Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of glyceraldehyde-3-phosphate dehydrogenase from *Plasmodium falciparum*

Jacqueline F. Satchell,[a,b] Robyn L. Malby,[a] Cindy S. Luo,[a] Akinola Adisa,[c,d] Aysun E. Alpyurek,[c,d] Nectarios Klonis,[c] Brian J. Smith,[a] Leann Tilley[c,d] and Peter M. Colman[a]*

[a]The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia, [b]Department of Medical Biology, The University of Melbourne, Victoria, Australia, [c]Department of Biochemistry, La Trobe University, Melbourne, Victoria, Australia, and [d]Co-operative Research Centre for Diagnostics, La Trobe University, Melbourne, Victoria, Australia

Correspondence e-mail: pcolman@wehi.edu.au

Received 3 May 2005
Accepted 9 June 2005

PDB Reference: glyceraldehyde-3-phosphate dehydrogenase, 1ywg, r1ywgsf.

The malaria parasite *Plasmodium falciparum* is responsible for about two million deaths annually, making it important to obtain information about enzymes from this organism that represent potential drug targets. The gene for *P. falciparum* glyceraldehyde-3-phosphate dehydrogenase (*Pf*GAPDH) has been cloned and the protein expressed as a hexahistidine-tagged recombinant protein in *Escherichia coli*. The recombinant protein has been crystallized and its three-dimensional structure determined. One molecule of the cofactor NAD^+ is bound to each of the four subunits in the tetrameric enzyme. The major structural feature distinguishing human GAPDH from *Pf*GAPDH is the insertion of a dipeptide (-KG-) in the so-called *S* loop. This insert, together with other characteristic single-amino-acid substitutions, alters the chemical environment of the groove that encompasses the *R* dyad and that links adjacent cofactor-binding sites and may be responsible for the selective inhibition of the enzyme by ferriprotoporphyrin IX.

1. Introduction

The malaria parasite *Plasmodium falciparum* resides for part of its lifecycle within a parasitophorous vacuole inside the mature red blood cells (RBC) of its human host. The mature RBC is one of the most highly terminally differentiated cells in the human body; it has no nucleus, no protein-synthesis capability and no protein-trafficking machinery. As a consequence, the host cell has relatively modest energy needs, consuming only 1–2 mmol glucose per hour per litre of cells (Roth *et al.*, 1988). Life within this quiescent cell offers the parasite protection from the host's immune system, but presents it with a major challenge with respect to the acquisition of sufficient nutrients.

For example, the intraerythrocytic parasite needs significant energy to power its rapid growth and division. However, the parasite itself lacks many of the biosynthetic pathways of higher eukaryotes. There appears to be little if any fatty-acid oxidation or oxidative phosphorylation in the blood stages of infection and the parasite relies almost entirely on glycolysis for the production of ATP. Seminal studies by Roth and coworkers (Roth, 1990; Roth *et al.*, 1988) demonstrated elevated levels of glycolysis in parasite-infected RBCs and showed that this was a consequence, at least in part, of parasite-derived glycolytic enzymes. These findings have stimulated interest in glycolytic enzymes as targets for therapeutic intervention in malaria (Brady & Cameron, 2004).

In the glycolytic pathway, ten catalytic reactions convert glucose to pyruvate with the production of ATP; glyceraldehyde-3-phosphate dehydrogenase (GAPDH) catalyses the sixth of these. More specifically, GAPDH reduces glycer-

aldehyde-3-phosphate to 1,3-bisphosphoglycerate in the presence of NAD^+ and inorganic phosphate. Interrupting glycolysis by inhibition of GAPDH would not only block production of ATP but would also consume ATP in the generation of the substrate (Daubenberger *et al.*, 2000).

Several considerations arise in contemplating the disruption of parasite glycolysis in the RBC by inhibition of parasite glycolytic enzymes. These include the problem of discovering inhibitors that are specific for the parasite enzyme, the degree of toxicity that might ensue if the host enzyme is also (even partially) inhibited, the potential capacity of the RBC enzyme to support an adequate level of glycolysis for parasite survival and the likely emergence of resistance to enzyme inhibitors that by their very design may be exploiting functionally nonessential components of the enzyme active-site structure.

The issue of selectivity has been addressed and successfully answered in a number of studies using structure-based drug design. For example, inhibitors of influenza virus neuraminidase designed from the active site of that enzyme do not inhibit mammalian or other viral neuraminidases (von Itzstein *et al.*, 1993) and specific trypanasomatid GAPDH inhibitors have been identified based on structural knowledge of the targets (Aronov *et al.*, 1999; Callens & Hannaert, 1995). These and other studies provide strong evidence that structure-based approaches can generate drugs or drug candidates that interact selectively with their targets.

Mammalian GAPDH is believed to play additional roles in diverse cellular functions (reviewed in Sirover, 1999), including apoptosis (Ishitani & Chuang, 1996). Up-regulation of GAPDH expression has been shown to be coordinated with induction of apoptosis in cultured cerebellar neurons, while GAPDH antisense DNA was shown to arrest chemically induced apoptosis. One GAPDH inhibitor (Kragten *et al.*, 1998) is in phase II clinical trials for neuroprotection in Parkinson's disease (Reed, 2002). It is still not clear whether *P. falciparum* GAPDH (*Pf*GAPDH) has any important function other than in glycolysis, although it may play a role in vesicular transport (Daubenberger *et al.*, 2000).

A recent novel finding is that ferriprotoporphyrin IX (FP), a toxic byproduct of haemoglobin degradation in infected RBCs, binds to a number of *P. falciparum* proteins, including GAPDH and glutathione reductase (Campanale *et al.*, 2003). Furthermore, FP has a K_i of 200 n*M* for turnover of substrate by *Pf*GAPDH, but shows little inhibition of RBC GAPDH. The biological consequences of the specific inhibition of *Pf*GAPDH remain unclear, but it has been suggested that FP-induced inhibition of glycolysis might enhance the flux of glucose through the hexose monophosphate shunt with concomitant production of NADPH, a requisite cofactor for antioxidant pathways (Campanale *et al.*, 2003). Indeed, hexose monophosphate shunt activity is elevated more than 50 times in parasite-infected RBCs (Atamna *et al.*, 1994).

Following the pioneering studies of Rossmann and co-workers (Buehner *et al.*, 1973, 1974), the three-dimensional structures of GAPDHs from 15 different organisms with a range of different cofactors, substrate analogues and inhibitors have been described and more than 40 entries are currently found in the Protein Data Bank (Berman *et al.*, 2002). The enzyme structure is bilobal, with an N-terminal cofactor-binding domain and a C-terminal catalytic domain, and in all cases described to date is a homotetramer with dihedral symmetry (reviewed in Nagradova, 2001). A number of recent studies have focused on the enzyme from trypanosomatids (Kim *et al.*, 1995; Souza *et al.*, 1998; Vellieux *et al.*, 1993) with the aim of exploiting structural features of the protozoan enzyme active site that distinguish it from that of the host. Selective inhibitors of *Trypanosomatidae* GAPDH have subsequently been reported (Aronov *et al.*, 1999; Bressi *et al.*, 2001) and their crystal structures analysed in complex with the target enzyme (Suresh *et al.*, 2001). These adenosine analogues induce structural changes in the adenine-binding pocket of *Leishmania mexicana* GAPDH, the specific site to which they were targeted, owing to the presence of novel amino-acid sequence motifs in that region of the protein sequence of trypanosomatids.

