Protein Science (1996), 5:52–61. Cambridge University Press. Printed in the USA.

The crystal structure of trypanothione reductase from the human pathogen *Trypanosoma cruzi* at 2.3 Å resolution

YIHONG ZHANG,[1,3] CHARLES S. BOND,[1] SUSAN BAILEY,[1,4] MARK L. CUNNINGHAM,[2] ALAN H. FAIRLAMB,[2] AND WILLIAM N. HUNTER[1]

[1] Department of Chemistry, University of Manchester, Manchester, M13 9PL, United Kingdom
[2] Department of Medical Parasitology, London School of Hygiene and Tropical Medicine, Keppel St, London, WC1E 7HT, United Kingdom

(RECEIVED August 8, 1995; ACCEPTED October 26, 1995)

Abstract

Trypanothione reductase (TR) is an NADPH-dependent flavoprotein unique to protozoan parasites from the genera *Trypanosoma* and *Leishmania* and is an important target for the design of improved trypanocidal drugs. We present details of the structure of TR from the human pathogen *Trypanosoma cruzi*, the agent responsible for Chagas' disease or South American trypanosomiasis. The structure has been solved by molecular replacement, using as the starting model the structure of the enzyme from the nonpathogenic *Crithidia fasciculata*, and refined to an *R*-factor of 18.9% for 53,868 reflections with $F \geq \sigma F$ between 8.0 and 2.3 Å resolution. The model comprises two subunits (968 residues), two FAD prosthetic groups, two maleate ions, and 419 water molecules. The accuracy and geometry of the enzyme model is improved with respect to the *C. fasciculata* enzyme model. The new structure is described and specific features of the enzyme involved in substrate interactions are compared with previous models of TR and related glutathione reductases from human and *Escherichia coli*. Structural differences at the edge of the active sites suggest an explanation for the differing specificities toward glutathionylspermidine disulfide.

Keywords: disulfide oxidoreductase; FAD; glutathione reductase; NADPH; trypanosomiasis; trypanothione reductase; X-ray structure

Trypanothione reductase (EC 1.6.4.8) is a dimeric disulfide oxidoreductase unique to parasitic flagellated protozoa of the order Kinetoplastida, suborder Trypanosomatina (Fairlamb & Cerami, 1992). Some trypanosomatids are the causal agents of several diseases of medical and veterinary importance in the tropical and subtropical regions of the world. Of particular note is Chagas' disease, or American trypanosomiasis, of which the etiological agent is *Trypanosoma cruzi*. This disease is endemic in South and Central America (World Health Organisation, 1991; Moncayo, 1993) and is now also found in the United States due to contaminated blood supplies (Kirchhoff, 1993). Despite some efforts, safe and effective treatments for *T. cruzi* infection have not yet been identified (de Castro, 1993). The characterization of trypanosome metabolism has suggested biochemical pathways sufficiently different from human metabolic pathways where chemical intervention might prove a viable route to the control of infection. One such pathway uses TR in the maintenance of thiol-redox balance and the protection against oxidative stress.

Trypanosomes use the peptide-polyamine conjugate metabolites, trypanothione ($(N^1,N^8$-*bis*(glutathionyl)spermidine) and N^1-glutathionylspermidine, to regulate an intracellular reducing environment and to protect biological molecules from the oxidizing byproducts of aerobic metabolism (Fairlamb, 1989;

Reprint requests to: William N. Hunter, Department of Chemistry, University of Manchester, Oxford Rd., Manchester, M13 9PL, United Kingdom; e-mail: wnh@spec.ch.man.ac.uk.

[3] Present address: N.I.A.I.D./N.I.H. Rockville, Maryland 20852.

[4] Present address: Daresbury Laboratory, Daresbury, Warrington, WA4 4AD, United Kingdom.

Abbreviations: TR, trypanothione reductase; GR, glutathione reductase; FAD, flavin adenine dinucleotide; GSH and GSSG, glutathione (L-γ-glutamyl-L-cysteinylglycine) and the disulfide, respectively; GspdSH and $[GspdS]_2$, N^1-glutathionyl spermidine and its disulfide, respectively; $T[SH]_2$ and $T(S)_2$, N^1,N^8-*bis*(glutathionyl)spermidine and its disulfide, respectively; SA, simulated annealing; NCS, noncrystallographic symmetry; σ, standard deviation; *B*-factor, crystallographic isotropic temperature factor; *R*-factor, crystallographic residual; *F*, structure factors; F_{obs}, experimentally observed structure factors; F_{calc}, calculated structure factors; α_{calc}, calculated phases; *I*, diffraction intensity; RMSD, RMS deviation.

Schirmer et al., 1995). These thiols are oxidized to the respective disulfides TS_2 and $[GspdS]_2$ in the course of this protective role. The flavoenzyme trypanothione reductase maintains the reducing environment in the protozoan cell by reduction of the disulfides to maintain high levels of thiols. Interestingly, a third metabolite, involved in this pathway in *T. cruzi*, termed homotrypanothione (N^1,N^9-bis(glutathionyl)aminopropylcadaverine), has recently been characterized (Hunter et al., 1994).

Humans utilize a different, although related, system that has been investigated thoroughly by a range of biophysical techniques (Meister, 1989; Schirmer et al., 1995). The protective thiol is glutathione (L-γ-glutamyl-L-cysteinylglycine), which is oxidized to glutathione disulfide and the cognate enzyme is glutathione reductase (EC 1.6.4.2).

TR and GR are closely related enzymes. They are homodimeric, with a subunit molecular weight of approximately 52 kDa, and they catalyze the transfer of electrons from NADPH to their specific substrates via an FAD prosthetic group and a redox active cysteine disulfide (Ghisla & Massey, 1989; Borges et al., 1995). Although TR's and mammalian GR's share approximately 40% sequence identity and the residues involved in catalysis are conserved, the enzymes are mutually exclusive with respect to disulfide substrate specificity (Fairlamb & Cerami, 1992). The difference in substrate specificity, the absence of trypanothione and TR in mammalian cells, and the observation that many chemicals with anti-trypanosomal properties show activity against trypanothione metabolism, including therapeutic arsenicals and antimonials (Cunningham et al., 1994; Schirmer et al., 1995), make this aspect of trypanosomatid biochemistry an excellent target for the development of new drugs either through screening methods or a rational design process (Verlinde & Hol, 1994). Access to accurate structural detail of TR and an understanding of structure-function relationships is required in support of such a process.

