2.0 Å X-ray structure of the ternary complex of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase from *Escherichia coli* with ATP and a substrate analogue

David K. Stammers[1], Aniruddha Achari, Donald O'N. Somers*, Patrick K. Bryant[2], Jane Rosemond[3], David L. Scott[4], John N. Champness[5]

Glaxo Wellcome R&D, Medicines Research Centre, Gunnell's Wood Road, Stevenage SG1 2NY, UK

Received 3 May 1999; received in revised form 15 June 1999

Abstract The X-ray crystal structure of 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase (PPPK) in a ternary complex with ATP and a pterin analogue has been solved to 2.0 Å resolution, giving, for the first time, detailed information of the PPPK/ATP intermolecular interactions and the accompanying conformational change. The first 100 residues of the 158 residue peptide contain a βαββαβ motif present in several other proteins including nucleoside diphosphate kinase. Comparative sequence examination of a wide range of prokaryotic and lower eukaryotic species confirms the conservation of the PPPK active site, indicating the value of this de novo folate biosynthesis pathway enzyme as a potential target for the development of novel broad-spectrum anti-infective agents.

Key words: Anti-microbial; Folate pathway; Drug target; X-ray crystallography; 7,8-Dihydro-6-hydroxymethylpterinpyrophosphokinase

1. Introduction

Mutsio 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase (PPPK, EC.2.7.6.3) catalyses pyrophosphorylation of 7,8-dihydro-6-hydroxymethylpterin (HP) to 7,8-dihydro-6-hydroxymethylpterinpyrophosphate. The PPPK reaction is one of a series that form the de novo folate biosynthesis pathway which, starting with dihydroneopterin and involving incorporation of *p*ABA and glutamic acid, results in the formation of dihydrofolate. Folate cofactors are required in many one-carbon transfer reactions, such as biosynthesis of purines, thymidylate, certain amino acids as well as formyl-tRNA. The folate biosynthesis pathway is generally only present in prokaryotes and lower eukaryotes [1]. Higher eukaryotes do not synthesise folates but depend on dietary folates that are taken up by a carrier system. Thus, enzymes of the folate biosynthesis pathway represent selective sites for drug action and are attractive targets for development of anti-microbials, as exemplified by the extensive use of sulfonamide anti-bacterial drugs which act as inhibitors of dihydropteroate synthase (DHPS), the enzyme following PPPK in the pathway. Owing to the development and wide-spread occurrence of resistance to many commonly used anti-infective drugs, it has become necessary to consider alternative targets for the development of novel anti-microbial drugs. PPPK represents one such potential target.

Inhibitors of PPPK that are close analogues of the pterin substrate HP have been reported [2]. Such compounds, for example 40Y67 and 87Y76 (Fig. 1), are a valuable aid to structural characterisation of the enzyme active site. However, their relatively low potency and unfavourable physical properties make them unsuitable candidates for development as anti-infectives.

PPPK from *Escherichia coli* is a monomeric 158 residue protein [3]. Amongst the PPPKs cloned and sequenced from both prokaryotic and eukaryotic organisms [3–7], there is an overall sequence conservation of approximately 30%, indicating considerable similarity of three-dimensional structures.

Crystal structures for the adjacent enzymes in the folate pathway, dihydroneopterin aldolase [8] and DHPS [9,10], have been determined. Recent structural studies of PPPKs describe *Haemophilus influenzae* enzyme [11] in complex with a degradation product of 40Y67 and *E. coli* apoenzyme [12], a model for HP and ATP binding being proposed for the latter. This report describes X-ray crystallographic details of ligand binding to *E. coli* PPPK in a ternary complex with ATP and 87Y76 and in binary complex with 40Y67.

