Lowest shell (15–3.3 Å)	2.1
Highest shell (0.98–0.97 Å)	24.5
$\langle I \rangle / \langle \sigma \rangle$	
Overall	54.4
Lowest shell (15–3.3 Å)	75.4
Highest shell (0.98–0.97 Å)	4.2

In total, 494742 observations from the four data sets were merged together, giving 116232 unique reflections with an overall completeness of 98.5% and an overall R_{merge} of 2.5%. A summary of data processing is given in Table 1.

2.2. Structure refinement

The structure of DTBS refined at 1.65 Å resolution with diffraction data collected at 298 K (PDB code 1DTS; Huang *et al.*, 1994) was used as the starting model for the refinement. All reflections were included in the refinement; no sigma cut-off was used. 5% of the reflections were set aside for cross-validation analysis by means of R_{free} (Brünger, 1992).

The difference in unit-cell dimensions resulted in large initial values of R and R_{free} for model 1DTS, 36.6 and 39.9%, respectively. After 20 steps of rigid-body refinement using *X-PLOR* (Brünger *et al.*, 1987) in the resolution interval 15.0–2.0 Å, R and R_{free} decreased to 29.9 and 32%, respectively. Several rounds of positional refinement including a solvent mask for bulk water followed by isotropic refinement of individual B factors were performed. In this procedure, the resolution was increased stepwise from 2.0 to 0.97 Å. At the end of this stage of refinement, R and R_{free} were 24.7 and 26.8%, respectively.

Several rounds of *REFMAC* (Murshudov *et al.*, 1997), during which eight residues were modelled in double conformations and more water molecules were added, resulted in a model with R and R_{free} of 17.6 and 19.1%, respectively. Conventional isotropic refinement did not decrease the R factors any further. The $3F_o - 2F_c$ electron-density map was of high quality and most atoms were very well defined, with the exception of two amino-acid residues in a flexible loop and several side chains at the protein surface. However, the electron density for many atoms, in particular the S atoms, indicated that isotropic treatment of the thermal atomic displacement parameters was not optimal for DTBS at this resolution.

At this stage, the refinement was continued with *SHELX*97 (Sheldrick, 1997) based on intensities rather than on structure-factor amplitudes. Each step of the refinement with *SHELX*97 consisted of 8–12 cycles of restrained conjugate-gradient least-squares minimization. After each step, $3F_o - 2F_c$ and $F_o - F_c$ electron-density maps were calculated and the model was rebuilt if necessary using the program *O* (Jones *et al.*, 1991). After every run, the 20 highest peaks in the $F_o - F_c$ electron-density map which were found not closer than 2 Å to protein atoms were added to the model as solvent molecules. These were manually checked after every run, and those solvent molecules which had either electron density lower than 1 e Å^{-3} (1.5σ) in the $3F_o - 2F_c$ map or a *B* factor of higher than 60 Å^2 were removed. During the first *SHELX*97 refinement runs, default values were used for all distances (σ = 0.02 Å for bonded atoms and 0.04 Å for atoms bonded to one atom), planarity and chiral restraints (σ = 0.1 Å^3). The final steps of the refinement were carried out without any restraints on the planarity of peptide bonds or the chiral volume of Cα atoms.

All H atoms were added according to geometrical criteria. Their atomic displacement parameters (ADPs) were set to 1.2 times the value of the atom they were bound to. Bonded atoms were restrained to have the same ADPs with a standard deviation of 4 Å^2. During anisotropic refinement, six parameters describing the anisotropy of the thermal vibrations were refined for each non-H atom. The components of the ADPs of bonded atoms were restrained to be the same in the direction of the bond within a standard deviation of 0.01. Atoms which are located closer than 1.7 Å were restrained to have equal components of the ADPs within an effective standard deviation of 0.04. The ADPs of water atoms were restrained to be approximately isotropic.

The first run of *SHELX* was performed with isotropic refinement for all atoms including H atoms and with diffuse solvent parameters. The *R* and $R_{\rm free}$ factors calculated after this run were 16.3 and 18.6%, respectively. Refinement of ADPs decreased *R* and $R_{\rm free}$ to 12.5 and 15.8%, respectively. After this step, the test and working sets of reflections were merged and all data were included in the refinement. Addition of more residues in multiple conformations and more water molecules resulted in the final model with *R* factors of 11.6% for all reflections and 11.2% for reflections with $I > 4\sigma$.

Fully unrestrained refinement resulted in serious deterioration of the protein geometry. After five cycles of *SHELXL*, the *R* factor dropped by 0.2%; however, 33% of the covalent bonds and 27% of the bond angles were outside the limits of standard uncertainties. Many residues with double conformations had 5–7σ deviations from average values of bond lengths and 10–12σ deviations in bond angles. Thus, unrestrained refinement cannot be used for this data set despite the high ratio of observations to parameters (Table 2).

Estimates for the coordinate errors were obtained from a *SIGMAA* plot (Read, 1986) and calculated using *REFMAC* or *SHELX*. The conjugate-gradient algorithm which was used in the refinement using *SHELX* does not provide estimated standard uncertainties. In order to calculate these, a final run of blocked-matrix refinement was performed. The protein was split into 20 overlapping blocks of 15 residues, each with a two-residue overlap. Positional and anisotropic displacement parameters were refined in each of these blocks.

Table 2
Refinement parameters and statistics of the final model.

Resolution range (Å)	15.0–0.97
Number of reflections used in refinement	116232
Number of parameters in the final run of *SHELXL*	20551
Protein atoms (including partially occupied)	2284
Water sites	436
R factor for $F > 4\sigma$ (%)	11.2
R factor for all data (%)	11.6
Goodness of fit	2.6
Mean isotropic equivalent *B* factor (Å^2)	14.8
All protein atoms	12.0
Main-chain atoms	10.2
Side-chain atoms	13.9
Water molecules	27.9
R.m.s. deviation from ideal geometry†	
Bond length (Å)	0.015
Bond angles (°)	2.345
Dihedral angles (°)	23.06
Impropers (°)	1.691
R.m.s. deviation in *B* factors (Å^2)‡	
B factors of bonded atoms	3.0
B factors of non-bonded atoms	7.5
Ramachandran plot statistics§	
Residues in most favoured regions (%)	89.5
Residues in allowed regions (%)	10.5
Residues in disallowed regions (%)	0.0

† Calculated by *X-PLOR*. ‡ Calculated by *MOLEMAN*. § Calculated by *PROCHECK*.

The stereochemistry of the model was analyzed with *PROCHECK* (Laskowski *et al.*, 1993) and *X-PLOR* (Brünger *et al.*, 1987). In order to calculate the solvent accessibility, all symmetry-related molecules were generated using crystallographic symmetry operators. The accessible surface area was calculated by *X-PLOR* for the whole molecule, the hydrophilic part of the surface (sum of the accessible surface area of all O and N atoms) and the hydrophobic area (sum of the atomic accessible surface area for all C and S atoms). Structural comparisons were performed using the *LSQ* option in *O* (Jones *et al.*, 1991) or using the *LSQMAN* program (Kleywegt & Jones, 1997). The observed intensities and the refined atomic coordinates have been deposited with the Protein Data Bank.

3. Results and discussion

3.1. Model and electron-density map

Table 2 shows the refinement parameters and statistics of the final model of dethiobiotin synthetase. This model consists of 2284 non-H atoms: 1848 protein atoms (54 main-chain and 95 side-chain atoms were modelled in double or triple conformations with partial occupancies) and 436 solvent sites. 1859 protein H atoms were built according to geometrical criteria.

From the plot of the final *R* factor as a function of resolution (Luzzati, 1952) (not shown), an overall coordinate error of 0.05–0.07 Å for the final model can be estimated. The

SIGMAA plot (Read, 1986) resulted in a similar value (0.07 Å) for the overall coordinate error. This value should, however, be used only for the comparison of the models rather than as a real coordinate error estimate. Estimated overall coordinate standard uncertainties calculated using *REFMAC* (Murshudov *et al.*, 1997) were 0.023 and 0.018 Å, based on the least-squares and maximum-likelihood residuals, respectively. These values compare well with the estimated standard uncertainties in atomic positions from blocked-matrix refinement in *SHELX*97, which gave a mean estimated positional standard uncertainty of 0.022 Å for all protein non-H atoms. The coordinate error averaged for all atoms of each residue is shown in Fig. 1(*a*). Some of the main-chain atoms have coordinate standard uncertainty less than 0.01 Å, with a mean value for all main-chain atoms of 0.017 Å. The largest coordinate errors were found for the terminal atoms of the side chains of Glu12, Arg32, Glu44, Glu89, Gln107, Gln108, Met156, Lys182 and Asn209, which are poorly defined in the electron-density map.

