Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Refined structure of *Pyrococcus furiosus* ornithine carbamoyltransferase at 1.87 A

Jan Massant,[a]* Johan Wouters[b] and Nicolas Glansdorff[a]

[a]Department of Microbiology, Vrije Universiteit Brussel, B-1070 Brussels, Belgium, and [b]J. M. Wiame Research Institute for Microbiology, B-1070 Brussels, Belgium

Correspondence e-mail: jan.massant@vub.ac.be

Received 9 July 2003
Accepted 1 September 2003

PDB Reference: ornithine carbamoyltransferase, 1pvv, r1pvvsf.

Using synchrotron radiation, X-ray data have been collected from *Pyrococcus furiosus* ornithine carbamoyltransferase (*Pfu* OTCase) to a maximal resolution of 1.87 Å, allowing the refinement of a previous structure at 2.7 Å [Villeret *et al.* (1998), *Proc. Natl Acad. Sci. USA*, **95**, 2801–2806]. Thanks to the high resolution of this refined structure, two sulfate ions and 191 water molecules could be localized directly from the electron-density maps. The identification of these molecules allowed a more rigorous description of the active site and the identification of residues involved in binding carbamoyl phosphate. The improved quality of the model resulted in a better definition of several loops and the various interfaces. The dodecameric protein is composed of four catalytic trimers disposed in a tetrahedral manner. The extreme thermal stability of *Pfu* OTCase is mainly the result of the strengthening of the intersubunit interactions in a trimer and oligomerization of the trimers into a dodecamer. Interfaces between monomers in a catalytic trimer are characterized by an increase in ion-pair networks compared with mesophilic OTCases. However, the interfaces between catalytic trimers in the dodecameric oligomer are mainly hydrophobic and also involve aromatic–aromatic and cation–π interactions.

1. Introduction

Anabolic ornithine carbamoyltransferase (OTCase) catalyzes the carbamoylation of L-ornithine to form citrulline in the sixth step of the arginine biosynthetic pathway or in the first reaction of the urea cycle. Catabolic ornithine carbamoyltransferases are found when the arginine deiminase pathway is present, as in *Pseudomonas aeruginosa* (Cunin *et al.*, 1986); they catalyze the formation of carbamoyl phosphate (CP) and L-ornithine from citrulline and phosphate.

Most anabolic OTCases are trimers of about 105 kDa displaying Michaelis–Menten kinetics. *P. aeruginosa* catabolic OTCase, however, is a protein of 456 kDa composed of four trimers disposed in a tetrahedral manner (Villeret *et al.*, 1995). The allosteric properties of this catabolic OTCase can be correlated with its highly symmetrical oligomeric structure (Nguyen *et al.*, 1996). The anabolic OTCase of the hyperthermophilic archaeon *P. furiosus* (*Pfu* OTCase) is also a dodecamer composed of four catalytic trimers, but displays Michaelis–Menten kinetics like the other anabolic OTCases (Legrain *et al.*, 1997).

The structure of *Pfu* OTCase has been determined at 2.7 Å (Villeret *et al.*, 1998). The structures of several other OTCases have also been solved: unliganded *Escherichia coli* OTCase at 2.8 Å (Jin *et al.*, 1997), *E. coli* OTCase in complex with the bisubstrate analogue δ-*N*-phosphonacetyl-L-ornithine (PALO) at 2.8 Å (Ha *et al.*, 1997), *P. aeruginosa* catabolic OTCase at

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

3.0 Å (Villeret *et al.*, 1995), human OTCase complexed with PALO at 1.85 Å (Shi *et al.*, 1998), with CP plus L-norvaline, an analogue of L-ornithine, at 1.9 Å (Shi *et al.*, 2000) and with CP alone at 2.4 and 2.6 Å (Shi *et al.*, 2001).

In all these enzymes, the monomers and trimers have similar structures. The polypeptide chain folds into an N-terminal CP-binding domain and a C-terminal ornithine-binding domain. Each domain consists of a central parallel β-pleated sheet, with α-helices and loops flanking this central motif. All OTCases form similar catalytic trimers, with the active sites shared between monomers. The active site is located in the cleft between the two domains and contains additional residues from an adjacent chain.

Comparison of the structures of human OTCase complexed with different substrates or substrate analogues revealed that binding of the first substrate, CP, induces a global conformational change involving relative domain movement, whereas the binding of the second substrate, L-ornithine, brings the flexible SMG-loop, containing a conserved Ser-Met-Gly motif in all OTCases, into the active site (Shi *et al.*, 2001).

Extensive experimental evidence on hyperthermophilic proteins has led to the conclusion that no single mechanism is responsible for their remarkable stability (Vieille & Zeikus, 2001). In many hyperthermophilic proteins, intersubunit interactions and oligomerization appear to be potential major stabilization mechanisms. For example, 3-isopropylmalate dehydrogenase from *Thermus thermophilus* shows a more hydrophobic subunit interface than its mesophilic counterparts (Kirino *et al.*, 1994). Oligomerization was found to be a prerequisite for the activity of *Methanopyrus kandleri* formylmethanofuran:tetrahydromethanopterin formyltransferase; in this enzyme, the subunit interfaces are mostly hydrophobic (Shima *et al.*, 1998). On the other hand, the hexameric glutamate dehydrogenase from *P. furiosus* has achieved hyperthermostability by the formation of extensive ion-pair networks between its different subunits (Vetriani *et al.*, 1998). A genomic study and analysis of the solvent-accessible surface of a large set of protein structures suggest that the replacement of polar non-charged surface residues by charged residues constitutes a major stabilization mechanism in proteins from hyperthermophilic organisms (Cambillau & Claverie, 2000). Based on the analysis of several structural parameters of high-quality structures of thermophilic proteins and their mesophilic homologues, an increase in the number of ion pairs could be correlated with increasing growth temperature (Szilágyi & Závodszky, 2000). Ionic interactions seem to become more important at extreme high temperatures (Kumar & Nussinov, 2001). In *Pfu* OTCase, a predominantly hydrophobic interface around each threefold symmetry axis at the four tops of the tetrahedral oligomer (Villeret *et al.*, 1998) was shown to play a major role in the quaternary structure and thermostability of the protein (Clantin *et al.*, 2001).

Pfu OTCase was also shown to interact physically *in vitro* and *in vivo* with the carbamate kinase-like carbamoyl phosphate synthetase (CKase) from *P. furiosus* (Massant *et al.*, 2002). Indeed, *Pfu* CKase and *Pfu* OTCase form an efficient channelling cluster for CP, an extremely thermolabile and potentially toxic metabolic intermediate (Legrain *et al.*, 1995). Molecular analysis of the protein–protein interactions assembling this channeling complex requires a detailed structural study of the two enzymes. A high-resolution structure of the CKase from *P. furiosus* is already available (Ramón-Maiques *et al.*, 2000). Here, we report the high-resolution refinement of the *Pfu* OTCase structure to 1.87 Å.

Table 1
Crystal characteristics and data-collection statistics.

Values in parentheses are for the highest resolution shell.

