A potential target enzyme for trypanocidal drugs revealed by the crystal structure of NAD-dependent glycerol-3-phosphate dehydrogenase from *Leishmania mexicana*

Stephen Suresh[1], Stewart Turley[1], Fred R Opperdoes[2], Paul AM Michels[2] and Wim GJ Hol[1,3]*

Background: NAD-dependent glycerol-3-phosphate dehydrogenase (GPDH) catalyzes the interconversion of dihydroxyacetone phosphate and L-glycerol-3-phosphate. Although the enzyme has been characterized and cloned from a number of sources, until now no three-dimensional structure has been determined for this enzyme. Although the utility of this enzyme as a drug target against *Leishmania mexicana* is yet to be established, the critical role played by GPDH in the long slender bloodstream form of the related kinetoplastid *Trypanosoma brucei* makes it a viable drug target against sleeping sickness.

Results: The 1.75 Å crystal structure of apo GPDH from *L. mexicana* was determined by multiwavelength anomalous diffraction (MAD) techniques, and used to solve the 2.8 Å holo structure in complex with NADH. Each 39 kDa subunit of the dimeric enzyme contains a 189-residue N-terminal NAD-binding domain and a 156-residue C-terminal substrate-binding domain. Significant parts of both domains share structural similarity with plant acetohydroxyacid isomeroreductase. The discovery of extra, fatty-acid like, density buried inside the C-terminal domain indicates a possible post-translational modification with an associated biological function.

Conclusions: The crystal structure of GPDH from *L. mexicana* is the first structure of this enzyme from any source and, in view of the sequence identity of 63%, serves as a valid model for the *T. brucei* enzyme. The differences between the human and trypanosomal enzymes are extensive, with only 29% sequence identity between the parasite and host enzyme, and support the feasibility of exploiting the NADH-binding site to develop selective inhibitors against trypanosomal GPDH. The structure also offers a plausible explanation for the observed inhibition of the *T. brucei* enzyme by melarsen oxide, the active form of the trypanocidal drugs melarsoprol and cymelarsan.

Addresses: [1]Department of Biological Structure, Biomolecular Structure Center, Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA, [2]Research Unit for Tropical Diseases and Laboratory of Biochemistry, Christian de Duve Institute of Cellular Pathology and Catholic University of Louvain, ICP-TROP 74.39, Avenue Hippocrate 74, B-1200 Brussels, Belgium and [3]Department of Biochemistry, University of Washington, Seattle, WA 98195, USA.

*Corresponding author.
E-mail: hol@gouda.bmsc.washington.edu

Key words: glycerol-3-phosphate dehydrogenase, glycosome, isomeroreductase, *Leishmania mexicana*, trypanosomiasis

Received: 8 December 1999
Revisions requested: 14 January 2000
Revisions received: 13 March 2000
Accepted: 15 March 2000

Published: 3 May 2000

Structure 2000, 8:541–552

0969-2126/00/$ – see front matter

Introduction

Trypanosomatids are the cause of severe diseases in both humans and livestock, causing considerable suffering especially in developing countries [1,2]. Several *Leishmania* species are responsible for a variety of debilitating diseases, collectively called leishmaniases, throughout the tropics and subtropics, whereas two closely related trypanosomatids are the causative agents of two other major global diseases. *Trypanosoma brucei gambiense* and *T. brucei rhodesiense* cause sleeping sickness in Africa and *T. cruzi* is responsible for Chagas' disease in Latin America. The health problems caused by these protozoa are currently reaching disastrous proportions with the resurgence of sleeping sickness across central Africa [3,4]. There has also been an increase in the frequency of HIV–Leishmania co-infections in southern Europe [5]. The situation is further compounded by the fact that organisms related to trypanosomatids that are not normally infectious to humans (e.g., herpetomonads and leptomonads) are emerging in immunosuppressed patients [6]. Despite the grave seriousness of the situation, the drugs in use are outmoded and grossly inadequate. For example, melarsoprol, the only drug against late stage *T. brucei rhodesiense* infections, has such toxic side-effects that 5–10% of the patients treated die due to reactive encephalopathy [7]. There is, therefore, an urgent need for better drugs, especially as the effectiveness of the few available drugs is being eroded by increasing resistance [8].

Many biochemical aspects of the different life stages of these parasites have been studied in detail and several possible targets have been suggested for antitrypanosomal chemotherapy [7,9]. One aspect of special relevance is that the parasites are uniquely adapted to their complex lifestyle in both a vertebrate host and an insect vector, and

their energy metabolism changes depending on the availability of nutrients [10]. The bloodstream form of *T. brucei* is totally dependent on glycolysis for energy generation, and the first seven glycolytic enzymes are sequestered in a specialized organelle called the glycosome [11], where glycolysis proceeds at a feverish pace [12,13]. Computer modeling of the glycolytic flux supports the feasibility of inhibiting the glycolytic enzymes and thereby blocking glycolysis [14]. This Achilles heel is being exploited in our laboratories by targeting a number of these glycosomal glycolytic enzymes from both *L. mexicana* and *T. brucei* for structure-based drug design [11,15–18].

NAD-dependent glycerol-3-phosphate dehydrogenase (GPDH; EC 1.1.1.8) catalyzes the interconversion of dihydroxyacetone phosphate (DHAP) and L-glycerol-3-phosphate (G3P) using NAD as a cofactor. In the bloodstream form of *T. brucei*, GPDH is especially important as it is directly responsible for maintaining the NAD/NADH balance in the glycosome by the reoxidation of the NADH produced by glyceraldehyde-3-phosphate dehydrogenase during glycolysis. The G3P generated is then transported across the glycosomal membrane and reoxidized via the G3P:DHAP shuttle that functions in conjunction with a mitochondrial glycerol-3-phosphate oxidase [11]. Blocking GPDH activity would seriously jeopardize ATP generation by the trypanosome. The inhibition of GPDH would, in addition, lead to the accumulation of DHAP in the glycosome. This would be very harmful to the trypanosome as DHAP is spontaneously converted to methylglyoxylate, a toxic compound that reacts aspecifically with proteins [19,20]. In contrast to the vital role of GPDH in *T. brucei*, the importance of this enzyme in *L. mexicana* and its value as a drug target are yet to be evaluated, because in *Leishmania* species glycolysis is less important for the energy supply and glycolytically produced NADH is thought to be reoxidized mainly via a glycosomal malate dehydrogenase [11,21,22].

Glycosomal GPDH from *L. mexicana* (*Lm*GPDH) the main subject of the present paper, is a polypeptide of 366 amino acids that forms a homodimer of 78 kDa [21]. Like many of the other glycolytic glycosomal enzymes, *Lm*GPDH has a very high isoelectric point of 9.5. GDPH from *T. brucei* is a polypeptide of 354 amino acids, forms a dimer of 76 kDa [21] and shares 63% sequence identity with the *L. mexicana* enzyme. In both organisms the enzyme is expressed in the cytosol and then targeted to the glycosome with the aid of a type 1 C-terminal tripeptide peroxisomal targeting signal, or PTS1 [13]. This tripeptide has the sequence SKL (in single-letter amino acid code) in *Lm*GPDH and SKM in the *T. brucei* homolog. Here we present the crystal structure of *Lm*GPDH, both in its apo form and as a holo enzyme in complex with NADH. To the best of our knowledge, this is the first report of the crystal structure of this enzyme from any source. The two-domain GPDH subunit shares significant structural similarity with plant acetohydroxyacid isomeroreductase. Another surprise of the structure determination was the discovery of extra fatty-acid-like density buried inside the C-terminal domain.