The gene encoding *Pf*GAPDH has recently been cloned and the protein expressed in *Escherichia coli* (Campanale *et al.*, 2003; Daubenberger *et al.*, 2000). It has been selected as a target in a drug-development project supported by the Medicines for Malaria Venture; however, attempts to grow crystals of the protein have to date been unsuccessful (Cowan-Jacob *et al.*, 2003). We have now crystallized recombinant *Pf*GAPDH and undertaken a study of its three-dimensional structure in order to provide an experimental basis for understanding and potentially exploiting structural features that are unique to the plasmodial enzyme, especially those that might be linked to survival of the parasite in the specialized environment of the red blood cell.

2. Materials and methods

2.1. Protein expression and purification

Following a previous report (Daubenberger *et al.*, 2000), the gene for GAPDH was amplified by PCR from the cDNA of ring-stage *P. falciparum* (3D7 strain) and cloned into the pET15b vector (Novagen). The construct was sequenced in both directions using fluorescent dideoxynucleotide termination. Cells from 1 l cultures of transformed *E. coli* BL21 (DE3) were frozen and resuspended in 25 ml 50 m*M* NaH_2PO_4 pH 7.2, 500 m*M* NaCl, 3 m*M* β-mercaptoethanol and one Complete protease-inhibitor tablet (Roche) and lysed by repeated flash-freezing and thawing followed by mild sonication. The cell lysate was centrifuged and the supernatant filtered and applied onto 3 ml Talon resin (Clontech). After washing in 50 m*M* NaH_2PO_4 pH 7.2, 500 m*M* NaCl, 3 m*M* β-mercaptoethanol and 50 m*M* imidazole, the protein was eluted with 50 m*M* NaH_2PO_4 pH 7.2, 500 m*M* NaCl, 150 m*M* imidazole, 3 m*M* β-mercaptoethanol. The purified protein was shown to be enzymatically active (Campanale *et al.*, 2003).

The protein was concentrated to 10 mg ml^{-1} and the buffer exchanged for 10 m*M* Tris pH 7.0, 10 m*M* imidazole and 3 m*M* dithiothreitol (DTT) for crystallization trials. Analysis of the resulting low-quality crystals suggested that the purification

tag should be removed. Protein was purified as above except protease inhibitors were not used. The tag was cleaved with 10 units of thrombin (Roche) and this material applied onto a further 3 ml of Talon resin (Clontech) and the untagged GAPDH was collected from the flowthrough. The protein was further purified by gel filtration (Superdex 200; Amersham Biosciences; equilibrated with 50 m*M* NaH_2PO_4, 500 m*M* NaCl and 3 m*M* DTT), where it migrated as a species with an apparent molecular weight of 130 kDa, consistent with a tetramer of 36 kDa subunits. The protein was concentrated to 10 mg ml^{-1} and the buffer exchanged as above for crystallization trials.

2.2. Crystallization and data collection

Crystals of the recombinant His-tagged protein were grown under a variety of conditions containing polyethylene glycols of varying molecular weights. The best of these grew from 0.2 *M* KBr, 16% PEG 3350, 20 m*M* HEPES pH 7.2 with a fourfold molar excess of NAD^+. The crystals belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 71.9$, $b = 104.5$, $c = 90.8$ Å, $\beta = 108.7^\circ$, one tetramer in the asymmetric unit and a solvent content of 48%. The crystals were transferred to cryoprotectant (0.2 *M* KBr, 16% PEG 3350, 20 m*M* HEPES pH 7.2 and 20% glycerol) prior to flash-freezing and data collection. Diffraction data were collected using an R-AXIS image-plate detector on a rotating-anode generator fitted with a capillary optic (AXCO) and were processed using *d*TREK* (Pflugrath, 1999), but extended only to 4 Å resolution.

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses refer to the outer shell.

Space group	$P6_1$
Unit-cell parameters (Å)	$a = b = 79.7$, $c = 407.1$
Fraction of unit cell occupied by protein	0.41
Crystal mosaicity (°)	0.3
Total No. of reflections	531565
No. of unique reflections	44207
Average redundancy	4.7 (3.85)
Completeness (%)	99 (95.7)
R_{merge}	0.14 (0.45)
Reduced χ^2	0.66 (0.53)
$I/\sigma(I)$	9.7 (2.0)
Resolution (Å)	2.6
No. of atoms	10340
No. of water molecules	314
Mean *B* value, protein (Å^2)	31
Mean *B* value, NAD (Å^2)	37
NCS restraint (kJ mol Å^{-2})	1256
R.m.s.d. bonds (Å)	0.007
R.m.s.d. bond angles (°)	1.3
R.m.s.d. dihedral angles (°)	5.6
R_{work}	0.231
R_{free}	0.256

Crystals of the tag-free protein were grown from 0.2 *M* calcium acetate hydrate, 20% PEG 3350 pH 7.3 with a fourfold molar excess of NAD^+. They belong to the hexagonal space group $P6_1$, with one tetramer per asymmetric unit and a solvent content of 59% (see Table 1). These crystals were flash-frozen in 25% glycerol, 23% PEG 3350, 0.2 *M* calcium acetate hydrate and data were collected from a single crystal at the Swiss Light Source on a microcrystal diffractometer equipped with a MAR 165 CCD detector. Data were