	Pfu OTCase (Villeret *et al.*, 1998)	*Pfu* OTCase (this work)
Crystallization solution	0.08 *M* Na acetate pH 4.6, 1 *M* NaCl	0.08 *M* Na acetate pH 4.6, 2.4 *M* ammonium sulfate, 20% glycerol
Protein concentration	10 mg ml^{-1} Pfu OTCase	8.3 mg ml^{-1} *Pfu* OTCase, 1.7 mg ml^{-1} *Pfu* CKase, (OTCase:CKase 1:1)
Unit-cell parameters (Å, °)	$a = b = c = 186.8$, $\alpha = \beta = \gamma = 90.00$	$a = b = c = 184.8$, $\alpha = \beta = \gamma = 90.00$
Space group	*F*23	*F*23
Molecules per AU	1	1
X-ray source	Nonius FR591	ESRF BM30A FIP beamline
Wavelength (Å)	1.54179	0.98900
Temperature	Room temperature	100 K
Resolution (Å)	30.0–2.7	99.00–1.87 (1.91–1.87)
No. of observations	184193	304009
$R_{\rm merge}$ (%)	6.4	3.3 (18.7)
No. unique reflections	14853	42928
Completeness (%)	99.3	98.5 (86.5)

2. Materials and methods

2.1. Purification

Pfu OTCase was purified from recombinant *Saccharomyces cerevisiae* cells as described in Legrain *et al.* (2001) using a first thermodenauration step followed by ion-exchange chromatography on Resource Q and Mono Q columns (Pharmacia-Amersham Biotech) and arginine–Sepharose chromatography for the last purification step.

2.2. Crystallization and X-ray data collection

Crystals were grown by the hanging-drop vapour-diffusion method. 5 µl of protein solution containing 8.3 mg ml^{-1} *Pfu* OTCase and 1.7 mg ml^{-1} *Pfu* CKase (1:1 *Pfu* OTCase:*Pfu* CKase) was mixed with 5 µl of crystallization solution containing 2.4 *M* ammonium sulfate, 0.08 *M* sodium acetate buffer pH 4.6 and 20% glycerol, *i.e.* condition 47 of the Crystal Screen Cryo Kit from Hampton Research (Table 1).

High-resolution X-ray diffraction intensity data were collected at 100 K at the FIP beamline BM30A at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France at 0.98900 Å wavelength using a MAR Research CCD detector. Data were processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997) and resulted in a 98.5% complete data set extending to 1.87 Å (Table1). The space group is *F*23, with unit-cell parameters

$a = b = c = 184.8$ Å. There is one monomer in the asymmetric unit (Fig. 1*a*) and four dodecamers in the unit cell.

2.3. Refinement procedure

Refinement procedures were initiated using the software package *SHELX*97 (Sheldrick, 1990; Sheldrick & Schneider, 1997). 5% of all measured reflections were randomly selected for use in the calculation of R_{free}. Initial phasing was started from the previously determined structure of *Pfu* OTCase (PDB code 1a1s; Villeret *et al.*, 1998). Several cycles of *SHELX* restrained conjugate-gradient least-squares refinement (CGLS) and manual modelling using *TURBO-FRODO* (Roussel & Cambillau, 1992) resulted in a first model. Further refinement, in particular of the 80s loop, the 240s loop (or SMG-loop) and the C-terminal residues, was performed with initial simulated annealing with *CNS* (Brünger *et al.*, 1998) followed by consecutive cycles of refinement with *SHELX* and model building with *XtalView*/*Xfit* (McRee, 1999) and *TURBO-FRODO*. However, residues 233–242 in the SMG-loop and the C-terminal residues 311–313 show very weak electron densities caused by the high flexibility of these structural elements. The average *B* factors for the main- and side-chain atoms have been calculated after omitting residues 233–242 and 311–313 from refinement of the model. A final R_{cryst} of 21.3% and R_{free} of 26.9% were obtained for the model containing all residues as well as for the model with residues 233–242 and 311–313 omitted. Two sulfate ions and 191 water molecules have been incorporated in the model. The refinement statistics of the final model are given in Table 2.

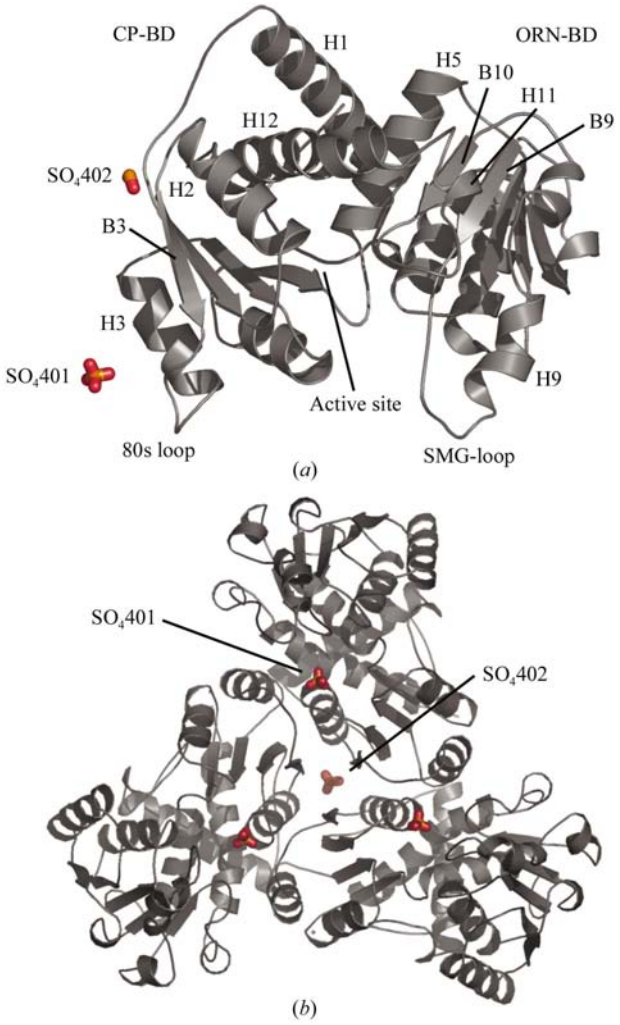

Figure 1
Cartoon representation of the monomeric and trimeric structure. (*a*) Structure of a monomer, with a CP-binding domain (CP-BD) and ornithine-binding domain (ORN-BD). The most important secondary structures are indicated: H, α-helices; B, β-sheets. The asymmetric unit contains one monomer and two sulfate ions. $SO_4$401 (S, O1, O2, O3, O4) is located in the active site of a neighbouring monomer. $SO_4$402 (S, O1, O2) is located along the threefold crystallographic symmetry axis at the interface between three monomers in a trimer. $SO_4$402 S and $SO_4$402 O1 occupy special positions on the axis, with occupancies of 0.33. Only one other O atom ($SO_4$402 O2) is in the asymmetric unit, generating two O atoms by crystallographic symmetry in the catalytic trimer. (*b*) View of a catalytic trimer along the threefold symmetry axis. Sulfate ions at the active sites ($SO_4$401) and on the threefold symmetry axis ($SO_4$402) are indicated.

Table 2
Model and refinement statistics.

Resolution range (Å)	99.00–1.87
No. of reflections	42928 [36075 $F_o > 4\sigma(F_o)$]
In working set	40775 [34266 $F_o > 4\sigma(F_o)$]
In test set (5%)	2153 [1809 $F_o > 4\sigma(F_o)$]
No. of amino-acid residues	313
No. of protein atoms	2454
No. of water molecules	191
No. of heteroatoms (sulfate ions)	8 (2)
R_{cryst} (%)	21.3 [19.5 for $F_o > 4\sigma(F_o)$]
R_{free} (%)	26.9 [24.9 for $F_o > 4\sigma(F_o)$]
Root-mean-square deviations	
Bond lengths (Å)	0.007
Bond angles (Å)	0.023
B factors† (Å^2)	
Average, main-chain atoms	29.58
Average, side-chain atoms	35.39
Solvent	41.50
Ramachandran plot statistics‡ (%)	
Most favoured regions	87.5 (245 residues)
Allowed regions	10.0 (28 residues)
Generously allowed	2.1 (6 residues)
Disallowed regions	0.4 (1 residue)

† Average *B* factors were calculated after omitting residues 233–243 (SMG-loop) and 311–313 (C-terminus) from the model refinement. ‡ Gly and Pro excluded.

