Tryparedoxin-I is a recently discovered thiol-disulfide oxidoreductase involved in the regulation of oxidative stress in parasitic trypanosomatids. The crystal structure of recombinant Crithidia fasciculata tryparedoxin-I in the oxidized state has been determined using multi-wavelength anomalous dispersion methods applied to a selenomethionyl derivative. The model comprises residues 3 to 145 with 236 water molecules and has been refined using all data between a 19- and 1.4-Å resolution to an R-factor and R-free of 19.1 and 22.3%, respectively. Despite sharing only about 20% sequence identity, tryparedoxin-I presents a five-stranded twisted β-sheet and two elements of helical structure in the same type of fold as displayed by thioredoxin, the archetypal thiol-disulfide oxidoreductase. However, the relationship of secondary structure with the linear amino acid sequences is different for each protein, producing a distinctive topology. The β-sheet core is extended in the trypanosomatid protein with an N-terminal β-hairpin. There are also differences in the content and orientation of helical elements of secondary structure positioned at the surface of the proteins, which leads to different shapes and charge distributions between human thioredoxin and tryparedoxin-I. A right-handed redox-active disulfide is formed between Cys-40 and Cys-43 at the N-terminal region of a distorted α-helix (α1). Cys-40 is solvent-accessible, and Cys-43 is positioned in a hydrophilic cavity. Three C-H...O hydrogen bonds donated from two proline residues serve to stabilize the disulfide-carrying helix and support the correct alignment of active site residues. The accurate model for tryparedoxin-I allows for comparisons with the family of thiol-disulfide oxidoreductases and provides a template for the discovery or design of selective inhibitors of hydroperoxide metabolism in trypanosomes. Such inhibitors are sought as potential therapies against a range of human pathogens.

Parasitic trypanosomatids, belonging to the order Kinetoplastrida, cause debilitating and life-threatening human diseases such as African sleeping sickness, Chagas’ disease, and the leishmaniasis (1). The current therapies against these infections are inadequate due to poor drug efficacy and toxicity combined with increasing drug resistance (2). There is therefore an urgent need to understand how drugs already in use function so that they might be improved and to identify new targets for chemotherapeutic attack. The ideal target is an enzyme of a metabolic pathway that is essential for the survival of the parasite and either absent in the human host or one that presents differing substrate specificities (3). Because trypanosomatids are susceptible to oxidative stress, this aspect of their metabolism represents an attractive target for the development of new trypanocidal agents (4–6).

As with other organisms living in an aerobic environment, trypanosomes are exposed to reactive oxygen intermediates such as superoxide anion and hydrogen peroxide. These potentially destructive chemicals are eliminated in most eukaryotic and prokaryotic cells by means of a combination of superoxide dismutases, catalase, and a variety of peroxidases. In mammalian cells, the principal route of hydrogen peroxide detoxification involves glutathione peroxidase working in concert with NADPH, glutathione, and glutathione reductase (7). The medically important trypanosomatids do not contain catalase, glutathione peroxidase, or glutathione reductase but rely on an analogous system to regulate oxidative stress. The details of this trypanothione peroxidase system have been elucidated in the model trypanosome Crithidia fasciculata (6, 8, 10, 11) and are shown in Fig. 1. NADPH provides reducing input to the flavoprotein trypanothione reductase, which in turn maintains high levels of reduced trypanothione (N4,N8-bis(glutathionyl)spermidine), a polypeptide conjugate unique to trypanosomatids (5). Trypanothione reduces tryparedoxin-I, which then reduces tryparedoxin peroxidase. The peroxidase then catalyzes the final reduction of hydrogen peroxide and organic hydroperoxides to water or alcohols. Hydroperoxide metabolism in mammals, yeasts, and some plants uses thioredoxin peroxidase to reduce H2O2 and alkyl hydroperoxides with reducing equivalents provided from a pathway involving thioredoxin, thioredoxin reductase, and NADPH (Refs. 12–15, Fig. 1).

Tryparedoxin-I and tryparedoxin peroxidase are analogous to thioredoxin and thioredoxin peroxidase, respectively. These proteins all utilize a reactive Cys-sulfhydryl group to reduce disulfides or, alternatively, to oxidize sulfhydryl groups. Tryparedoxin-I is a thiol-disulfide oxidoreductase, a family of proteins that participates in a diverse range of biological processes, which include the formation of or isomerization of disulfides during protein folding events, deoxyribonucleotide synthesis, repair of oxidative damage, activation of transcription factors, modulation of protein-DNA interactions, and the
regulation of cell growth (see Refs. 12, 15–19). The archetypal thiol-disulfide oxidoreductases is thioredoxin, a small globular protein of approximate molecular mass 12 kDa (15, 17). The three-dimensional structures of oxidized and reduced thioredoxins, including *Escherichia coli* and human proteins have been extensively characterized by NMR spectroscopy and x-ray crystallography, as have thioredoxin-ligand interactions with peptides that represent fragments of naturally occurring substrates (20–25). Structures of other members of the thioredoxin superfamily have been determined, and these include glutaredoxins (26–28), the bacterial oxidant proteins DebA (disulfide bond formation protein A (29)) and TcpG (toxin-coregulated colonization pilus gene product (30)), a human peroxiredoxin hORF6 (31), *Pyrococcus furiosus* disulfide oxidoreductase (32) and the thioredoxin-like domain of eukaryotic protein disulfide isomerase (33). Wide-ranging biochemical, biophysical, and theoretical studies have been carried out, in particular on thioredoxin and DebA, seeking to understand structure-reactivity-function relationships (15, 17, 34–38).