Results and discussion

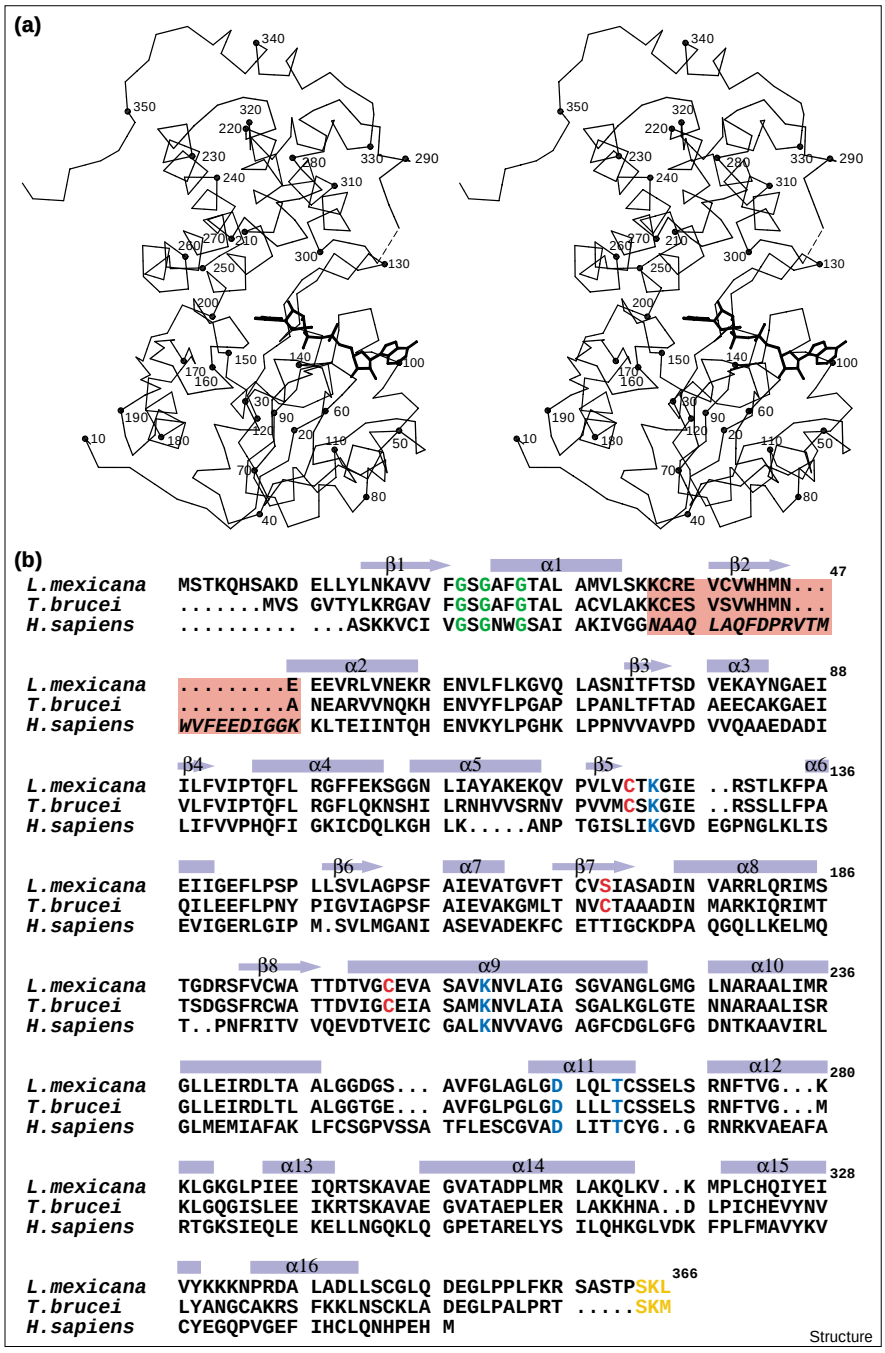
Description of the overall structure

The crystal structure of *Lm*GPDH at 1.75 Å resolution has been refined to an R factor of 19.5%, with 93.3% of the residues in the most favorable region of the Ramachandran map [23]. The remaining 6.7% of the residues are in the additional allowed regions, and there are no residues in the generously allowed or disallowed regions. The asymmetric unit contains a monomer composed of two domains (Figures 1,2), as observed in other NAD(P)-dependent dehydrogenases. The N-terminal NAD-binding domain (residues 9–197) contains a canonical six-stranded parallel β sheet [24] with helices on either side. In *Lm*GPDH these six strands are part of a larger eight-stranded β sheet with an extra pair of strands running antiparallel to the six-stranded Rossmann fold. The first β-α-β unit (Figures 1b,2) contains the highly conserved GxGxxG NAD-binding motif [25] that is present in all NAD-dependent GPDHs studied so far [26]. The C-terminal substrate-binding and catalytic domain (residues 200–357) is exclusively α-helical. This domain contains a total of eight helices, which comprise 64.6% of the chain of this domain, and one particularly long helix (helix 9; Figure 1b) consisting of residues 200–222. The two domains are linked by a single three-residue loop.

The biological dimer is generated by crystallographic twofold symmetry. Residues 343–357 of the C terminus of each monomer wrap around the other monomer of the biological dimer and the residues of the two C termini approach each other to within 10 Å (Figure 2). Although the density for most of the structure was well defined (Figure 3), the last nine residues (358–366) of the C terminus, containing the PTS1 glycosomal targeting signal SKL, are missing in density, as are the initial eight residues and five residues of loop 293–297 in the C-terminal domain.

*Lm*GPDH has a large dimer interaction surface, burying 1200 Å^2 and 1924 Å^2 out of 7888 Å^2 and 7472 Å^2 accessible surface area on the N- and C-terminal domains, respectively. The subunit–subunit interface is 70% nonpolar, but is also stabilized by a number of hydrogen bonds and interactions between sidechains of opposite charge. Although the majority of the dimer contacts come from the C-terminal α-helical domain, residues 156 to 166 of the N-terminal domain contribute some important specific interactions. In particular, the mainchain amino group of Phe156 is involved in an intermonomer hydrogen bond with the sidechain of Asn228. Asn228 is conserved among 23 of the 24 known GPDH sequences (the only exception being GPDH from *Helicobacter pylori*, where this residue is a serine). Also Glu159, which is totally conserved among the 24 known GPDH sequences, is involved in intermonomer

Figure 1



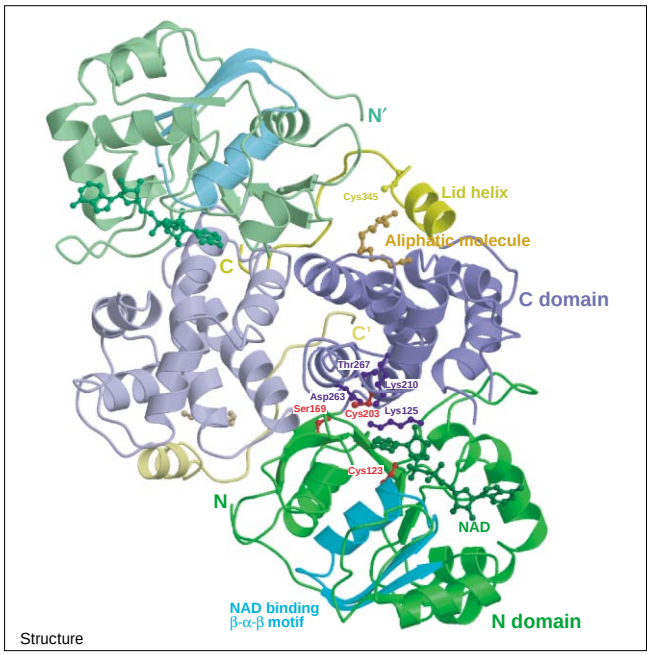
The tertiary and secondary structure of *L. mexicana* GPDH. (a) Stereoview Cα trace of one monomer of holo *Lm*GPDH. The bound cofactor NAD is depicted in bold as a stick model. The initial eight residues and last nine residues are missing in density, as are the residues in the loop 293–297, which is depicted by dotted lines. This figure was made with the program Molscript [50]. (b) Amino acid sequence alignment of GPDH from *L. mexicana, T. brucei* and *Homo sapiens* (SWISS-PROT accession numbers P90551, P90593 and P21695, respectively). The NAD-binding GxGxxG sequence motif is colored green, putative active-site residues are shown in blue, residues implicated in interactions with the trypanocidal drug cymelarsan in *T. brucei* GPDH (and the equivalent residues in *Lm*GPDH) are in red, and the C-terminal PTS1 glycosomal targeting signal is colored gold. Shaded in salmon is the region of the sequence that could not be aligned with confidence, due to the very low sequence similarity between the human and the two trypanosomatid sequences; the corresponding residues (Asn24 to Lys47) of the human sequence have been italicized.

hydrogen bonds to mainchain amino groups of Leu227 and Asn228. The large dimer interface and the highly conserved nature of the interface residues Gln159 and Asn228 suggest that the *Lm*GPDH structure represents the general dimeric picture of this enzyme across species.