2.4. Structure analysis

The quality of the model was checked with the program *PROCHECK* (Laskowski *et al.*, 1993). Hydrogen bonds between protein atoms were calculated with the HBPLUS routine (McDonald & Thornton, 1994) with the default parameters for distance and angles. The presence of salt bridges was inferred when Asp or Glu side-chain carbonyl O atoms were found to be within a 4.0 Å distance of the N atoms in Arg, Lys and His side chains. Superposition of C_α atoms was performed with the program *LSQMAN* (Kleywegt & Jones, 1994). Accessible surface areas were calculated with *NACCESS* (Hubbard & Thornton, 1993). Molecular graphics were drawn with *PyMOL* (DeLano, 2002).

2.5. Molecular-dynamics simulations

Energy minimizations (molecular dynamics and molecular mechanics) were performed with the *DISCOVER* program (Accelrys Inc., San Diego) using the cvff force field. Graphical displays were generated with the *INSIGHT*II molecular-modelling system (Accelrys Inc., San Diego). Computations were performed on a Silicon Graphics O2 workstation running Irix6. The three-dimensional coordinates of *Pfu* OTCase were taken from the refined crystallographic structure. H atoms were automatically assigned. Water molecules were removed and a distance-dependent (1.r) dielectric constant was used. All coordinates (main chain and side chain) were fixed except those of the SMG-loop (residues 228–250). A combination of molecular-mechanics runs (steepest descent plus conjugate gradient plus Newton–Raphson) was first applied to the structure in order to define a starting input for the molecular-dynamics simulations. Molecular dynamics then followed at 300 K for a total of 100 ps (20 runs of 5000 fs), generating 20 conformations. Those resulting conformations were further minimized (annealing to a final derivative cutoff of 0.2 kJ mol^{-1}) and analyzed.

3. Results and discussion

3.1. Overview

A previous 2.7 Å structure of *Pfu* OTCase was determined in a cubic space group (Villeret *et al.*, 1998). New crystals grown under a new condition in the presence of *Pfu* CKase (see §2) were shown to diffract in the same space group with the same unit-cell parameters. Subsequently, the same crystals were also obtained without *Pfu* CKase. Although a physical interaction between the *Pfu* CKase and the *Pfu* OTCase has already been demonstrated (Massant *et al.*, 2002), co-crystallization of both proteins has not yet been obtained. Crystal characteristics and data-collection statistics are given in Table 1. The new model could be refined to 1.87 Å with a crystallographic R factor of 21.3% [19.5% for $F_o > 4\sigma(F_o)$] and an $R_{\rm free}$ of 26.9% [24.9% for $F_o > 4\sigma(F_o)$] calculated on 5% of all the measured reflections. The loop comprising residues 233–242 (the SMG-loop) and the three C-terminal residues show very weak electron densities, explaining in part why a lower final R factor is not obtained. Model and refinement statistics are given in Table 2. The refined model had an acceptable overall geometry, with r.m.s. deviations for bond lengths and angle distances of 0.007 and 0.023 Å, respectively. The Ramachandran statistics were determined using the program *PROCHECK* (Laskowski *et al.*, 1993). 87.5% of the dihedral angles were found to be in the most favoured region, 10% in the additional allowed region, 2.1% in the generously allowed region and 0.4% (Leu270) in the disallowed region. Leu270 is part of the active site and is followed by Pro271, which may explain its unusual geometry. Most of the residues in the generously allowed region are residues from the SMG-loop. In the uncomplexed form of *Pfu* OTCase, the SMG-loop appears to be too flexible to observe electron densities for the residues that make up part of the loop, even at higher resolution. Molecular-dynamics simulations have confirmed the flexibility of this loop (data not shown). Residues 243–248 from the SMG loop, which show better defined electron

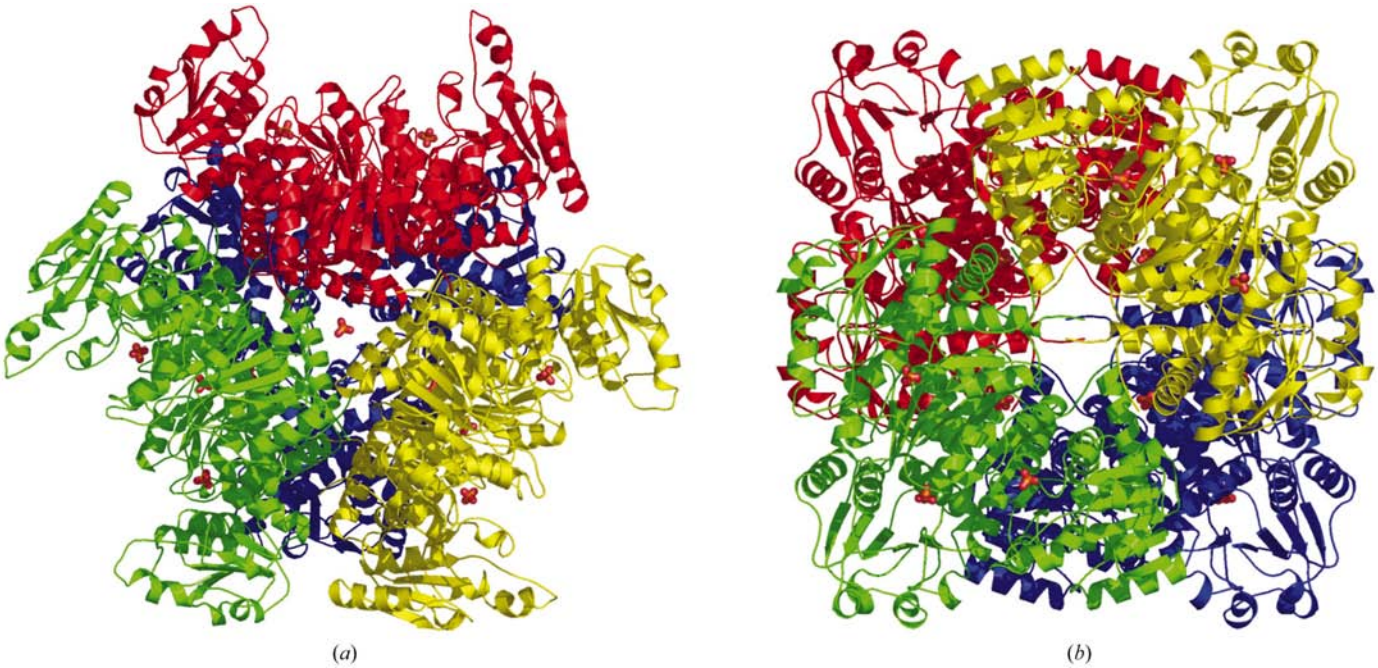

Figure 2
Cartoon representation of the dodecameric structure of *Pfu* OTCase. (*a*) View along the threefold symmetry axis at one top of the dodecamer. (*b*) View along a twofold symmetry axis. Each trimer is represented by a different colour. The structure can be viewed as a tetrahedron, each face of which is constituted of a trimer. Four threefold symmetry axes pass through the four tops of the tetrahedron and the centre of the opposite faces; three twofold symmetry axes pass through the middles of opposite edges. All sulfate ions are shown.