Our initial crystallographic studies utilized TR isolated from *Crithidia fasciculata* clone H56 (Hunter et al., 1990; 1992; Bailey et al., 1993; 1994). A number of complications emerged from the early studies with this enzyme. The difficulties associated with the growth and culturing of this nonpathogenic trypanosomatid, combined with the low yields of enzyme that could be isolated and purified, were a limiting factor in terms of quantity of material for crystallographic study. The presence of at least three isozymes was also a complicating factor (Field et al., 1992). In addition, we decided that it was a priority to characterize and work with the correct target, namely a TR from a human pathogen. The sequence identity between the enzymes isolated from *C. fasciculata* and *T. cruzi* is around 69% (Aboagye-Kwarteng et al., 1992). Although the residues important for catalysis and those implicated in substrate binding are conserved, only an accurate structure would help identify any important differences. For inhibitor design or identification studies, it is desirable to obtain as accurate a structure as possible, and we sought to improve our TR model. This led us to consider a recombinant source of TR. An *Escherichia coli* overexpression system for TR, now from the causal agent of Chagas' disease, *T. cruzi*, has been constructed (Borges et al., 1995) and provides large quantities of enzyme for our studies. Access to such quantities of TR enabled a rigorous screening of crystallization conditions and this has produced tetragonal crystals that diffract to high resolution (Zhang et al., 1993). We now report the crystal structure analysis of this TR and present some comparisons with the *C. fasciculata* TR and related glutathione reductases. While our study was under way, Lantwin and co-workers (1994) published details of a monoclinic crystal form of *T. cruzi* TR. The low resolution of that structure (e.g., see Table 1 in Lantwin et al. [1994]), limits the usefulness of their model for comparative purposes. The structure we now report is a significant improvement in resolution for *T. cruzi* TR.

Results and discussion

Quality of the structure

The final model comprises 968 residues, 7,451 non-hydrogen protein atoms, Ser 3A to Pro 487A of the first subunit, Ile 5B to Pro 487B of the second subunit. In addition, there are two FAD molecules (106 atoms), two maleate ions (16 atoms), and 419 solvent positions modeled as oxygen atoms. The first two residues in subunit A, the first four residues in subunit B, and the last six in both subunits were excluded from the structure because there was no convincing electron density to model. The *R*-factor is 18.9% for the 53,868 reflections (91.5% of the possible unique data) with $F_{obs} \geq \sigma F$ in the resolution range 8.0–2.3 Å. A Luzzati plot (Luzzati, 1952; not shown) suggests an error level of 0.25–0.3 Å. An example of electron density is presented in Figure 1. At a level of 0.7σ, the final $2F_{obs} - F_{calc}$ map shows continuous electron density for all main-chain atoms. Density is fragmented or absent for 34 side chains of residues located primarily on loops at the molecular surface or at the termini. These residues are mainly lysines (12), arginines (8), and glutamic acids (8), and are remote from the active sites. The electron and difference density maps suggest multiple conformations for the following arginines, 75A, 90A, 151A, 223A, 75B, 358B, His 359A, Lys 486B, and, in both subunits, Trp 22. This last example will be discussed in detail later. No attempt was made to incorporate multiple conformations in the refinement.

The real-space *R*-factor (Jones et al., 1991) ranges from 15.7 to 43.7%, with an average of 22.8%. The mean *B*-factors are 25.2 and 33.8 $Å^2$ for subunits A and B, respectively, with an RMS of 1.8 $Å^2$. Mean values for side-chain atoms are 30.2 and 39.0 $Å^2$, with RMS values of 4.4 and 4.1 $Å^2$ per subunit. The average *B*-factor for FAD atoms of subunit A is 13.4; for subunit B, 23.6 $Å^2$; and for the maleate ions and solvents, 53.9 and 44.9 $Å^2$, respectively. Figure 2 presents both thermal parameters and the real-space *R*-factor per residue for each subunit. The subunits have a similar *B*-factor profile, although the increased thermal parameters associated with subunit B suggest a larger degree of static disorder of that subunit. Some of the lowest thermal parameters are observed in and around the catalytic center and dinucleotide-binding regions, implying that rigidity may be important in the redox chemistry of these enzymes (see Bailey et al., 1994).

The model has RMS deviations from ideality of 0.01 Å in bond lengths, 2.6° in bond angles, 1.4° for improper angles, and 24.1° for dihedral angles. The Ramachandran plot (Ramachandran & Sasisekharan, 1968) shows 91.1% of non-glycine residues in most favored regions with 19 residues in the left-handed helical section. A further 8.8% of these residues are in allowed regions. A notable outlier in the Ramachandran plot is Phe 46 in both subunits. The electron density associated with this residue in subunit A is presented in Figure 1. As with the corresponding residue, Tyr 45 in *C. fasciculata* TR, a hydrogen bond

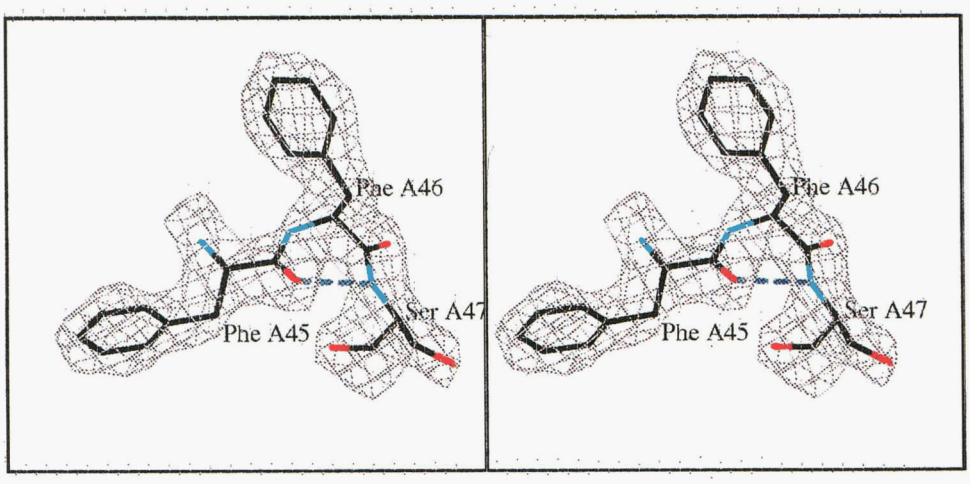

Fig. 1. An example of the final $2F_{obs} - F_{calc}$ electron density map. The map, thistle-colored chicken wire, is contoured at a 1.5σ level. Atomic positions are colored according to atom type: C, black; N, cyan; O, red. Blue dashed lines represent a hydrogen bond. This figure, like Figures 2 and 8, was prepared using the program O (Jones et al., 1991).

with a preceding residue holds this peptide in a strained conformation (Bailey et al., 1994). In excess of 90% of side-chain dihedral angles are within 2.5σ of the ideal *gauche*$^{\pm}$ and *trans* conformations. Noncrystallographic symmetry is indicated by least-squares superposition of the two subunits that constitute the asymmetric unit. This gives an RMSD of 0.31 Å for main-chain atoms and 0.82 Å for all atoms in common.

A least-squares fit of *C. fasciculata* TR on *T. cruzi* TR provides an RMSD of 0.67 Å for 962 Cα atoms, a clear indication of main-chain similarity. However, the accuracy of the TR structure has been improved over our previous studies with the *C. fasciculata* enzyme (Bailey et al., 1994). We have extended the resolution from 2.6 to 2.3 Å, improved the completeness of the data, and incorporated about 19,000 more reflections into the refinement calculations of an asymmetric unit of similar size. The geometry of the model has been improved. For example, the RMSDs from ideality on bond lengths and angles reduced from 0.02 Å and 3.4° for *C. fasciculata* TR to 0.01 Å and 2.6° for the *T. cruzi* enzyme. Moreover, these improvements relate to the actual biochemical target for drug discovery, TR from a human pathogen.