The NAD-binding site

The holo structure of *Lm*GPDH in complex with NADH was determined from crystals grown in the presence of 15 mM NADH. Crystals grown in the presence of NAD^+ under similar conditions did not contain the cofactor, presumably because the K_m for NAD^+ is 420 μM as opposed to 24 μM for NADH. The structures of apo and holo *Lm*GPDH are very similar to each other with 346 Cα atoms superposing with a root mean square deviation (rmsd) of 0.4 Å. The binding mode of NADH to *Lm*GPDH was clearly visible in the 2.8 Å holo enzyme electron-density map (Figure 4a). The interactions of the

Figure 2



Tertiary structure of dimeric *Lm*GPDH with the reference monomer shown in brighter colors than the other. Within each monomer the N-terminal NAD-binding domain is colored green and the C-terminal substrate-binding domain is violet. Highlighted in cyan is the nucleotide-binding β-α-β structural motif; bound NAD is depicted as a ball-and-stick model in green. The aliphatic molecule near Cys345 is shown in gold. Putative active-site residues are highlighted in blue and residues corresponding to those reported to interact with the trypanocidal drug cymelarsan in *T. brucei* are in red [20]. The last residues (335–357) visible in the electron-density map, which precede the C-terminal tail with residues 358–366 containing the PTS1 SKL glycosomal import signal, are depicted in bright yellow. This figure was made with Molscript [50] and Raster3D [51].

enzyme with the cofactor are depicted in Figure 4b. An unusual feature of the NAD-binding site is the absence of an acidic residue at the end of the second β strand of the β-α-β motif. It is characteristic of NAD-binding dehydrogenases (as opposed to NADP-binding dehydrogenases) that there is an aspartate or a glutamate residue at this position interacting with the 2′-hydroxyl group of the adenosine ribose [25,26], although exceptions to this pattern have been observed before [27]. In *Lm*GPDH (Figures 4a,b), which is an NAD-dependent enzyme, this position is occupied by a histidine residue (His45). Compared to other dehydrogenases, the mainchain of His45 is moved away from the NAD with the nearest atom, the imidazole Cγ, being at a distance of 4.7 Å from the 2′-hydroxyl. Instead of the usual acidic residue hydrogen bonding to the 2′-hydroxyl of the adenosine sugar, in the *Lm*GPDH structure this hydroxyl hydrogen bonds to a water molecule, which in turn interacts with the mainchain amino group of Met46. The 3′-hydroxyl group is stabilized by a hydrogen bond to Ser23. The adenine ring packs against Phe97 and is shielded from solvent by Met46, while the adenosine sugar sits snugly against Ile93 and Pro94. At the other end of NADH, the nicotinamide makes hydrophobic contacts with Phe26, whereas the nicotinamide ribose packs against Thr124 with its 3′-hydroxyl making a hydrogen bond to the mainchain amino group of Lys125. The 2′-hydroxyl of the nicotinamide sugar is stabilized by a hydrogen bond to the sidechain carboxylate of Glu300, provided by helix α14. The rotational isomer of the glycosyl bond around the nicotinamide is *syn*, with the hydride being added to the *pro*S position of the pyridine ring based on the putative substrate-binding site (discussed below).

The active site

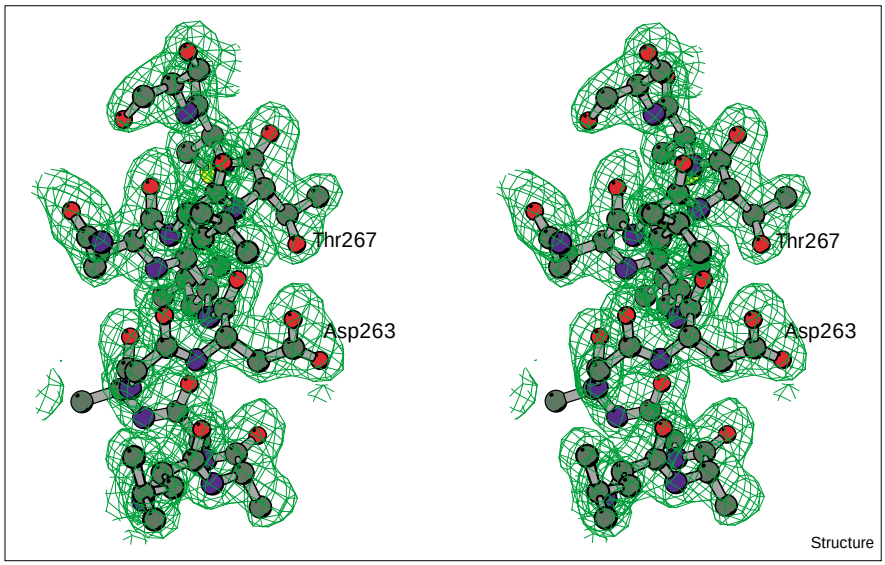
The structure of *Lm*GPDH in complex with NADH permits us to delineate the putative active site. Residues Lys125, Lys210, Asp263 and Thr267 are located in the vicinity of the nicotinamide ring (Figure 5), suggesting that they might be involved in substrate binding or catalysis. This idea is supported by the fact that Lys125, Lys210 and Glu263 are absolutely conserved in the 24 available sequences, whereas Thr267 is conserved in 23 sequences and is a serine residue in *Saccharomyces pombe* A [28]. The basic sidechains of the two lysine residues create a phosphate-binding site (as determined using the program GRID [29]) that presumably accommodates the phosphate of the substrate G3P during catalysis. The carboxylate oxygens of Asp263 are 6.3 Å from C4 of the nicotinamide and the hydroxyl oxygen of Thr267 is 7.2 Å from C4. It is likely that both Asp263 and Thr267 move towards the nicotinamide during catalysis, as the domains may come closer upon substrate binding. Such domain motions during substrate and cofactor binding have been observed in a number of dehydrogenases including glutamate dehydrogenase [30,31], liver alcohol dehydrogenase [32] and glyceraldehyde-3-phosphate dehydrogenase [33]. Other residues relatively close to the nicotinamide that could possibly have a role in substrate binding are Asn211, Arg274 and Asn275, of which the latter two are totally conserved, whereas Asn211 is conserved in 23 known sequences and is an aspartate in rabbit GPDH [34]. The sidechain amide of Asn211 is 7.8 Å from the nicotinamide C4 and the amide of Asn275 is 9.6 Å from C4 in the holo complex, while the sidechain of Arg274 appears flexible. Although the elucidation of the reaction mechanism clearly requires more structural and biochemical investigation, it is likely that Asp263 and Thr267 are critical catalytic residues as these are closest to the C4 of the nicotinamide.

Structural similarities and evolutionary links

Although the similarity of the N-terminal domain of GPDH to other NAD-binding enzymes was anticipated, it was surprising to discover, using the program DALI [35], that the α-helical C-terminal domain has a partial but striking similarity to the bifunctional enzyme acetohydroxyacid isomeroreductase (EC 1.1.1.86) from plants. The *Spinacia*

Figure 3

Representative section of the experimental electron-density map after density modification of apo *Lm*GPDH at 1.75 Å. The highly conserved putative active-site residues Asp263 and Thr267 are labeled. The figure was made with Bobscript [52].



oleracea isomeroreductase [36] is a dimer, but has entirely different subunit–subunit interactions to those of *Lm*GPDH (data not shown). However, the similarity in structure of the individual subunits of these two enzymes is especially interesting considering that acetohydroxyacid isomeroreductase catalyzes an alkyl migration followed by a reduction, uses NADPH as a cofactor, and needs a bivalent metal cation for activity, whereas, in contrast, *Lm*GPDH does not migrate an alkyl function, uses NAD as cofactor, and does not need a metal cation for activity.