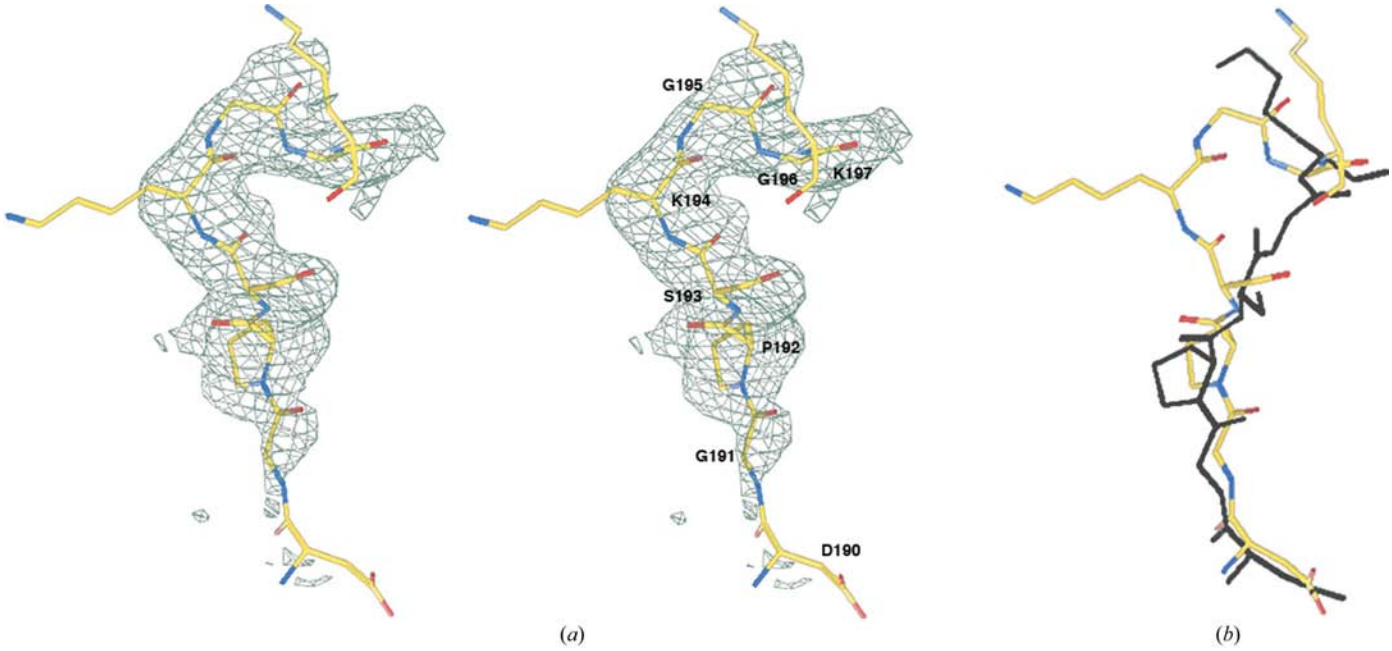

Figure 1
(*a*) Stereo representation of the fourfold-averaged electron density of a $F_o - F_c$ omit map contoured at 3σ. The model is shown for residues D^{190}GPSKGGK^{197}, where residues 192–196 were omitted from the phasing. The density for the carbonyl of residue 193 suggests a hydrogen bond with the backbone amide of residue 196. (*b*) Overlay of the *P. falciparum* model with the rabbit muscle GAPDH structure (PDB code 1j0x) D^{186}GPSGK^{191}, shown in black, illustrating that the additional residues (-KG-) are best modelled as a single dipeptide insert. This segment is part of the *S* loop, which is shown in its entirety in Fig. 4(*b*).

processed using the *HKL* program suite (Otwinowski & Minor, 1996).

2.3. Structure determination and refinement

A molecular-replacement solution for crystals of tagged protein was found with *MOLREP* (Vagin & Teplyakov, 1997), searching with the whole tetramer from PDB entry 1gd1. Examination of this solution suggested that the purification tag was located near a lattice contact and accordingly the tag-free material was prepared and subsequently crystallized.

A homology model of *Pf*GAPDH, without the NAD^+ cofactor, was constructed in *MODELLER* (Sali & Blundell, 1993) and used to solve the structure of the tag-free protein crystals, again using *MOLREP*. Following rigid-body refinement in *REFMAC* (Murshudov *et al.*, 1997), an electron-density map at 3.5 Å resolution was calculated. Phase extension was performed in *DM* (Cowtan, 1994) to 2.6 Å resolution over 600 cycles, applying solvent flattening, histogram matching and non-crystallographic symmetry averaging. An atomic model was constructed into the resulting electron-density map using *O* (Jones, 1978).

Atomic refinement was performed using *CNS* (Brünger *et al.*, 1998; Luthy *et al.*, 1992) (see Table 1 for details) with manual model-building intervention being performed in *O* (Jones, 1978). Non-crystallographic symmetry restraints were maintained throughout the refinement and *B*-factor restraints were slightly relaxed from the default for side-chain bonds and angles (σ target for bonded side-chain atoms = 4.0 Å and σ target for angle side-chain atoms = 5.0 Å). $R_{\rm free}$ (Brünger, 1992) was monitored during the refinement and the final refined model was assessed using *PROCHECK* (Eisenberg *et al.*, 1997; Luthy *et al.*, 1992) and *WHATIF* (Vriend, 1990). An unbiased electron-density map highlighting the region of the inserted dipeptide after residue 193 was calculated by omitting residues 193–196 from the model and running a simulated-annealing protocol (at 4000 K) using *CNS*. The difference omit map density for residues 193–196 is shown in Fig. 1.

2.4. Docking of FP

Docking studies were conducted in a potential binding groove that is flanked by residues Lys125 and Lys194 of the *O* and *R* subunits (Fig. 2). This groove was selected as it harbours the most significant amino-acid substitutions between rabbit muscle GAPDH and *P. falciparum* GAPDH and could be responsible for the observed specific inhibition (Campanale *et al.*, 2003). This docking study was performed to determine whether haem could be spatially accommodated in this groove both in the presence and absence of NAD^+. Owing to the high *B* factors of these lysine side chains, both Lys125 and Lys194 were mutated to alanine residues for the purposes of docking. Water molecules were removed and H atoms were added to the protein using *SYBYL* (Tripos). The positions of the H atoms were optimized while the remainder of the protein was held fixed. Kollmann all-atom charges were assigned for all atoms of the protein and for the NAD^+ cofactor (where present). The active site was defined as being all residues within 21 Å of the backbone N atom of Val188 and a molecular surface (Connolly, 1983) was calculated over these residues. The program *ATPTS* (Moreno & Leon, 2002) was used to fill the active site with sphere points and electrostatic and van der Waals energies were calculated on a grid with a 0.25 Å spacing with an additional 6 Å margin beyond the location of the spheres (Meng *et al.*, 1992). The *DOCK* program (Ewing, 1998) was used to dock FP to the protein, allowing conformational flexibility of the ligand and minimization of each docked configuration. The 50 highest scoring orientations were found to cluster into two groups; a representative of each of these groups was chosen as the orientation with the lowest sum of r.m.s.d. values with each of the other orientations in that cluster.

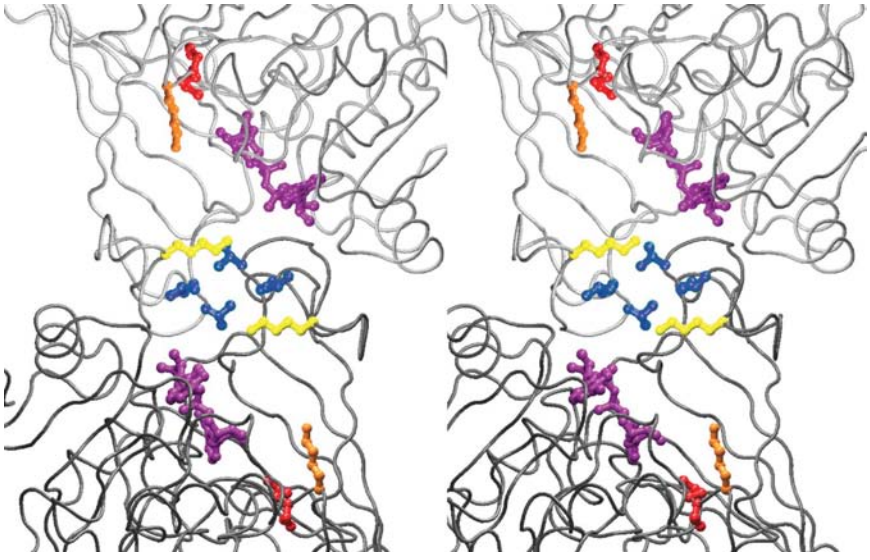

Figure 2
Stereo representation of the interface of *O* and *R* subunits (shown in grey) as seen down the *R* dyad. The residues Lys194 and Gly195, which make up the novel insertion present in *Pf*GAPDH, are coloured yellow. Note that no electron density is visible for Lys194 beyond C^α. Leu187 and Val188, which are Lys and Thr in the human enzyme, respectively, are shown in blue. Lys126 is shown in orange and the two consecutive proline residues (Pro124 and Pro125) are shown in red. NAD^+ is represented in purple. This figure was prepared using *VMD* (Humphrey *et al.*, 1996) and *RASTER3D* (Merritt & Bacon, 1997).