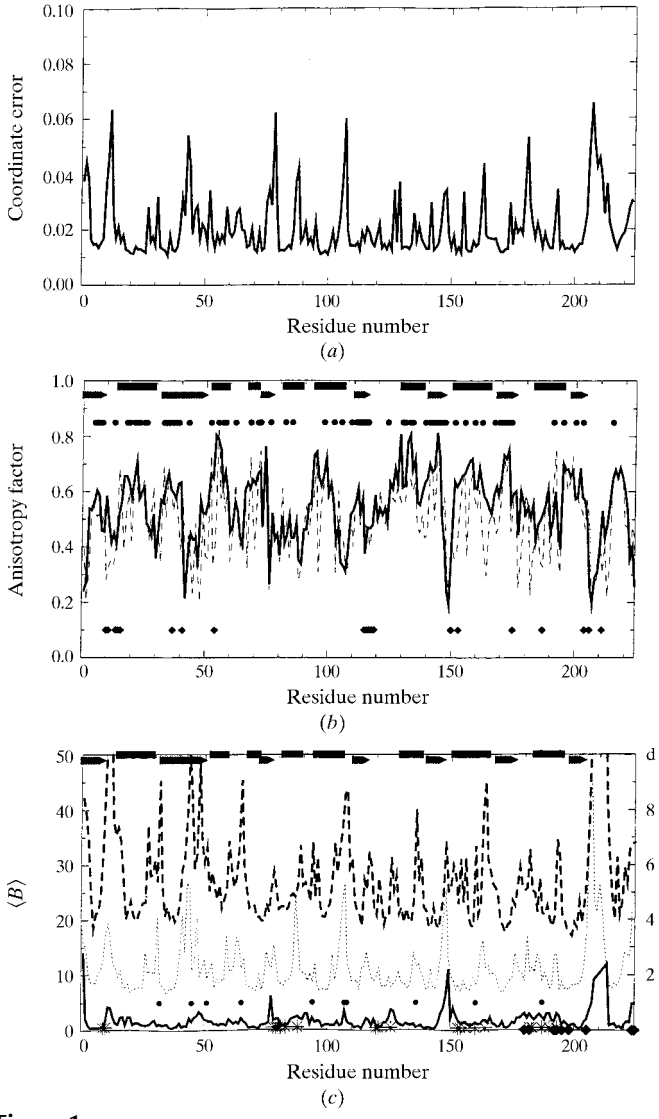


Figure 1
(*a*) Mean coordinate error of residues estimated from block-matrix refinement in *SHELX*. (*b*) Anisotropy factors (maximum eigenvalue of the atomic displacement parameters divided by maximal eigenvalue) for all main-chain and side-chain atoms are shown as solid and dashed lines, respectively. All buried residues are shown as a filled circle above the lines. Residues from the active site are denoted as filled diamonds below the plot. Residues in helices and in β-strands are shown as filled rectangles and triangles at the top of the graph. (*c*) Difference between the 298 and 100 K models of DTBS. The temperature factor $\langle B \rangle$ averaged for main-chain atoms of every residue for room temperature (dashed line) and liquid-nitrogen temperature (dotted line) structures is shown as a function of residue number. Deviation in Cα position, *d*, for each residue is shown as a solid line. Residues from the dimer interface are denoted as stars and residues involved in dimer–dimer packing as filled diamonds at the bottom. Residues in helices and in β-strands are shown as filled rectangles and triangles at the top of the graph. Residues with different side-chain conformation are shown as circles above the solid line.

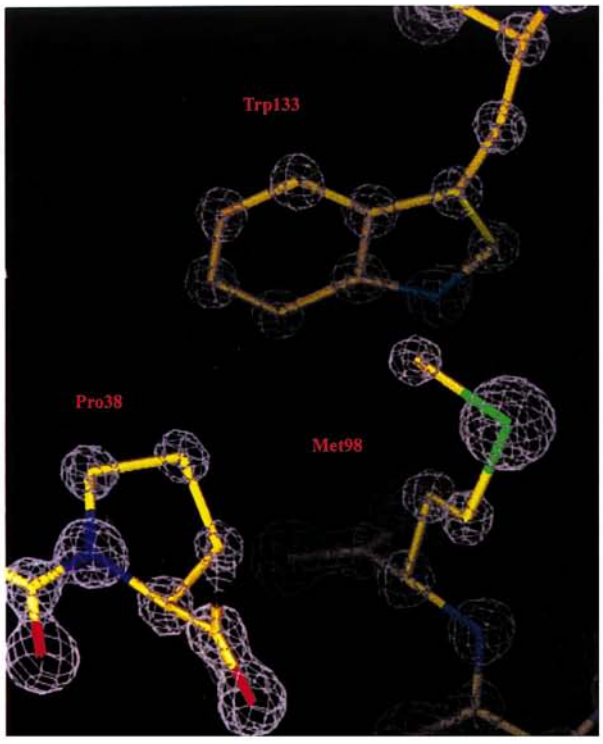
Refinement of DTBS at 0.97 Å resolution resulted in an electron-density map of excellent quality. For the majority of the molecule, individual atoms have spherical electron densities in the $3F_o - F_c$ map contoured at 2.7 e Å^{-3} (4.0 r.m.s.). Fig. 2(*a*) shows a representative part of the $3F_o - F_c$ electron-density map contoured at 2.7 e Å^{-3} (4.0 r.m.s.). The electron-density map allowed the localization of H atoms. Figs. 2(*b*) and 2(*c*) show parts of an omit map [contoured at 0.24 e Å^{-3} (2.7 r.m.s.)] where H atoms were removed from the model and two cycles of refinement were performed to remove bias; clear peaks indicating H atoms can be seen. About 75% of the main-chain H atoms can be seen in this $F_o - F_c$ map at a contour level of 2.7 r.m.s. Fig. 2(*c*) shows that also many side-chain H atoms can also be positioned in the electron density, especially for residues with low ADPs. In some (but not all) cases, the $F_o - F_c$ map allowed the localization of hydroxyl H atoms (Fig. 2*c*).

In the 298 K structure of DTBS , there is a flexible loop with weak electron density and four residues (residues 209–212) could not be seen at all in the electron density. This loop is better defined at 100 K. However, for one residue, Ala207, the electron density is still very poor. There are also several side chains which are poorly defined in the final electron-density map.

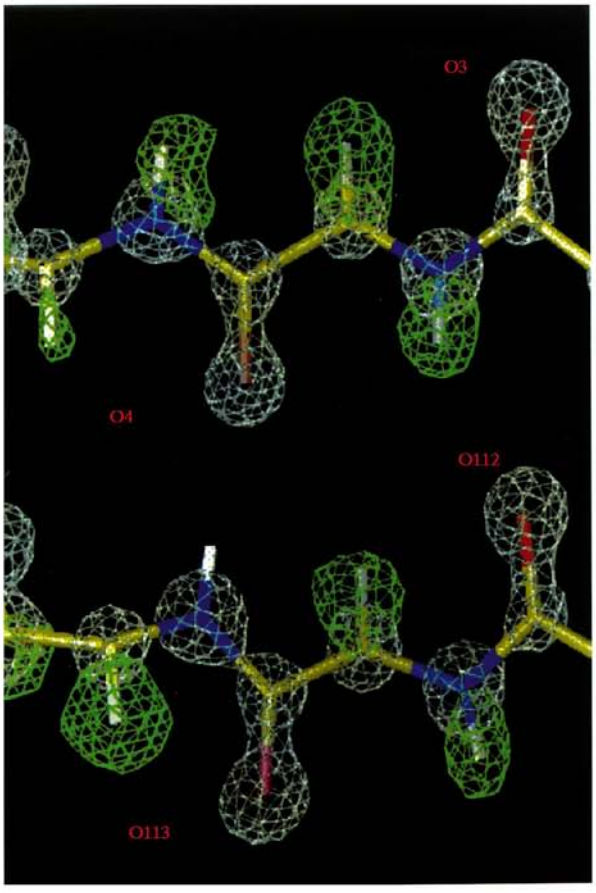
The three-dimensional structure of DTBS at 100 K is essentially the same as was found at 298 K (Fig. 3). However, several loop regions differ in conformation, as discussed below.

3.2. Stereochemistry

The analysis of the stereochemistry of the model of DTBS shows that all geometric parameters are within the limits expected for this resolution. The mean deviation from the ideal geometry is larger than that often present in the final refinement statistics of proteins refined at lower resolution because of tighter restraints on geometry in these cases (Table 2). Several close contacts (2.2–2.4 Å) between water and protein atoms and between water molecules were found, despite anti-bumping restraints in *SHELXL*. Similar contacts were reported for other proteins refined at atomic resolution. The Ramachandran plot shows 89.5% of residues in the most favoured regions; no residues were found in the disallowed or generously allowed regions.



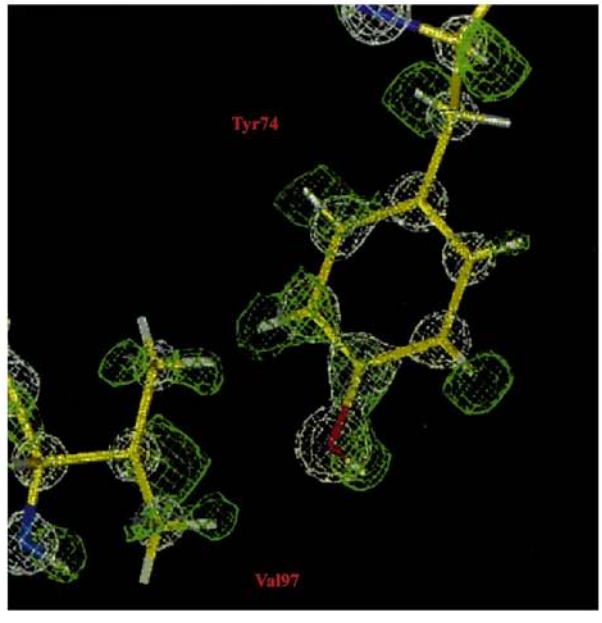
(*a*)



(*b*)

3.2.1. Bond lengths and bond angles. During refinement, default restraints as implemented in *SHELX* were applied for all 1–2 and 1–3 distances with $\sigma = 0.02$ and 0.04 Å, respectively. The final model has r.m.s. deviations of 0.017 and 0.035 Å for these values. The crystallographic contribution dominates at atomic resolution; nevertheless, the stereochemical quality of the resulting protein model, as assessed by the usual criteria, is rather high. The mean values of the bond lengths are very similar to those used in *PROCHECK*, but some deviations greater than 3σ were found, mostly between atoms with partial occupancies. Such large deviations from ideal geometries were also found in other proteins refined at atomic resolution (EU 3-D Validation Network, 1998). In DTBS as well as in cutinase (Longhi *et al.*, 1997), the largest differences between observed bond lengths and their target values were found for C—N distances in proline and Cα—C distances in glycine residues.