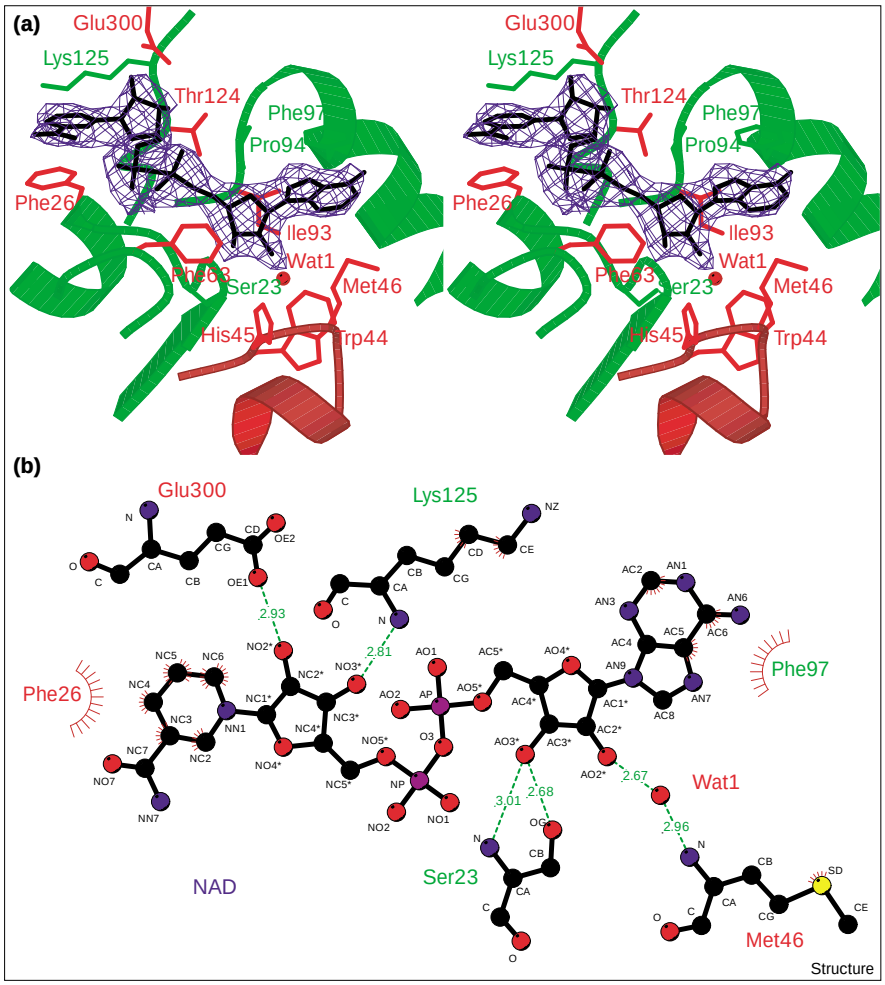
The structural similarity encompasses both the topology as well as the positioning of the secondary structure elements in both the canonical NAD(P)-binding N-terminal domain as well as parts of the substrate-binding C-terminal domain (Figure 6). The 158 equivalent Cα atoms in the two structures superpose with an rmsd of 3.2 Å despite a sequence identity of only 13.3%. The cofactor-binding domains have 103 residues in common with a sequence identity of 16.5% and an rmsd of 2.9 Å. In the C-terminal domains, 55 residues superimpose with an rmsd of 3.4 Å (sequence identity as low as 7.3%). The common feature in the C-terminal domains of both enzymes is a three-helix bundle (Figure 6), two helices of which provide key residues to the active site. In the isomeroreductase the helices comprising residues Leu308–Arg320 and Leu324–Glu340 contribute Asp315 and Glu319 to chelate the Mg^{2+} ions at the active site [36], whereas in *Lm*GPDH the structurally equivalent long helix (Thr200–Gly222) contributes Lys210 to the putative active site. In addition, the consecutive helices Tyr486–Ser493 and Ile495–Asp499 of the isomeroreductase contribute Glu492 and Glu496 to the putative active site, and in *Lm*GPDH the equivalent helix (Leu261–Cys268) contributes Asp263 and Thr267 to the active site. Although the two enzymes bind different substrates, the specific catalytic residues are in a similar spatial juxtaposition. After superposition of the 158 equivalent Cα atoms of the two enzymes, the Cα atom of Lys210 in *Lm*GPDH is 3.1 Å from the Cα atom of Glu319 in the isomeroreductase, the Cα atom of Asp263 of *Lm*GPDH is 3.0 Å from Glu492 of the isomeroreductase, and the Cα atom of Thr267 of *Lm*GPDH is 4.9 Å from the Cα atom of Glu496 of the isomeroreductase. Similar distances between catalytically equivalent residues in distantly related proteins have been reported elsewhere (see e.g. [37]). This extensive similarity of *Lm*GPDH to a functionally distantly related plant enzyme raises interesting questions regarding its evolutionary history.

One possible explanation for the structural similarity between GPDH and the isomeroreductase is that they are descendants of the same bifunctional protein and have widely diverged during the course of evolution, while maintaining essentially the same relative orientation of the three common C-terminal helices with respect to the N-terminal nucleotide-binding domain. Another possibility is that the C-terminal domains of these two enzymes evolved independently in different kingdoms and were fused to one of the many genes coding for nucleotide-binding proteins. The similarity in the orientation of the C-terminal helices with respect to the N-terminal domain, would in this scenario, either be fortuitous or imposed by the constraint of placing critical residues in the active site located in-between the two domains.

Post-translational modification of trypanosomatid GPDHs

A most surprising discovery of our structure determination was the observation of extra electron density in the unbiased experimental map at 1.75 Å resolution

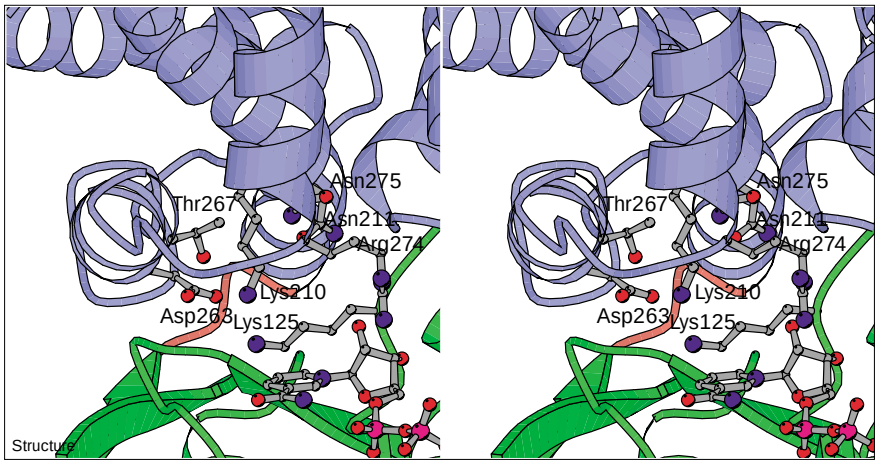
Figure 4



Cofactor binding by GPDH. **(a)** Simulated annealing omit map [53] of the *Lm*GPDH–NADH holo complex, contoured at 2.8σ, superposed on the refined model of NADH. Residues that differ between the human and the two trypanosomatid enzymes (*L. mexicana* and *T. brucei*) are in red; conserved residues are in green. Also indicated in red are the water molecule that forms a hydrogen bond to the 2′ oxygen of the adenosine ribose in *Lm*GPDH, instead of the usual acidic residue that is found in most NAD-dependent dehydrogenases [25,26], and the region around the β2 loop that shows considerable sequence difference between human and trypanosomatid enzymes (salmon-shaded in Figure 1b). **(b)** A LIGPLOT [54] representation of the cofactor-binding site, showing the interactions between NADH and *Lm*GPDH. Residues that are conserved between the human and two trypanosomatid enzymes are labeled in green; those that are different are labeled in red. Atoms are shown in standard colors.