Table 3
Interactions at interfaces between subunits in a catalytic trimer.

Interactions between different monomers are in italic.

Refined *Pfu* OTCase		*E. coli* OTCase (PDB code 1duv)		Human OTCase (PDB code 1c9y)	
Interaction	Distance (Å)	Interaction	Distance (Å)	Interaction	Distance (Å)
(*a*) Ion-pair networks between two monomers in a trimer					
R95 NH2–D92 OD1	2.50	R94 NH2–D91 OD1	2.96		
R274^{viii} NE–D92 OD2	*3.09*	*H278H NE2–D91G OD2*	*2.62*		
R274^{viii} NH2–D92 OD1	*3.03*				
R274 NH1–D280 OD1	2.99	H278 ND1–D302 OD1	2.56	R306 NH1–D313 OD1	2.85
R274 NH1–D284 OD2	2.72				
R274 NH2–D284 OD1	2.81			R306 NH2–D313 OD1	2.94
R95 NH2–D284^{viii} OD1	*4.20*	*R94G NH2–E306H OE1*	*2.81*		
		R94G NH2–E306H OE2	*2.69*		
R95–W291^{viii}†		*R94G–F313H* †		*R129–F324^{viii}*†	
R99 NH1–D292^{viii} OD1	*2.74*	*R98G NH1–D314H OD1*	*3.18*		
R99 NH2–D292^{viii} OD1	*3.08*	*R98G NH2–D314H OD1*	*3.38*		
R99 NH1–E295^{viii} OE1	*2.86*	*R98G NH1–E317H OE1*	*2.68*		
H299 NE2–E295 OE2	2.80	H321 NE2–E317 OE2	2.77	W332 NE1–E328 OE2	2.89
K35 NZ–E295 OE2	2.79	K34 NZ–E317 OE2	2.86	K68 NZ–E328 OE2	2.93
K148 NZ–D292 OD2	2.91	H147 NE2–D314 OD2	2.78		
K148 NZ–E18^{vi} OE1‡	*2.57*				
(*b*) Ion-pair networks around the threefold symmetry axis					
H73 NE2–E64^{viii} OE2	*2.60*			*H107 NE2–E98^{viii} OE2*	2.80
R60 NE–E64 OE1	2.82	R59 NE–E63 OE2	2.75	R94 NE–E98 OE1	2.84
R60 NH1–E64 OE1	3.12	R59 NH2–E63 OE2	3.01	R94 NH2–E98 OE1	3.24
H73 ND1–$SO_4$402 O1	2.88				
H73 ND1–$SO_4$402 O2	3.29				
R60–Y76†		R59–Y75†		R94–F110†	
R41 NE–E44^{iv} OE2	*3.27*				
R41 NH2–E44^{iv} OE2	*3.76*				
R41 NH1–E44 OE1	2.52			K70 NZ–E72 OE2	3.12
R41 NH1–E44 OE27	3.17				
R41 NH2–H_2O1067	2.57				
(*c*) Interface at the active site					
S56 N–D81^{iv} O	2.74	S55G N–G80I O	2.73	S90 N–D115^{iv} O	2.84
T57 OG1–Q83^{iv} N	3.02	T56G OG1–Q82I N	3.21	T91 OG1–H117^{iv} N	3.07
R58 NH1–E88^{iv} OE2	*2.93*	*R57G NH1–E87I OE2*	*2.66*	*R92 NH1–E122^{iv} OE2*	*2.71*
R58 NH2–E88^{iv} OE1	*2.75*	*R57G NH2–E87I OE1*	*2.90*	*R92 NH2–E122^{iv} OE1*	*2.73*

Crystallographic symmetry codes: (iv) z, x, y; (vi) $-z, -x, y$; (viii) y, z, x. † Cation–π interaction. ‡ Interaction between two catalytic trimers in the pyrococcal dodecamer.

densities, are close to the same stretch of residues from a neighbouring enzyme in the unit cell.

A least-squares rigid-body superimposition of the current high-resolution structure and the previously reported structure of *Pfu* OTCase (PDB code 1a1s) was conducted with *LSQMAN* (Kleywegt & Jones, 1994). The r.m.s. deviation for 291 equivalent C^{α} atoms excluding residues 79–89 (80s loop) and residues 233–243 (part of the SMG-loop) was 0.38 Å. When comparing the main-chain trace of this refined structure with the previously solved structure, the main differences are observed in three loops: the loop between B2 and H2 (residues 53–58), the loop between B3 and H3 (residues 78–90) and the loop between B9 and H9 (residues 232–248). Residues in these loops are functionally important since they constitute the binding motifs for the substrates or are involved in interactions at the interfaces between monomers in a catalytic trimer. Binding of a sulfate ion at the active site has caused the conformational change observed in the first two loops. The difference in the main-chain trace for the SMG-loop is a consequence of the flexibility of this loop.

The current structure represents an increase in resolution of about 1 Å, which allowed a better assignment of side-chain orientation and description of the solvent model. Thanks to the high resolution of this refined structure, 191 water molecules and two sulfate ions could be localized from the electron-density maps. The identification of these molecules allows a more rigorous description of the active site and the different interfaces in this oligomeric structure; these interfaces play a key role in the extreme thermal stability of this hyperthermophilic protein. The enzyme is a dodecamer with exact 23 point-group symmetry (Villeret *et al.*, 1998). The oligomer is built from four catalytic trimers disposed in a tetrahedral manner, with the convex faces of the trimers pointing to the inner side (Fig. 2). The four threefold symmetry axes are perpendicular to the faces in their centre and the three twofold symmetry axes pass through the middle of opposite edges.

3.2. Structural integrity of the monomer

A polypeptide chain of 314 amino acids folds into a carbamoyl phosphate-binding or polar domain and an ornithine-binding or equatorial domain. Each domain is organized around a β-pleated sheet of five parallel strands

surrounded by α-helices, with helices H5 and H12 linking the two domains (Fig. 1*a*). Charged interactions stabilize the relative positions of helices H1, H5 and H12, which constitute an important structural motif around which the dodecamer is built. Lys31, Glu147 and Glu295 are involved in these interactions and are conserved in all known OTCases. These residues thus appear to be important for the structural integrity of the monomer. Glu295 (H12) forms salt bridges with Lys35 (H1) and His299 (H12); Asn296 (H12) interacts with Lys31 (H1) and Glu147 (H5). Arg7 (B1) forms a salt bridge with Glu19 (H1). Lys31 and Glu147 also interact with Val1 from an adjacent chain to stabilize the N-terminal residue. A cation–π interaction between Arg60 (H3) and Tyr76 (B3) and a salt bridge between Arg60 and Glu64 (H2) are also found in most of the other OTCases and are probably important for the stabilization of helix H2 (Table 3*b*; Fig. 3*a*). Residues from H2 and B3 are part of the active site or form the interface with an adjacent chain.