3.2.2. Planarity of peptide bonds. During the refinement with *SHELX*97, the planarity of the peptide bonds was maintained by the chiral volume of main-chain C atoms which were restrained to zero (with a standard deviation of 0.10 Å^3). With these restraints the average value for ω is equal to 179.7° (σ = 5.3°). The distribution is slightly asymmetric and the value is in accordance with those calculated previously, *e.g.* for the cutinase (179.6°, σ = 5.2°; Longhi *et al.*, 1997). Towards the final stages of the refinement these restraints were removed, and in the last rounds no restraints for the planarity of peptide bonds were applied. In order to remove the possible bias which could be imposed by previous stages of the refinement, a random shift of 0.04 Å was applied to the coordinates. This change in the protocol did not result in significant changes in ω or the model. After 35 cycles of *SHELX* an average ω of 179.6° was obtained, which is very close to the previous value,



(*c*)

Figure 2
(*a*) $3F_o - 2F_c$ electron-density map around Trp133 in the core of DTBS. The contour level is 2.7 e Å^{-3} (4.0 r.m.s.) and the electron-density map clearly shows the position of all non-H atoms. (*b*) $3F_o - 2F_c$ (in white) and hydrogen omit $F_o - 2F_c$ (in green) electron-density maps of the central β-sheet of DTBS. The $F_o - 2F_c$ map is contoured at 0.24 e Å^{-3} (2.7 r.m.s.). (*c*) The same electron-density map as in (*b*), in the region of Tyr78.

with a minor increase in the standard deviation to 5.9°. Moreover, the extreme ω values change by 0.5° at most. Removal of the restraint did not really affect the model: neither *R* factor nor the goodness of fit changed significantly after this stage of refinement. We conclude, at least for DTBS, that the planarity of the peptide bonds is inherent to the polypeptide chain and not a property imposed by the restraints used in the refinement programs.

In order to compare these results with those obtained for other proteins refined at atomic resolution, the average ω value for 1201 residues of DTBS and seven other proteins available in Protein Data Bank was calculated (Table 3). ω varies in the range 159–198° with an average value for all 1201 residues of 178.97° (σ = 5.8°). We note that the standard deviation is only 0.2° higher than accepted in *PROCHECK*; however, the average value is 1° lower. The value of the mean ω depends on the α-helix content of the protein (MacArthur & Thornton, 1996). 'Ideal' hydrogen-bonding geometry of main-chain atoms in α-helices results in enhancing the double-bond character of the peptide bond, thus preventing large deviations from planarity. Indeed, the lowest value, $\langle\omega\rangle$ = 177.7°, was found for concanavalin (PDB code 1JBC), which has less than 5% of residues in α-helical conformation. The mean values of ω calculated for DTBS and cutinase are very close to 180°, consistent with the high content of α-helices (35–40%) in these proteins. Thus, the lower mean value of ω obtained for the eight proteins probably reflects the low proportion of α-helical structure in these proteins.

In order to elucidate whether different types of residues have different propensities to disturb the planarity of the peptide bonds, the average ω angle was calculated for the different amino acids (Table 3). Threonine, serine and tyrosine residues have their average ω values lowered by more than 1.5° from the mean value. The low values of $\langle\omega\rangle$ for tryptophan and methionine might be an artifact arising from the small number of these residues in the database. The other residues with statistically significant numbers of observations show no correlation between residue type and planarity of the peptide bond.

In all proteins refined at atomic resolution, some residues were found with ω deviating by more than 3σ from the average value. There seems to be no correlation between the frequency of large deviations in ω and the residue type, but this conclusion might be biased by the limited size of the database of proteins at atomic resolution.

3.3. Analysis of anisotropic displacement

The drop of 3.8% in the *R* factor after anisotropic refinement confirmed that DTBS, like other proteins refined at atomic resolution, is better described in terms of anisotropic displacement. Similar to the results for other proteins refined at atomic resolution, the anisotropy does not result in a large shift of the atomic positions: the root-mean-square coordinate distance between Cα atoms of the DTBS model before and after anisotropic refinement is about 0.1 Å. However, the shift is significant if compared with the mean coordinate error (0.02 Å). The equivalent *B* factor calculated after anisotropic refinement (Table 2) is slightly larger than the *B* factor of the isotropic model, which is 9.5 and 12.0 Å^2 for main-chain and side-chain atoms, respectively.

The mean anisotropy factor (minimum eigenvalue divided by maximum eigenvalue) is equal to 0.55 and 0.50 for main-chain and side-chain atoms of DTBS, respectively. This means that all ellipsoids approximating the thermal motion of atoms are twice as large in one direction than in the other. Analysis of the anisotropic thermal motion of crambin (Stec *et al.*, 1995) and lysozyme (Harata *et al.*, 1998) showed that a considerable part of the motion (60–67%) corresponds to motion of the protein molecule as a rigid body. However, the local motion of atoms, especially those located on the surface of the protein, are responsible for the difference in anisotropy of residues. The anisotropy factor, averaged for all main-chain and side-chain atoms of each residue, is shown in Fig. 1(*b*). There is a correlation between the accessibility of the residue and its anisotropy. The mean anisotropy factor of main-chain atoms of all buried residues [accessible surface area, calculated by *DSSP* (Kabsch & Sander, 1984), does not exceed 10 Å^2] is 0.59. Among the 48 most isotropic residues (anisotropy factor >0.65), 34 residues are completely buried. Nevertheless, there are several buried residues which are rather anisotropic: Phe76 is the last residue in strand β3 and is completely buried, but its anisotropy factor is 0.27. Several active-site residues are rather anisotropic: for instance Glu115 and the side chain of Tyr187 both have anisotropy factors of 0.37.

The majority of the most anisotropic residues are located close to each other, forming four clusters. Three of them are in the vicinity of the active site. The first

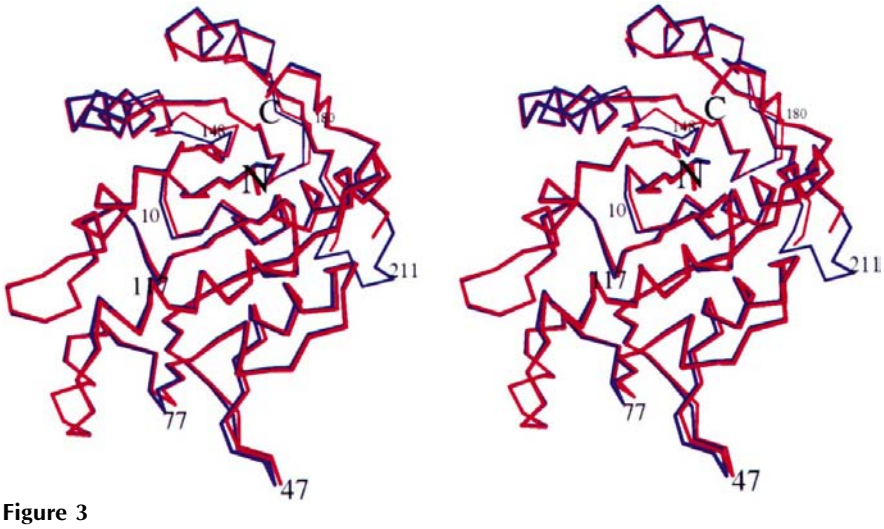


Figure 3
Stereoview of superposition of the 298 K (red) and 100 K (blue) structures of DTBS. Residues with the largest deviation are indicated.

Table 3
Average ω angle (°) for different residue types.

	DTBS			Average from database†		
Resid	ω	σ	N_{res}‡	ω	σ	N_{res}‡
Ala	178.94	5.71	32	178.16	5.10	74
Asp	183.65	6.52	8	179.54	5.16	69
Asn	179.77	5.14	9	180.80	7.03	39
Arg	183.79	7.21	9	178.47	6.8	51
Cys	179.01	4.41	2	180.51	5.24	23
Gln	183.17	3.28	10	179.08	5.4	31
Glu	177.11	5.93	13	180.36	6.36	40
Gly	181.82	5.34	19	179.91	5.77	81
His	179.08	5.06	4	177.52	8.98	16
Ile	180.48	4.94	9	178.75	6.2	50
Leu	178.25	5.76	24	178.63	5.78	64
Lys	181.67	7.08	9	179.10	3.80	39
Met	179.38	2.91	4	177.10	4.03	12
Phe	180.21	5.65	6	179.63	6.15	37
Pro	178.28	5.73	15	180.21	4.34	51
Ser	180.79	5.47	13	176.13	5.80	75
Thr	177.71	7.22	22	177.20	6.3	85
Trp	183.10	4.41	5	174.75	5.33	11
Tyr	177.80	7.36	7	176.19	6.6	58
Val	177.55	5.18	16	179.21	5.51	57
All	179.6	5.96	238	178.79	5.93	963
Total§				178.97	5.82	1201

† The following seven proteins were used in the analysis: eye lens γ-crystallin (Kumaraswamy *et al.*, 1996), crambin (Stec *et al.*, 1995), cutinase (Longhi *et al.*, 1997), immunoglobulin binding protein (Derrick & Wigley, 1994), concanavalin (Min *et al.*, 1992), ribonuclease (Sevcik *et al.*, 1996) and rubredoxin (Dauter *et al.*, 1992). ‡ Number of residues. § DTBS and seven proteins used in analysis.

cluster includes residues from the loop connecting strand β5 and helix α5, Lys148 and Leu149, which are close to Tyr187 of the same subunit and Thr80 of another subunit of the dimer.