Overall, domain, and secondary structure

The TR dimer has approximate dimensions of 95 × 66 × 93 Å (Fig. 3). The disulfide substrate-binding cleft, a distinctive V-shaped crevice, is formed at the dimer interface, as will be detailed later. The assignment of domain structure of the disulfide oxidoreductase family of enzymes has been set by the work on GR and there have been changes over the years, (see Mittl & Schulz, 1994). A four-domain model was convenient because it had domains consecutive along the amino acid chain. However, because domains can be inserted in other domains (Schulz, 1992), a three-domain model has been adopted. Swindells

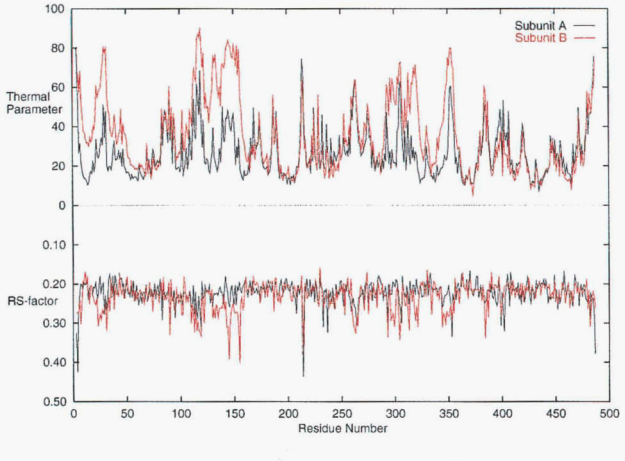

Fig. 2. Plots of thermal parameters (top), averaged over all atoms, and the real space *R*-factor (bottom) per residue. Values for subunit A are depicted with a black line, those for subunit B with a red line.

Fig. 3. Ribbon drawing of the trypanothione reductase dimer. Selected elements of secondary structure are shown as spirals for α and 3_{10} helices, arrows for β-strands. Domains are colored as follows: yellow, FAD-bonding domain (I); green, NADPH-binding domain (II); red, interface domain (III). This diagram, together with Figures 4, 6, 9, and 10, was prepared with MOLSCRIPT (Kraulis, 1991).

(1995a, 1995b) has recently described a method for domain characterization based on the identification of hydrophobic cores and the results are consistent with a three-domain structure for GR. We now adopt the three-domain structure for TR and assign the domains on the basis of comparison with GR. Domain I comprises two segments of the chain, namely residues 1–163 and 293–352. Domains II and III are formed from residues 164–292 and 353–497, respectively. Domains are shown in Figure 4 and Kinemage 1. Domain I binds the FAD, domain II binds NADPH, and domain III forms the interface with the partner subunit. Domains I and II each have β-α-β-α-β motifs, which will be discussed later.

The dimer is formed through domain III interactions with the partner subunit domain III. There are 25 direct (mostly side-chain–main-chain interactions) and 17 solvent-mediated hydrogen bonds between them. To complement these interactions, there are 373 van der Waals contacts less than 4.0 Å. Residue C444 is worthy of comment. This residue is positioned close to the partner subunit C444. The electron density unambiguously indicates that there is no disulfide link rather than that the residues are in the reduced form. A similar observation has been made with respect to the *C. fasciculata* TR (Bailey et al., 1994).

Solvent accessibility calculations (Lee & Richards, 1971) indicate that dimer formation buries about 6,400 Å^2 of protein surface. According to Chothia and Jain (1975) this represents a considerable hydrophobic free energy contribution of around 160 kcal mol^{-1} to dimer stabilization.

The *T. cruzi* TR secondary structure is about 28% α-helix, 4% 3_{10}-helix, 24% β-strands, and 10% turns. Figure 5 shows the amino acid sequence with secondary structure assignments that have been made using the hydrogen bonding criteria of Kabsch and Sander (1983), the analysis software PROMOTIF (G. Hutchinson & J. Thornton, pers. comm.), and by visual checks. There are some slight differences compared to the previous assignments on the *C. fasciculata* TR (Bailey et al., 1994) and in part this is due to an improvement in the resolution and accuracy of the structure. In particular, we have assigned an additional small segment of two-stranded antiparallel β-sheet. This is formed by residues 163–165 and 288–290. There are thus six β-sheets labeled A–F. Sheets A and D are five- and four-stranded parallel sheets, respectively. The remaining four sheets are antiparallel, comprising three (B), two (C), three (E), and seven (F) strands.

Fig. 4. Monomer of trypanothione reductase. The FAD group bound to domain I is depicted in ball-and-stick mode. View into the active site is similar to that employed further on. Note that the partner subunit contributes to the active site with the interface domain.

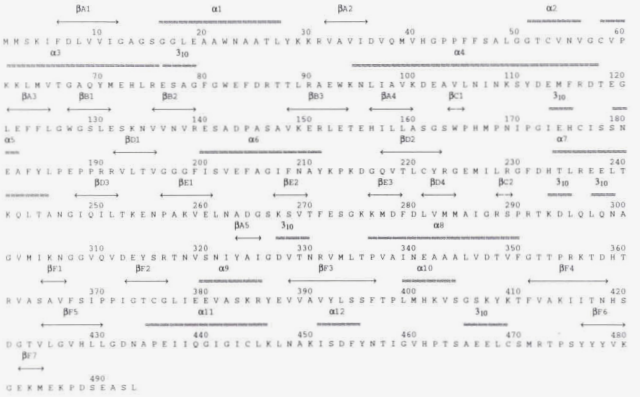

Fig. 5. Amino acid sequence of *T. cruzi* trypanothione reductase with assigned secondary structure. ≈≈ indicates α and 3_{10} helices; ←→, β-strands. Naming of sheets and numbering of α-helices is also given.

Disulfide substrate-binding site

The catalytic center involving the redox active disulfide (Cys 53 with Cys 58) and active site base (His 461) lie at the bottom of a cleft formed by domains I and III (of the partner subunit; Fig. 6, Kinemage 2). The cleft has dimensions of approximately 20 Å in length, 15 Å deep and wide. A number of previous publications have addressed the question of why TR and GR are mutually exclusive with respect to substrate specificity (reviewed by Fairlamb & Cerami, 1992). High-resolution crystal structures of human erythrocyte GR and the complex with GSSG are available (Karplus & Schulz, 1987, 1989) for comparative purposes. Analysis of the residues involved in binding GSSG reveal five nonconservative changes between GR and TR; a Glu, Trp, Ser, Met, and Ala in TR replace Ala 34, Arg 37, Ile 113, Asn 117, and Arg 347, respectively in human GR. Site-directed mutagenesis experiments have confirmed a contribution from these residues in conferring substrate specificity (Bradley et al., 1991; Henderson et al., 1991; Sullivan et al., 1991).

The comparison of enzyme–substrate complexes characterized crystallographically provides the most reliable detail (Bailey et al., 1993) and indicates that specificity is determined by a combination of steric and electrostatic factors that can be summarized as follows. TR has an enlarged active site in comparison with GR in order to accommodate a larger substrate. The human GR active site is too small for $T(S)_2$ or $[GspdS]_2$. The difference in size of the active sites is due to a combination of domain alterations and specific amino acid variations (Hunter et al., 1992). The overall negative charge of the TR active site complements the positively charged substrates and a hydrophobic patch formed by a Trp and a Met serve to bind the aliphatic moiety of substrate. The positive GR active site binds the negative GSSG.