2. Materials and methods

2.1. PPPK crystallisation and data collection

The cloning, expression and purification of *E. coli* PPPK have been described previously [3]. The enzyme was crystallised under a number of different conditions. Crystals suitable for X-ray studies were obtained when PPPK (5 mg/ml) was incubated for 3 h in the presence of 2 mM DTT, 6 mM $MgCl_2$, 0.7 mM ATP and 0.5 mM pterin analogue (40Y67 or 87Y76). The enzyme was set up as sitting drops using 5 μl of enzyme with 5 μl of reservoir consisting of 0.1 M sodium citrate pH 5.6, 0.2 M ammonium acetate and PEG 3400 in the range of 27–31%(w/v) at 4°C. Cuboidal crystals grew in a few weeks, in some cases following seeding, having maximum dimensions of 0.3 mm and diffracting to 1.5 Å resolution. They belong to the orthorhombic space

*Corresponding author. Fax: (44) (1438) 764918.
E-mail: ds44295@ggr.co.uk

[1] Current address: Laboratory of Molecular Biophysics, University of Oxford, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK.
[2] Current address: School of Biochemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.
[3] Glaxo Wellcome, Research Triangle Park, NC, USA.
[4] Current address: Departments of Medicine and Orthopaedic Surgery, Massachusetts General Hospital/Harvard Medical School, 149 13th Street, MZ 1498000, Charlestown, MA 02129-2000, USA.
[5] Current address: Division of Biomedical Sciences, Randall Institute, King's College, 26-29 Drury Lane, London WC2B 5RL, UK.

Abbreviations: PPPK, 7,8-dihydro-6-hydroxymethylpterinpyrophosphokinase; HP, 7,8-dihydro-6-hydroxymethylpterin; DHPS, dihydropteroate synthase; NDP, nucleoside diphosphate

0014-5793/99/$20.00
PII: S0014-5793(99)00860-1

group $C222_1$ with $a = 41.0$ Å, $b = 69.2$ Å, $c = 115.7$ Å and contain one molecule in the asymmetric unit. Most X-ray data were collected with a Siemens type P1000 area detector system with Cu K_α radiation from a MacScience rotating anode generator running at 5.4 kW. A single crystal was used to collect a first native PPPK/40Y67 data set and was also used as the vehicle for the subsequent evaluation of several heavy atom soaks to about 2.5 Å resolution. Isomorphous X-ray data were collected from crystals of enzyme which bound lead acetate, trimethyl lead acetate, potassium platinum tetranitrite, mercury chloride and uranyl nitrate. Native and derivative data were processed using the FRAMBO and XENGEN software [13]. A second native data set, PPPK complexed with 87Y76 and ATP, was collected on a Molecular Structure Corporation R-AXIS II image plate system with Cu K_α radiation from a Rigaku generator running at 5 kW and processed with the program DENZO [14] (Table 1).

2.2. Structure solution and refinement

Heavy atom sites of the lead acetate derivative were located from difference Patterson calculations using the program HASSP [15] and refined with a modified version of the CCP-4 program MLPHARE [16]. Sites of other derivatives were obtained from difference Fourier syntheses and cross-checked against difference Patterson maps. MIR phases from five heavy atom derivatives were used to calculate an electron density map at 2.5 Å resolution, from which ligand binding and secondary structure features could be identified. Density modification and phase improvement was done with the program SOLOMON [17]. The structure was modelled with the program O [18]. As ATP appeared to be not present in the 40Y67-bound structure, coordinates were refined against the second, 87Y76-bound, native data set (which showed the presence of ATP) using PROFFT [19] followed by re-building with O. The final round of refinement was done using the maximum likelihood option of the program REFMAC [20]. Coordinates will be deposited with the Protein Data Bank at Brookhaven.

2.3. Structure alignments

The PPPK coordinates were aligned with protein structures available in the Protein Data Bank using the distance matrix alignment algorithm, DALI [21]. Two structures, ribosomal S6 protein and NDP kinase, were subjected to more detailed comparison with PPPK. Both were aligned using the structure homology program SHP [22] and displayed on a Silicon Graphics Octane workstation using O [18]. Active site residues of *E. coli* PPPK were aligned by multiple amino acid sequence alignment with those in other PPPKs (Table 2) using the program PILEUP included in the GCG suite [23].