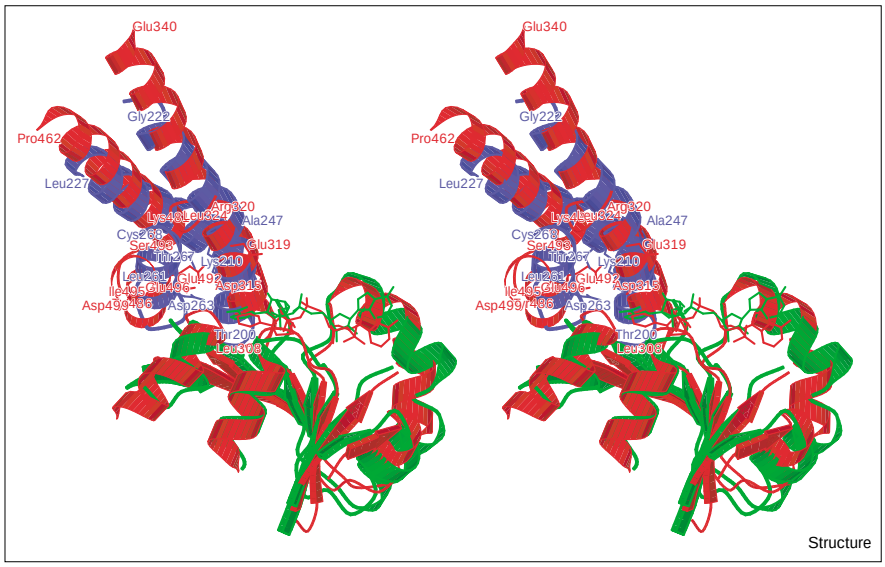
(Figure 7). The density is compatible with an extended aliphatic molecule with a length of approximately 16 carbon atoms. The molecule is located in a well defined, entirely hydrophobic pocket created by sidechains from

Figure 5



Close-up stereoview of the active site of the *Lm*GPDH–NADH holo complex. Residues from the N-terminal NAD-binding domain are shown in green, residues from the C-terminal substrate-binding domain in violet, and the three-residue linker between the two domains is in red. Active-site residues Lys125, Lys210 and Asp263 are totally conserved among the 24 known GPDH sequences, whereas Thr267 is a serine in *Saccharomyces pombe* A and a threonine in the other 23 sequences.

Figure 6

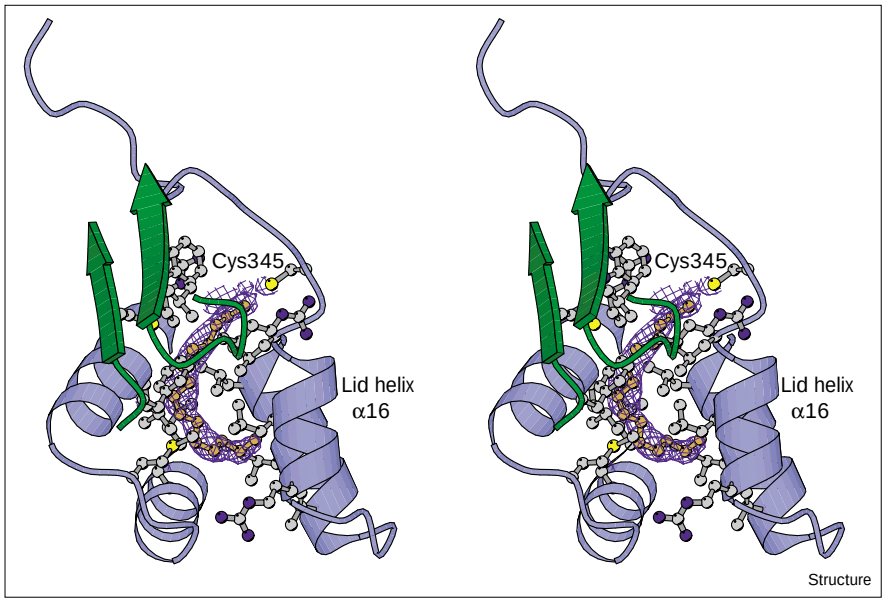


Structurally conserved regions between *Lm*GPDH and *Spinacia oleracea* acetohydroxyacid isomeroreductase. In addition to the similarity of the N-terminal NAD-binding region (in green) of *Lm*GPDH to the isomeroreductase (in red), a three-helix bundle (in violet) in the C-terminal substrate-binding domain of *Lm*GPDH is structurally conserved in the isomeroreductase (in red). Specifically, helix 9 (Thr200–Gly222) of *Lm*GPDH is structurally equivalent to the helices Leu308–Arg320 and Leu324–Glu340 of the isomeroreductase, helix 11 (Leu261–Cys268) of *Lm*GPDH is equivalent to the consecutive helices Thr486–Ser493 and Ile495–Asp499 of the isomeroreductase, and helix 10 (Leu227–Ala247) of GPDH is equivalent to Pro462–Lys481 of the isomeroreductase. The bound cofactors are depicted as stick models: NAD in *Lm*GPDH is in green, and NADPH in the isomeroreductase is in red. Also indicated are the residues from the critical three-helical bundle of the C-terminal domain that are believed to contribute to the active site. In performing the superposition, segments 16–35, 40–46, 52–58, 74–108, 117–124, 147–163, 165–172, 199–213, 215–224, 228–244 and 258–269 of *Lm*GPDH were superposed onto segments 126–145, 156–162, 168–174, 182–217, 218–225, 245–261, 272–279, 307–321, 325–334, 466–482 and 487–498 of the isomeroreductase.

the C-terminal domain of one monomer and the N-terminal domain of the other monomer of the biological dimer (Figure 7). One end of the density approaches Cys345. Although the density did not indicate a covalent linkage to this cysteine in the recombinant enzyme produced in *Escherichia coli*, it is plausible that the native enzyme in the trypanosome is post-translationally modified by a covalent linkage to this cysteine which is conserved among the *L. mexicana* and *T. brucei* GPDH sequences.

Covalent modification of cysteine residues has been known to occur in *T. brucei*, with reversible thiomyristoylation being implicated in the regulation of the integral membrane protein glycosylphosphatidylinositol-specific

Figure 7

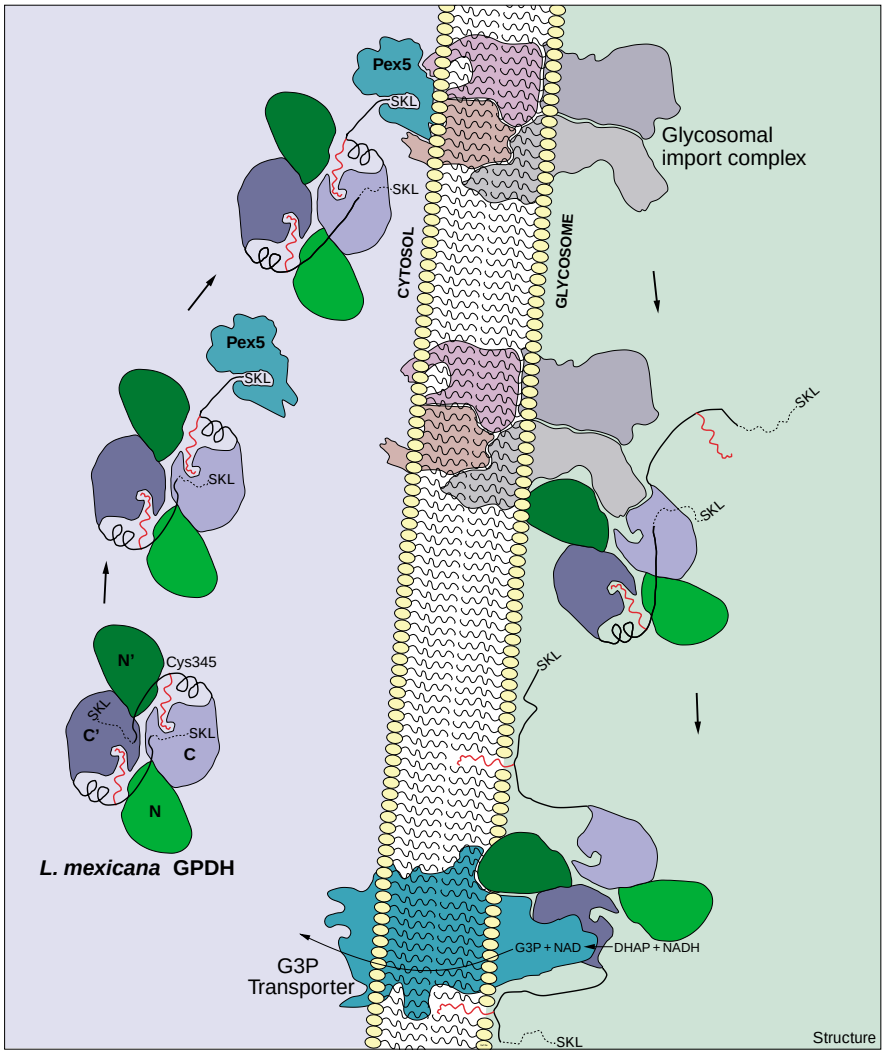


Density-modified MAD map, prior to model building, showing the extra density in the hydrophobic pocket of recombinant apo *Lm*GPDH. This density was subsequently modeled as a linear aliphatic molecule. Secondary structure elements from the C-terminal domain are colored violet and those from the N-terminal domain of the symmetry-related monomer are colored dark green. The cysteine residue (Cys345) that is believed to be post-translationally modified in the native enzyme is labeled, as is the helix (the lid helix) that is postulated to move during glycosomal import, exposing the aliphatic modification and facilitating glycosomal membrane association.

phospholipase C (GPI-PLC) [38]. More recently it has been demonstrated that GPI-PLC introduced as a reporter protein in *Leishmania major* also undergoes S-myristoylation [39]. In *Trypanosoma cruzi* it has been observed that palmitoylation of a cysteine residue from a flagellar calcium-binding protein, coupled with myristoylation of the N-terminal glycine, is necessary for the association of the protein to the flagellar membrane [40]. Hence, the extra density near Cys345 suggests the occurrence of a post-translational modification of this cysteine in trypanosomatids, even though the electron density does not indicate a covalent link between the extra density and this cysteine in the recombinant enzyme.