The electrostatic properties in and around the disulfide substrate-binding active sites of TR and GR can be visualized

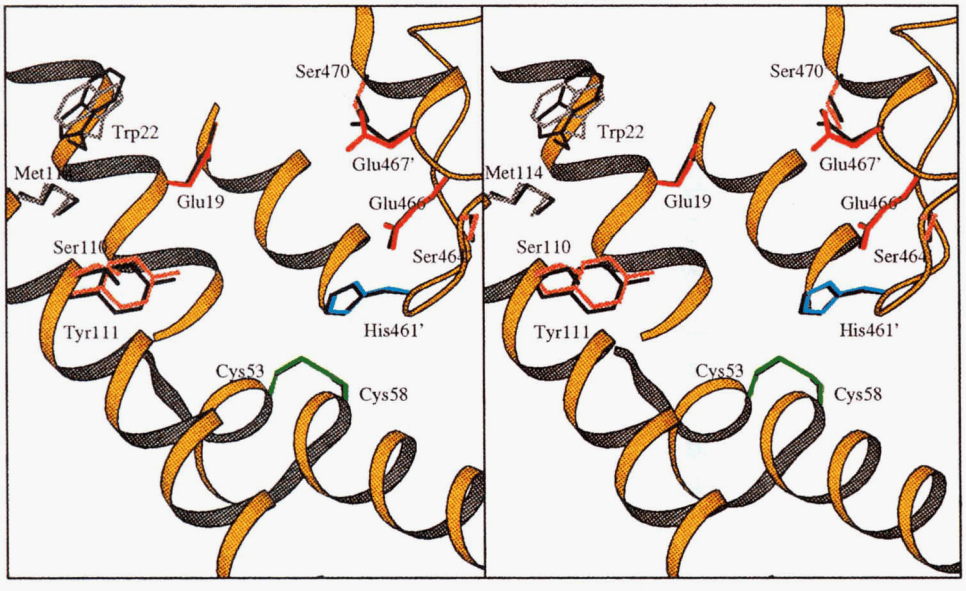

Fig. 6. View into the disulfide substrate binding site and catalytic center. Residues implicated in binding of substrate or involved in catalysis are shown and labeled. For comparative purposes, conserved *C. fasciculata* TR side chains are drawn in black.

using the GRASP program (Nicholls, 1993; Fig. 7). There is a marked difference between the human GR and TR. The active site charges are, as discussed above, designed for and complement the charge of the respective substrates. The charge complementarity is not limited to the actual binding site, but rather extends out and around the clefts. This suggests that longer-range electrostatic attraction and repulsion contributes to the enzyme's specificity for its own substrate. Such long-range electrostatic effects will need to be considered in inhibitor design.

A comparison of TR from *C. fasciculata* in a monoclinic (Kuriyan et al., 1991) and a tetragonal crystal form (Bailey et al., 1994) has identified multiple conformations for the Trp 21 and Met 113 that form the hydrophobic patch at the rim of the active site cleft (Bond et al., 1995). This heterogeneity was not identified from the individual analyses, but only became apparent upon detailed comparison. We note that, in the *T. cruzi* TR, the electron density maps are of a quality to clearly indicate conformational flexibility of the Trp 22 side chain. (Fig. 8). Two conformations are clearly present. We have modeled the predominant conformation as judged by fit to electron density and omit maps, but decided that, without high-resolution data, we are not justified in refining both conformations. It will be instructive to see if the binding of substrate to the *T. cruzi* enzyme produces an ordering of this residue. In any case, conformational flexibility may be used by TR in the binding of cognate substrates and should be considered when characterising the steric and electrostatic properties of the active site.

Our structural characterization of *T. cruzi* TR and detailed comparison with the *C. fasciculata* enzyme in complex with the physiological substrate $[GspdS]_2$ suggests important differences on the rim of the active site involving residues on a helix, $\alpha 4$, (Figs. 6, 9, Kinemage 2). This part of the active site seems, so far, to have escaped attention in the literature.

One spermidine moiety of $[GspdS]_2$ is observed to interact with this helix in the *C. fasciculata* TR complex (Bailey et al., 1993), where Ser 109 can participate in a hydrogen bond with substrate. This serine is conserved in *T. cruzi* TR (Ser 110). Two other residues on this helix interact with van der Waals contacts and are also conserved, namely Ile 106 and Tyr 110, which correspond to Ile 107 and Tyr 111 in the pathogen TR. A third residue, which interacts with $[GspdS]_2$ in *C. fasciculata* TR, is Gly 112. In the *T. cruzi* structure, there is a nonconservative sequence alteration here to Glu 113. This introduces a negatively charged residue into an area where a positively charged part of one of the enzyme substrates could be located. There are three other changes along this helix to note. Gly 105 and Asp 108 in *C. fasciculata* TR are replaced by Asn and Lys, respectively. A polar substitution accompanies a switch from negative to positive charge. A switch from a positive Lys 101 to a negative Glu 102 also occurs. These differences between the pathogenic *T. cruzi* and *C. fasciculata* TR alter the size, shape, and electrostatic properties of this part of the active site. These residues are probably too distant from the binding site to influence enzyme–$T(S)_2$ interactions other than by long-range electrostatic effects. However, the situation alters when the other physiological substrate, $[GspdS]_2$, which is larger and more positive, is considered. The importance of these amino acid differences between the two TR molecules will only be known when the structure of the enzyme–$[GspdS]_2$ complex is determined. This work is in progress.

This area of the active site may also contribute to the specificities displayed by GR's. In contrast to the human enzyme, GR from *E. coli* shows measurable activity to process $T(S)_2$ (Henderson et al., 1991). The high-resolution crystal structure of the bacterial enzyme has recently been reported (Mittl & Schulz, 1994) and some comment on this point is made. The ability to accommodate $T(S)_2$ is attributed to changes of two residues that bind GSSG carboxylates in human GR. These are Arg 37 and Asn 117 of human GR, which become Asn 21 and Val 102 in the bacterial protein. These changes bring about a reduction in positive charge (Fig. 7) and an increase in space in the active site, both of which would promote $T(S)_2$ activity. *E. coli* GR