3. Results

3.1. Crystals, structure determination and model quality

His-tagged protein did not generate good-quality crystals. Data could only be collected to 4 Å resolution, with an $R_{\rm merge}$ of 0.14 (0.35 in the outermost shell) for 37 844 measurements of 10 584 independent reflections. The data completeness was 97.5% (97.8% for the highest resolution shell) and the mean $I/\sigma(I)$ was 8.0 (1.7 for the highest resolution shell). The molecular-

replacement solution based on the tetrameric model of *Bacillus stearothermophilus* GAPDH (PDB code 1gd1) was unambiguous and analysis of the molecular packing revealed that the purification tag on one of the four subunits was located at a crystal contact site, providing a plausible rationale for the poor quality of the crystals.

Crystals of tag-free protein diffracted X-rays to 2.6 Å resolution at the Swiss Light Source. Statistics from data processing and reduction are presented in Table 1. Unambiguous rotation and translation functions using the homology model of the tetramer confirmed the space group to be $P6_1$. Density modification and phase extension resulted in a figure of merit of 0.66. The resulting electron-density map at 2.6 Å resolution showed clear evidence for the bound cofactor and for many features of the structure consistent with the *Pf*GAPDH sequence; however, no density was apparent for amino-acid residues 193–198, consistent with the high thermal parameters (>50 Å^2) in this region of the final refined model. The final refinement statistics are also presented in Table 1.

Analysis of the final model of *Pf*GAPDH showed that 85% of residues were located in the most favourable regions of the Ramachandran plot, while nine residues were located in generously allowed regions. These nine residues were Ala151 from each chain, Val243 from each chain and His291 from the *R* chain. All nine residues are located on turns and have well defined electron density in the final Fourier map.

3.2. Structure description

The structures of the four subunits of *P. falciparum* GAPDH are essentially identical. One NAD$^+$ cofactor is bound per monomer and the average r.m.s.d. between monomers calculated over all atoms is 0.03 Å, calculated using *WHATIF* (Vriend, 1990). Analysis of differences between subunits is limited by the resolution (2.6 Å) and the maintenance of non-crystallographic symmetry restraints throughout refinement. The overall structure of the tetramer is shown in Fig. 3.

The electron-density map is consistent with the active-site cysteine (Cys153) being reduced, in contrast to the findings of Daubenberger and coworkers, who demonstrated by mass spectrometry that this cysteine was oxidized in *Pf*GAPDH (Daubenberger *et al.*, 2000). In their study, *Pf*GAPDH was purified and analysed in the absence of DTT, while the protein here was crystallized in a reducing environment. The active-site geometry is consistent with other GAPDH structures and the proposed catalytic mechanism (Harrigan & Trentham, 1974; Moras *et al.*, 1975; Polgar, 1975; Soukri *et al.*, 1989; Trentham, 1971).

The NAD$^+$ cofactor binds in an elongated cleft, with the nicotinamide moiety deeply embedded in the protein, consistent with other GAPDH structures. Three residues involved in cofactor binding are completely conserved (Asp35, Phe102 and Asn319), while Phe37 is unique to human and rabbit muscle GAPDH. Lys80 is conservatively substituted in *Pf*GAPDH from an arginine in human and rabbit muscle GAPDH; however, the hydrogen-bonding interaction is mediated through the backbone O atom of this residue.

Table 2
R.m.s.d. values (Å) of *Pf*GAPDH (*R* subunit) against GAPDH structures from other organisms.

The rabbit muscle GAPDH (1jox, *R* subunit), *E. coli* GAPDH (1gad, *P* subunit), Chinese lobster GAPDH (1szj, *A* subunit), *L. mexicana* GAPDH (1gyp, *A* subunit) and *T. brucei* GAPDH (1gga, *R* subunit) were all taken from the PDB. The r.m.s.d. values were calculated over all C^α atoms using the program *DALI* (Holm & Sander, 1995).

	Rabbit	*E. coli*	Lobster	*L. mexicana*	*T. brucei*
*Pf*GAPDH	0.8	0.6	0.7	1.3	1.3

3.3. Structure comparisons

Three-dimensional alignment of *Pf*GAPDH with other GAPDH structures illustrates that the structure is more similar to that of rabbit muscle, *E. coli* and lobster GAPDHs than to *L. mexicana* or *Trypanosoma brucei* GAPDHs (Table 2). For this group of structures, the amino acids around the active site are highly conserved both in sequence and in conformation. The *T. brucei* and of *L. mexicana* structures, however, showed some significant deviations, particularly near the adenosine moiety (residues 36–44 in the *L. mexicana* structure), where the backbone of these structures traces a different path (Fig. 4*a*).

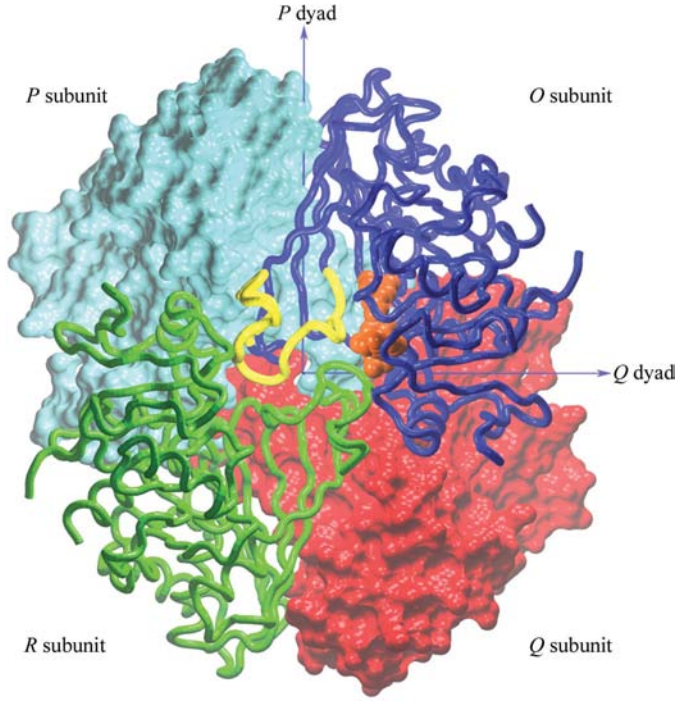

Figure 3
Structure of the tetramer as viewed down the *R* dyad. The *O* subunit is shown in blue and the *R* subunit in green. The *P* and *Q* subunits are shown as surfaces in cyan and red, respectively. The *S* loops meet in pairs (*O* with *R* and *P* with *Q*) about the *R* dyad. The *S* loop of the *O* subunit is shown in yellow. The NAD$^+$ of the *O* subunit is shown in orange. This figure was prepared using *VMD* (Humphrey *et al.*, 1996) and *RASTER3D* (Merritt & Bacon, 1997).