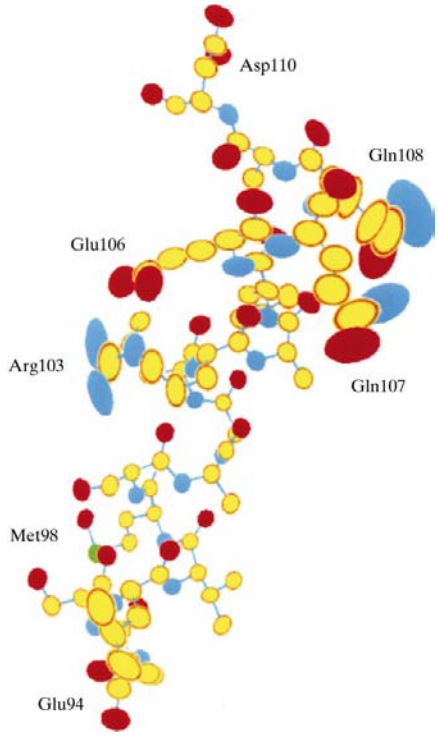


Figure 4
View of one of the anisotropic clusters in DTBS, at the end of helix α3. Thermal ellipsoids are drawn with 80% probability using *ORTEX* (McArdle, 1995).

Table 4
Residues with alternative conformations of side chains in DTBS.

Residue	Atoms	Occupancy
Lys37	C^{ε}, N^{ζ}	0.58, 0.42
Ser43	O^{γ}	0.57, 0.43
Ser53	O^{γ}	0.42, 0.31, 0.27
Thr70	O^{γ}	0.66, 0.34
Pro73	C^{α}, C^{β}, C^{γ}, C^{δ}	0.68, 0.32
Gln88	From C^{β}	0.61, 0.39
Glu89	From C^{γ}	0.76, 0.24
Phe128	All atoms	0.64, 0.36
Phe130	All atoms	0.67, 0.33
Leu143	From C^{γ}	0.81, 0.19
Met156	From C^{γ}	0.34, 0.35, 0.31
His164	From C^{γ}	0.53, 0.47
Asn175	All atoms	0.68, 0.32
Met194	From C^{β}	0.77, 0.23
Lys217	C^{ε}, N^{ζ}	0.70, 0.30

The second cluster consists of the residues from the C-termini of strands β2 and β3, residues Ser41 and Phe76. Gln88 and Glu89 from the helix α3A are located nearby. These residues were modelled in alternate conformations but the high anisotropy implies that this part of the molecule is rather disordered. Residues 207–215 from the loop which is disordered at room temperature form the third cluster. In all these clusters the longest axis of ellipsoids is more than three times larger than the shortest. The last cluster of anisotropic residues is on the opposite side of the molecule, at the C-terminus of helix α3 (Fig. 4).

Water molecules were refined with *ISOR* restraints in *SHELX* to keep them isotropic. Despite this restraint, water molecules turned out to be rather anisotropic: the mean anisotropy factor of water molecules is 0.45, indicating that water molecules are in average more anisotropic than protein atoms. There is no selected direction of anisotropy for water molecules.

3.4. Alternate conformations and disordered residues

Double or triple conformations of side chains were found for 15 residues in DTBS (Table 4). Electron-density maps calculated before the anisotropic refinement indicated that a large number of residues (in particular, half of all proline residues) might adopt multiple conformations. However, anisotropic ADP refinement led to the disappearance of most of these densities. Nevertheless, at 15 sites the electron-density maps and the large anisotropy of the atomic displacement parameters led us to model the residues with double or triple conformations.

In most cases, two main rotamers with angles χ_i from the favourable region suffice to model the observed electron densities. However, there are some exceptions where one conformation of the side chain has unusual angles χ_1 and χ_2. One particularly striking example is two adjacent residues, Phe128 and Phe130, which both have to be modelled with alternate positions (Fig. 5). The closest distance between side-chain atoms of these residues is 4 Å. In both cases, rotations around the C^{α}–C^{β} and C^{β}–C^{γ} bonds result in a shift of the

positions of the benzene rings of up to 1 Å, with the χ_1 and χ_2 angles outside the favourable regions for one of the conformations.

12 out of 15 residues modelled with multiple side-chain conformations are exposed residues without steric restrictions on the position of the side chain. Three residues are completely buried in the interior but only one of them, Leu143, is located in the hydrophobic core of the molecule. In spite of this, the two different conformers do not strongly overlap with other atoms. The two other buried residues with bifurcate side-chain conformations, Lys37 and Ser43, are found at the active site. Lys37 is very important both for substrate binding and catalytic activity (Huang *et al.*, 1995; Yang *et al.*, 1997; Käck *et al.*, 1998). In the native enzyme, Lys37 forms hydrogen bonds with main-chain atoms. The terminal N^{ζ} atom of Lys37 was found in two conformations with equal occupancy. In the first conformation, N^{ζ} forms hydrogen bonds with the backbone O atom of residue 116 and a water molecule. In the second conformation, the hydrogen bond with O116 is retained and an additional salt bridge with the side-chain O atom $O^{\varepsilon 2}$ of Glu115 is formed. The side chain of Ser43 points into the active-site cavity, but it is not directly involved in the binding of the substrate or in catalysis. In both conformations there is a hydrogen bond with the main-chain O atom of residue 76. In addition, a hydrogen bond with $N^{\delta 2}$ of Asn52 is formed in the second conformation.

Two methionine residues which are located on the surface of the molecule can adopt two (Met194) or three (Met156) alternative conformations. They are the only examples of non-polar exposed residues with multiple side-chain conformations in DTBS.

In the structure of DTBS determined at room temperature, double conformations were modelled for four amino acids: Arg60, Arg91, Leu113 and Asn175. In the 100 K structure, only one of these residues, Asn175, is observed with multiple conformation.

Table 5
Main-chain angles of the residues from three loops which were modelled in two discrete conformations.

Residue	φ (°)	ψ (°)
Thr11	−74	−6
	−71	155
Glu12	−120	54
	56	70
Ala77	−70	−33
	−73	−52
Glu78	−115	146
	−70	125
Lys182	−47	−38
	−55	0
Arg183	−105	24
	68	22

In addition to these 15 residues with side chains modelled in discrete alternate conformations, the $3F_o - 2F_c$ electron-density map showed clearly that three loop regions were discretely disordered and could be modelled in two conformations with approximately equal occupancy (Table 5). The electron density for these regions is not as good as for the core of the protein; nevertheless, in the $3F_o - 2F_c$ electron-density map contoured at the level 1 e $Å^{-3}$ (1.5σ) most of the main-chain atoms are inside the electron density. All attempts to model these loops with one conformation of the main-chain resulted in positive density in the difference electron-density map.

One of these loops is the phosphate-binding loop (residues 9–15), which is flexible in the unliganded enzyme. In the room-temperature structure, the mean B factor for main-chain atoms in this region is almost twice the average B factor. The 100 K structure shows two conformations of this loop. Neither of them coincides with the room-temperature conformation, which in turn does not fit to the low-temperature electron-density map. In the second conformation, a peptide flip at residue Thr11 has occurred in the main chain (Table 5). In addition, residues Thr11 and Glu12 have multiple positions of all their main-chain atoms; the distance between the Cα atoms of Thr11 in the two conformations is 1.15 Å.

The other regions with two discrete conformations of the main-chain atoms, comprising residues 77–78 and 182–183, participate in the packing of the DTBS molecules in the crystal. In the room-temperature structure these loops are well defined in the electron density, and the B factors of the corresponding atoms are equal to the average. Shrinkage of the cell owing to flash-freezing resulted in closer contacts between dimers, especially in the region of the loop between strand β3 and helix α3a (residues 77–80) of one molecule and the loop between β6 and α6 (residues 182–184) of another molecule. The 100 K structure shows that overly

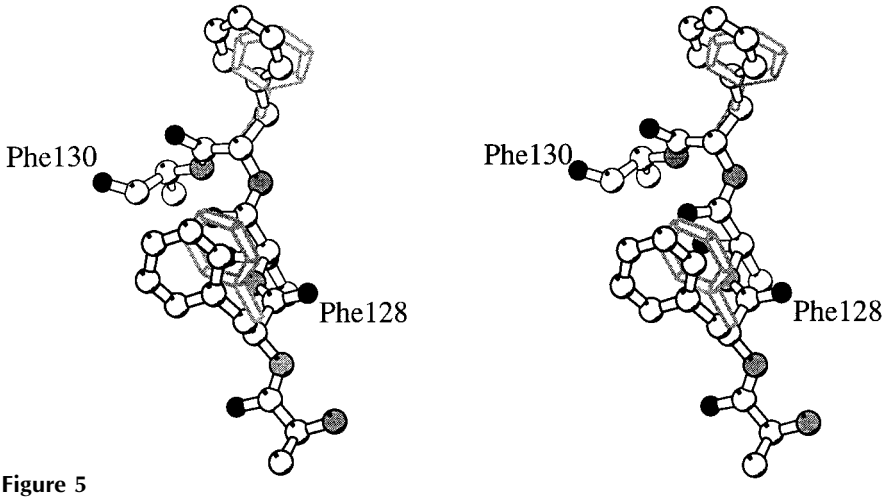


Figure 5
Stereoview of unusual double conformations of the side chains of two adjacent phenylalanines, Phe128 and Phe130. The χ_1 and χ_2 angles are far from the favourable regions for one of conformations: (χ_1, χ_2) are equal to (175, 36°) and (−148, 72°), respectively, for the first and second conformation of Phe128 and (−170, 70°) and (−124, −63°), respectively, for the first and second conformation of Phe130.

Table 6
Hydrogen-bond network in DTBS.