3. Results and discussion

3.1. The PPPK fold

The main features of the *E. coli* enzyme structure (Fig. 2) are as described earlier for the apoenzyme [12]. However, residues showing conformational variability in the apoenzyme structure now appear well-defined in this ternary complex, one such residue, Leu-45, being involved in interaction with the HP substrate analogue. The folding at the core of PPPK, a four-stranded anti-parallel β-sheet with helical interconnections, β1-α1-β2-β3-α2-β4, is not unique. The first 101 residues fit the βαββαβ topology already characterised [24]. The structure closest-matched to PPPK by distance matrix alignment [21] appears to be the ribosomal S6 protein [25]. Overlapping this with SHP software [22] gives a root mean square (r.m.s.) distance for 82 equivalent α-carbons of 2.3 Å. One other structure matched to PPPK, a kinase of similar size, nucleoside diphosphate (NDP) kinase [26], gave a r.m.s. of 2.7 Å for 82 equivalent residues. Comparison of the superimposed βαββαβ topologies of PPPK and NDP kinase (Fig. 3) emphasises their similar secondary structure, with NDP kinase nevertheless having longer β2-β3 and α2-β4 interconnecting loops and a different folding of the C-terminal region. Also, whilst the ATP of the PPPK ternary complex is held between the α2-β4 interconnecting loop and the initial loop of the C-terminal region, the ADP bound to NDP kinase is instead located between the α2-β4 and the β2-β3 interconnecting loops and is orientated differently. Thus, only one phosphate from each of the superimposed structures coincides, substrate binding being defined more by the connecting loop structure than by the secondary structure.

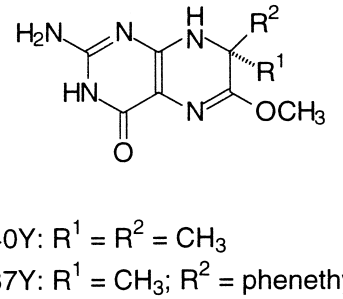

Fig. 1. Structure of the pterin analogues 40Y67 and 87Y76.

3.2. Interaction of ligands with active site residues

In the binary complex of PPPK with 40Y67, the orientation of the ligand is clearly defined in difference electron density maps by the prominent gem-dimethyl group. Ring stacking interactions of the aromatic side chains of Tyr-53 and Phe-123 sandwich the pterin of the ligand. Additional hydrophobic contacts are made by Arg-92 and Trp-89, the latter side chain interacting with the gem-dimethyl group. Hydrogen bonding contacts (as for 87Y76, see Fig. 4) are made by interactions between the pterin and Thr-42, Pro-43, Leu-45, Asn-55 and

Table 1
X-ray data collection and phasing

Data set	Resolution (Å)	Number of observations	Unique data (% possible)	R_{sym} (%)	R_{merge} (%)	Phasing power (number of sites)
Native-1	2.5	22 063	5 319 (87.3)	4.35	–	
Native-2	2.0	47 660	10 879 (94.4)	5.00	–	
Lead acetate	2.0	16 758	10 111 (84.5)	5.42	25.0	2.19 (2)
Tetra-methyl lead	2.0	22 605	9 248 (74.8)	4.84	28.0	1.66 (4)
Platinum nitrite	2.5	15 692	5 942 (94.9)	5.42	15.1	0.47 (4)
Mercury chloride	2.5	3 038	2 421 (38.7)	6.93	10.8	0.23 (3)
Uranyl nitrate	2.5	10 521	4 345 (69.7)	3.59	12.0	0.69 (7)

$R_{sym} = \Sigma|I_i - \langle I\rangle|/\Sigma I_i$, where I_i is the measured intensity of an individual reflection and $\langle I\rangle$ is the mean intensity of repeated measurements of this reflection. The highest resolution shell of native-2 data (2.07–2.00 Å) gave $R_{sym} = 10.2\%$ and was 92.2% complete. $R_{merge} = \Sigma|F_{ph} - F_p|/\Sigma F_p$, where F_{ph} is the derivative amplitude and F_p the native amplitude. In multiple isomorphous replacement phasing, data to 2.5 Å were used throughout. The phasing power is the ratio of the r.m.s. heavy atom F to the r.m.s. residual lack of closure. In structure refinement (see text), R-factor $= (\Sigma||F_o| - |F_c||/\Sigma F_o)$, where F_o ($F > 2\sigma$) and F_c are the observed and calculated structure factors, respectively.