Multiple GAPDH sequence alignments point to regions of special interest in the *Pf*GAPDH structure (Fig. 5), some of which have already been highlighted in sequence (Daubenberger *et al.*, 2000) and structural (Cowan-Jacob *et al.*, 2003) analyses. In only one place does the *Pf*GAPDH sequence show an insertion or deletion with respect to other known structures and that is following residue 193, where two amino acids, -KG-, are added (shown aligned with rabbit muscle GAPDH in Fig. 1*b*). This insert, located in the so-called *S* loop, is accommodated as a type I β-turn, although the temperature factors for the main-chain atoms of residues 193–198 are >50 Å^2. The *B* values for atoms at the extremities of Lys194 and the conserved Lys197 are >70 Å^2. An unbiased fourfold-averaged omit map showing the inserted dipeptide (Fig. 1*a*) confirms the trace and the peptide-plane orientations in this region. The insert projects into the groove created by the *S* loop of the subunit (Fig. 2) and the C$^\alpha$ atom of Lys194 is only 8 Å away from the molecular *R* dyad. Structurally, the insert is as described above and not as two one-residue inserts as indicated in a proposed alternative alignment (Cowan-Jacob *et al.*, 2003). This is further illustrated in Figs. 1(*b*) and 4(*b*), which show structural alignments of the *O* polypeptides of *Pf*GAPDH and rabbit muscle GAPDH in the vicinity of the dipeptide insert.

Amino-acid sequence differences between human and trypanasomatid GAPDHs that have been exploited for designing specific trypanasomatid GAPDH inhibitors are (referring to *L. mexicana* sequence numbers) Val38 (Asn in human), Met40 (Pro), Leu114 (Val) and Val207 (Pro) (Suresh *et al.*, 2001). At these positions, the *P. falciparum* sequence (residues 34, 36, 101 and 192) is identical to the human sequence (Fig. 5). Furthermore, the trace of the polypeptide following Asp35 is similar to that seen in mammalian GAPDHs rather than that observed in *L. mexicana* or *T. brucei* GADPH (Fig. 4*a*). The course of the polypeptide in this region seems to be determined by the proline residue that follows the conserved aspartic acid.

Nearby, also in the *S* loop, Leu187 in the *Pf*GAPDH sequence replaces a Lys or an Arg in other sequences. The C$^\alpha$ atom of Leu187 is 3.5 Å from the *R*-dyad symmetry axis and the side chain extends outwards from the centre of the tetramer parallel to that dyad axis (Fig. 2). The structurally homologous lysine residue (Lys183) in rabbit muscle GAPDH (PDB code 1j0x) is similarly oriented with a symmetry-related partner parallel to the *R* dyad. However, there is considerable uncertainty in the position of the ε-amino group of Lys194 in *Pf*GAPDH owing to the high thermal parameters described above. Despite this uncertainty, the ε-amino group of the side chain of Lys194 can occupy the same general space near the *R* dyad as is occupied by its counterpart Lys183 in the rabbit muscle enzyme and by the homologous basic amino acid in GAPDH from other sources.

Relative to human GAPDH, which is of special interest when considering regions to target parasite-selective ligands, *Pf*GAPDH has just one other length mismatch and this is the insertion of Asp128 into a surface-exposed loop far from the active site, also with elevated temperature factors. Among the

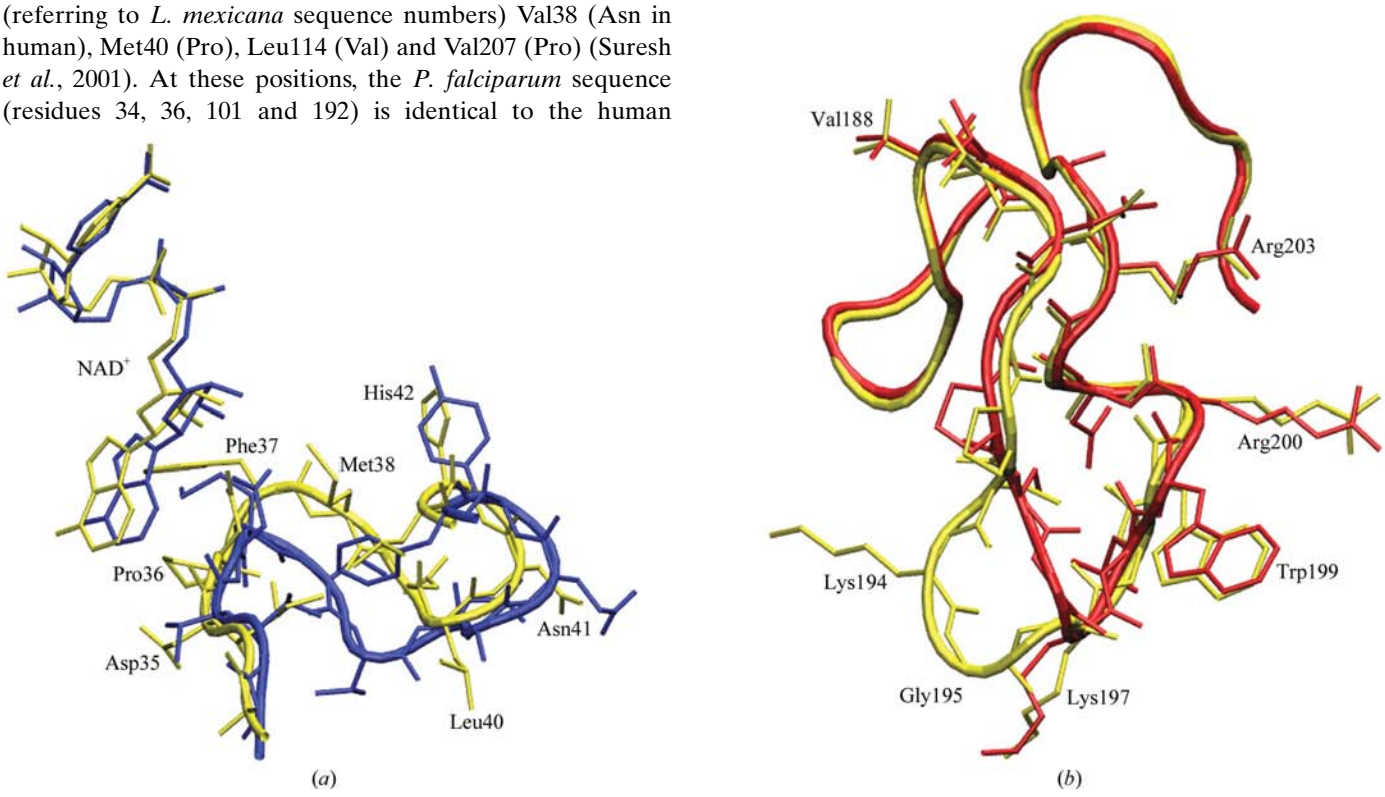

Figure 4
(*a*) Alignment of the loop near the adenosine moiety, which has been the target of extensive drug-design studies. *Pf*GAPDH (residues 32–43) is shown in yellow and the equivalent region from the GAPDH of another parasite, *L. mexicana* (residues 36–44), is shown in blue. In this area *Pf*GAPDH appears more like mammalian GAPDH. (*b*) Alignment of the *S* loops of *Pf*GAPDH (residues 188–203) shown in yellow and rabbit muscle GAPDH (residues 184–198) shown in red. The additional lysine (Lys194) and glycine residues (Gly195) are present as a dipeptide insertion. Residues are labelled using numbering from *Pf*GAPDH. This figure was prepared using *VMD* (Humphrey *et al.*, 1996).