Type of bond†	N‡	$\langle d \rangle$ (Å)§	d_{min} (Å)¶
MM O···HN	156	2.96	2.8
MM O···HC	13	3.25	3.0
MS O···HN	13	2.85	2.7
MS NH···O	22	2.98	2.8
MS O···HO	14	2.74	2.5
MS O···HC	1	3.15	3.15
S O···HN	9	2.93	2.7
S OH···O	8	2.75	2.5
MW O···HO_W	104	2.77	2.6
MW NH···O_W	51	2.96	2.7
MW CH···O_W	5	3.10	3.0
SW O···HO_W	40	2.78	2.4
SW OH···O_W	31	2.80	2.6
SW NH···O_W	44	2.83	2.5
SW CH···O_W	33	3.05	2.8

† M, main-chain atom; S, side-chain atoms; W, water. ‡ Number of bonds. § Average distance for this type of bond. ¶ Minimal distance for this type of bond.

close contacts in this region can be avoided in two different ways, both resulting in the disorder of the two loops. The difference between the two conformations of these loops is rather large; the distance between Cα atoms is 1.26 Å (Ala77) and 1.08 Å (Arg183). In addition to the shift of the main-chain atoms in the loops, a peptide flip is found for the second conformation of Arg183 (Table 5). The conformation of these loops in the room-temperature structure coincides with one of the discrete conformations found in the low-temperature structure.

3.5. Hydrogen bonds in DTBS

An overview of all hydrogen bonds in DTBS is given in Table 6. DTBS has 446 backbone O and N atoms which can form hydrogen bonds. Only four of them (O35, N78, O121 and N201) are buried in hydrophobic regions of the protein and have no contacts closer than 3.5 Å with any other polar atoms. These residues are two buried glycines from the edge of the central sheet (Gly35 and Gly201) and two residues from loops participating in the intersubunit contacts, Glu78 and Phe121. These two residues have extensive accessible surface areas but their backbone atoms are buried.

All other backbone atoms participate in hydrogen bonds. 156 main-chain O···H—N hydrogen bonds were found in DTBS using the program *DSSP* (Kabsch & Sander, 1984). 34 of these bonds are found in the central β-sheet ($\langle d \rangle$ = 2.92 Å, σ = 0.11 Å; hydrogen-bond angle 160°, σ = 12°) and 64 in α-helices ($\langle d \rangle$ = 2.97 Å, σ = 0.11 Å; hydrogen-bond angle 159.8°, σ = 10°). Besides these main-chain–main-chain interactions, there are 42 hydrogen bonds between main-chain atoms and side chains. Thus, 312 out of 446 polar backbone atoms form main-chain–main-chain hydrogen bonds, 42 atoms form hydrogen bond to side chains and four atoms do not form any hydrogen bonds at all. All other polar backbone atoms are accessible to solvent and form hydrogen bonds with ordered water molecules. In addition, there are 17 side-chain–side-chain hydrogen bonds or salt bridges with an average O—N distance of 2.79 Å (σ = 0.17 Å).

3.5.1. Helix capping. DTBS has eight α-helices and most of them have an N-cap residue with high N-capping propensity (Doig *et al.*, 1997), standard geometry and hydrogen-bond pattern. However, the conformation of the N-cap residue Glu94 of helix α6 is rather unusual. Glutamate is not often found among N-cap residues in proteins; only seven helices out of 393 from 85 high-resolution structures have glutamate as an N-cap residue (Doig *et al.*, 1997). None of them has the *trans* conformation of the N-cap side chain that Glu94 has in DTBS with χ_1 = 167°. All seven N-cap glutamates have g^+ conformation, with a carbonyl O atom bound to the N3 residue and the glutamate side chain pointing outwards. Unlike all previously found N-caps, Glu94 forms two hydrogen bonds with N-terminal residues of helix 6. Its backbone O atom forms an (i, i + 3) hydrogen bond with the N3 residue as in other proteins, whereas a *trans* side-chain conformation allows formation of a hydrogen bond with the backbone N atom of the N2 residue.

3.5.2. Main-chain–side-chain hydrogen bonds. Hydrogen bonds between side-chain and main-chain atoms are important for maintaining the integrity of the DTBS globule, since they form at least 25% of the total number of hydrogen bonds. In addition, the shorter mean distance of these bonds suggest that they are stronger than main-chain–main-chain bonds. The contribution of these bonds to the DTBS hydrogen-bond network exceeds that of β-sheet hydrogen bonds. All types of polar uncharged residues contribute to main-chain–side-chain hydrogen bonds. 50% of all threonine, serine, tyrosine, asparagine and glutamine residues form hydrogen bonds between their side-chain and backbone atoms.

Aspartate, asparagine and threonine can form hydrogen bonds between their side-chain O atom and backbone atoms of the adjacent residue (i + 1) or (i + 2). There are nine such bonds in DTBS. In a few cases this bond is part of a more complex hydrogen-bonding network including several residues; an example is shown in Fig. 6, where the side chain of Asn72 participates in the complex hydrogen-bond network including both main-chain and side-

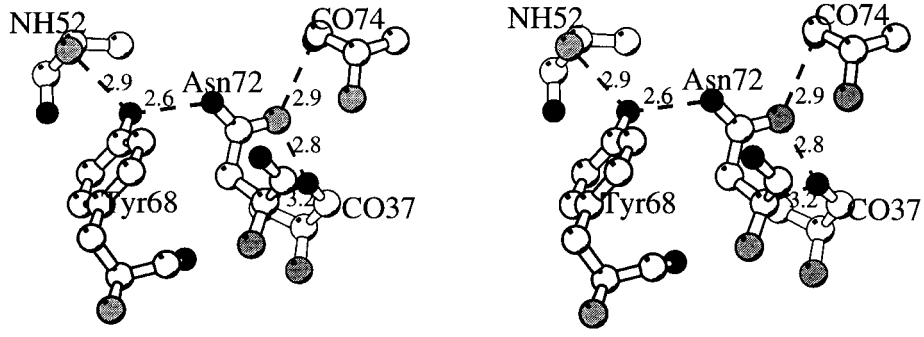


Figure 6
Stereoview of the region around Asn72. The side chain of Asn72 participates in hydrogen bonds with both main-chain and side-chain atoms.

chain atoms from different parts of the molecule. In total, there are 17 hydrogen bonds connecting different secondary structure elements in DTBS.

Threonine residues from helices can form a hydrogen bond between their side chain and the backbone O atom of residue $i + 3$ (or 4), thus competing with the usual main-chain–main-chain hydrogen bond. In DTBS, all threonines from helical regions are engaged in such bonds.

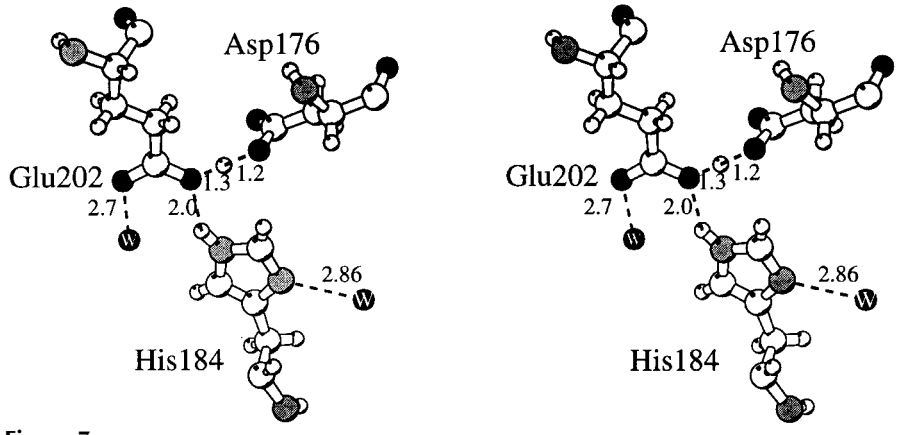


Figure 7
Hydrogen-bonding pattern around His184, including a proton between Glu202 and Asp176. The position of the proton was determined by a peak in an omit-hydrogen map.

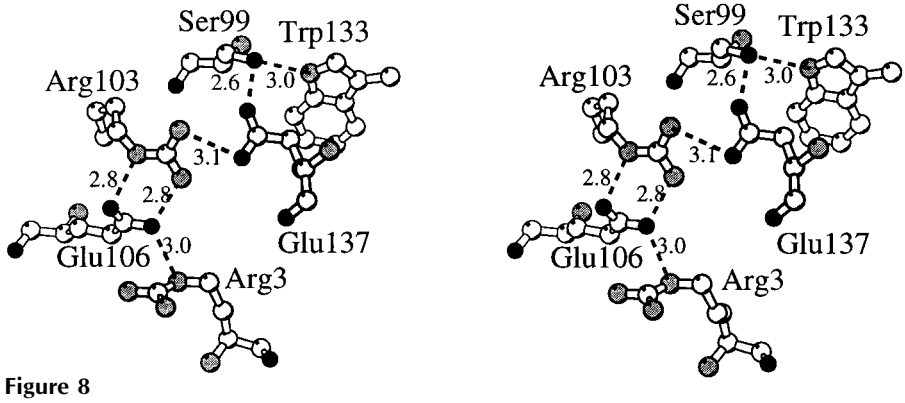


Figure 8
Stereoview of the complex salt bridge close to Trp133 and Arg103.