Table 2
Comparison of active site residues for *E. coli* PPPK (Ec) [3] and other PPPKs

Ec	Bs	Pc	Sp	Pf	Tg	Mt	Hp
R41	E	E	S	E	E	E	I
T42	*	S	D	*	*	A	N
P43	D	K	G	V	V	D	*
P44	*	*	A	*	*	*	*
L45	V	M	*	A	F	W	F
G46	*	Y	*	E	Y	*	*
D52	Q	A	C	Y	A	Q	N
Y53	F	F	F	F	A	F	F
L54	*	Y	A	Y	*	*	Y
N55	*	*	*	*	*	*	*
A56	M	*	Q	L	D	*	*
R88	*	D	H	K	*	*	K
W89	*	K	*	K	G	*	D
G90	*	*	*	E	A	*	A
P91	*	*	*	N	*	*	*
R92	*	*	*	*	*	*	*
T93	*	C	L	I	L	N	*
L94	A	I	I	I	*	*	*
D95	*	*	*	*	*	*	*
L96	*	*	*	I	V	V	I
D97	*	*	*	*	*	*	*
I98	*	*	L	*	L	L	*
H115	*	*	*	*	*	*	*
Y116	P	P	P	V	P	P	P
D117	R	R	Y	Y	R	L	K
N120	E	E	E	H	T	L	E
R121	*	*	*	*	*	*	*
G122	L	S	L	Y	N	A	D
F123	*	*	*	S	*	*	S
M124	V	V	V	I	V	V	V

Bacillus subtilis (Bs) [27], *P. carinii* (Pc) [7], *Streptococcus pneumoniae* (Sp) [5], *P. falciparum* (Pf) [28], *T. gondii* (Tg) [29], *Mycobacterium tuberculosis* (Mt) [30] and *Helicobacter pylori* (Hp) [31]. Identical residues with respect to Ec are denoted by * and residues in direct contact with the pterin ligand and with the β- and γ-phosphates are emboldened. Alignment was prepared using the program PILEUP, included in the GCG software suite [23].

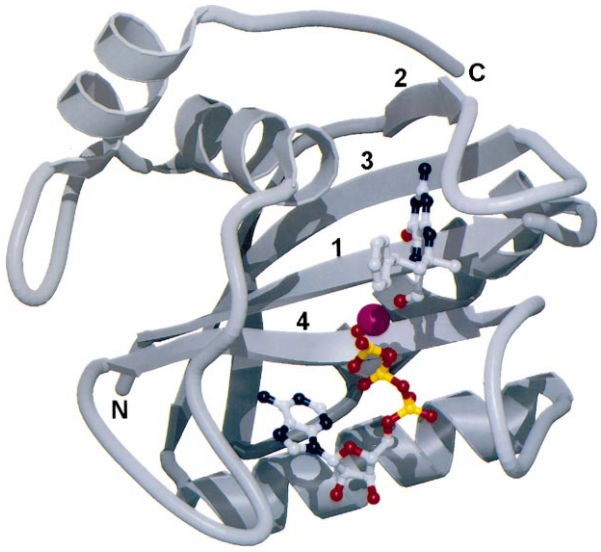

Fig. 2. Diagrammatic view of the fold of PPPK (the N- and C-termini and β-strands 1–4 are arrowed). The active site is represented by the presence of ligands Mg-ATP and 87Y76.

several water molecules. A tetrahedrally coordinated anion defines the α-phosphate site of the product 7,8-dihydro-6-hydroxymethylpterinpyrophosphate. There was no further density visible for the ATP molecule in this complex, suggesting it had degraded.