In contrast, much less is known about the thioredoxin homologs recently discovered in trypanosomes. The first of these was identified in *C. fasciculata* and is called tryparedoxin-I (6, 8, 11). This is a protein of 146 amino acids with a molecular mass of approximately 16 kDa; hence, significantly larger than typical thioredoxins. Tryparedoxin-I contains a Trp<sup>39</sup>-Cys-Pro-Pro-Cys<sup>43</sup> sequence near the N terminus that resembles the thioredoxin-type Trp-Cys-Gly (or Ala)-Pro-Cys active-site motif (12), in which the vicinal cysteine residues form a redox-active disulfide. Tryparedoxin has been found in *Trypanosoma brucei* (39) and *Trypanosoma cruzi* but not so far in *Leishmania* species. However, tryparedoxins may be ubiquitous for the trypanosomatids since a functional tryparedoxin peroxidase has been identified in *Leishmania* major (40). A second tryparedoxin (II) has recently been cloned from *C. fasciculata* and also serves as a physiological electron donor for tryparedoxin peroxidase (41).

A crystallographic study of tryparedoxin-I has been initiated to delineate structure-function relationships in the Kinetoplastida-specific trypanothione peroxidase system to enable comparisons with thioredoxins and to investigate how this pathway might be used for the development of trypanocidal agents. Attempts to solve the structure of tryparedoxin-I by molecular replacement using thioredoxin as a search model were unsuccessful (42), and we also encountered problems of nonisomorphism following cryo-protection/freezing of the crystals and difficulties in the preparation of heavy-atom derivatives. It was therefore decided to obtain experimental phases using multielemental anomalous dispersion (MAD (43)) targeting a selenomethionine derivative. We now describe the preparation of the selenomethionyl protein, the structure solution and refinement using diffraction terms to a resolution of 1.4 Å. The high resolution model reveals the fold and conserved core structure of tryparedoxin-I together with the molecular details in and around the active site. Comparisons with other thiol-disulfide oxidoreductases, in particular human thioredoxin, are presented as appropriate.

**EXPERIMENTAL PROCEDURES**

**Sample Preparation**—The methionine auxotrophic strain of *E. coli* B384(DE3) was heat shock-transformed with the plasmid carrying the gene for *C. fasciculata* tryparedoxin-I (pET-TryX(11)) and selected on Luria-Bertani agar plates containing 100 μg ml<sup>−1</sup> ampicillin. Bacteria were cultured in M9 media supplemented with the usual amino acids except that selenomethionine (100 mg l<sup>−1</sup>) was substituted for l-methionine. Expression of tryparedoxin-I was induced at mid-log phase with 0.4 mM isopropyl-β-D-thiogalactopyranoside, and cell growth was allowed to continue for a further 4 h. Cells were harvested by centrifugation (2500 × g) at 4 °C, resuspended in binding buffer (20 mM Bis-Tris propane, pH 7.5, 500 mM NaCl, 5 mM imidazole, 5 mM benzamidine), and lysed in a French press. The insoluble debris was separated by centrifugation (27,000 × g) at 4 °C for 20 min, the supernatant filtered then applied to a Ni<sup>2+</sup>-resin column (Poros 4.6 mm/100 mm) pre-equilibrated with binding buffer (20 mM Bis-Tris propane, pH 7.5, 500 mM NaCl, 5 mM imidazole, 5 mM benzamidine), and eluted with 60 mM imidazole to 60 mM in a single step. Fractions were analyzed by SDS-polyacrylamide gel electrophoresis and those containing tryparedoxin-I pooled, dialyzed against binding buffer, and subjected to a second purification step on the Ni<sup>2+</sup>-resin column. Pooled fractions of tryparedoxin-I were dialyzed against 50 mM HEPES, pH 7.5, and the histidine tag was cleaved with biotinylated-thrombin (Novagen). The protease was removed by use of a streptavidin-garose resin (Novagen), and the histidine tag was removed by passage through the Ni<sup>2+</sup>-resin column. Pool fractions of tryparedoxin-I were dialyzed against 50 mM HEPES, pH 7.5, and concentrated (Centricron-10, Amicon) to approximately 10 mg ml<sup>−1</sup> for use in crystalization experiments. Matrix-assisted laser desorption ionization time of flight mass spectroscopy (PerSeptive Biosystems, Voyager) was used to assess purity and to confirm the full incorporation of selenomethionine. Orthorhombic crystals in space group P<sub>2</sub>_<sub>2</sub>_<sub>1</sub> were obtained under similar conditions to those employed on the wild-type protein (42). The selenomethionine tryparedoxin-I crystals have unit cell dimensions of a = 38.39, b = 50.70, c = 70.82 Å with a single copy of the protein in the asymmetric unit. The N-terminal hexahistidine affinity tag is cleaved by thrombin during the purification procedure, leaving a Gly-Ser-His extension that ensures that the initiating methionine is preserved on the recombinant protein.