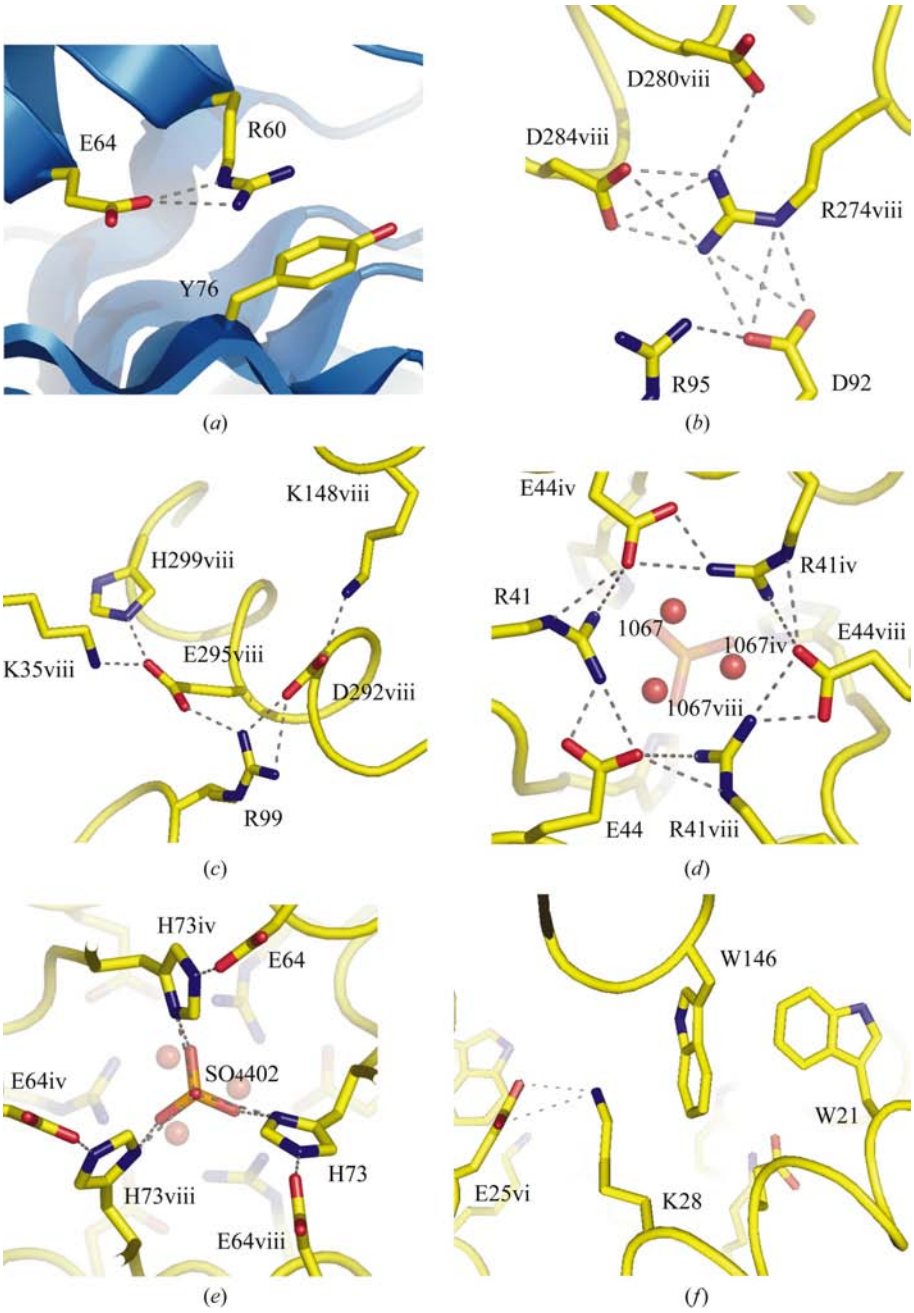

Figure 3
Interactions at and near the different interfaces. (*a*) Conserved salt bridge and cation–π interaction near the active site. (*b*) and (*c*) Ion-pair networks at the interface between two monomers in a catalytic trimer. (*d*) Ion-pair network in a catalytic trimer around the threefold symmetry axis. Three water molecules are displayed as 1067 and its symmetrical relatives. (*e*) Complexation of a sulfate ion on the threefold axis and salt bridges between monomers around this axis. (*f*) Aromatic–aromatic, cation–π and ion-pair interactions near the hydrophobic dodecameric interface. Crystallographic symmetry codes: (iv) z, x, y; (vi) $-z, -x, y$; (viii) y, z, x.

3.3. The catalytic trimer

The basic structure is a homotrimer similar to those found in other carbamoyltransferases, with one active site per monomer (Fig. 1*b*). The interfaces between monomers within catalytic trimers involve mainly the CP-binding domains and are characterized by ion-pair networks (Table 3). Helix H3 from one monomer is positioned in front of helix H12 of the second monomer, with the loop comprising residues 78–83 located near the catalytic site of the adjacent monomer. Several salt bridges are formed between the residues of these two helices: the first ion-pair network (Table 3*a*; Fig. 3*b*) involves Asp92 and Arg95 of H3 from one chain and Arg274 and Asp284 of H12 from the adjacent chain, while the second ion-pair network (Table 3*a*; Fig. 3*c*) consists of Arg99 (H3) and Asp292 and Glu295 (H12). A major ion-pair network is found at the centre of the catalytic trimer around the threefold symmetry axis. Arg41 and Glu44 (in the loop between helix H1 and sheet B2) from three monomers form an ion-pair network around this symmetry axis (Table 3*b*; Fig. 3*d*). The S atom and one O atom from $SO_4$402 are positioned along this symmetry axis. The sulfate ion is complexed by three histidine residues, His73 (B3) from three different chains (Table 3*b*; Fig. 3*e*). This His73 further forms a salt bridge with Glu64 (H2) from an adjacent chain. Interactions between the three chains in a catalytic trimer in the current model were compared with those in the mesophilic OTCases from *E. coli* complexed with PALO at 1.70 Å (PDB code 1duv) and from human complexed with CP and L-norvaline at 1.90 Å

(PDB code 1c9y) (Table 3). An increase in ion-pair networks between monomers in a catalytic trimer was observed in *Pfu* OTCase. Two of these networks can also be found in the *E. coli* OTCase structure but not in the human enzyme and one, involving His73, can also be found in the human OTCase structure but not in the *E. coli* enzyme. The ion-pair network involving residues Arg41 and Glu44 could not be found in either of the two mesophilic enzymes.

A final interface is formed between the N-terminal residues of helix H2 from one chain and residues Asp81 and Glu88 in the 80s loop from an adjacent chain (Table 3*c*). The salt bridge between Arg58 and Glu88 was not observed in the previous reported structure of *Pfu* OTCase (Villeret *et al.*, 1998). The interaction occurs close to the binding site of the phosphate group of CP and the loop comprising residue 88 undergoes a conformational change upon substrate binding. In the refined structure, this conformational change occurred upon binding of a sulfate ion at the CP-binding site. Arg58 and Glu88 are conserved in all OTCases and ATCases (Labédan *et al.*, 1999; Allewell *et al.*, 1999) and are thought to be important for maintaining the trimeric structure and catalytic activity. Several of the other residues at this interface are part of the active site.

Table 4
Interactions at the active site.