A possible connection between recognition of the PTS1 signal and a function of the putative covalent modification of Cys345 might exist. Although the last nine residues (358–366) of *Lm*GPDH, including the PTS1 import signal, are not visible in the density, the preceding residues are well ordered. Of these, residues 343 to 357 wrap around the second monomer of the dimer (Figure 2), while helix α16, comprising the preceding residues 335–342, is involved in forming the hydrophobic pocket containing the extra fatty acid density (Figure 7). Inspection of the structure suggests that movement of the 'lid helix' α16 would open up the hydrophobic pocket and expose the covalently linked aliphatic molecule, which could then associate with the glycosomal membrane. Such an opening of the pocket is postulated to take place during the membrane translocation process and hence after the interaction of the C-terminal PTS1 tripeptide with its Pex5 receptor (Figure 8). This sequence of events would permit this trypanosomal enzyme to be soluble prior to transport into the glycosome, and once imported into the glycosome to become associated with the inner leaflet of the glycosomal membrane. Such a membrane association of this enzyme

Figure 8



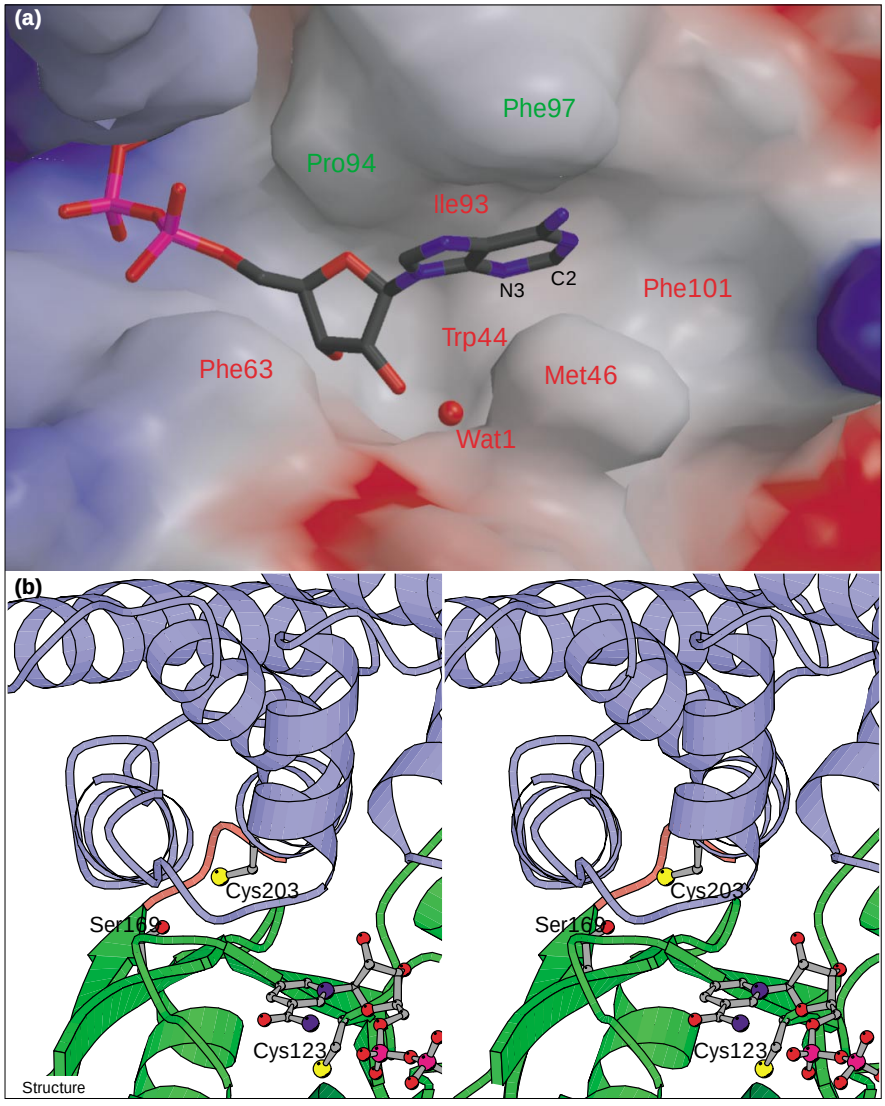
Hypothetical scheme for the import of *L. mexicana* and other trypanosomatid GPDHs into the glycosome, mediated by Pex5, and subsequent membrane association through an aliphatic modification of a cysteine residue. It is hypothesized (see text) that prior to targeting to the glycosome, the putative aliphatic modification is enclosed in a hydrophobic pocket within the soluble protein, and once imported into the glycosome movement of a helix that caps the hydrophobic pocket would open up the pocket, exposing the aliphatic molecule and permitting membrane association.

would be consistent with the fact that the product of the reaction it catalyzes, namely G3P, has to be transported across the glycosomal membrane to be reoxidized by the mitochondrial G3P oxidase. Membrane association of trypanosomatid GPDH would also be in accordance with the observation of Stebeck *et al.* [41], that during the purification of *T. brucei* GPDH from African trypanosomes, the enzyme is present in the glycosomal membrane fraction. It is possible that GPDH could be somehow associated with the G3P transporter (Figure 8). Clearly, further experimental studies are needed to confirm this hypothesis.

Drug design

The vital role played by GPDH in the glycosome of the bloodstream form of *T. brucei*, coupled with the low degree of sequence conservation between the trypanosomal and the human enzyme, makes it a promising target for drug design. Computer modeling of the glycolytic flux, based on experimentally determined kinetic parameters of the trypanosomal enzymes, corroborates the feasibility of blocking glycolysis in *T. brucei* by inhibiting GPDH [14]. While GPDH may or may not be a good drug target in *Leishmania* species, the enzyme from *L. mexicana* presented here provides us with an excellent model for the trypanosomal enzyme. The fact that the active sites of enzymes are usually extremely well conserved, makes the development of inhibitors with high affinity for the active site of the parasite enzyme, and with low affinity for the active site of the homologous host enzyme, often an intractable problem. It is possible to circumvent this obstacle when the enzyme utilizes a large cofactor like NAD, as usually there are substantial differences between the parasite and host enzymes in regions that bind parts of the cofactor that are distant from the active site. Selective inhibitors can then be designed

Figure 9



Drug target areas in trypanosomatid GPDH. (a) View of the hydrophobic pocket near the adenosine-binding site that can potentially be exploited for structure-based drug design. Labeled in red are residues that differ between the human and *T. brucei* sequences; residues that are identical are labeled in green. This figure was made with GRASP [55] and Molscript [50]. (b) Residues in *Lm*GPDH that correspond to those implicated in interactions with the trypanocidal drug cymelarsan in *T. brucei*. Cys123 and Cys203 are conserved between *L. mexicana* and *T. brucei*, whereas Ser169 is a cysteine residue that is unique to *T. brucei*. The Oγ atom of Ser169 is 4.0 Å from the Sγ atom of Cys203.

that bind these regions of the cofactor-binding site on the parasite enzyme with higher affinity than those of the host enzyme. This approach has been successful in the design and synthesis of selective inhibitors of trypanosomatid glyceraldehyde-3-phosphate dehydrogenase [18].