**Data Collection**—A single crystal (0.2 × 0.2 × 0.5 mm) was cryo-protected with 40% polyethylene glycol monomethyl ether 2000 then flash-frozen in a nylon loop with a nitrogen stream at −170 °C (Oxford Cryo-stream), maintained at low temperature and transported dry to
The European Synchrotron Radiation Facility BM14 in Grenoble (State-bourne Cytogenetics Biotrek 3), then used for MAD data collection. The x-ray detector was a charge coupled device (MAR Research), and the choice of wavelengths for the peak ($\lambda_1$), inflection point ($\lambda_2$), and remote data ($\lambda_3$) were derived from a XANES (x-ray absorption near edge structure) scan of the Se K-absorption edge from the crystal. The three wavelengths were selected to provide the largest value for the anomalous difference, $f^R(L)$, and the minimum value for $f^R(L_\lambda)$ or inflection point. The remote wavelength ($\lambda_3$) maximized the dispersive difference, $\Delta f^R$ ($\lambda_3$ - $\lambda_2$) and, by virtue of being at a higher energy, extended the resolution to which data were recorded. The STRATEGY program (44) was used to determine the range for data collection, and each wavelength data set was subsequently collected in a sweep of 0.5° oscillations through 87.5°. All data were processed, reduced, and scaled using the HKL suite of programs (45) then analyzed using the CCP4 software package (46). Relevant statistics are provided in Table I.

### Table I

| Wavelength (Å) | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ |
|----------------|-------------|-------------|-------------|
| 0.9790         | 0.9792      | 0.8265      |
| Resolution range (Å) | 21.1    | 21.1        | 19.1        |
| No. of measurements | 186,894 | 187,704     | 330,283     |
| No. unique reflections | 15,812  | 15,774      | 28,003      |
| Coverage overall (%) | 97.0 (79.4) | 97.0 (82.7) | 95.3 (66.1) |
| I/|d|all|data |21.52 (5.68) |19.45 (5.52) |20.19 (5.77) |
| Rmerge (%) | 2.9 (12.6) | 3.1 (13.8) | 3.0 (17.3) |
| Rfree (%) | 6.0 | 5.6 | 3.5 |
| $R_{int}$ (%) | 2.7 | 5.5 |

The resolution of the data was extended to 1.4 Å with the $\lambda_3$ data set measured at a shorter wavelength, and refinement was accomplished using REFMAC (51) interspersed with computer graphics map inspection (Silicon Graphics) and model manipulation combined with ARP (52) for water placement. The R-free was monitored as a guide for the refinement (53). Several residues on the surface of the protein were poorly defined by the MAD phased electron density map and truncated to alanines in the first model; however, in the course of refinement, the complete side chains could be built in. Similarly, as the refinement progressed, the maps indicated dual conformations for seven residues, which were successfully modeled as such. The stereochemistry of the tryparedoxin-I model was assessed with PROCHECK (54), and further details are presented in Table II.

### RESULTS AND DISCUSSION

#### Shape, Fold, and Secondary Structure of Tryparedoxin

A ribbon diagram showing the secondary structure and fold of C. fasciculata tryparedoxin-I is presented in Fig. 3, and the secondary structure is mapped onto the amino acid sequence in Fig. 4a. The conservation of sequence among the three tryparedoxins is color-coded on a Cα trace of tryparedoxin-I in Fig. 4b. Secondary structure was assigned using a combination of automated methods in the programs DSSP (55), PROCHECK (54), and PROMOTIF (56) and by visual inspection. The three tryparedoxins for which sequence information is available all share high levels of sequence identity with approximately 45% of residues conserved across all three proteins (Fig. 4). The two examples from C. fasciculata, namely tryparedoxin-I and II, are 52% identical, and the T. brucei protein is 57% identical with tryparedoxin-II. The sequence conservation extends over three dimensions as shown in Fig. 4b, so the structure we present can be considered as representative of this subset of the thioredoxin superfamily. The sequence conservation also provides an indication of which residues are likely to contribute to the stability and reactivity of tryparedoxins, and in the analysis of the structural model, we pay particular attention to those residues.

Tryparedoxin-I is an ellipsoid of approximate dimensions 40 × 30 × 30 Å constructed around a seven-stranded twisted β-sheet scaffold using both parallel and anti-parallel alignments (Fig. 3). The sheet starts with a β-hairpin formed by β1 and β2, then a β3-α1-β4-α2-β5 motif that can be considered as a combination of two right-handed β-α-β units sharing strand β4. Another β-hairpin formed by strands β6 and β7 completes the twisted sheet. The first hairpin turn linking β1 and β2 is on the periphery of the β-sheet, is solvent-accessible, and carries seven charged residues, Glu13, Lys13, Arg15, Arg16, Asp18, Glu20, and Glu22. This segment of the structure is poorly conserved among tryparedoxins. The longest helical stretch is provided by a distorted α1, which is about five turns in length and curves over the surface of the protein. There are three segments of the polypeptide in a 310 helix conformation that involve residues between β2 and β3 (α1 in Fig. 3) and also the N-terminal sections leading into helices α2 and α3 (α2 and β3 in Fig. 3).