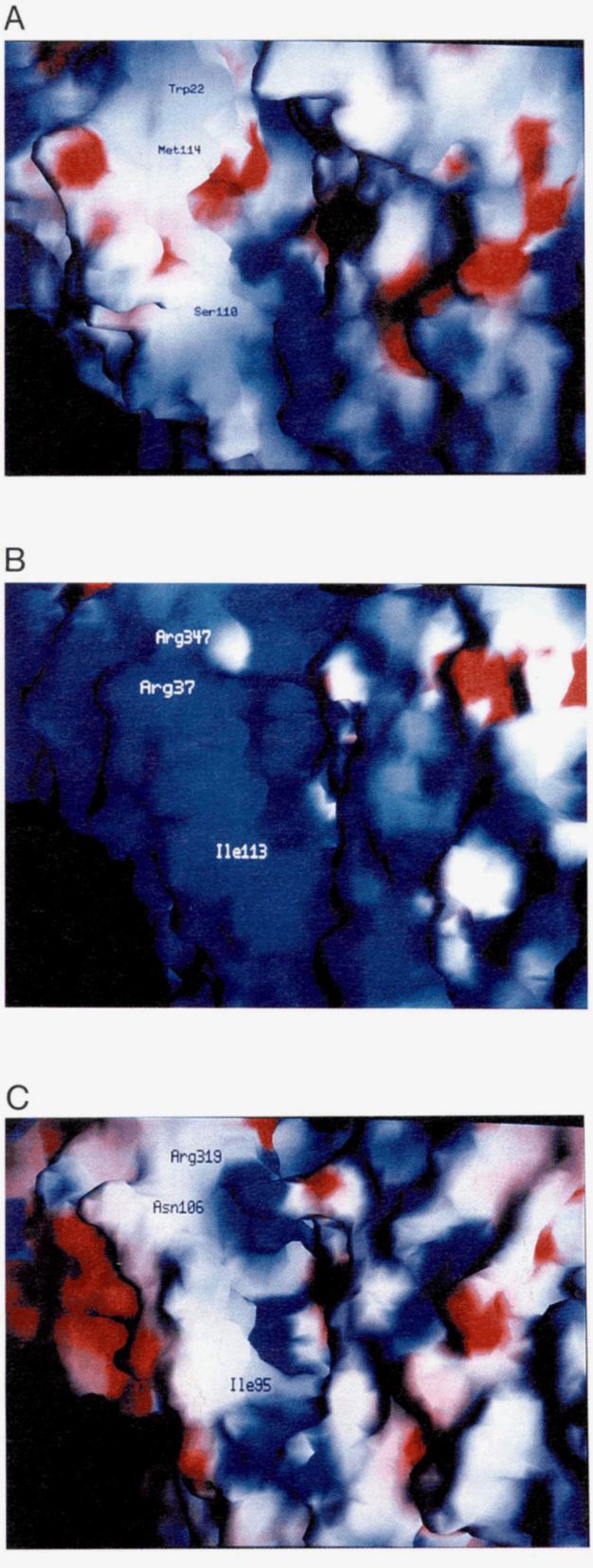

Fig. 7. Distribution of charge in and around the disulfide binding sites of (**A**) *T. cruzi* TR; (**B**) human GR; and (**C**) *E. coli* GR. Figures produced with GRASP (Nicholls, 1993) and colored as follows: white, neutral; red, negative charges; blue, positive charges. Coordinates for GR structures were retrieved from the Brookhaven Protein Data Bank.

cannot efficiently process $[GspdS]_2$ (Smith et al., 1995). We note that in *E. coli* GR, Arg 94 corresponds to Asn 106, located on $\alpha 4$, the helix providing residues that interact with $[GspdS]_2$. Arg 94 is positioned so that it might repulse the spermidine component of $[GspdS]_2$. Site-directed mutagenesis experiments would allow this point to be addressed.

In each disulfide substrate-binding site, a water molecule is located within hydrogen bonding distance of N^{ϵ} His 461 (Fig. 10, Kinemage 2). In active site A, the distance is 2.6 Å; in active site B, it is 2.9 Å. A water molecule has also been located at a similar position in human GR (Karplus & Schulz, 1989). The geometry mimics a catalytic triad and bears a striking similarity with the Glu-His-Water system identified the active site of O^6-methylguanine–DNA methyltransferase (Moore et al., 1994) and the Asp-His-Water in the *Thermus aquaticus* elongation factor catalytic site (Kjeldgaard et al., 1993). In TR, this may be the water that provides the proton to charge the active site base as a component of catalysis, although, as discussed above with respect to Trp 22, it will be instructive to see what happens when substrate is bound. The chemistry could be very similar to that utilized by serine proteases (Carter & Wells, 1988), except that the water molecule replaces the serine hydroxyl group in acting as a proton donor.

FAD and NADP binding sites

The nucleotides, FAD and NADP, bind to domains I and II, respectively (Kinemage 1). Each domain contains the Rossmann-fold motif commonly utilized in nucleotide binding (Rao & Rossmann, 1973; Schulz, 1992; Mittl et al., 1994). In domain I, this involves βA1-α1-βA2-α3-βA3; in domain II, βD1-α6-βD2-α7-βD3. Two fingerprints for binding the dinucleotide FAD, previously identified in other enzymes, are conserved. The first is the consensus sequence Gly-X-Gly-X-X-Gly at residues 12–17, which is part of the β-α-β-α-β Rossmann fold (Schulz et al., 1982; Schulz, 1992). X represents any amino acid. The second is the 11-amino acid stretch Thr-X-X-X-X-h-y-h-h-Gly-Asp at residues 317–327, where h indicates a small nonpolar amino acid, and y an aromatic residue. This motif was first identified from studies on rubredoxin reductase (Eggink et al., 1990). In the former case, Ser 15 and Gly 16 form hydrogen bonds with the pyrophosphate segment of the FAD. In the latter case, Asp 327 terminates the motif and hydrogen bonds to the ribityl chain. The preceding Gly 326 introduces a turn in the polypeptide chain that helps to accommodate the pyrophosphate group. As in *E. coli* GR, any residue other than glycine at this position would lead to steric clash between the side-chain atoms and the FAD pyrophosphate group.

Our earlier study using the TR from *C. fasciculata* determined details of NADP binding (Bailey et al., 1993). Four amino acids that were identified as important in this respect are conserved in the *T. cruzi* enzyme. These are Met 333, Gly 197, Arg 223, and Arg 229. Phe 199 in *T. cruzi* TR replaces a tyrosine in the *C. fasciculata* structure. This is the residue that, in the absence of NADPH, covers the isoalloxazine ring system of the FAD and protects it from the aqueous environment. As first shown for human GR, when NADPH binds, this side chain swings out of the way (Karplus & Schulz, 1989). The key oxidoreductase–NADP interactions, first reported for TR by Bailey et al. (1994) and subsequently by Lantwin et al. (1994), remain conserved. The similarities between human GR and TR with respect to

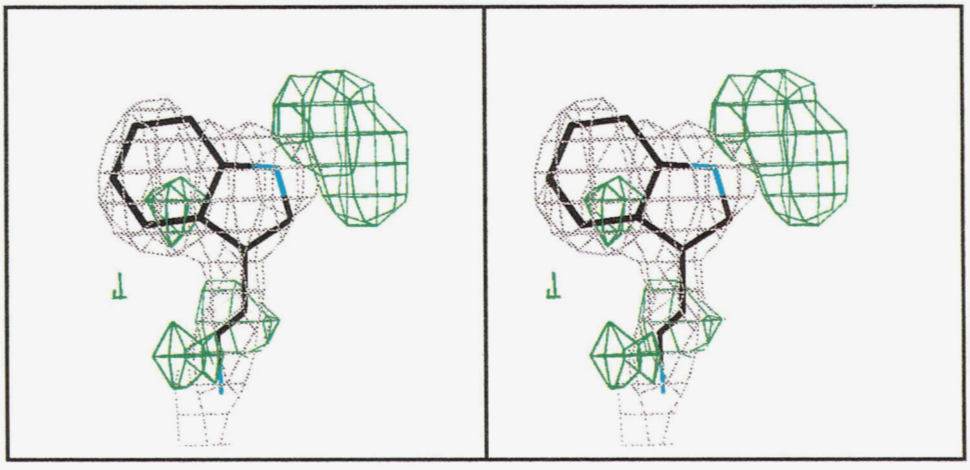

Fig. 8. Electron density ($2F_{obs} - F_{calc}$) and difference density ($F_{obs} - F_{calc}$) maps in the vicinity of Trp 22A. The same color scheme as used in Figure 1 is adopted with the addition that positive difference density is displayed in green. Both maps are contoured at a 1.5σ level.