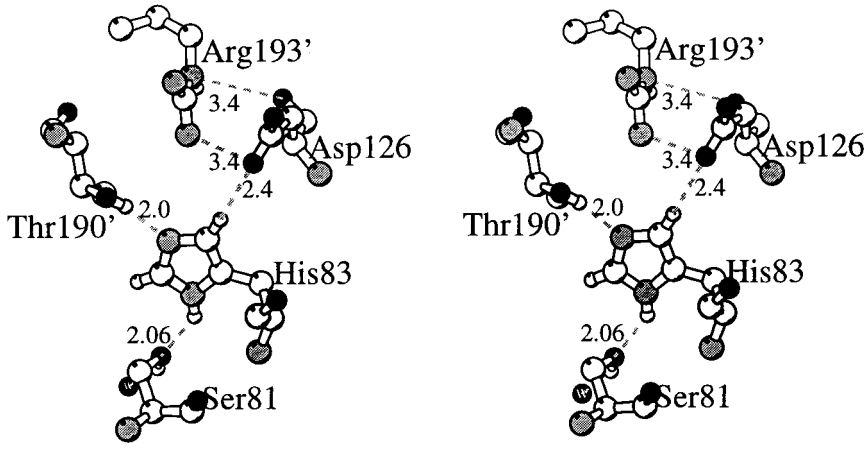


Figure 9
Stereoview of hydrogen-bonding pattern of residue His83 involving a $C^{\delta 2}$—H···O bond.

3.5.3. Charged residues and salt bridges. There are 36 charged residues in DTBS: eight arginines, eight lysines, seven aspartates and 13 glutamates. All these residues are located on the surface of the molecule, except two residues from the active site, Lys37 and Glu115, which are not accessible to the solvent. These two residues, which are involved in substrate binding and catalysis (Huang *et al.*, 1995; Yang *et al.*, 1997), form hydrogen bonds to protein atoms.

Residues exposed to the solvent form hydrogen bonds with ordered water molecules and some of them form hydrogen bonds with backbone atoms or with side chains of uncharged polar residues. The side chains of 11 exposed charged residues do not interact with other protein atoms but form hydrogen bonds to solvent molecules only.

There are 15 short intramolecular contacts (less than 4 Å) between exposed charged residues, and these residues form five networks of salt bridges of different complexity. The two simplest salt bridges are formed by the Lys148/Asp10 and Glu48/Arg51 pairs. There are no other polar atoms from the protein which are close to these bridges. Another interaction includes three residues, Asp176, His184 and Glu202 (Fig. 7). In this interaction, the distance between Glu202 $O^{\varepsilon 2}$ and Asp176 $O^{\delta 1}$ is 2.56 Å, indicating that one of the carboxyl groups must be uncharged. Indeed, electron density at 2.5σ is found in between the two O atoms in a difference omit-hydrogen map, indicating the presence of a proton (Fig. 7). The third network of charged interactions is centered around Arg103, and includes ten atoms from six different residues: Arg3, Arg103, Glu137, Glu106, Ser99 and Trp133 (Fig. 8).

In addition to these intramolecular salt bridges, there are two salt bridges at the subunit–subunit interface. One is formed between Arg193 and two residues of the second subunit, Asp 126 and His83 (Fig. 9). The second salt bridge is very complex, and is formed by Lys182, Arg183 and Glu186 from one subunit and Glu78 and Gln88 from the second subunit. This network of salt bridges is extended in the crystal lattice through participation of the C-terminal carboxyl group from another dimer. It proved very difficult to model this region: two alternative conformations were modelled for Glu78 and Arg183 and

Table 7
Hydrogen bonds C—H···O in DTBS.

Main-chain–main-chain bonds.

O	HCα	d† (Å)	dh‡ (Å)	C—H···O§ (°)	H···O—C§ (°)	χ¶ (°)
5	V141	3.22	2.37	144	137	9
34	V112	3.28	2.38	152	145	12
36	V114	3.19	2.32	148	137	15
111	T33	3.24	2.35	154	143	2
113	G35	3.35	2.53	143	140	5
114	Y5	3.32	2.47	145	128	24
142	V6	3.35	2.46	137	152	14
143	W172	3.04	2.33	128	133	12
144	G8	3.35	2.52	143	126	3
174	E202	3.18	2.31	148	130	13
176	W205	3.20	2.46	134	156	15
198	G171	3.22	2.58	124	130	22
219	G201	3.14	2.25	152	151	25
Average		3.26	2.42	142	139	13

Main-chain–side-chain bonds.

Main-chain atom	Side-chain atom	d† (Å)	dh‡ (Å)	C—H···O§ (°)	H···O—C§ (°)	χ¶ (°)
O146	Cys151 C$^{\beta}$	3.15	2.57	119	161	16

Side-chain–side-chain bonds.

Atom 1	Atom 2	d† (Å)	dh‡ (Å)	C—H···O§ (°)	H···O—C§ (°)	χ¶ (°)
Tyr74 O$^{\zeta}$	Val97 C$^{\beta}$	3.32	2.64	127	116	24
His83 C$^{\delta2}$	Asp126 O$^{\delta1}$	3.34	2.42	171	170	4

† Donor–acceptor distance. ‡ Hydrogen bond H···O distance. § Hydrogen-bond angles. ¶ Dihedral angle H···O—C—N/C.

three conformations for Lys182. There is still some residual positive electron density in the difference map, indicating that the situation is even more complex.

3.6. Hydrogen bonds involving carbon-bound H atoms, C—H···O

Traditionally, close contacts (<3.2 Å) of CH_n groups with polar atoms have been considered as artifacts of particular refinement protocols. Recently, hydrogen bonds involving C—H···O bonds have attracted some attention (Derewenda *et al.*, 1995; Fabiola *et al.*, 1997) and have even been implicated in enzyme catalysis. In order to analyse such bonds in DTBS all close non-bonded C—H···O contacts were identified and the stereochemistry of these contacts was studied. If the protein conformation was not well determined (*i.e.* double conformations or ADP > 40 Å^2) those atoms were not included into this analysis.

The sum of the van der Waals radii of the C and O atoms is 3.25 Å. Thus, contacts closer than 3.3 Å might be candidates for C—H···O hydrogen bonds. Because of the highly directional nature of the hydrogen bond, the local geometry of all candidates was studied. Five different parameters (Derewenda *et al.*, 1994) were calculated for every contact: donor–acceptor distance C···O (d), hydrogen-bond distance C—H···O (dh), C—H···O angle (ζ), H···O—C angle (ξ) and dihedral angle H···O—C—N (χ).

16 hydrogen bonds of the type C—H···O with d < 3.5 Å, dh < 2.6 Å and χ < 30° were found in DTBS (Table 7). The average distance between the Cα H atom and the carbonyl O atom is 2.42 Å, about 0.4 Å longer than a N—H···O bond. Nevertheless, this distance is significantly shorter than the sum of the van der Waals radii of Cα, H and O (2.7 Å). The longer distance suggests, however, that the energy contribution of such bonds is considerably smaller than hydrogen bonds involving polar atoms. The calculated energy of these bonds ranged from 0.5 to 2.0 kcal mol^{-1} (Umeyama & Marakuma, 1977).

3.6.1. Main-chain–main-chain C—H···O bonds. Cα—H···O hydrogen bonds are typical for parallel β-sheets with

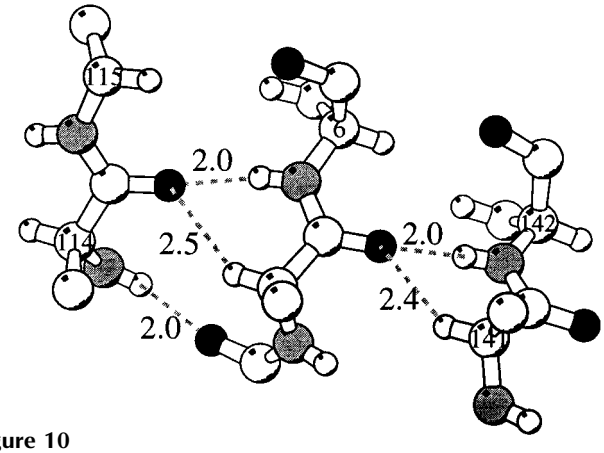


Figure 10
Main chain Cα—H···O hydrogen bonds in the central parallel β-sheet of DTBS. Only Cβ atoms of side chains are drawn. Residue numbers are shown on Cα atoms.

the backbone carbonyl O atom involved in a bifurcated hydrogen-bonding arrangement (Derewenda *et al.*, 1995; Fabiola *et al.*, 1997). DTBS contains a seven-stranded parallel β-sheet containing 34 N—H···O hydrogen bonds. 13 additional Cα—H···O hydrogen bonds involving residues from the β-sheet were found (Table 7). In most cases, one backbone O atom is involved in two hydrogen bonds: a bond to the NH group of residue i and a hydrogen bond with the H atom of the Cα atom of residue $(i-1)$ (Fig. 10). Not all carbonyl O atoms and not all Cα atoms from the central sheet form this type of bonds. No such close contacts were found between strands β2 and β3 on the edge of the β-sheet, but there are two to three Cα—H···O bonds between all other strands including β7 (at the other edge of the sheet).

The arrangement of atoms forming Cα—H···O bonds can differ from that shown in Fig. 10. When the residues participating in a Cα—H···O hydrogen bond are not located in the middle of the strand, their backbone O atoms can form bifurcated hydrogen bonds, but the second hydrogen bond is formed with water rather than with an amide group. For instance, the carbonyl O atom of residue 176 is located just outside strand β6 and forms a hydrogen bond to a water molecule (d = 2.88 Å) and to the Cα—H of Trp205 adjacent to strand β7. Tyr218 does not belong to the β-sheet region; however, its carbonyl O atom forms a hydrogen bond with the Cα—H of Gly201. In addition, a well defined water molecule is found at the distance of 2.88 Å from the carbonyl O atom of Tyr218. Thus, in all cases where backbone O atoms are engaged in a Cα—H···O hydrogen bond they also form a stronger interaction with another atom.

Five out of seven valine residues found in the β-sheet region in DTBS form Cα—H···O bonds, consistent with the observed preference of this residue to form such bonds (Fabiola *et al.*, 1997). However, unlike other proteins analyzed in the study of Fabiola and coworkers, in DTBS glycine residues were found to participate frequently in Cα—H···O hydrogen bonds. All four glycine residues which are located in the β-sheet form hydrogen bonds with carbonyl O atoms (Table 7).