The residues that comprise strands β3 and β4 are mostly buried and conserved among tryparedoxins. These elements of secondary structure provide a significant component of the hydrophobic core of the protein. The core of tryparedoxin-I is mainly formed from the aromatic tyrosines (residue numbers 33, 54, and 80), tryptophans (residues 70 and 86), and phenylalanines (residues 32, 33, 35, 46, 53, 57, 63, 67, 77, 81, 91, and 104). Thirteen of these aromatic residues are strictly conserved in tryparedoxins, the exceptions being three phenylalalines (residues 33, 67, and 81), which are replaced by leucine or tyrosine, and Phe57, which is replaced by histidine in the other proteins.

#### Comparison of Tryparedoxin and Thioredoxin

Topology diagrams for tryparedoxin-I and human thioredoxin are compared in Fig. 5a. Thioredoxin presents a characteristic α/β-fold that is a five-stranded twisted β-sheet assembly surrounded by four helices. This folding unit occurs as a globular entity in glutaredoxin or as a component of multi-domain systems exemplified by DabA, TcpG and glutathione-S-transferase (57), or peroxiredoxins (51) and duplicated in the case of P. furiosus disulfide oxidoreductase (32).

Thioredoxins share only about 20% sequence identity with tryparedoxin-I (11, 39), yet as we will show, the proteins are structurally related, sharing the three-dimensional alignment of seven elements of secondary structure. The sequence homology is only significant in two regions of the protein, namely the N-terminal end of the disulfide-carrying helix, α1, and the loop that links α3 with β6 (Fig. 4a). This loop is adjacent in three-dimensional space to the active-site disulfide (Figs. 3 and 5c). Least-squares superposition of a human thioredoxin monomer and tryparedoxin-I were carried out using the algorithms implemented in the program O (50). There are 39 Cα atoms, corresponding to the residues of the β3-α1-β4 segment, that
align with an r.m.s.$^2$ of 1.26 Å (Fig. 5b). Although a slightly lower r.m.s. of 1.14 Å is obtained for the same alignment of tryparedoxin-I with *E. coli* thioredoxin, this suggests that there is no significant difference, and either thioredoxin model can be used for comparative purposes. Based on these alignments we observe a close structural relationship between the core of thioredoxins and five of the seven β-strands in tryparedoxin-I, although the order of the strands is different, producing a distinctive topology for each protein (Fig. 5a). β-Strands β5, β4, β3, β6, and β7 of tryparedoxin-I align with β1, β3, β2, β4, and β5 of thioredoxin, respectively. Two helical sections also superimpose well with this alignment. The first is the redox disulfide-carrying helix, α1 of tryparedoxin-I, which overlaps with α2 of human thioredoxin for almost their entire lengths of 5 turns. The second is α3 of both proteins, which are positioned near the active sites. The superposition overlaps at the C ter-

---

**FIG. 2.** A stereoview of the MAD phased electron density map, at 1.7-Å resolution, with a section of the refined model corresponding to residues at the N-terminal region of β4. The density is contoured at the 1σ level of that observed in the unit cell and is depicted in *pink chicken wire*, amide nitrogens are colored *cyan*, carbon atoms are *purple*, and oxygen atoms are *red*. This figure was produced with the program O (50).

**TABLE II**

| Refinement and model geometry statistics of selenomethionyl-tryparedoxin-I |
|-----------------------------------------------|
| Protein residues | 143 |
| Protein atoms | 1145 |
| Solvents | |
| R-work(%)| 19.3/25857 |
| R-free(%) | 22.4/1379 |
| Average isotropic thermal parameters (Å$^2$) | 12.9 |
| Overall | 12.9 |
| Main chain | 12.0 |
| Side chains | 13.7 |
| Solvents | 26.1 |
| Residues in dual conformations | |
| Glu$^{73}$, Glu$^{74}$, Asp$^{75}$, Gln$^{79}$, Asn$^{103}$, Arg$^{123}$, Gln$^{139}$ | |
| r.m.s. bond lengths (Å) | 0.014 |
| r.m.s. bond angle associated distances (Å) | 0.030 |
| r.m.s. planarity (Å) | 0.120 |
| Overall G factor$^a$ | −0.1 |
| Ramachandran analysis | |
| Favorable | 94.3 (115) |
| % and No. of residues | |
| Additional | 5.7 (7) |

$^a$G-factor and Ramachandran analysis were derived from PROCHECK (53).

---

**FIG. 3.** A ribbon diagram showing the secondary structure of tryparedoxin-I. α-Helices are colored red, 3$_{10}$ helices (θ) are green, β-strands are cyan, and the redox-active disulfide is yellow. The main chain of residues in random coil are gray. Figs. 3 and 4b, 5b, 6, and 7b were produced with MOLSCRIPT (68) and RASTER-3D (69).
minus of tryparedoxin-I α3, which is extended by approximately one turn compared with human thioredoxin.

The Redox-active Disulfide and Environment—This crystal structure is of the oxidized form of tryparedoxin-I, where a right-handed redox-active disulfide is formed between residues Cys40 and Cys43 at the N terminus of α1. This part of the helix protrudes out from the molecule. The disulfide environment is constructed from the C-terminal residues of β3, the turn that links β3 to the N terminus of α1, and residues of the loop following the C terminus of α3 and the N-terminal section of β6. A complex network of hydrogen bonds occurs with and around the redox-active disulfide as will be described.