NADPH interactions suggests that this component of TR is not suitable as a target area for drug design.

Summary

We have completed an accurate structure for TR from the pathogenic *T. cruzi*. The final model is an improvement on previously reported TR structures and represents a sound template for computational studies of inhibitor design. Our structural comparisons with the TR from the nonpathogenic *C. fasciculata* highlight differences between the enzymes located on a helix (α4) lining one side of the active site and confirms the importance of characterizing the structure of the actual target for drug discovery. The identification of structural differences in this part of the TR active site led us to compare the TR structures with the related glutathione reductase's from human and *E. coli*. There are indications that alterations on α4 may be responsible for the different specificities of these GR's toward glutathionylspermidine disulfide.

With access to large quantities of TR that provide good quality crystals, we are now in a position to pursue details of enzyme–substrate and enzyme–inhibitor complexes to provide more structural detail in support of a search for TR inhibitors. In addition, with a recombinant source of material, we plan to

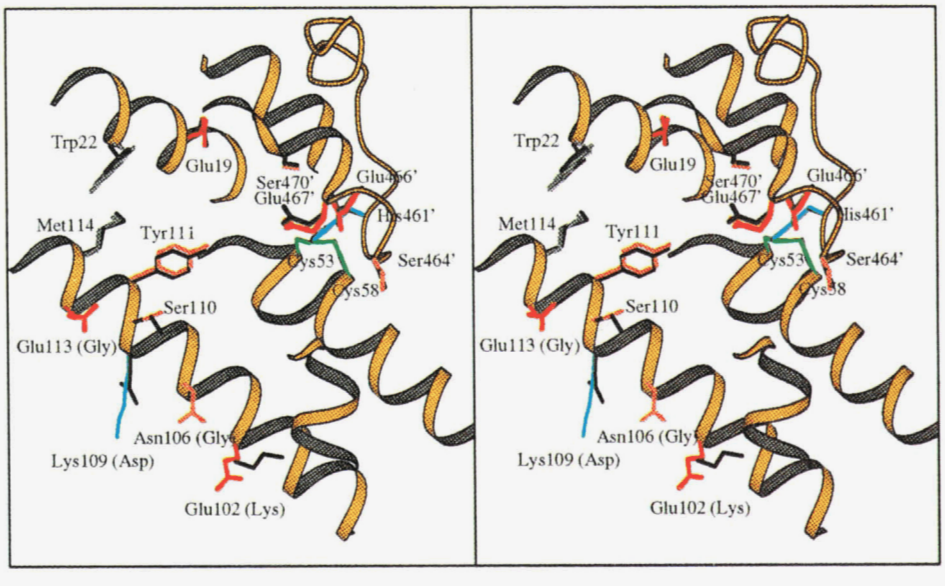

Fig. 9. Comparison of the disulfide substrate-binding sites of *T. cruzi* and *C. fasciculata* trypanothione reductases. The stretch of α-helix in the foreground (α4) harbors a number of sequence differences between the enzymes. Side chains in black correspond to the *C. fasciculata* structure with residue labels in brackets. *T. cruzi* structural components denoted as follows: red, acidic glutamic acids; cyan, basic lysines and histidine; grey, tryptophan and methionine; pink, tyrosine and serine; green, redox active disulfide.

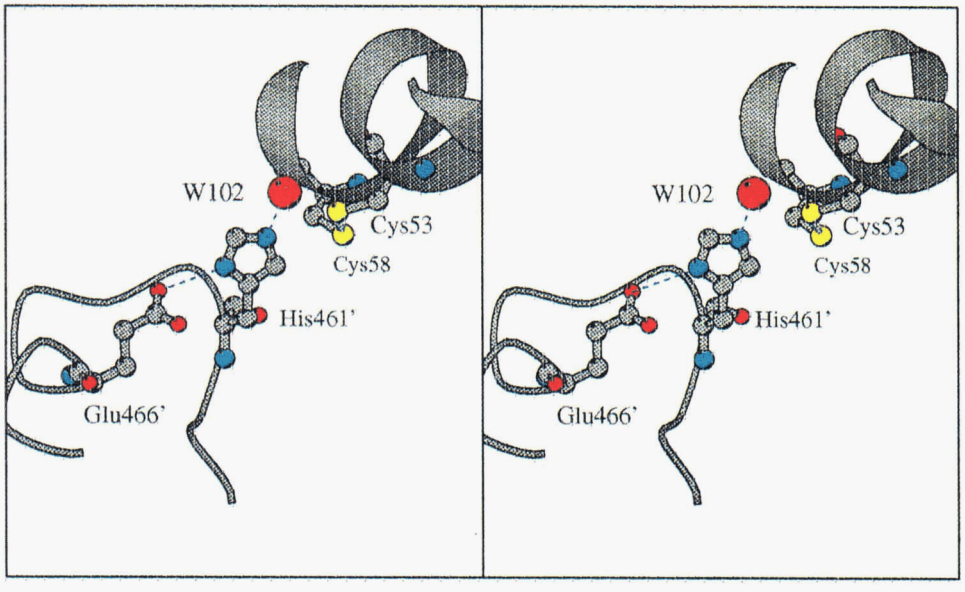

Fig. 10. Stereo view to show the triad Glu-His-Water system in active site A that may assist protonation of the active site base His 461. Blue dashed lines represent hydrogen bonds. Atoms are colored as follows: N, cyan; O, red; C, grey.

combine site-directed mutagenesis with structural analysis to fully delineate enzyme structure–activity relationships.

Materials and methods

Crystallization and data collection

Recombinant enzyme corresponding to the trypanothione reductase from the Silvio strain of *T. cruzi* was obtained as described by Borges et al. (1995). Crystallization was effected from a wide range of conditions using either hanging or sitting drops. For example, 7.5 mg/mL protein solution in 100 mM maleic acid buffer, pH 6.0, 0.2% (v/v) dioxane, 1% sodium dihydrogen phosphate, and approximately 10% PEG 8000 vapor-diffused against 20% PEG 8000 at 6–8 °C (Zhang et al., 1993). Yellow crystals in the form of elongated tetragonal prisms with dimensions of 0.25 × 0.25 × 1.5 mm grow in about 10 days. The crystals have unit cell dimensions of $a = b = 92.8$, $c = 156.6$ Å, and space group $P4_3$.