The sequences of GPDH from *T. brucei* and *L. mexicana* are 63% identical and there are no insertions or deletions. The two trypanosomatid sequences differ considerably from the human enzyme, with only 28.5% and 29.8% identity, respectively (Figure 1b). As can be seen from Figure 9a, the differences are significant especially in the binding site of the adenosine portion of NAD. In fact, the residues in the loop after the second β strand of the canonical β-α-β motif (after His45) are so different (Figure 1b) that it is not possible to be confident about the alignment. As mentioned previously, in most NAD-binding dehydrogenases the residue equivalent to His45 of *Lm*GPDH is occupied by an acidic residue that makes a hydrogen bond to the 2′ oxygen of the adenosine sugar. In the human enzyme there is an insertion in this region (Figure 1b), indicating substantial differences in the tertiary structure. We hope to exploit these differences in the adenosine-binding region to develop adenine analogs as selective inhibitors against the parasite enzyme.

It is also anticipated that the affinity of these analogs could be improved by exploiting the hydrophobic cavity below the adenine ring formed by Ile93, Trp44 and Met46 (Figure 9a). Filling this pocket by attaching hydrophobic groups to the C2 position of the adenine would be likely to create favorable hydrophobic interactions between the ligand and target proteins and result in tight-binding inhibitors. Alternatively, adding substituents to the 2′ position of the adenosine sugar, which would displace the hydrogen bonded water molecule and interact with hydrophobic groups of nearby residues, might increase the affinity for the parasite enzyme but decrease the affinity for the human homolog.

The role of GPDH dehydrogenase as a possible trypanosomal drug target is underscored by the observed inhibition of the enzyme by cymelarsan (an analog of melarsoprol), suramin, pentamidine and berenil [20,42,43], drugs in use against African trypanosomiasis. The inhibition by cymelarsan is especially interesting as this arsenical, which is the only available treatment against late stage *T. brucei rhodesiense* infections, forms covalent links with *T. brucei* GPDH [20]. Of the two cysteine residues implicated in interactions with this compound, Cys123 in *Lm*GPDH (Cys116 in *T. brucei*) is close to the NAD-binding site, with its Sγ atom 7.7 Å from C6 of the nicotinamide (Figure 9b). Although the Sγ atom of Cys123 points towards the hydrophobic interior of the protein in both the native and holo structures, a change in the sidechain χ_1 dihedral angle could expose the Sγ atom to the solvent. The other cysteine, Cys203 in *L. mexicana* (Cys196 in *T. brucei*), is located near the active site with its Sγ atom positioned 8.8 Å and 8.9 Å from the nearest sidechain atoms of the active-site residues Lys125 and Asp263, respectively. This cysteine occurs at the interface between the N- and C-terminal domains and is, in the present structures, not accessible to solvent. The Sγ atoms of Cys123 and Cys203 are 11.5 Å apart in the *Lm*GPDH structure. Interestingly, a third cysteine (Cys162) occurs in the *T. brucei* enzyme, which is a serine (Ser169) in *Lm*GPDH. This residue also occurs at the domain interface near the active site with its Oγ atom 4.0 Å from the Sγ of Cys203 (Cys196 in *T. brucei*). This suggests the possibility that an interdomain motion may open up the interface and that subsequently a divalent interaction of Cys162 and Cys196 with cymelarsan occurs in *T. brucei*. This is consistent with the observation that cymelarsan immobilized on a column binds *T. brucei* GPDH [20] but does not bind *Lm*GPDH (T Baltz, personal communication), presumably because the latter enzyme does not have two cysteine residues in the active site sufficiently close together. Although it might be difficult to alleviate the toxicity of this arsenical drug, and derivatives thereof, further studies on its precise interaction with *Lm*GPDH, and in particular with the *T. brucei* enzyme, would be of value for the development of novel inhibitors.

Biological implications

The crystal structure of NAD-dependent glycerol-3-phosphate dehydrogenase (GPDH) from *Leishmania mexicana* presented here is the first structure of the enzyme from any source. The structure reveals that the enzyme has the canonical two-domain NAD-dependent dehydrogenase architecture, with an N-terminal NAD-binding domain and a C-terminal substrate-binding domain connected by a single short loop. The biological molecule is a dimer with residues from both the N- and C-terminal domains contributing to the extensive dimer interface. The complex with NADH allows identification of the residues involved in NAD binding and those forming the putative active site. The structure also provides a plausible explanation for the observed inhibition of the *T. brucei* enzyme by the trypanocidal drug cymelarsan. The residues in the *L. mexicana* structure corresponding to those implicated in interactions with the drug in the *T. brucei* enzyme are near the active site and cofactor-binding site.

A surprising observation was the extensive structural similarity of *L. mexicana* GPDH to the enzyme acetohydroxy-acid isomeroreductase from *Spinacia oleracea*. The similarity extends not only to the N-terminal NAD-binding domain, but also to the C-terminal substrate-binding domain. Specifically, the C-terminal domains of both structures contain a crucial three-helix bundle, two helices of which contribute active-site residues. This unexpected similarity to a functionally distantly related plant enzyme suggests several possible scenarios with respect to the evolutionary relationship of these two proteins.

An interesting finding of the structure elucidation was the presence of extra electron density compatible with an extended aliphatic molecule in a well defined, entirely hydrophobic pocket with one end of the density close to the Sγ atom of Cys345. This observation suggests a post-translational modification of this cysteine in trypanosomatid GPDH, which may be of relevance for the localization of the enzyme inside the glycosome.

Materials and methods

Preparation and purification of Lm*GPDH*

*Lm*GPDH overproducing *E. coli* were essentially grown and lysed as previously described [21]. The cell lysate was clarified by centrifugation at 10,000 g for 30 min and the supernatant with the soluble protein was passed through a 0.45 μm filter. The high pI of the enzyme enabled the purification of the crude lysate by a single pass over a SPRINT 20HS cation exchange column. The enzyme was then concentrated to 3.5 mg/ml and buffer exchanged into 20 mM TEA pH 7.2, 4 mM NAD, 4 mM DTT, 1 mM EDTA and 0.2 mM PMSF.

Crystallization and data collection

A number of different crystal forms of *Lm*GPDH were grown employing the sitting drop vapor diffusion method with a variety of precipitants including sodium formate, sodium citrate and PEG 5000 mme (mono methyl ether). Crystals grown at room temperature by mixing 4 μl of the protein solution with 2 μl of the reservoir solution composed of 0.9 M sodium citrate were the most robust and were utilized for structure solution. These crystals were transferred to a cryo-solution containing 0.9 M sodium citrate, 25 mM TEA pH 7.2, 4 mM NAD, 5 mM DTT, 0.2 mM EDTA and 15% v/v ethylene glycol, and frozen under a stream of nitrogen prior to data collection. The single platinum derivative used in the MAD phasing experiment was grown by equilibrating 3 μl of the protein solution against 2 μl of the reservoir solution containing 0.6 mM $K_2Pt(CN)_6$ and 0.95 M sodium citrate, and the crystal obtained was frozen as described above. Despite the presence of 4 mM NAD in the protein solution, the native crystal structure did not contain any density for the dinucleotide. Crystals in complex with NADH were grown by substituting 15 mM NADH for the NAD and setting up crystallization experiments as before. The holo crystals were transferred to the above mentioned cryo-solution that contained 15 mM NADH (instead of 4 mM NAD) and frozen as before.

Structure determination

The native apo and holo data sets were collected at the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 7-1 and processed with DENZO and SCALEPACK (Table 1). The platinum MAD data set at four wavelengths was collected on SBC beamline 19-ID at the Advanced Photon Source (APS) in the Argonne National Laboratory. After an Extended X-ray Absorption Fine Structure (EXAFS) scan on a frozen crystal to determine the peak and inflection point, a four-wavelength MAD experiment was conducted with a low energy and a high energy reference. The MAD data set was essentially complete with high redundancy to 2.5 Å, but the completeness fell off rapidly between 2.5 and 2.0 Å. A single platinum site was determined from difference Pattersons and its parameters refined with SHARP [44], before locating two weaker sites from residual Fouriers and including them in the phasing process. A 2.5 Å resolution map, calculated after solvent flipping with the SOLOMON option in SHARP, was easily interpretable. This map was then used to perform further density modification and phase extension to 1.75 Å with the isomorphous native (apo) data set using the program WARP [45]. The density at this stage was simply superb (Figure 3). The model was built with the program O [46]. After initial refinement with X-PLOR [47] the model was refined with TNT [48]. The holo crystals were essentially isomorphous to the apo form and after initial rigid-body positioning the structure was refined with CNS [49]. Refinement statistics are provided in Table 1.