3.6.2. C—H···O bonds with side chains. Carbonyl O atoms can form close interactions with other carbon hydrogen-bond donors, not just Cα atoms. One main-chain–side-chain and two side-chain–side-chain hydrogen bonds of type C—H···O were found (Table 7). Again, oxygen forms two hydrogen bonds, a strong bond with some polar hydrogen donor and a weaker one with C—H. One out of the four histidine residues in DTBS forms hydrogen bonds *via* its $C^{\delta2}$ atom. A close contact of an imidazole C atom with an acidic residue was observed between His83 and Asp126 (Fig. 9). Three atoms of the imidazole ring are involved in hydrogen bonding. Atom $N^{\delta1}$ is located at a distance of 2.78 Å from Ser81 O^{γ}. $N^{\delta1}$ acts as a hydrogen donor and the hydrogen of the hydroxyl group of Ser81 interacts with a well defined water molecule. The position of this H atom was confirmed by a difference hydrogen-omit map. His83 $N^{\varepsilon2}$ is at 2.78 Å distance from Thr190 $O^{\gamma1}$ from the second subunit of the dimer. These two hydrogen bonds define the position and orientation of the imidazole ring. The distance between the $C^{\delta2}$ H atom of His83 and Asp126 $O^{\delta1}$ is 2.42 Å, which is significantly less than the sum of their van der Waals radii (2.7 Å). In addition, Asp126 $O^{\delta1}$ lies in the plane of the imidazole ring and is thus perfectly positioned for hydrogen bonding.

Formation of C—H···O hydrogen bonds from side-chain atoms is not the exclusive property of the imidazole ring. Another close contact with hydrogen-bond features was found between the hydroxyl O atom of Tyr74 and the Cβ—H group of Val97. Both residues are well defined in the $3F_o - 2F_c$ electron density and positions of many H atoms are defined in the hydrogen omit $F_o - 2F_c$ map (Fig. 2*c*). The hydroxyl H atom points towards a water molecule (not shown in Fig. 2*c*) and the hydroxyl O atom forms a C—H···O bond with the Cβ—H of Val97.

3.7. Solvent structure

The crystal of DTBS contains 39% solvent. The cell volume is 2×10^5 Å^3, which translates into 870–900 water molecules per asymmetric unit of the DTBS crystal, if we assume that the solvent in the crystal has the same density as in the box containing 125 water molecules equilibrated by Monte Carlo as supplied in *X-PLOR* for molecular-dynamics calculations. In total, 436 ordered water sites were located in DTBS, which is about 47–49% of the whole solvent content of the crystal. There are 99 water sites with partial occupancies: five of them are in special positions and 68 sites form pairs with a distance less than 2.0 Å, indicating alternate water networks in the crystal. The other solvent sites with partial occupancy are found close to side chains with alternative conformation.

The average B factor for all 436 water sites is 27 Å^2. There are five water molecules with $B < 9$ Å^2. Inspection of their local geometry and surroundings suggests that they cannot be ions such as chloride or Mg^{2+}. All water molecules with low temperature factors are in contact with the rigid part of the protein and form at least one hydrogen bond with a backbone O atom. 233 water molecules have at least one hydrogen bond with a protein atom, forming the first hydration shell around the protein molecule. 33 water molecules form more than two contacts with protein atoms and their mobility is reduced ($\langle B \rangle$ for these water molecules is 20 Å^2). 184 water molecules have no contacts with protein, but have at least one contact shorter than 3.2 Å with another water molecule. This set of solvent molecules represents the ordered part of the second hydration shell. The mobility of the water molecules from the first and second hydration shells differs, with mean temperature factors of 25 and 30 Å^2, respectively.

The ordered water molecules do not cover the whole surface of the protein. The surface area of the monomer, calculated with a probe radius of 1.4 Å, is 10300 Å^2. Part of this area, 1570 Å^2 (15% of the whole surface of the monomer), participates in the dimer-interface region and crystal contacts. Of the remaining accessible surface area, 3060 Å^2 (35%) is not covered by ordered water molecules and is thus accessible to bulk solvent.

Contrary to what might have been expected, the proportion of the protein surface not covered by ordered water is about the same for the hydrophobic and hydrophilic regions of the protein surface. A calculation taking into account crystal-packing interactions showed that 64% of the hydrophobic and 68% of the polar accessible surface area is covered by ordered water molecules. In particular, this means that polar regions are not fully covered with ordered water molecules: for example, only half of the accessible surface area of carboxyl groups is covered by ordered water molecules, with the remaining area accessible to bulk solvent

The characteristics of the protein–solvent hydrogen bonds are shown in Table 6. Most of the water molecules form hydrogen bonds (a total of 164) with main-chain atoms. Backbone O atoms participate in twice as many bonds with water as backbone N atoms. This is consistent with the fact that backbone O atoms are in general more accessible to the solvent than backbone N atoms (the sums of the accessible surface area calculated by *X-PLOR* with probe radius 1.4 Å for all backbone O atoms and N atoms/amide H atoms are 975 and 35 Å^2, respectively). The mean distance between water and backbone O atoms is almost 0.2 Å smaller than between water and peptide N atoms (Table 6).

There is a strong correlation between the mobility of a protein side chain and the number of ordered water molecules making contact. In general, aspartate residues have a rather low side-chain *B* factor, $\langle B \rangle$ = 10 Å^2, and they bind 3–5 water molecules on average. The mobility of glutamates is higher, $\langle B \rangle$ for their side chains being about 30 Å^2. On average there are 1–2 water molecules around carboxyl groups of glutamates. This number can increase if additional hydrogen bonds or participation in a salt bridges lowers the mobility of the glutamate side chain. For instance, Glu94 O$^{\varepsilon1}$ forms a hydrogen bond with main-chain atoms, resulting in decreased mobility of the side chain as reflected by the low temperature factor, $\langle B \rangle$ = 12 Å^2. Three hydrogen-bonded water molecules were found close to the O$^{\varepsilon2}$ atom of this residue.

Two arginines are present in DTBS with their side chains not involved in hydrogen bonds or salt bridges to other protein atoms. Three bonded water molecules are at the N$^{\eta1}$ positions and two waters are at the N$^{\eta2}$ positions. This fully hydrated arginine shows a pattern of hydrogen bonds very similar to that found in in a survey of 16 other proteins (Thanki *et al.*, 1988). All exposed lysines (not involved in hydrogen bonds or salt bridges with protein atoms) interact with 2–3 water molecules.

3.8. Comparison of the DTBS structure at room temperature and at low temperature

A number of crystal structures of DTBS have been determined – these include the free enzyme (Huang *et al.*, 1994; Alexeev *et al.*, 1994) and various complexes with substrates, analogues and reaction intermediates (Huang *et al.*, 1995; Alexeev *et al.*, 1995; Käck *et al.*, 1998). Of particular interest is the structure of the unliganded enzyme determined at room temperature at 1.65 Å resolution, using crystals obtained under the same conditions as in this study. In the following, this coordinate set (PDB code 1DTS) is used in the comparison, unless otherwise stated.

Flash-freezing of DTBS crystals resulted in a shrinkage of the volume of the unit cell by about 5%. Calculation of the volume of the protein globules in the two crystal forms using *VOIDOO* (Kleywegt & Jones, 1994) showed that the difference in volume does not exceed 0.5%, hence freezing of the crystals does not result in a shrinkage of the globules, but the decrease in unit-cell parameters mostly arises from a reduced distance between protein molecules. DTBS forms most crystal contacts with one symmetry-related molecule, forming the active dimer. However, the position of the subunits in the dimer does not change very much with temperature, but the dimers change their relative position during the temperature jump. This is possible as there are only a few contacts between different dimers.

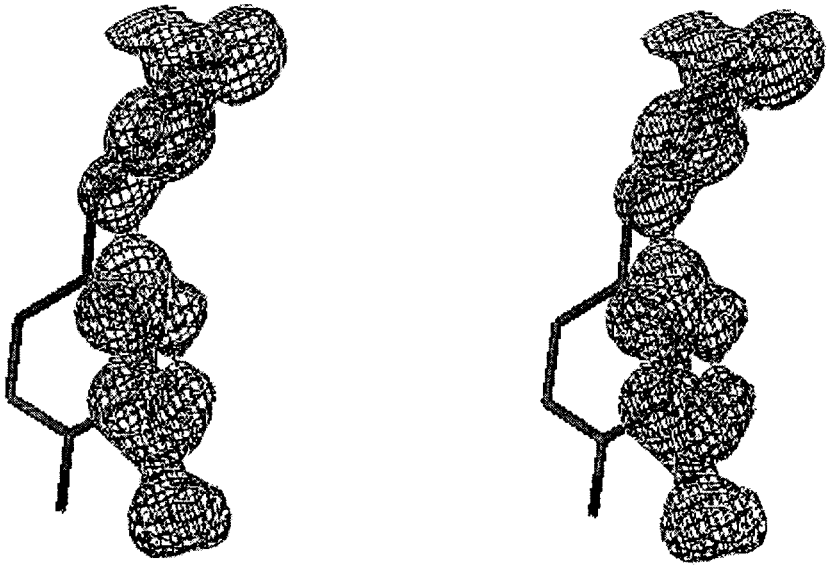


Figure 11
Stereoview of the conformational change of the side chain of Tyr187 in DTBS upon flash-freezing. The 298 K model of DTBS is shown together with a $3F_o - 2F_c$ electron-density map (contoured at 1.5σ) calculated using data collected at 100 K.

The most obvious difference between the 100 and 298 K structures is the difference in the thermal motion of the protein. The average temperature factors for all non-H atoms are 27.4 and 11.5 Å^2 for the 298 and 100 K models, respectively. The comparison of the *B* factor as a function of residue number (Fig. 1*c*) clearly shows that the regions that are flexible in the room-temperature structure remain rather mobile even at low temperature but the amplitude of the motion is lower.