In the Cys-Xaa2-Xaa3-Cys motif of thiol-disulfide oxidoreductases, the N-terminal cysteine is solvent-accessible, but dimerization and crystal packing effects occlude solvent binding to these cysteines in human, E. coli, and Anabaena thioredoxin structures (20–22). In tryparedoxin-I Cys40, Sγ is more accessible to solvent than its partner Cys43 Sγ; indeed it forms a hydrogen bond of length 3.09 Å with a water molecule (Fig. 6). This solvent molecule also hydrogen bonds to the main chain carbonyl of Ile109 (distance 3.03 Å) and to two other water molecules (2.73 and 2.87 Å distant, not shown). Such a pattern of hydration is similar to that observed for the amino-proximal cysteine in the active site of Vibrio cholerae TcpG (30). Cys40 is most likely to be the cysteine that interacts directly with the substrates, and the hydrogen-bonding solvent marks the space that substrate may occupy during the reaction of tryparedoxin-I with either trypanothione or tryparedoxin peroxidase. By comparison with thioredoxin, it is thought that Cys40 of tryparedoxin-I will be more reactive than Cys43. This is consistent with the observations made by Gommel et al. (8), based on chemical modification studies combined with mass spectrometry, that the amino-proximal cysteine is indeed the more reactive of the two. There may be a contribution toward this reactivity from the positive dipole that arises from helix α1. However, as pointed out by Weichsel et al. (21) when describing the structure of the human thioredoxin, the disulfide is actually displaced relative to the optimal alignment with the helix dipole (Fig. 6a). The position of α1 relative to the Cys40-Cys43 disulfide suggests to us that this dipole may contribute more to

![Fig. 4.](image_url)

**a**, the amino acid sequence of *C. fasciculata* tryparedoxin-I with the assignment of secondary structure based on the crystallographic model. Residues encased in gray are conserved in either tryparedoxin-II (41) or T. brucei tryparedoxin-I (39), whereas residues in black are strictly conserved in all three proteins. Residues underlined with ~ and # are homologous and strictly conserved with human thioredoxin, respectively. b, stereoview Ca trace, colored according to sequence homology with other tryparedoxins. The Cα-Cα links are colored depending on the conservation or not of the C-terminal residue of the pair. **Black** indicates that a residue is strictly conserved, **cyan** is conserved with one other tryparedoxin, and **red** is not conserved.

![Structure of Tryparedoxin](image_url)
tryparedoxin-substrate interactions than to the chemistry of the redox-active disulfide.

Cys43 lies below Cys40, buried within the protein structure, and occupies a hydrophilic pocket lined by hydroxyl groups from Tyr34, Ser36, Thr47, and Tyr80. Cys43 accepts a hydrogen bond from Ser36 (3.19 Å), which in turn accepts two hydrogen bonds donated from the amide of Ala37 (2.95 Å) and the hydroxyl of Tyr80 (2.77 Å). The hydroxyl groups of Tyr34 and Thr47 are hydrogen-bonded to each other at a distance of 2.74 Å. The pocket is filled with well-ordered solvent molecules and guarded by the basic residues, Arg44 and Lys83. In most thioredoxins this pocket contains a highly conserved aspartate residue (Asp26 in the human protein) linked via solvent to a lysine side chain at the edge of the molecule (21), but like tryparedoxin-I, Anabaena thioredoxin has a tyrosine rather than an aspartate at this position, and the high-resolution structure of the latter suggests that this substitution alters the position of the redox-active disulfide relative to the central β-sheet (22). In our structural alignments we see no evidence to support a shift in position of the disulfide but conclude that the substitution of the tyrosine hydroxyl (Tyr34) for the aspartate carboxylate (Asp26 in human thioredoxin) may influence the environment and reactivity of the buried cysteines.

The redox-active disulfide of tryparedoxin-I resides in the pentapeptide sequence Trp39-Cys-Pro-Pro-Cys43. As first noted by Gommel et al. (8), several thioredoxins of plant origin, the thioredoxin from Caenorhabditis elegans, as well as mouse nucleoredoxin also carry this pentapeptide sequence. In thioredoxins the redox motif is Cys-Gly-Pro-Cys, in glutaredoxins it is Cys-Pro-Tyr-Cys, and in protein disulfide oxidoreductases it is Cys-Gly-His-Cys (12). The variance in the intervening residues of the reactive cysteines is implicated in determining the redox potential of each type of protein, although a complete understanding of all the factors that are involved is lacking (35).

The loop and N-terminal end of strand β6 of tryparedoxin-I is well conserved in terms of both sequence and structure with thioredoxins (Figs. 4a and 5b). The sequence in tryparedoxin-I is Val109-Glu-Ser-Ile-Pro-Thr114, and in a human thioredoxin it is Val71-Lys-Cys-Met-Pro-Thr76 (see Fig. 2 in Ref. 11). The proline in this hexapeptide sequence is strictly conserved and adopts a cis conformation. Further discussion on this point will be made later. Crystalllographic studies show that human thioredoxin is a homodimer linked by a disulfide bond involving Cys73 of each monomer. Dimer formation occludes the active site, and a conformational change must occur for the protein to function (21). Site-directed mutagenesis and analysis of the C73S mutant human thioredoxin still showed a dimer with a hydrogen bond linking the serines. Our structural alignment places Ser108 of tryparedoxin-I on Cys73 of the human protein.
but this serine is directed to bulk solvent, and we see no structural evidence for dimer formation by tryparedoxin-I.