Human OTCase, complex with CP (PDB code 1ep9)			Refined *Pfu* OTCase		
Heteroatoms	Protein atoms	Distance (Å)	Heteroatoms	Protein atoms	Distance (Å)
CP355 O1P	H117B NE2	3.33	$SO_4$401 O4	Q83^{iv} NE2	2.79
	R141 NH1	2.71		H_2O1001	2.97
			H_2O1001	R107 NH2	3.09
CP355 O2P	T93 N	2.84	$SO_4$401 O3	T59 N	2.93
	T93 OG1	2.84		T59 OG1	3.57
	S90 OG	2.67		S56 OG	2.80
	R141 NH2	3.14			
CP355 O3P	T91 N	2.84	$SO_4$401 O2	T57 N	2.90
	R92 N	2.89			
	R92 NE	2.98		R58 NE	3.25
	R92 NH2	2.92		R58 NH2	2.87
CP355 O4P	R92 NE	2.83	$SO_4$401 O1	R58 NE	2.97
CP355 O	T93 OG1	2.92	H_2O1026	T59 OG1	
	H168 NE2	2.89		H134 NE2	3.19
	R141 NH2	2.92			
	R330 NH1	2.98		R297 NH2	2.79
CP355 N	Q171 OE1	2.80	H_2O1150	Q137 OE	2.68
	C303 O	2.96		C269 O	2.93
	L304 O	3.27		L270 O	3.18
	R330 NH1	3.23		R297 NH2	3.75
H_2O359	CP355 O4P	3.53	H_2O1003	$SO_4$401 O1	2.98
	CP355 O1P	3.04		$SO_4$401 O4	3.92
	L304 O	2.97		L270 O	2.73
	H_2O460	3.05		H_2O1005	3.69
H_2O460	C303 SG	3.16	H_2O1005	C269 SG	4.02
	H168 CD2	3.47			
	H_2O 439	2.93		H_2O1054	3.95
H_2O439	D263 OD2	2.81	H_2O1054	D229 OD2	2.70
	N199 OD1	3.09		N165 OD1	3.44

3.4. Active site and solvent network

The active site is located in a pocket between the two domains. It is delimited by helices H2, H12, H5 and the loop between strand B10 and helix H11. A sulfate-binding site was located in the active site of *Pfu* OTCase; it is similar to the CP-binding site in human OTCase complexed with CP (Shi *et al.*, 2001). Residues that are involved in binding of the sulfate ion are shown in Fig. 4 and are listed in Table 4. Ser56, Arg58 and Thr59 from one chain and Gln83 from an adjacent chain interact with the four O atoms of the sulfate ion. Ser56, Thr57, Arg58, Thr59 and Arg60 constitute the conserved binding motif for the phosphate moiety of CP. Gln83 is a histidine in the human enzyme. Arg107 does not interact directly with $SO_4$401 O4 but *via* a water molecule (H_2O1001). Not all residues involved in binding of the phosphate group of CP in the human structure are involved in binding of the sulfate ion in the pyrococcal structure. The conformational change after binding of the sulfate ion is probably an intermediate stage between the native structure and the CP binary structure.

Several water molecules were found in the active site that interact with the sulfate ion and residues known to be involved in the binding of CP and ornithine in *E. coli* and human OTCases (Ha *et al.*, 1997; Shi *et al.*, 1998, 2001). Water molecule 1150 interacts with O1 from the sulfate, with the side chains of His134 and Gln137 and the carbonyl O atoms of Cys269 and Leu270. In the human CP-complexed structure, these residues bind to the N atom of CP (Shi *et al.*, 2001). Leu270 O interacts with the NE of PALO in the bisubstrate complex of human OTCase (Shi *et al.*, 1998). Water molecule 1026 also interacts with O1 from the sulfate ion and with residues Thr59, His134 and Arg297. These residues bind to the oxygen of the carbamoyl moiety in the human binary structure (Shi *et al.*, 2001). The side chains of Asn165 and Asp229 interact with water molecules 1005 and 1054. Both residues are part of the ornithine-binding site in human OTCase. Water molecule 1005 is hydrogen bonded to water molecule 1003, which is itself hydrogen bonded to O1 of the sulfate

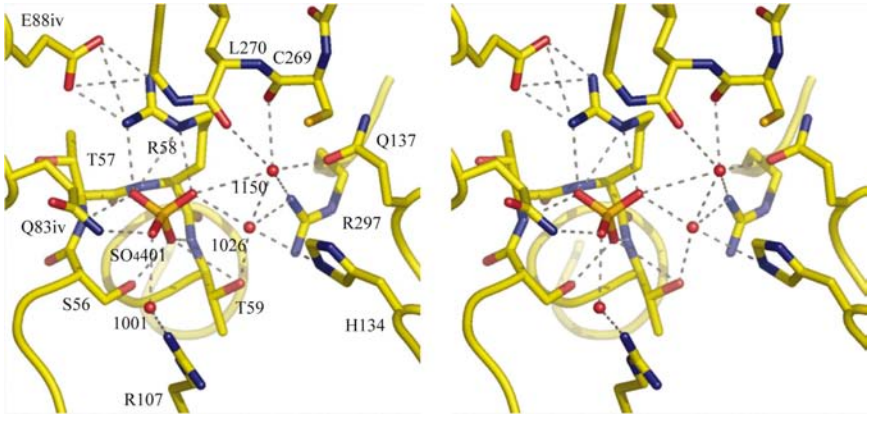

Figure 4
Stereoview of the binding of a sulfate ion at the CP-binding site. 1001, 1026 and 1150 are water molecules. Crystallographic symmetry code: (iv) *z*, *x*, *y*.

ion and the carbonyl O atom of Leu270. This solvent network appears to fill the CP and ornithine-binding sites. These last three water molecules, 1003, 1005 and 1054, are equivalent to water molecules 359, 439 and 460 found in the human OTCase complexed with CP (Shi *et al.*, 2001; PDB code 1ep9). In the structure of human OTCase complexed with PALO, 17 water molecules form a complex hydrogen-bonded network near the active site, with two water molecules directly involved in binding PALO (Shi *et al.*, 1998). The network around these two water molecules is thought to act as a proton-transfer pathway in the catalytic mechanism, whereas most of the remaining water molecules are thought to provide structural stabilization (Shi *et al.*, 1998). The available crystal structures indicate that the main binding site for ornithine is formed by the conserved D229*xxx*SMG motif in the SMG-loop. Superposition of the human OTCase structures complexed with CP and with PALO demonstrates that the SMG-loop shifts by more than 8.0 Å on binding the second substrate (Shi *et al.*, 2001). The movement of this loop seems to be essential for binding the second substrate and release of the products (Shi *et al.*, 1998). The conserved H268CLP probably stabilizes the tetrahedral intermediate rather than being essential for binding of ornithine (Shi *et al.*, 2001).

3.5. Dodecameric interfaces

The main interfaces between the trimers in the dodecameric assembly are located at the four tops of the tetrahedron (Fig. 2). Helices H1 from three monomers from different catalytic trimers face each other around the threefold symmetry axis and constitute a hydrophobic interface composed of residues Trp21, Met29, Ile32, Trp33 and Ile36 (Fig. 5). The side-chain C atom of Lys38 makes a cation–π interaction with the aromatic ring of Trp33. The ammonium group of this lysine is exposed to the solvent. Positioning the carbon of the lysine side chain closest to the ring may contribute favourable van der Waals binding and exposing the ammonium group may lead to better interactions with solvent or hydrogen-bonding groups (Gallivan & Dougherty, 1999). Ile32 forms a hydrophobic cluster with Met29 and Val1 from an adjacent chain. Trp33 is stabilized by hydrophobic interactions with His40 and Phe30. Trp21 is involved in an aromatic–aromatic interaction with Trp146 of helix H5 within the same monomer (Fig. 3*f*). Lys28 forms a cation–π interaction with Trp146. Lys28 and Glu25 from an adjacent chain form a salt bridge across the interface, thus resulting in three salt bridges around each threefold symmetry axis. The cluster of interactions formed by these residues plays a major role in the packing of the trimers into the dodecameric structure, an arrangement from which hyperthermostability emerges (Clantin *et al.*, 2001).