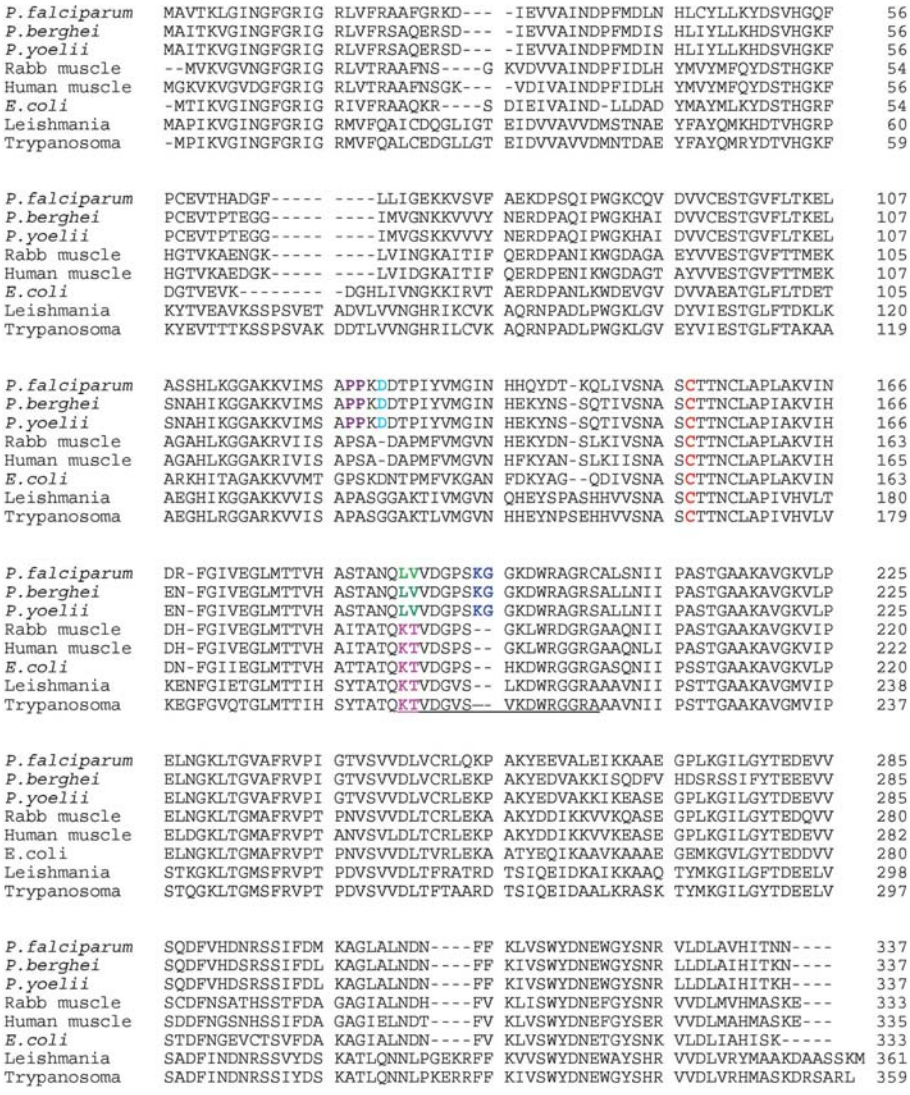

Figure 5
Sequence alignment of GAPDHs. Sequence alignment of GAPDH from the plasmodial species *P. falciparum*, *P. berghei* and *P. yoelii* and from human and rabbit muscle (equivalent to erythrocyte), *E. coli*, *L. mexicana* and *T. brucei*. The active-site Cys is depicted in red. The *S* loop is underlined. The -KG- insert, which is conserved across plasmodial species, is in blue. Also in the *S* loop, Leu187 and Val188 in the *Pf*GAPDH sequence (green) replace Lys and Thr (pink) in other species. The two consecutive proline residues (Pro124 and Pro125) of plasmodial GAPDH are depicted in purple. The inserted Asp128 is shown in aqua. The alignment was performed using *CLUSTALW*1.8 (Higgins *et al.*, 1994).

remaining sequence anomalies in *Pf*GAPDH, two (Met38 and Pro125) are of potential interest because of their proximity to the active site. Met38, which replaces Ile36 in the rabbit muscle sequence (Ile38, human), is close to the conserved aspartic acid Asp35 that is hydrogen bonded to the adenine ribose of the cofactor. Pro125 immediately follows the conserved proline residue Pro124 that is near the nicotinamide ribose and is thus also close to the reactive cysteine Cys153. The Pro-Pro sequence at this point may confer anomalous rigidity to this region of the *Pf*GAPDH active site.

3.4. Docking of FP

Docking studies showed that the groove around the *R* dyad between Lys126 and Lys194 (of the *O* and *R* subunits) could spatially accommodate haem binding. On inspection of the top 50 docked orientations of FP, it was found that two molecules of FP could bind in this groove when NAD^+ was either present or absent. In both cases, the docked orientations were clustered into two symmetry-related groups (two docked orientations representative of each of these clusters are depicted in Fig. 6) and were positioned near the NAD^+-binding site. When the cofactor was omitted, one of the propionic acid substituents of FP was placed in the NAD^+-binding site. All of the docked orientations are placed very close to the side chain of Lys126 (see Figs. 2 and 6), which was replaced by alanine during the docking procedure. Noteably, an excess of NAD^+ has been shown to inhibit the interaction of FP with *Pf*GAPDH (A. E. Alpyurek, N. Klonis, A. Adisa, J. F. Satchell, P. M. Colman and L. Tilley, unpublished data).

4. Discussion

Crystallization of *Pf*GAPDH was possible both with and without the histidine tag, but only in the latter case were the crystals suitable for obtaining a solution at moderate resolution. The largest structural divergence of *Pf*GAPDH from the many other GAPDH structures described to date is the addition of two amino acids, -KG-, following residue Ser193. This insert is within a renowned feature of the three-dimensional structure of GAPDH, the *S* loop, which is an elaboration of the polypeptide fold between residues 187 and 205 (*Pf* numbering). The loop contributes to a diverse set of interactions that include hydrophobic and ionic contacts around the *R* dyad and ionic contacts around the *P* dyad (Biesecker *et al.*, 1977; Skarzynski *et al.*, 1987). Furthermore, Pro192 forms one side of the binding pocket for the adenine ribose of the cofactor bound to a neighbouring subunit related by the *R* dyad. The dipeptide insert takes the form of a reverse turn which projects across the cleft created by the *S* loop and towards residue Lys80 of the same subunit. Residues Lys80 and Lys194 are at the entrance to one end of the cleft that harbours the NAD^+.