The redox potentials of the thiol-disulfide oxidoreductases and pKa values for the individual cysteines in the Cys-Xaa1-Xaa2-Cys motif have been extensively studied, and this motif has been likened to a rheostat, a particularly appropriate analogy (58). There are several major factors that determine the redox properties and the role they play for the in vivo activity of the proteins concerned. These factors include the local environment of the cysteine sulfurs and their proximity to functional groups, so that they can participate in hydrogen bonding and van der Waals interactions or be influenced by longer range effects such as a helix dipole as discussed earlier. The nature of the amino acids that constitute Xaa1 and Xaa2 and also the influence of the substrates that these proteins must interact with can also influence the redox potentials. Despite the wealth of data, a complete understanding of what determines the redox potentials of the Cys-Xaa1-Xaa2-Cys motif eludes us, but given the complexity of the system, this is perhaps not surprising.

Two tryptophans, Trp39 and Trp70, create a flat ledge above the redox-active disulfide that leads to a wall created by helix a3 and the loop leading into b6 and the C-terminal end of b4 (Figs. 6 and 7a). These residues are held in place by hydrogen bonds shared between their Ne1 atoms with the carbonyl oxygen of Trp70 and the Oγ Ser101, with distances of 2.93 and 2.92 Å, respectively. This helps to form a lid covering the redox-active disulfide (Fig. 6). The interaction with Ser101 is omitted from Fig. 6 for the purpose of clarity. Just below the tryptophans are the side chains of Ala37 and Ile109. The disulfide structure is stabilized by direct hydrogen bonds to the S atoms as described earlier and by van der Waals interactions with surrounding residues. For example Cys40 Sγ has contacts to Ile109 Cγ1 and Cδ1.

**Fig. 6. Two stereoviews of the redox-active disulfide and its immediate environment.** a, side view; b, top view. Atomic positions are colored according to type: black, C; cyan, N; red, O; yellow, S and the disulfide bond. The solvent hydrogen bonding to Sy Cα is depicted as a red cross in a sphere. The single-letter amino acid code is used for labeling purposes. Hydrogen bonding interactions are shown as green dashed lines, C-H-O interactions are depicted as black dashed lines. The cylinder in a indicates helix a1.
Prolines and C-H–O Hydrogen Bonding in Tryparedoxin—
The amino acid sequence of C. fasciculata tryparedoxin-I includes 10 proline residues, of which 9 are in well defined electron density. The missing proline is Pro(OH), the C-terminal residue for which we are unable to see any electron density. The nine ordered prolines are strictly conserved in the three tryparedoxin sequences that are available (Fig. 4a), and they are all positioned on the surface of the molecule and are accessible to solvent. Seven are in a trans conformation, whereas Pro(Cα) and Pro are cis. Pro(Cα) is important because it is adjacent to and participates in van der Waals contacts with the redox-active disulfide. This cis-proline is a common structural feature of the thiol-disulfide oxidoreductases. In DisA, the mutation of the cis-proline to trans-alanine destabilized the structure, reducing activity to about 50% of the normal level, and structural analysis indicated that the loop from a to b adopted a different structure and lost the van der Waals interactions with the redox-active disulfide (36). Clearly the unique structural contributions from a cis-proline in this position are necessary to optimize activity and stability of this protein family.

Pro and Pro are the intervening residues of the disulfide-forming cysteines, and so tryparedoxin-I contains a reactive disulfide embedded in a unique environment created by the three prolines. As discussed earlier, it is recognized that in the Cys-Xaa-Cys redox-active disulfide motif, the intervening residues modulate the redox potentials of the thiol-disulfide oxidoreductases (12, 58).

The existence of C-H–O interactions and the classification as hydrogen bonds has provoked discussion in the literature (59, 60). It is now widely accepted that they do form and contribute to the stability of macromolecules (61) and to enzyme catalysis as discussed in the context of serine proteases (62) or disulfide oxidoreductases (63). With respect to tryparedoxin-I, there are three C-H–O hydrogen-bonding interactions involving only protein atoms that are of note. The importance of C-H–O hydrogen-bonding interactions is not yet known, but they may alleviate the destabilizing effect of having an unsatisfied hydrogen bond acceptor in the structure. Two of the C-H–O hydrogen bonds are donated from the methylene group Pro(Cα) Cα to the main chain carbonyls of Arg and Gly, with C-O distances of 2.93 and 3.44 Å, respectively (Fig. 6). Pro(Cα) is positioned almost three turns along helix α1, facing out toward solvent and causing a distortion that redirects the final two turns of the helix. Such a feature of prolines in helices is well documented (64). This C-H–O hydrogen bonding interaction is conserved in crystal structures of thioredoxins, as discussed by Chakrabarti and Chakrabarti (65), and may contribute to the distortion of helix α1, which carries the redox-active disulfide. We observe another C-H–O interaction involving a proline, namely between Pro(Cα) Cα, which is part of the active site-conserved motif Trp-Pro-Cys-Pro-Cys, and the main chain carbonyl of Trp, with a distance of 3.16 Å (Fig. 6). This interaction is unique to tryparedoxin-I, since in other thiol-disulfide oxidoreductases for which structures are available, the proline position is occupied by either a glycine, an alanine, or a histidine.