The ternary complex data from PPPK with 87Y76 and ATP were refined in the resolution range 10–2 Å and gave a *R*-factor (see Table 1, footnote) of 0.169. The r.m.s. errors in the model of bond-length and valence-angle were 0.013 Å and 1.6°, respectively. 130 Waters and two magnesium ions were identified in addition to ligands 87Y76 and ATP. Intermolecular hydrogen bonding for the ternary complex is shown in Fig. 4. The pterin-based hydrogen bond interactions as well as the hydrophobic interactions are similar to those described for

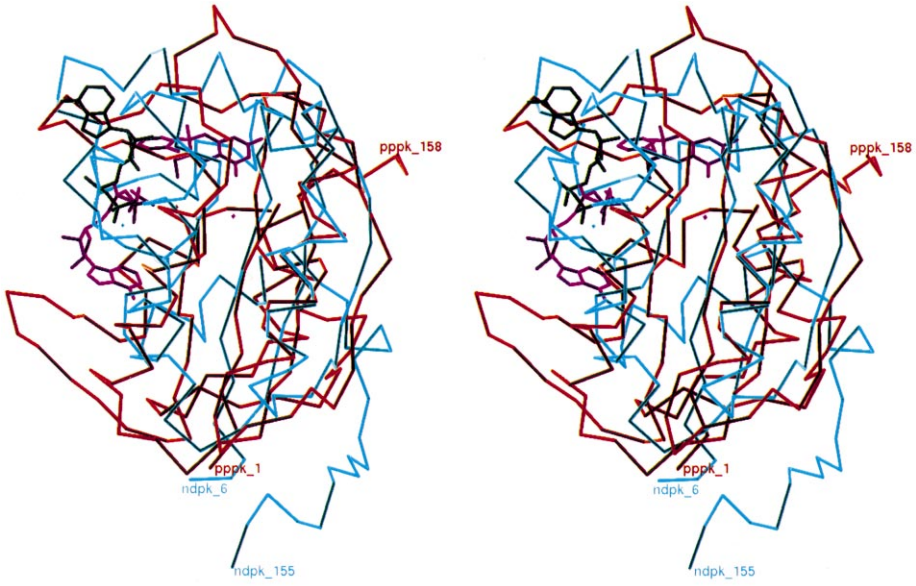

Fig. 3. Overlap of Cα traces of NDP kinase (blue) with PPPK (red). Skeletal models of the respective ligands of each, ADP (black) and 87Y76 with ATP (magenta) are also shown. The display was generated by the program QUANTA (Molecular Simulations).

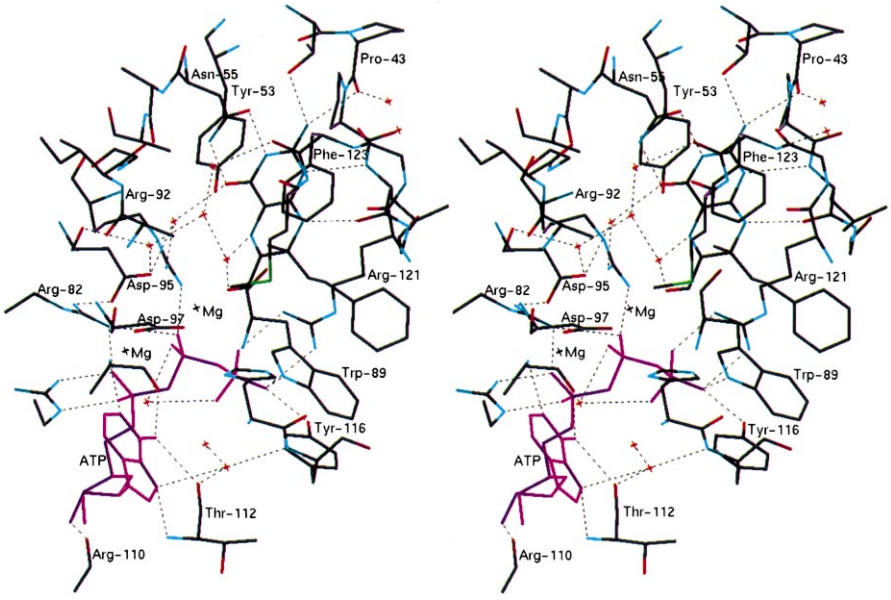

Fig. 4. Stereo view of the active site of the PPPK/ATP/87Y76 complex, showing hydrogen bond interaction. The display was generated by the program QUANTA (Molecular Simulations).