The model of DTBS refined at room temperature (Huang *et al.*, 1994) contains 189 water molecules, whereas the 100 K structure contains 436 solvent sites. The comparison of the solvent structure in both models, carried out using *LSQMAN* (Kleywegt & Jones, 1997),

Table 8
Amino-acid residues with different conformations at 100 and 298 K in DTBS crystals.

Residue	Position	Temperature (K)	χ_1	χ_2	χ_3	χ_4	B_{mc}	B_{sc}
Arg32	α1-β2	100	−64	162	50	93	9.0	27.4
		298	−64	−179	−104	90	31.0	44.3
Lys45	β2a	100	67	−170	−172	20	14.0	30.0
		298	−60	85	168	68	35.1	45.0
Arg51		100	−64	−70	177	−97	9.0	14.9
		298	−54	−77	113	133	27.5	36.5
Gln65	α2-α2a	100	−168	165	−104		13.7	17.3
		298	51	−55	−40		33.6	54.7
Glu94	α3a-α3	100	167	57	65		8.0	12.7
		298	−152	64	40		22.7	47.5
Gln107	α3-β4	100	−62	169	14		17.2	35.6
		298	74	−150	−155		37.2	46.5
Gln108	α3-β4	100	17	−152	−30		16.1	35.6
		298	−47	174	−111		34.7	50.7
Gln136	α4	100	−68	169	1		9.7	17.9
		298	−174	−155	−48		28.8	38.5
Gln160	α5	100	178	169	−50		7.9	11.7
		298	−107	−139	38		23.3	45.0
Tyr187		100	−83	77			9.1	15.5
		298	−69	−18			18.5	25.3

showed that 136 water molecules are within 1.5 Å from each other, with an r.m.s. of 0.7 Å. In general, these matched water molecules have lower *B* factors in both structures, the average *B* factors for these solvent molecules being 20.6 and 43 Å^2 at 100 and 298 K, respectively; for non-matched solvent molecules the corresponding *B* factors are 30.0 and 50.2 Å^2, respectively.

The superposition of the two protein models gives an r.m.s. for all Cα atoms of 0.48 Å, which is comparable to those found for other proteins with known structure at low and room temperatures. These differences are not evenly distributed along the polypeptide chain (Fig. 1*c*). The largest movements occur at the N-terminal residue Ser1 (3 Å between Cα atoms) and three flexible loops which cover the active site of DTBS. These regions are loop 1, the phosphate-binding loop comprising residues 8–15, loop 5 (residues 147–150) and loop 7 (residues 206–214). Loop 7 is very flexible at room temperature; four residues (207–211) from this loop had no electron density and were therefore not included in the model (Huang *et al.*, 1994). At 100 K, the electron density for these residue is more defined and the loop can be modeled, albeit with difficulties. This is reflected by the temperature factors – the highest *B* factors of the whole structure are found for this region of the protein. One residue, 207, is still not defined in electron density even at this resolution. This loop is well defined in complexes of DTBS with substrates and substrate analogues (Huang *et al.*, 1995).

Two other loops, loop 1 and loop 5, have well defined main-chain atoms in both structures and the difference in their conformation is very clear. They do not directly participate in crystal contacts, although they are rather close to the dimer interface. Both loops change their position during binding of substrates. From a biochemical point of view, the most interesting difference between the room-temperature and low-temperature structures occurs in loop 5. The conformation of this loop in the low-temperature model of DTBS is much closer to that found in the ternary complex with the non-hydrolysable ATP analogue AMPPCP and DAPA (Huang *et al.*, 1995) than that observed in the room-temperature structure. For instance, the distance between the Cα atoms of Leu148 is 2.2 Å in the superposed models of the unliganded enzyme, whereas it is 0.2 Å in the superposed models of the low-temperature unliganded enzyme and the 298 K model of the complex. Loop 5 is part of the active site; amide groups of two residues from this loop (NH150 and NH153) are directly involved in binding of the carboxyl group of the DAPA substrate (Huang *et al.*, 1995).

There is one more difference in the conformation of the residues participating in the binding of the substrate. The hydroxyl group of Tyr187 forms a hydrogen bond with one of the carboxyl groups of the DAPA substrate (Huang *et al.*, 1995). The distance between the Cα positions of Tyr187 in the 100 and 298 K models is only 0.47 Å. However, the side chain of Tyr187 rotates by about 90° with respect to its position in the room-temperature structure (Fig. 11).

In addition to Tyr187, several other residues have a different conformation in the two models (Table 8). All of these residues are polar and exposed to the solvent; however, they are well defined in the electron density. The location of these residues along the chain is shown in Fig. 1(*c*). It can be seen that they are located rather far from the crystal contacts.

Flash-freezing is increasingly used as a means to trap catalytic intermediates in the crystal (Moffat & Henderson, 1995). One of the underlying assumptions is that conformational changes which might produce a species not of catalytic relevance are not likely to occur because of the rapid temperature jump. The comparison of the structures of DTBS determined at 100 and 298 K shows, however, that during flash-freezing significant conformational changes at the active site of the enzyme did occur. This result seems to challenge the underlying assumption of this approach in kinetic crystallography. Because of the possible implications for mechanistic enzymology, such comparisons must therefore be extended to complexes of the enzyme with substrate and substrate analogues analyzed both at room temperature and under cryo-conditions.

We thank Dr Katharine Gibson for the gift of purified enzyme. We are grateful for access to synchrotron radiation at beamline X11, EMBL outstation, c/o DESY, Hamburg and thank the staff, in particular Dr Victor Lamzin, for helpful discussions during data collection. This work was supported by the Swedish Natural Science Research Council and the Swedish Agricultural Research Council.

References

Alexeev, D., Baxter, R. L. & Sawyer, L. (1994). *Structure*, **2**, 1061–1072.
Alexeev, D., Baxter, R. L., Smekal, O. & Sawyer, L. (1995). *Structure*, **3**, 1207–15.
Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1995). *Curr. Opin. Struct. Biol.* **5**, 784–790.
Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1997). *Curr. Opin. Struct. Biol.* **7**, 681–688.
Dauter, Z., Sieker, L. C. & Wilson, K. S. (1992). *Acta Cryst.* B**48**, 42–59.
Derewenda, Z. S., Derewenda, U. & Kobos, P. M. (1994). *J. Mol. Biol.* **241**, 83–93.
Derewenda, Z. S., Lee, L. & Derewenda, U. (1995). *J. Mol. Biol.* **252**, 248–262.
Derrick, J. P & Wigley, D. B. (1994). *J. Mol. Biol.* **243**, 906–918.
Doig, A. J., MacArthur, M. W., Stapley, B. J. & Thornton, J. M. (1997). *Protein Sci.* **6**, 147–155.
Eisenberg, M. A. & Krell, K. (1979). *Methods Enzymol.* **62**, 348–352.
EU 3-D Validation Network (1998). *J. Mol. Biol.* **276**, 417–436.
Fabiola, G. F, Krishnaswamy, S., Nagarajan, V. & Pattabhi, V. (1997). *Acta Cryst.* D**53**, 316–320.
Harata, K., Abe, Y., & Muraki, M. (1998). *Proteins*, **30**, 232–243.
Huang, W., Jia, J., Gibson, K. J., Taylor, W. S., Rendina, A. R., Schneider, G. & Lindqvist Y. (1995). *Biochemistry*, **34**, 10985–10995.
Huang, W., Lindqvist, Y., Schneider, G., Gibson, K. J., Flint, D. & Lorimer, G. (1994). *Structure*, **2**, 407–414.
Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabsch, W. & Sander, C. (1984). *Biopolymers*, **22**, 2577–2637.
Käck, H., Gibson, K. J., Lindqvist, Y. & Schneider, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 5495–5500.
Kleywegt, G. J. & Jones, T. A. (1994). *Acta Cryst.* D**50**, 178–185.
Kleywegt, G. J. & Jones, T. A. (1997). *Methods. Enzymol.* **277**, 525–545.
Kumaraswamy, V. S., Lindley, P. F., Slingsby, C. & Glover, I. D. (1996). *Acta Cryst.* D**52**, 611–622.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. & Cambillau, C. (1997). *J. Mol. Biol.* **268**, 779–799.
Luzzati, V. (1952). *Acta Cryst.* **5**, 801–810.
McArdle, P. (1995). *J. Appl. Cryst.* **28**, 65.
MacArthur, M. W. & Thornton, J. M. (1996). *J. Mol. Biol.* **264**, 1180–1195.
Min, W., Dunn, A. J. & Jones, D. H. (1992). *EMBO J.* **11**, 1303–1307.
Moffat, K. & Henderson, R. (1995). *Curr. Opin. Struct. Biol.* **5**, 656–663.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–63. Warrington: Daresbury Laboratory,.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Schneider, G. & Lindqvist, Y. (1997). *Methods Enzymol.* **279**, 376–85.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. (1997). *SHELX*97 *Manual.* University of Göttingen, Germany.
Sevcik, J., Dauter, Z., Lamzin, V. S. & Wilson, K. S. (1996). *Acta Cryst.* D**52**, 327–344.
Stec, B., Zhou, R. & Teeter, M. M. (1995). *Acta Cryst.* D**51**, 663–681.
Thanki, N., Thornton, J. M. & Goodfellow, J. M. (1988). *J. Mol. Biol.* **202**, 637–657.
Umeyama, H. & Marakuma, K. (1977). *J. Am. Chem. Soc.* **99**, 1316–1332.
Yang, G., Sandalova, T., Lohman, K., Lindqvist, Y. & Rendina, A. R. (1997). *Biochemistry*, **36**, 4751–4760.