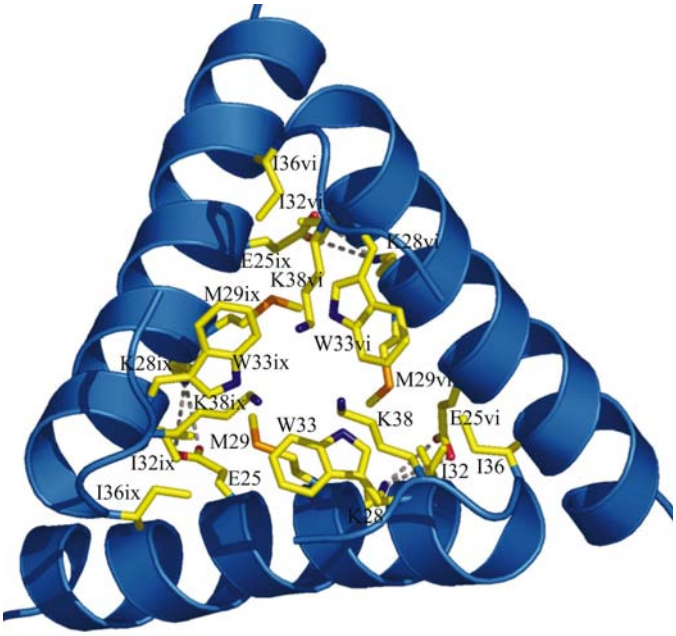

Figure 5
Interface between trimers located around a threefold symmetry axis at one top of the dodecamer. The three H1 helices from three monomers are shown in blue. The interface is mainly composed of hydrophobic residues: Met29, Ile32, Trp33 and Ile36. Lys38 makes a cation–π interaction with the aromatic ring of Trp33 and Glu25 forms a salt bridge with Lys28 from an adjacent chain. Crystallographic symmetry codes: (vi) $-z, -x, y$; ix, $-y, z, -x$.

Each monomer is also involved in an interface between two catalytic trimers that mainly involves the N- and C-terminal residues. In addition to the hydrophobic interaction of the side chain of Val1 with Ile32 from an adjacent chain, Val1 is also stabilized by hydrogen bonds with Lys31 and Glu147 from an adjacent chain. Lys31 interacts with Glu147 and Asn296. Glu18 forms a salt bridge with Lys148 across the interface; Lys148 forms a salt bridge with Asp292 from the same chain.

The C-terminal residues of two monomers belonging to different catalytic trimers face each other at the interface located around a twofold symmetry axis. They cover a 15 Å channel penetrating into the structure and joining the internal cavity at the centre of the protein (Villeret *et al.*, 1998). However, these C-terminal residues have a high flexibility.

3.6. Thermal stability

Villeret *et al.* (1998) and Clantin *et al.* (2001) already have shown that oligomerization plays a major role in the stabilization of this hyperthermophilic OTCase. A decrease in solvent-accessible surface area for *Pfu* OTCase can be observed compared with mesophilic OTCases (Table 5) and oligomerization in *Pfu* OTCase leads to an enlarged percentage of buried surface area. Refining this structure at higher resolution allowed a more rigorous description of the different interfaces and the interactions important for the extreme thermal stability of the protein. An increase in ion-pair networks compared with mesophilic OTCase was observed between monomers in a catalytic trimer. An increase in salt links at the interfaces between monomers in a catalytic trimer was also observed in the structure of the catalytic subunit of *P. abyssi* aspartate carbamoyltransferase (Van Boxstael *et al.*, 2003). Further stabilization of the *P. abyssi* ATCase is provided by the association with regulatory subunits, whereas *Pfu* OTCase is further stabilized by oligomerization of the

Table 5
Solvent-accessible surface area (ASA).

Comparison of the structure of *Pfu* OTCase with the structures of OTCase from *P. aeruginosa* (PDB code 1ort), *E. coli* (PDB code 1duv) and human (PDB code 1c9y). ASA was calculated with *NACCESS* (Hubbard & Thornton, 1993). Values in parentheses are percentages of the total surface area.

	Pfu OTC	1ort	1duv	1c9y
ASA (Å^2)	9326 (70.3)	11225 (72.4)	10884 (81.2)	11507 (83.1)
Hydrophobic	5279 (56.6)	6479 (57.7)	6395 (58.8)	6307 (54.8)
Polar	4047 (43.4)	4746 (42.3)	4489 (41.2)	5200 (45.2)
ASA buried in a trimer (Å^2)	2462 (18.5)	2232 (14.4)	2512 (18.8)	2334 (16.9)
Hydrophobic	1425 (57.9)	1315 (58.9)	1452 (57.8)	1522 (65.2)
Polar	1037 (42.1)	917 (41.1)	1060 (42.2)	812 (34.8)
ASA buried in a dodecamer (Å^2)	1490 (11.2)	2044 (13.2)		
Hydrophobic	1008 (67.7)	984 (48.1)		
Polar	482 (32.3)	1060 (51.9)		

catalytic trimers into a dodecamer. The interfaces between catalytic trimers are mainly hydrophobic (Table 4), but also involve aromatic–aromatic and cation–π interactions. Aromatic clusters (Kannan & Vishveshwara, 2000) and previously overlooked cation–π interactions are also thought to be possible stabilizing factors in thermophiles (Vieille & Zeikus, 2001; Chakravarty & Varadarajan, 2002; Gromiha *et al.*, 2002). Mutagenesis studies of Trp21, Glu25, Met29 and Trp33 showed their importance for the dodecameric structure and stability (Clantin *et al.*, 2001). The triple variant (E25Q/M29A/W33A) showed a very large decrease in thermal stability and was found to be a trimer. However, this trimeric variant was still as stable as the trimeric OTCase from the less extreme thermophile *Thermus thermophilus* (Sanchez *et al.*, 1997). This underlines the importance of the interactions between monomers in a trimer as a basic stabilization mechanism. In addition to the hydrophobic interfaces between catalytic trimers, the N-termini are also stabilized by interactions with neigbouring trimers. In *M. kandleri* methenyltetrahydromethanopterin cyclohydrolase not only are the hydrophobic interactions between subunits strengthened, but several loops as well as the N- and C-termini are also fixed by contacts to neighbouring subunits (Grabarse *et al.*, 1999).

The difficulty in identifying the factors responsible for increased thermal stability is illustrated by two low-temperature-adapted mutants of *Pfu* OTCase. The double mutants Y227C/E277G and A240D/E277G, which were selected *in vivo* by experimental evolution, show a dramatic decrease in thermal stability (Roovers *et al.*, 2001). In the absence of the structures of these mutant enzymes, the role of these residues remains uncertain.

4. Conclusion

Refinement of the *Pfu* OTCase structure to a resolution of 1.87 Å allowed a detailed description of the active site and the different interfaces in this dodecameric hyperthermophilic enzyme. The localization of a sulfate ion at the binding site of the phosphate group of CP led to the identification of the residues involved in binding the first substrate. However, compared with the human enzyme complexed with CP, not all the residues possibly involved in binding to CP were already in place in the sulfate-bonded structure, which appears to have a conformation intermediate between the structure of the unliganded enzyme and the CP binary complex. Since co-crystallization of *Pfu* OTCase and *Pfu* CKase has not yet been obtained, high-resolution structures of these enzymes in complex with different substrates or substrate analogues would be useful to investigate possible interactions between the two proteins.