Since the ordered prolines are all accessible to solvent, we investigated the possibility that C-H–O interactions with water molecules might be observed. We restricted our distance criteria to a cutoff of 3.5 Å and note three C-H–O possible interactions involving Pro and Pro(Cα) Pro(OH) interacts with two solvents that are 3.27 and 3.39 Å away, whereas Pro(Cα) Cβ contacts a solvent 3.36 Å distant.

Tryparedoxin-I Surface for Substrate Recognition—The shape and charge at and around the redox-active disulfide of tryparedoxin-I allow it to interact with both a small molecule metabolite, trypanothione, and a protein, tryparedoxin peroxidase (Fig. 1). Fig. 7a shows the surface of tryparedoxin-I colored according to electrostatic properties and highlights the positions of the conserved Cys-Pro-Pro-Cys motif together with residues that might interact with substrates. A similar view of human thioredoxin is shown for comparison where the differences in size, charge, and shape around the conserved redox-active site motif are evident.

The crystal structure of T7 DNA polymerase in complex with thioredoxin (66) and NMR studies of thioredoxin and glutaredoxin complexes with peptides or glutathione (23, 24, 28) indicate that these proteins interact with ligands using the surface around the redox-active disulfide. In particular, the loop carrying the cis-proline adjacent to the redox-active disulfide participates in β-strand type associations. In tryparedoxin-I, the amide and carbonyl of Ile are positioned to accomplish just such an interaction. On the other side of the Cys-Cys disulfide, the carbonyl of Cys(40) is also directed out to accept hydrogen bonds from substrate (Fig. 6).

There is currently no structural information on tryparedoxin peroxidase, but there is a model of trypanothione based on the crystal structure of the complex with trypanothione reductase (63). We used these coordinates to construct a model for the tryparedoxin-tryparedoxine complex using the computer graphics program O (50). One of the metabolite sulfur atoms was overlapped with the water molecule that is hydrogen bonded to S Cys(40), and trypanothione was rotated about this point, seeking to minimize atomic collisions and produce sensible chemical interactions of the type previously observed for this ligand (63). The model (not shown) serves to identify residues within the range of the redox-active disulfide of tryparedoxin, which might be important in the recognition and binding of trypanothione.

When the polyamine substrates glutathionylpermidine disulfide and trypanothione bind in the active site of trypanothione reductase, they are positioned beneath a tryptophan and to a certain extent shielded from solvent (63, 67). When trypanothione binds to tryparedoxin-I, it would remain on the surface of the protein, and the secondary amine of the spermidine moiety would be solvent-accessible. Trp of tryparedoxin-I may participate in cation-π interactions with the secondary amine of trypanothione as observed for Trp of C. fasciculata trypanothione reductase when in complex with glutathionylpermidine disulfide (66). van der Waals interactions with the spermidine and cysteine components of trypanothione are also an important component of binding (63). In tryparedoxin-I, Trp, Pro, Pro(Cα), Pro(Cα), and Ile are near the redox disulfide and may interact in such a manner with trypanothione.

A cluster of acidic residues (Asp, Glu, Glu, Glu, Asp, and Glu) are located adjacent to the ledge created by Trp and Trp. These residues create a negatively charged patch on the surface of the protein that could attract then bind the positively charged substrate trypanothione (Fig. 7a). However, of these residues, only Asp and Glu are strictly conserved with the other tryparedoxins. Asp accepts hydrogen bonds from the hydroxyl group of Ser(O) to Asp(O) separation of 2.59 Å and the main chain amide of Ser(O) (N to O distance 2.73 Å). Although on the periphery of the active site, these interactions (not shown in any of the figures) serve to stabilize the position of the N-terminal sections of the parallel strands β and β. Ser is conserved in other tryparedoxins, and so this appears to be a common structural feature of this group of proteins. Glu is directed toward Trp and the closest contact is 5.81 Å between Glu(O) and Trp(O) N. This glutamic acid
is a likely candidate to interact with trypanothione and or tryparedoxin peroxidase.

Glutathione and trypanothione share γ-glutamyl groups, and the complex of glutaredoxin with glutathione indicates that the carboxylates interact with positively charged residues, including a lysine and an arginine (28). In tryparedoxin-I there are three charged residues just below the active site, Arg₄₄ Lys₈₃, and Arg₁₂₈ (Fig. 6). These residues are strictly conserved in the tryparedoxins and could form salt-bridges with the glutamyl α-carboxylates of trypanothione. We can use the solvent molecule that is hydrogen-bonded to Sy Cys⁴⁰ as a marker for the position of substrate interacting with tryparedoxin-I, as described above (Figs. 6 and 7b). Arg⁴⁴, Arg₁₂₈ and Lys⁸₃ are 7.9, 11.2, and 16.1 Å away, respectively, from this solvent position. In the complex of T. cruzi trypanothione reductase with trypanothione, there is a single charge-charge interaction formed between substrate and the enzyme, and this involves Lys⁷² with a γ-Glu carboxylate. The distance between the trypanothione disulfide and Lys⁷² Nᵢ is 12.2 Å (63). On the basis of the model described above, we predict that Arg⁴⁴ and Arg₁₂₈ of tryparedoxin-I are definitely within range to interact directly with the carboxylates of trypanothione. Mutagenesis studies and further crystallographic work will now be used to test these predictions.