The data used for the molecular replacement calculations to solve the structure were recorded at room temperature on station PX9.5 at the Synchrotron Radiation Source, Daresbury. A MAR image plate was used in conjunction with X-rays of λ 1.07 Å and a 0.3-mm collimator. Four deliberately misset crystals provided 75,108 measurements of 29,021 independent reflections to 2.7 Å resolution. The *R*-merge in Laue group 4/m of 11.5% compared to 46% in 4/mmm and systematic absences indicated space group $P4_1$ or enantiomorph with a dimer in the asymmetric unit.

Station PX9.6 with a Rigaku RAXIS image plate detector, X-rays of λ 0.89 Å, 0.2-mm collimator, provided more data from a further three crystals, also at room temperature. First exposures displayed diffraction well beyond 2.0 Å, but the high-resolution diffraction quickly degraded due to radiation damage. With this in mind, and due to time constraints, data to only 2.3 Å were measured. We plan to extend the resolution at some later point. In total, the seven crystals provided 170,106 measurements of 55,392 independent reflections, redundancy 3.1, with *R*-merge 10.9% for $I \geq \sigma(I)$ data. This represents 94.1% of the data in the range 8.0–2.3 Å, of which 64.6% have $I \geq 3\sigma(I)$. Greater than 80% of the data in the outermost bin, 2.4–2.3 Å, is observed at the $\geq \sigma(I)$ level. This combined data set was used for refinement. MOSFLM (A.W. Leslie, pers. comm.) was used for all data processing.

Structure solution and refinement

Molecular replacement calculations (MERLOT; Fitzgerald, 1988) positioned the TR dimer from *C. fasciculata* (Bailey et al., 1994) in the *T. cruzi* unit cell (Zhang et al., 1993). The enzymes share almost 70% sequence identity, but, in order to reduce model bias, side-chain atoms beyond Cβ were removed from structure factor calculations. A self-rotation function indicated the presence of a noncrystallographic twofold within 2° of a primary axis in the *a*/*b* plane. Cross-rotation functions using a range of Patterson radii and different resolution limits identified the major peak at a 3–4σ level. Translation functions and two-dimensional *R*-factor searches, again with a variety of parameters, provided a consistent solution in space group $P4_3$. The molecular packing in the unit cell was reasonable, and rigid-body refinement of the dimer reduced *R* from 50.7% for 3σ data in the range 9.0–6.0 Å, to 42.9%. Two further rounds of rigid-body refinement, extending the data to 4.0 Å, gave $R = 46.6\%$. Note that, at this stage, only the truncated model, which was about 60% of the asymmetric unit, had been used in the structure factor calculations.

The first electron density ($2F_o - F_c$, α_{calc}) and difference density maps ($F_o - F_c$, α_{calc}) were calculated using data to 2.7 Å and inspected on an ESV30 graphics workstation (FRODO, Jones, 1978; P.R. Evans, pers. comm.). These maps guided model manipulation and incorporation of the side-chains to construct the first *T. cruzi* TR model. The amino acid sequence in

use at this point was that derived from the work of Sullivan and Walsh (1991). The correct sequence for our enzyme became available after refinement had been initiated and several residues had to be changed in the course of the study. The final model comprises the sequence according to the study by Borges et al. (1995).

A total of 21 rounds of model-map inspection, model manipulation, and refinement were then carried out with a gradual extension of resolution. The refinement incorporated slow-cooling simulated annealing methods combined with restrained least-squares methods using XPLOR (Brünger, 1992). Strict NCS restraints were employed for the SA calculations. Weaker NCS restraints were employed for the restrained least-squares until the limit of the data set was reached, at which point they were released. Main-chain and side-chain geometry of the model was monitored during refinement using PROCHECK (Laskowski et al., 1993). This indicated areas of the molecule where particular attention was required. Restrained isotropic thermal parameters were only refined when all data were included, and, likewise, waters were carefully added only when the 2.3-Å data had been incorporated in the refinement. In particular, the solvent structure near disordered side chains was treated conservatively. A number of the solvent positions refined to geometrically unacceptable positions out of electron density peaks, or produced very high thermal parameters. When this was observed, the water was removed from the model. Two maleate ions were identified. The CCP4 suite of programs were used for assorted crystallographic calculations (Collaborative Computational Project Number 4, 1994). Coordinates and observed structure factors have been submitted to the Brookhaven Protein Data Bank (Bernstein et al., 1977).

Acknowledgments

We thank our colleagues for encouragement, many discussions, and contributions to this project, and the staff at Daresbury laboratory, in particular Sean McSweeney and Pierre Rizkallah, for excellent support. We also thank Mark Swindells for investigating aspects of domain structure. Funded by the Wellcome Trust, BBSRC/EPSRC, and the Daresbury Synchrotron Laboratory. Y.Z. was supported by an Overseas Research Scholarship and the University of Manchester. W.N.H. thanks the Nuffield Foundation and the Leverhulme Trust for science research fellowships.

References

Aboagye-Kwarteng T, Smith K, Fairlamb AH. 1992. Molecular characterisation of the trypanothione reductase gene from *Crithidia fasciculata* and *Trypanosoma brucei*: Comparison with other flavoprotein disulfide oxidoreductases with respect to substrate specificity and catalytic mechanism. *Mol Microbiol 6*:3089-3099.

Bailey S, Fairlamb AH, Hunter WNH. 1994. Structure of trypanothione reductase from *Crithidia fasciculata* at 2.6 Å resolution; enzyme-NADP interactions at 2.8 Å resolution. *Acta Crystallogr D 50*:139-154.

Bailey S, Smith K, Fairlamb AH, Hunter WN. 1993. Substrate interactions between trypanothione reductase and N^1-glutathionylspermidine disulfide at 0.28 nm resolution. *Eur J Biochem 213*:67-75.

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rogers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535-542.

Bond CS, Fairlamb AH, Hunter WN. 1995. A comparison of two independently determined structures of trypanothione reductase from *Crithidia fasciculata*. *Acta Crystallogr D 51*:567-574.

Borges A, Cunningham ML, Tovar J, Fairlamb AH. 1995. Site-directed mutagenesis of the redox-active cysteines of *Trypanosoma cruzi* trypanothione reductase. *Eur J Biochem 228*:745-752.

Bradley M, Bucheler US, Walsh CT. 1991. Redox enzyme engineering: Conversion of human glutathione reductase into trypanothione reductase. *Biochemistry 30*:6124-6127.

Brünger AT. 1990. *X-PLOR version 3.1. Manual.* New Haven, Connecticut: Yale University Press.

Carter P, Wells JA. 1988. Dissecting the catalytic triad of a serine protease. *Nature 237*:909-913.

Chothia C, Janin J. 1975. Principles of protein-protein recognition. *Nature 256*:705-708.

Collaborative Computational Project Number 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D 50*:760-763.

Cunningham ML, Zvelebil MJJM, Fairlamb AH. 1994. Mechanism of inhibition of trypanothione reductase and glutathione reductase by trivalent organic arsenicals. *Eur J Biochem 221*:285-295.

de Castro S. 1993. The challenge of Chagas' disease chemotherapy: An update of drugs assayed against *Trypanosoma cruzi*. *Acta Tropica 53*:83-98.