Extreme thermal stability in this hyperthermophilic enzyme is achieved by the strengthening of subunit interactions and oligomerization of catalytic trimers into a dodecameric enzyme. At the interfaces between monomers in catalytic trimers an increase in ion-pair interactions appears to be the main stabilizing mechanism. Such a trimer is already as stable as the trimeric enzyme from a less extreme thermophilic bacterium. Hyperthermostability emerges from the oligomerization of four trimers into a dodecameric enzyme. The interfaces between the trimers in the dodecameric enzyme are mainly hydrophobic.

From the point of view of evolution, this could mean that the ancestor of this hyperthermophilic enzyme was a more moderately thermophilic trimeric enzyme that was already partially stabilized by ion-pair networks at the interfaces between monomers. Hyperthermophilic archaea are thought to be ancestral in their group (Matte-Tailliez *et al.*, 2002), but may have been derived from non-extreme thermophilic primordial cells (Galtier *et al.*, 1999; Xu & Glansdorff, 2002).

We thank the staff of the FIP beamline at the European Synchrotron Radiation Facility for their support, Dany Van Elder for helping with X-ray diffraction data collection and Marc Demarez for his help in fermentation and protein purification. This work was supported in part by a grant made to JM as Aspirant at the Fonds voor Wetenschappelijk Onderzoek–Vlaanderen.

References

Allewell, N. M., Shi, D., Morizono, H. & Tuchman, M. (1999). *Acc. Chem. Res.* **32**, 885–894.
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.
Cambillau, C. & Claverie, J.-M. (2000). *J. Biol. Chem.* **275**, 32383–32386.
Chakravarty, S. & Varadarajan, R. (2002). *Biochemistry*, **41**, 8152–8161.
Clantin, B., Tricot, C., Lonhienne, T., Stalon, V. & Villeret, V. (2001). *Eur. J. Biochem.* **268**, 3937–3942.
Cunin, R., Glansdorff, N., Piérard, A. & Stalon, V. (1986). *Microbiol. Rev.* **50**, 314–352.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* http://www.pymol.org.

Gallivan, J. P. & Dougherty, D. A. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 9459–9464.
Galtier, N., Tourasse, N. & Gouy, N. (1999). *Science*, **283**, 220–221.
Grabarse, W., Vaupel, M., Vorholt, J. A., Shima, S., Thauer, R. K., Wittershagen, A., Bourenkov, G., Bartunik, H. D. & Ermler, U. (1999). *Struct. Fold. Des.* **7**, 1257–1268.
Gromiha, M. M., Thomas, S. & Santosh, C. (2002). *Prep. Biochem. Biotechnol.* **32**, 355–362.
Ha, Y., McCann, M. T., Tuchman, M. & Allewell, N. M. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 9550–9555.
Hubbard, S. J. & Thornton, J. M. (1993). *NACCESS* program. Department of Biochemistry and Molecular Biology, University College, London, England.
Jin, L., Seaton, B. A. & Head, J. F. (1997). *Nature Struct. Biol.* **4**, 622–625.
Kannan, N. & Vishveshwara, S. (2000). *Protein Eng.* **13**, 753–761.
Kirino, H., Aoki, M., Aoshima, M., Hayashi, Y., Ohba, M., Yamagishi, A., Wakagi, T. & Oshima, T. (1994). *Eur. J. Biochem.* **220**, 275–281.
Kleywegt, G. J. & Jones, T. A. (1994). *CCP4 Newsl.* **31**, 9–14.
Kumar, S. & Nussinov, R. (2001). *Cell. Mol. Life Sci.* **58**, 1216–1233.
Labédan, B., Boyen, A., Baetens, M., Charlier, D., Chen, P., Cunin, R., Durbecq, V., Glansdorff, N., Hervé, G., Legrain, C., Liang, Z., Purcarea, C., Roovers, M., Sanchez, R., Thia-Toong, T., L., Van de Casteele, M., Van Vliet, F., Xu, Y. & Zhang, J. (1999). *J. Mol. Evol.* **49**, 461–473.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Legrain, C., Demarez, M., Glansdorff, N & Piérard, A. (1995). *Microbiology*, **141**, 1093–1099.
Legrain, C., Villeret, V., Roovers, M., Gigot, D., Dideberg, O., Piérard, A. & Glansdorff, N. (1997). *Eur. J. Biochem.* **247**, 1046–1055.
Legrain, C., Villeret, V., Roovers, M., Tricot, C., Clantin, B., Van Beeumen, J., Stalon, V. & Glansdorff, N. (2001). *Methods Enzymol.* **331**, 227–235.
McDonald, I. & Thornton, J. (1994). *J. Mol. Biol.* **238**, 777–793.
McRee, D. E. (1999). *J. Struct. Biol.* **125**, 156–165.
Massant, J., Verstreken, P., Durbecq, V., Kholti, A., Legrain, C., Beeckmans, S., Cornelis, P. & Glansdorff, N. (2002). *J. Biol. Chem.* **277**, 18517–18522.
Matte-Tailliez, O., Brochier, C., Forterre, P. & Philippe, H. (2002). *Mol. Biol. Evol.* **19**, 631–639.
Nguyen, V. T., Baker, D. P., Tricot, C., Baur, H., Villeret, V., Dideberg, O., Gigot, D., Stalon, V. & Haas, D. (1996). *Eur. J. Biochem.* **236**, 283–293.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ramón-Maiques, S., Marina, A., Uriarte, M., Fita, I. & Rubio, V. (2000). *J. Mol. Biol.* **299**, 463–476.
Roovers, M., Sanchez, R., Legrain, C. & Glansdorff, N. (2001). *J. Bacteriol.* **183**, 1101–1105.
Roussel, A. & Cambillau, C. (1992). *TURBO-FRODO*. Biographics, AFMB, Marseille, France.
Sanchez, R., Baetens, M., Van de Casteele, M., Roovers, M., Legrain, C. & Glansdorff, N. (1997). *Eur. J. Biochem.* **248**, 466–474.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **227**, 319–343.
Shi, D., Morizono, H., Aoyagi, M., Tuchman, M. & Allewell, N. M. (2000). *Proteins*, **39**, 271–277.
Shi, D., Morizono, H., Ha, Y., Aoyagi, M., Tuchman, M. & Allewell, N. M. (1998). *J. Biol. Chem.* **273**, 34247–34254.
Shi, D., Morizono, H., Yu, X., Tong, L., Allewell, N. & Tuchman, M. (2001). *Biochem. J.* **354**, 501–509.
Shima, S., Tziatzio, C., Schubert, D., Fukuda, H., Takahashi, K., Ermler, U. & Thauer, R. K. (1998). *Eur. J. Biochem.* **258**, 85–92.
Szilágyi, A. & Závodszky, P. (2000). *Structure*, **8**, 493–504.
Van Boxstael, S., Cunin, R., Khan, S. & Maes, D. (2003). *J. Mol. Biol.* **326**, 203–216.
Vetriani, C., Maeder, D. L., Tolliday, N., Yip, K. S.-P., Stillman, T. J., Britton, K. L., Rice, D. W., Klump, H. H. & Robb, F. T. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12300–12305.
Vieille, C. & Zeikus, G. J. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 1–43.
Villeret, V., Clantin, B., Tricot, C., Legrain, C., Roovers, M., Stalon, V., Glansdorff, N. & Van Beeumen, J. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2801–2806.
Villeret, V., Tricot, C., Stalon, V. & Dideberg, O. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 10762–10766.
Xu, Y. & Glansdorff, N. (2002). *Comput. Biochem. Physiol. A*, **133**, 677–688.