40Y67. The presence of the phenethyl moiety of 87Y76 gives rise to further hydrophobic contacts involving the side chains of Trp-89 and Leu-45. Difference Fourier density clearly reveals the presence of ATP in this complex. The adenine interacts via hydrogen bonds with the carbonyl oxygens of Leu-98 and Thr-112, the ribose is hydrogen-bonded to Arg-110, the α- and β-phosphates interact with Arg-82, -84 and -92 and the γ-phosphate is hydrogen-bonded to Trp-89, Tyr-116 and Arg-121. The side chain of the highly conserved His-115 residue is positioned about 3.7 Å from a γ-phosphoryl oxygen and does not appear to interact with an α-phosphoryl oxygen, as earlier proposed from model-building [12]. Density peaks lying between the β- and γ-phosphates on the one hand and Asp-95 and Asp-97 on the other are identified as magnesiums. It seems likely that this arrangement of aspartates and magnesiums could stabilise a transition state in the course of the pyrophosphoryl transfer reaction. Absence of ATP (as in the 40Y67 complex) appears to result in a conformational change of a single side chain, Arg-84, to a position close to the ribose 3′-hydroxyl site.

3.3. Conservation of the PPPK active site

Multiple sequence alignments followed by mapping 30 active site residues on to the three-dimensional structure of *E. coli* PPPK reveal invariance in the region where the magnesium ions bind and lesser conservation beyond this. Variation is often limited to a conservative change, otherwise, the residue is involved in main-chain contact which allows greater diversity of the side chain (Table 2). The level of identity or conservative difference between *E. coli* PPPK and the eukaryotic (*Pneumocystis carinii, Toxoplasma gondii* and *Plasmodium falciparum*) PPPKs is 63–67%, within the prokaryotic PPPKs, the level is at least 73%.

3.4. Comparison of ligand sites in PPPK and DHPS

The adjacent enzyme in the pathway, DHPS [9,10], is an 8-stranded αβ barrel which binds pterin substrate at a location on the barrel axis near the C-terminal pole. In both PPPK and DHPS, the pterin moiety is similarly hydrogen-bonded around the edges with its faces sandwiched between hydrophobic groups. By contrast, the phosphate binding sites on PPPK and DHPS are significantly different. Whilst *E. coli* DHPS [9] interacts directly with ligand α- and β-phosphates via several basic residues, the corresponding phosphate sites of PPPK (the γ- and β-phosphates of ATP) include interactions involving cations, in addition to hydrogen bonding to basic residues. Thus, the design of a single inhibitor of both enzymes, a favourable objective in providing potent 'double sequential blockade' (reducing the likelihood of resistance development), should focus on the pterin site, avoiding the more structurally divergent phosphate sites.

The information we report on the refined X-ray structures of *E. coli* PPPK in ternary complex with ATP and a substrate analogue is a valuable starting point for characterisation of this enzyme in terms of its potential as a target for the development of anti-microbial drugs. The high resolution of this structure determination will provide a basis for the design of novel compounds with potential broad-spectrum anti-microbial activity.

Acknowledgements: We thank Carl Ross, Stella Fabiane and Kate Kirwan for their assistance.

References

[1] Warhurst, D.C. (1986) Parasitol. Today 2, 57–58.
[2] Wood, H.C.S. (1975) in: Chemistry and Biology of Pteridines (Pfleiderer, W., Ed.), Proc. 5th Int. Symp., pp. 27–49, W. de Gruyter, Berlin.
[3] Talarico, T.L., Ray, P.H., Dev, I.K., Merrill, B.M. and Dallas, W.S. (1992) J. Bacteriol. 174, 5971–5977.
[4] Volpe, F., Dyer, M., Scaife, J.G., Darby, G., Stammers, D.K. and Delves, C.J. (1992) Gene 112, 213–218.
[5] Lopez, P., Greenberg, B. and Lacks, S.A. (1990) J. Bacteriol. 172, 4766–4774.