Summary—We have determined an accurate crystal structure for tryparedoxin-I, a newly discovered thiol-disulfide oxidoreductase. The initial electron density map that was used to construct the first model of tryparedoxin-I was based on a MAD experiment at 1.7-Å resolution, targeting a selenomethionine derivative. Refinement of the model and incorporation of solvents was subsequently accomplished using diffraction terms to a resolution of 1.4 Å. Tryparedoxin displays a similar ar-
rangement of secondary structure elements to thioredoxin and can be classified as a member of the thioredoxin superfamily, although each protein presents a distinctive topology. The high degree of sequence homology among trypanothione reductases indicates that the structure of C. fasciculata tryparedoxin-I provides a model for this subset of the thioredoxin superfamily. Residues that might play an important, possibly essential role in the functions provided by the molecule are redox-active disulfide to accept and pass on reducing equivalents from and to its substrates. Although this is a common use of the disulfide, the structural details in and around the active site of tryparedoxin-I are distinct from other proteins for which structures have been determined. Comparisons with structures of trypanothione reductase, thioredoxin, and glutaredoxin provide a working model for tryparedoxin interactions with the reducing substrate and suggest which residues might be important for molecular recognition. The high resolution model now provides the framework for further experiments combing site-directed mutagenesis, biophysical methods, and enzyme kinetics to investigate the detailed structure-function relationships of trypanothione reductase. With the structural details now available, it may be possible to identify specific and potent inhibitors to disrupt the pathway detailed in Fig. 1. Such inhibitors could be useful tools to assist the characterization of the proteins function in vivo and might be widely applicable to studies of thiol-disulfide oxidoreductases. More important, however, is the potential use of the tryparedoxin-I model in a structure-based approach to identify new and improved therapies against a range of trypanosomal infections.

Acknowledgment—We thank the European Union Program and EMBL for their support allowing M. S. Alphey to carry out experiments at ESRF. We also thank A. Mehletr for carrying out the mass spectrometry measurements and C. S. Bond, D. R. Hall, and M. Peterson for contributions and support. Due to the extensive literature on thiol-disulfide oxidoreductases, it has not been possible to include references to all the primary literature, and we apologize to authors whose work has not been included.

REFERENCES

1. WHO (1993) Tropical Disease: Progress in Research 1989–1990, Programme report 10, World Health Organization, Geneva.

2. Webster, L. T. (1990) in The Pharmacological Basis of Therapeutics (Gilman, A. G., Ra1, T. W., Nies, A. S., and Taylor, P., eds) Section X, pp. 954–959, 978–988, and 1008–1017, 8th Ed, Pergamon Press, Inc., Tarrytown, NY.

3. Wang, C. C. (1997) Parasitology 114, 1–27.

4. Docampo, R. (1990) Annu. Rev. Microbiol. 44, 59–79.

5. Tetaud, E., and Fairlamb, A. H. (1998) J. Biol. Chem. 272, 21692–21699.

6. Holmgren, A. (1995) J. Mol. Biol. 254, 1–19.

7. Chang, P., Liu, B., Kang, S. W., Seo, M. S., Rhee, S. G., and Obeid, L. M. (1997) J. Biol. Chem. 272, 15661–15667.

8. Katti, S. K., Robbins, A. H., Yang, Y., and Wells, W. W. (1995) Protein Sci. 4, 1975–1985.

9. Sun, C., Berardi, M. J., and Bushwhacker, J. H. (1998) J. Mol. Biol. 280, 671–671.

10. Aslund, F., Berndt, K. D., and Holmgren, A. (1997) J. Biol. Chem. 272, 21692–21699.

11. Argon, A. S., and Taylor, P., eds) Section X, pp. 954–959, John Wiley & Sons, Inc., New York.

12. Raina, S., and Missiakas, D. (1997) J. Biol. Chem. 272, 21692–21699.

13. Barlow, D. J., and Thornton, J. M. (1988) Acta Crystallogr. Sec. D 44, 954–959.

14. Chae, H. Z., Chung, S. J., and Rhee, S. G. (1994) J. Biol. Chem. 269, 1097–1108.

15. Holmgren, A. (1995) J. Mol. Biol. 286, 541–552.

16. Raina, S., and Missiakas, D. (1997) J. Biol. Chem. 272, 21692–21699.

17. Barlow, D. J., and Thornton, J. M. (1988) Acta Crystallogr. Sec. D 44, 954–959.

18. Steiner, T., and Saenger, W. (1993) Biochem. Hoppe-Seyler 379, 1137–1142.

19. Zhang, P., Liu, B., Kang, S. W., Seo, M. S., Rhee, S. G., and Obeid, L. M. (1997) J. Biol. Chem. 272, 30952–30961.

20. Katti, S. K., Robbins, A. H., Yang, Y., and Wells, W. W. (1995) Protein Sci. 4, 1975–1985.

21. Dostal, J., and Rhee, S. G. (1991) J. Biol. Chem. 266, 290–296.

22. LeMaster, D. M., and Eklund, H. (1994) J. Mol. Biol. 248, 239–243.

23. Jin, D. Y., Chae, H. Z., Rhee, S. G., and Jeang, K.-T. (1997) J. Biol. Chem. 272, 30952–30961.