The insert is also adjacent to amino acids 37–40 of a neighbouring subunit in the tetramer (related by the *R* dyad). In *Pf*GAPDH, we have modelled a water molecule in the interface between the peptide carbonyl O atoms of residues

193 and 37 of the *O* and *R* subunits and Met38 packs close to Pro192 across the interface. Furthermore, as noted above, the insert is also close to the *R*-dyad axis and thus the inserts are found paired approximately 16 Å apart (see Fig. 2).

Electron density for the side chain of Lys194 is weak and is attributed to high thermal parameters, which are also evident in the backbone in this region. The C^{α} atom of Lys194 (*O* subunit) is almost exactly equidistant (13.5 Å) from the N7 atom of the adenine moieties in the cofactors bound to the *O* and *R* subunits. A homology model of *Pf*GAPDH (Cowan-Jacob *et al.*, 2003) has placed the side chain of Lys194 in the vicinity of the cofactor bound to the neighbouring subunit. Based on the omit map described above, the $C^{\alpha}-C^{\beta}$ bond projects the side chain in the direction of the NAD^{+} associated with the same subunit. Even there, if fully extended, the closest approach of Lys194 to this adenine ring is 7.5 Å.

Differences between the structures of human and trypanosomatid GAPDHs have been exploited for designing specific trypanasomatid GAPDH inhibitors (Suresh *et al.*, 2001). These substitutions are clustered around the adenosine-binding pocket and have inspired a number of selective designed adenosine analogues. Remarkably, at every one of these four positions the plasmodial GAPDH sequence is identical to the human sequence and the backbone traces are also very similar. These particular sites do not seem to offer possibilities for guiding the design of *P. falciparum*-specific ligands; however, the differences we have identified and discussed above also map to the vicinity of the adenosine pocket (Fig. 2).

Designing specific inhibitors of *Pf*GAPDH that would target the inserted dipeptide is complicated by the apparent disorder of the insert, especially of the side chain of Lys194. However, based on an extended conformation of Lys194, adenosine analogues derivatized at N7 with acidic functionality could potentially engage Lys194. The adenosine analogues described by Bressi *et al.* (2001) project substituents from Asn6 towards the novel Leu114 in *L. mexicana* GAPDH referred to above. Thus, it may be possible to generate similar analogues that specifically target *Pf*GAPDH but do not inhibit RBC GAPDH. Recently, inhibitors that target the binding sites of glyceraldehyde-3-phosphate and NAD^{+} have been described that are selective for the enzyme from *T. cruzi* over rabbit muscle (Ladame *et al.*, 2005).

It is interesting to consider which structural features of *Pf*GAPDH might result in preferential binding of FP to a site that inhibits the enzyme activity, thereby allowing *Pf*GAPDH to act as a metabolic sensor of FP levels. As the dipeptide insert in the *S* loop is the major structural feature distinguishing mammalian GAPDH from *Pf*GAPDH, it is an attractive candidate for the structural correlate of the selective inhibition of *Pf*GAPDH by FP. In this context, it is interesting to note that flexible loops located near active sites in other enzymes have been shown to be important in controlling substrate binding or allosteric activation (Buchbinder & Fletterick, 1996; Kai *et al.*, 2003; Schulte & Hengstenberg, 2000). Formed at the interface between two subunits is a long groove (see Fig. 2). Along the base of this cleft are located Ala184, Asn185 and Leu187 (close to the *R* dyad) and Ser213, Gly215 and Lys218 (about 20 Å from the *R* dyad). The walls of the groove are formed by the *S* loop (connecting the two floor segments) on one side and an adenosine moiety on the other. The disorder in the insert frustrates any rigorous *in silico* docking analysis; however, one or two FPs could be accommodated in this groove, thereby affecting access to the active site (Fig. 6).

The chemical character of this putative binding site is altered in *Pf*GAPDH with respect to human GAPDH by the substitution of Leu187 with Lys (in the human enzyme) and Val188 with Thr, both in the floor of the crevasse, and by the -KG- insert in the *S* loop, this notwithstanding our observation above that the ε-amino group of Lys194 could occupy a similar space to that of Lys183 in the rabbit muscle enzyme. It is possible that the repositioning of the positive charge (on lysine) might promote an interaction with a propionic acid side chain of FP and facilitate binding of the hydrophobic porphyrin ring within the crevasse. In this location, FP would also significantly hinder access to the NAD^{+}-binding site.

This suggestion regarding the binding site for FP would be most appropriately tested by studying *Pf*GAPDH crystals complexed with FP. Attempts to generate crystals of the *Pf*GAPDH–FP complex or to soak FP into the *Pf*GAPDH crystal have not yet proven successful, owing in part to the tendency of FP to slowly precipitate from aqueous solutions at neutral pH. Mutagenesis studies swapping the *S* loops

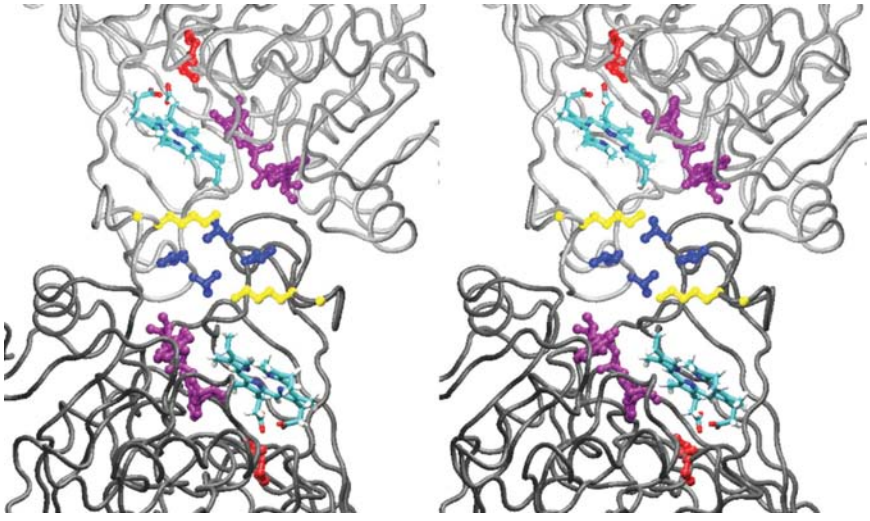

Figure 6
Stereo representation of the interface of the *O* and *R* subunits (shown in grey) as seen down the *R* dyad. Residues are coloured as in Fig. 2. Two representative (symmetry-related) docked orientations of FP are shown, with C atoms in cyan, N atoms in blue, O atoms in red and H atoms in white. This figure was prepared using *VMD* (Humphrey *et al.*, 1996) and *RASTER3D* (Merritt & Bacon, 1997).

of plasmodial and mammalian GAPDH are in progress to establish whether FP does indeed exert its inhibitor activity by binding to this region of the protein.