[6] Volpe, F., Ballantine, S.P. and Delves, C.J. (1993) Eur. J. Biochem. 216, 449–458.
[7] Ballantine, S.P., Volpe, F. and Delves, C.J. (1994) Protein Expr. Purif. 5, 371–378.
[8] Hennig, M., D'Arcy, A., Hampele, I.C., Page, M.G., Oefner, C. and Dale, G.E. (1998) Nat. Struct. Biol. 5, 357–362.
[9] Achari, A., Somers, D.O., Champness, J.N., Bryant, P.K., Rosemond, J. and Stammers, D.K. (1997) Nat. Struct. Biol. 4, 490–497.
[10] Hampele, I.C., D'Arcy, A., Dale, G.E., Kostrewa, D., Nielsen, J., Oefner, C., Page, M.G., Schonfeld, H.J., Stuber, D. and Then, R.L. (1997) J. Mol. Biol. 268, 21–30.
[11] Hennig, M., Dale, G.E., D'arcy, A., Danel, F., Fischer, S., Gray, C.P., Jolidon, S., Muller, F., Page, M.G.P., Pattison, P. and Oefner, C. (1999) J. Mol. Biol. 287, 211–219.
[12] Xiao, B., Shi, G., Chen, X., Yan, H. and Ji, X. (1999) Structure 7, 489–496.
[13] Howard, A.J., Gilliland, G.L., Finzel, B.C., Poulos, T.L., Ohlendorf, D.H. and Salemme, F.R. (1987) J. Appl. Cryst. 20, 383–387.
[14] Otwinowski, Z. (1993) in: Data Collection and Processing (Sawyer, L., Isaacs, N.W. and Bailey, S., Eds.), pp. 55–62, DL/SCI/R34, Daresbury Laboratory, Warrington.
[15] Terwilliger, T.C., Kim, S.-H. and Eisenberg, D. (1987) Acta Crystallogr. A43, 1–5.
[16] Otwinowski, Z. (1991) in: Isomorphous Replacement and Anomalous Scattering (Wolf, W., Evans, P. and Leslie, A.G.W., Eds.), pp. 80–86, DL/SCI/R32, Daresbury Laboratory, Warrington.
[17] Abrahams, J.P. and Leslie, A.G.W. (1996) Acta Crystallogr. D52, 30–42.
[18] Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M. (1991) Acta Crystallogr. A47, 110–119.
[19] Finzel, B.C. (1987) J. Appl. Cryst. 20, 53–55.
[20] Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Acta Crystallogr. D53, 240–255.
[21] Holm, L. and Sander, C. (1993) J. Mol. Biol. 233, 123–138.
[22] Stuart, D.I., Levine, M., Muirhead, H. and Stammers, D.K. (1979) J. Mol. Biol. 134, 109–142.
[23] Program Manual for the Wisconsin Package (1994) version 8, Genetics Computer Group, 575 Science Drive, Madison, WI.
[24] Orengo, C.A. and Thornton, J.M. (1993) Structure 1, 105–120.
[25] Lindahl, M., Svensson, L.A., Liljas, A., Sedelnikova, S.E., Eliseikina, I.A., Fomenkova, N.P., Nevskaya, N., Nikonov, S.V., Garber, M.B., Muranova, T.A. and Rykonova, A.I. (1994) EMBO J. 13, 1249–1254.
[26] Morera, S., Lacombe, M.-L., Yingwu, X., LeBras, G. and Janin, J. (1995) Structure 3, 1307–1314.
[27] Slock, J., Stahly, D.P., Han, C.-Y., Six, E.W. and Crawford, I.P. (1990) J. Bacteriol. 172, 7211–7226.
[28] Brooks, D.R., Wang, P., Read, M., Watkins, W.M., Sims, P.F.G. and Hyde, J.E. (1994) Eur. J. Biochem. 224, 397–405.
[29] Pashley, T.V., Volpe, F., Pudney, M., Hyde, J.E., Sims, P.F. and Delves, C.J. (1997) Mol. Biochem. Parasitol. 86, 37–47.
[30] Cole, S.T. et al. (1998) Nature 393, 537–544.
[31] Tomb, J.-F. et al. (1997) Nature 388, 539–547.