Eggink G, Engel H, Vriend G, Terpstra P, Witholt B. 1990. Rubredoxin reductase of *Pseudomonas oleovoras*: Structural relationship to other flavoprotein oxidoreductases based on one NAD and two FAD fingerprints. *J Mol Biol 212*:135-142.

Fairlamb AH. 1989. Novel biochemical pathways in parasitic protozoa. *Parasitology 99S*:93-112.

Fairlamb AH, Cerami A. 1992. Metabolism and functions of trypanothione in the kinetoplastida. *Annu Rev Microbiol 46*:695-729.

Field H, Cerami A, Henderson GB. 1992. Cloning, sequencing and demonstration of polymorphism in trypanothione reductase from *Crithidia fasciculata*. *Mol Biochem Parasitol 50*:47-56.

Fitzgerald PMD. 1988. Merlot, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J Appl Crystallogr 21*:273-278.

Ghisla SK, Massey V. 1989. Mechanisms of flavoprotein-catalysed reactions. *Eur J Biochem 181*:1-17.

Henderson GB, Murgolo NJ, Kuriyan J, Osapay K, Kominos D, Berry A, Scrutton NS, Hinchcliffe NW, Perham RN, Cerami A. 1991. Engineering the substrate specificity of glutathione reductase toward that of trypanothione reduction. *Proc Natl Acad Sci USA 88*:8769-8773.

Hunter KJ, Le Quesne S, Fairlamb AH. 1994. Identification and biosynthesis of N^1, N^9-bis(glutathionyl)aminopropylcadaverine (homotrypanothione) in *Trypanosoma cruzi*. *Eur J Biochem 226*:1019-1027.

Hunter WN, Bailey S, Habash J, Harrop SJ, Helliwell JR, Aboagye-Kwarteng T, Smith K, Fairlamb AH. 1992. Active site of trypanothione reductase; a target for rational drug design. *J Mol Biol 227*:322-333.

Hunter WN, Smith K, Derewenda Z, Harrop SJ, Habash J, Islam MS, Helliwell JR, Fairlamb AH. 1990. Initiating a crystallographic study of trypanothione reductase. *J Mol Bol 216*:235-237.

Jones TA. 1978. A graphics model building and refinement system for macromolecules. *J Appl Crystallogr 11*:268-272.

Jones TA, Zou JY, Cowan SW, Kjeldgaard M. 1991. Methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A 47*:110-119.

Kabsch WJ, Sander C. 1983. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers 22*:2577-2637.

Karplus PA, Schulz GE. 1987. Refined structure of glutathione reductase at 1.54 Å resolution. *J Mol Biol 195*:701-729.

Karplus PA, Schulz GE. 1989. Substrate binding and catalysis by glutathione reductase as derived from refined enzyme:substrate crystal structures at 2.0 Å resolution. *J Mol Biol 210*:163-180.

Kirchhoff LV. 1993. Current concepts—American Trypanosomiasis (Chagas' disease): A tropical disease now in the United States. *New England J Med 329*:639-644.

Kjeldgaard M, Nissen P, Thirup S, Nyborg J. 1993. The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure 1*:35-50.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946-950.

Kuriyan J, Kong XP, Krishna TSR, Sweet RM, Murgolo NJ, Field H, Cerami A, Henderson GB. 1991. X-ray structure of trypanothione reductase from *Crithidia fasciculata* at 2.4 Å resolution. *Proc Natl Acad Sci USA 88*:8764-8768.

Lantwin CB, Schlichting I, Kabsch K, Pai EF, Krauth-Siegel RL. 1994. The structure of *Trypanosoma cruzi* trypanothione reductase in the oxidised and NADPH reduced state. *Proteins Struct Funct Genet 18*:161-173.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283-291.

Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol 55*:379–400.

Luzzati V. 1952. Traitment statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr A 43*:134–136.

Meister A. 1989. Glutathione: General review of mechanism of action. In: Dolphin D, Poulson R, Avramovic O, eds. *Glutathione: Chemical, biochemical and medical aspects*. New York: John Wiley & Sons. pp 367–474.

Mittl PRE, Berry A, Scrutton NS, Perham RN, Schulz GE. 1994. Anatomy of an engineered NAD-binding site. *Protein Sci 3*:1504–1514.

Mittl PRE, Schulz GE. 1994. Structure of glutathione reductase from *Escherichia coli* at 1.86 Å resolution: Comparison with the enzyme from human erythrocytes. *Protein Sci 3*:799–809.

Moncayo A. 1993. Chagas' disease. In: *Tropical disease research: Progress 1991–92*. Geneva: World Health Organisation. pp 67–76.

Moore MH, Gulbis JM, Dodson EJ, Demple B, Moody PCE. 1994. Crystal structure of a suicidal DNA repair protein: The Ada O^6-methylguanine-DNA methyltransferase from *E. coli*. *EMBO J 13*:1495–1501.

Nicholls A. 1993. *GRASP Manual, V1.1*. New York: Columbia University.

Ramachandran GN, Sasiskeharan V. 1968. Conformation of polypeptides and proteins. *Adv Protein Chem 23*:283–437.

Rao ST, Rossmann MG. 1973. Comparison of super-secondary structure in proteins. *J Mol Biol 76*:241–256.

Schirmer RH, Muller JG, Krauth-Siegel RL. 1995. Disulfide-reductase inhibitors as chemotherapeutic agents: The design of drugs for trypanosomiasis and malaria. *Angew Chem Int Ed Engl 34*:141–154.

Schulz GE. 1992. The binding of nucleotides by proteins. *Curr Opin Struct Biol 2*:61–67.

Schulz GE, Schirmer RH, Pai EF. 1982. FAD binding-site of glutathione reductase. *J Mol Biol 160*:287–308.

Smith K, Borges A, Ariyanayagam MR, Fairlamb AH. 1995. Glutathionylspermidine metabolism in *Escherichia coli*. *Biochem J*. Forthcoming.

Sullivan FX, Sobolov SB, Bradley M, Walsh CT. 1991. Mutational analysis of the parasite trypanothione reductase: Acquisition of glutathione reductase activity in a triple mutant. *Biochemistry 30*:2761–2767.

Sullivan FX, Walsh CT. 1991. Cloning, sequencing, overproduction and purification of trypanothione reductase from *Trypanosoma cruzi*. *Mol Biochem Parasitol 44*:145–148.

Swindells MB. 1995a. A procedure for the automatic determination of hydrophobic cores in protein structures. *Protein Sci 4*:93–102.

Swindells MB. 1995b. A procedure for detecting structural domains in proteins. *Protein Sci 4*:103–112.

Verlinde CLMJ, Hol WGJ. 1994. Structure-based drug design: Progress, results and challenges. *Structure 2*:577–587.

World Health Organisation. 1991. *Tropical diseases: Progress in research*. Geneva: Tenth Programe Report of the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases.

Zhang Y, Bailey S, Naismith JH, Bond CS, Habash J, McLaughlin P, Papiz M, Borges A, Cunningham M, Fairlamb AH, Hunter WN. 1993. *Trypanosoma cruzi* trypanothione reductase, crystallisation, unit cell dimensions and structure solution. *J Mol Biol 233*:1217–1220.