This work was supported by the National Health and Medical Research Council, Australia. We thank Dr J. Gulbis of The Walter and Eliza Hall Institute of Medical Research for valuable discussions and assistance with the refinement of the structure. X-ray data collection was performed at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. We are grateful to the machine and beamline groups whose outstanding efforts have made these experiments possible. We especially thank Dr Clemens Schulze-Briese for his support during our time at the beamline.

References

Aronov, A. M., Suresh, S., Buckner, F. S., Van Voorhis, W. C., Verlinde, C. L., Opperdoes, F. R., Hol, W. G. & Gelb, M. H. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 4273–4278.

Atamna, H., Pascarmona, G. & Ginsburg, H. (1994). *Mol. Biochem. Parasitol.* **67**, 79–89.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D. & Zardecki, C. (2002). *Acta Cryst.* D**58**, 899–907.

Biesecker, G., Harris, J. I., Thierry, J. C., Walker, J. E. & Wonacott, A. J. (1977). *Nature (London)*, **266**, 328–333.

Brady, R. L. & Cameron, A. (2004). *Curr. Drug Targets*, **5**, 137–149.

Bressi, J. C., Verlinde, C. L., Aronov, A. M., Shaw, M. L., Shin, S. S., Nguyen, L. N., Suresh, S., Buckner, F. S., Van Voorhis, W. C., Kuntz, I. D., Hol, W. G. & Gelb, M. H. (2001). *J. Med. Chem.* **44**, 2080–2093.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L.M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Buchbinder, J. L. & Fletterick, R. J. (1996). *J. Biol. Chem.* **271**, 22305–22309.

Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 3052–3054.

Buehner, M., Ford, G. C., Moras, D., Olsen, K. W. & Rossmann, M. G. (1974). *J. Mol. Biol.* **82**, 563–585.

Callens, M. & Hannaert, V. (1995). *Ann. Trop. Med. Parasitol.* **89**, *Suppl.* 1, 23–30.

Campanale, N., Nickel, C., Daubenberger, C. A., Wehlan, D. A., Gorman, J. J., Klonis, N., Becker, K. & Tilley, L. (2003). *J. Biol. Chem.* **278**, 27354–27361.

Connolly, M. L. (1983). *Science*, **221**, 709–713.

Cowan-Jacob, S. W., Kaufmann, M., Anselmo, A. N., Stark, W. & Grutter, M. G. (2003). *Acta Cryst.* D**59**, 2218–2227.

Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

Daubenberger, C. A., Poltl-Frank, F., Jiang, G., Lipp, J., Certa, U. & Pluschke, G. (2000). *Gene*, **246**, 255–264.

Eisenberg, D., Luthy, R. & Bowie, J. U. (1997). *Methods Enzymol.* **277**, 396–402.

Ewing, T. J. (1998). *DOCK* v.4.0. The University of California, California, USA.

Harrigan, P. J. & Trentham, D. R. (1974). *Biochem. J.* **143**, 353–363.

Higgins, D. G., Thompson, J. & Gibson, T. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Holm, L. & Sander, C. (1995). *Trends Biochem. Sci.* **20**, 478–480.

Humphrey, W., Dalke, A. & Schulten, K. (1996). *J. Mol. Graph.* **14**, 33–38.

Ishitani, R. & Chuang, D. M. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 9937–9941.

Itzstein, M. von, Wu, W. Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Van Phan, T., Smythe, M. L., White, H. F. & Oliver, S. W. (1993). *Nature (London)*, **363**, 418–423.

Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.

Kai, Y., Matsumura, H. & Izui, K. (2003). *Arch. Biochem. Biophys.* **414**, 170–179.

Kim, H., Feil, I. K., Verlinde, C. L., Petra, P. H. & Hol, W. G. (1995). *Biochemistry*, **34**, 14975–14986.

Kragten, E., Lalande, I., Zimmermann, K., Roggo, S., Schindler, P., Muller, D., van Oostrum, J., Waldmeier, P. & Furst, P. (1998). *J. Biol. Chem.* **273**, 5821–5828.

Ladame, S., Faure, R., Denier, C., Lakhdar-Ghazal, F. & Willson, M. (2005). *Org. Biomol. Chem.* **3**, 2070–2072.

Luthy, R., Bowie, J. U. & Eisenberg, D. (1992). *Nature (London)*, **356**, 83–85.

Meng, E. C., Shoichet, B. K. & Kuntz, I. D. (1992). *J. Comput. Chem.* **13**, 505–524.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Moras, D., Olsen, K. W., Sabesan, M. N., Buehner, M., Ford, G. C. & Rossmann, M. G. (1975). *J. Biol. Chem.* **250**, 9137–9162.

Moreno, E. & Leon, K. (2002). *Proteins*, **47**, 1–13.

Murshudov, G., Vagin, A. & Dodson, E. (1997). *Acta Cryst.* D**53**, 240–255.

Nagradova, N. K. (2001). *Biochemistry (Mosc.)*, **66**, 1067–1076.

Otwinowski, Z. & Minor, W. (1996). *Methods Enzymol.* **276**, 307–326.

Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.

Polgar, L. (1975). *Eur. J. Biochem.* **51**, 63–71.

Reed, J. C. (2002). *Nature Rev. Drug Discov.* **1**, 111–121.

Roth, E. Jr (1990). *Blood Cells*, **16**, 453–466.

Roth, E. F. Jr, Calvin, M. C., Max-Audit, I., Rosa, J. & Rosa, R. (1988). *Blood*, **72**, 1922–1925.

Sali, A. & Blundell, T. L. (1993). *J. Mol. Biol.* **234**, 779–815.

Schulte, D. & Hengstenberg, W. (2000). *Protein Eng.* **13**, 515–518.

Sirover, M. A. (1999). *Biochim. Biophys. Acta*, **1432**, 159–184.

Skarzynski, T., Moody, P. C. & Wonacott, A. J. (1987). *J. Mol. Biol.* **193**, 171–187.

Soukri, A., Mougin, A., Corbier, C., Wonacott, A., Branlant, C. & Branlant, G. (1989). *Biochemistry*, **28**, 2586–2592.

Souza, D. H., Garratt, R. C., Araujo, A. P., Guimaraes, B. G., Jesus, W. D., Michels, P. A., Hannaert, V. & Oliva, G. (1998). *FEBS Lett.* **424**, 131–135.

Suresh, S., Bressi, J. C., Kennedy, K. J., Verlinde, C. L., Gelb, M. H. & Hol, W. G. (2001). *J. Mol. Biol.* **309**, 423–435.

Trentham, D. R. (1971). *Biochem. J.* **122**, 71–77.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Vellieux, F. M., Hajdu, J., Verlinde, C. L., Groendijk, H., Read, R. J., Greenhough, T. J., Campbell, J. W., Kalk, K. H., Littlechild, J. A., Watson, H. C. & Hol, W. G. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 2355–2359.

Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.