Table 1

Data collection and refinement statistics.

	Native	NADH co-crystal
Space group	$P4_12_12$	$P4_12_12$
Cell parameters (Å)	a = b = 70.3; c = 210.1	a = b = 69.9; c = 211.9
Resolution (Å)	1.75	2.8
Data processing		
Observations	229,249	93,667
Unique reflections	46,462	13,746
Completeness (%)*	85.4 (63.7)	99.6 (100.0)
R_{merge}*	0.044 (0.317)	0.066 (0.183)
Refinement		
Rmsd from ideality		
bond lengths (Å)	0.019	0.006
bond angles (°)	2.81	1.23
R_{cryst}*	0.195 (0.36)	0.183 (0.27)
R_{free}*†	0.257 (0.38)	0.255 (0.40)
Number of atoms		
protein	2608	2607
waters	342	80
aliphatic molecule	15	16
NADH	–	44
Average B factors (Å^2)		
mainchain	35.1	37.2
sidechain	44.9	43.0
water	59.8	52.1
aliphatic molecule	49.3	43.0
NADH	–	63.9

*Values in parentheses refer to the outermost shell: 1.75 to 1.80 Å for the native form of the enzyme and 2.80 to 2.93 Å for the NADH co-crystal.
†R_{free} was calculated for 5% of reflections omitted from the refinement.

Accession numbers

The coordinates have been deposited in the PDB with accession codes 1EVY and 1EVZ.

Acknowledgements

We thank T Drmota for his work on the cloning and sequencing of *Lm*GPDH and S Marché for developing the over-expression system. We thank F Athappilly for valuable help with computing, CLMJ Verlinde for assistance with modeling, and EA Merritt, A Ævarsson, R Li and M Feese for useful discussions. We also thank the staff members of the Stanford Synchrotron Radiation Laboratory, the Advanced Light Source, the Cornell High Energy Synchrotron Source, the Brookhaven National Laboratory and the Advanced Photon Source for their extensive support with data collection. We would like to acknowledge substantial instrumentation support from the Murdock Charitable Trust to the Biomolecular Structure Center. We thank one of the referees for critical yet stimulating comments.

References

1. WHO. (1997). Tropical Disease Research: Progress 1995–1996, Thirteenth Programme Report.
2. WHO. (1999). Tropical Disease Research: Progress 1997–1998, Fourteenth Programme Report.
3. Smith, D.H., Pepin, J. & Stich, A.H. (1998). Human African trypanosomiasis: an emerging public health crisis. *Br. Med. Bull.* **54**, 341-355.
4. Moore, A., Richer, M., Enrile, M., Losio, E., Roberts, J. & Levy, D. (1999). Resurgence of sleeping sickness in Tambura County, Sudan. *Am. J. Trop. Med. Hyg.* **61**, 315-318.
5. Alvar, J., *et al.*, & Moreno, J. (1997). Leishmania and human immunodeficiency virus coinfection: the first 10 years. *Clin. Microbiol. Rev.* **10**, 298-319.
6. Denise, H., Matthews, K., Lindergard, G., Croft, S. & Barrett, M.P. (1999). Trypanosomiasis and leishmaniasis: between the idea and and the reality of control. *Parasitol. Today* **15**, 43-45.

7. Wang, C.C. (1995). Molecular mechanisms and therapeutic approaches to the treatment of African trypanosomiasis. *Annu. Rev. Pharmacol. Toxicol.* 35, 93-127.
8. Barrett, M.P. (1999). The fall and rise of sleeping sickness. *Lancet* 353, 1113-1114.
9. Barrett, M.P., Mottram, J.C. & Coombs, G.H. (1999). Recent advances in identifying and validating drug targets in trypanosomes and leishmanias. *Trends Microbiol.* 7, 82-88.
10. Tielens, A.G.M. & Van Hellemond, J.J. (1998). Differences in energy metabolism between trypanosomatidae. *Parasitol. Today* 14, 265-271.
11. Opperdoes, F.R. & Borst, P. (1977). Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: the glycosome. *FEBS Lett.* 80, 360-364.
12. Opperdoes, F.R. (1987). Compartmentation of carbohydrate metabolism in trypanosomes. *Annu. Rev. Microbiol.* 41, 127-151.
13. Hannaert, V. & Michels, P.A. (1994). Structure, function, and biogenesis of glycosomes in kinetoplastida. *J. Bioenerg. Biomembr.* 26, 205-212.
14. Bakker, B.M., Michels, P.A., Opperdoes, F.R. & Westerhoff, H.V. (1999). What controls glycolysis in bloodstream form *Trypanosoma brucei*? *J. Biol. Chem.* 274, 14551-14559.
15. Opperdoes, F.R. & Michels, P.A. (1993). The glycosomes of the kinetoplastida. *Biochimie* 75, 231-234.
16. Verlinde, C.L., *et al.*, & Hol, W.G. (1994). Protein crystallography and infectious diseases. *Protein Sci.* 3, 1670-1686.
17. Bernstein, B.E., Michels, P.A. & Hol, W.G. (1997). Synergistic effects of substrate-induced conformational changes in phosphoglycerate kinase activation. *Nature* 385, 275-278.
18. Aronov, A.M., *et al.*, & Gelb, M.H. (1999). Structure-based design of submicromolar, biologically active inhibitors of trypanosomatid glyceraldehyde-3-phosphate dehydrogenase. *Proc. Natl Acad. Sci USA* 96, 4273-4278.
19. Lo, T.W., Westwood, M.E., McLellan, A.C., Selwood, T. & Thornalley, P.J. (1994). Binding and modification of proteins by methylglyoxal under physiological conditions. A kinetic and mechanistic study with N α-acetylarginine, N α-acetylcysteine, and N α-actyllysine, and bovine serum albumin. *J. Biol. Chem.* 269, 32299-32305.
20. Denise, H., Giroud, C., Barrett, M.P. & Baltz, T. (1999). Affinity chromatography using trypanocidal arsenical drugs identifies a specific interaction between glycerol-3-phosphate dehydrogenase from *Trypanosoma brucei* and Cymelarsan. *Eur. J. Biochem.* 259, 339-346.
21. Kohl, L., *et al.*, & Michels, P.A. (1996). Cloning and characterization of the NAD-linked glycerol-3-phosphate dehydrogenases of *Trypanosoma brucei brucei* and *Leishmania mexicana mexicana* and expression of the trypanosome enzyme in *Escherichia coli. Mol. Biochem. Parasitol.* 76, 159-173.
22. Marché, S., Michels, P.A.M. & Opperdoes, F.R. (2000). Comparative study of *Leishmania mexicana* and *Trypanosoma brucei* NAD dependent glycerol-3-phosphate dehydrogenase. *Mol. Biochem. Parasitol.* 106, 83-91.
23. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283-291.
24. Rossmann, M.G., Liljas, A., Brändén, C.-I. & Banaszak, L.J. (1975). Evolutionary and structural relationships among dehydrogenases. In *The Enzymes.* (Boyer, PD., ed.), Vol. XI Part A, pp. 61-102, Academic Press, New York.
25. Wierenga, R.K., Terpstra, P. & Hol, W.G. (1986). Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. *J. Mol. Biol.* 187, 101-107.
26. Carugo, O. & Argos, P. (1997). NADP-dependent enzymes. I: Conserved stereochemistry of cofactor binding. *Proteins* 28, 10-28.
27. Johansson, K., El-Ahmad, M., Ramaswamy, S., Hjelmqvist, L., Jornall, H. & Eklund, H. (1998). Structure of betaine aldehyde dehydrogenase at 2.1 Å resolution. *Protein Sci.* 7, 2106-2117.
28. Pidoux, A.L., Fawell, E.H. & Armstrong, J. (1990). Glycerol-3-phosphate dehydrogenase homologue from *Schizosaccharomyces pombe. Nucleic Acids Res.* 18, 7145.
29. Goodford, P.J. (1985). A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 28, 849-857.
30. Stillman, T.J., Baker, P.J., Britton, K.L. & Rice, D.W. (1993). Conformational flexibility in glutamate dehydrogenase. Role of water in substrate recognition and catalysis. *J. Mol. Biol.* 234, 1131-1139.
31. Stillman, T.J., *et al.*, & Rice, D.W. (1999). Insights into the mechanism of domain closure and substrate specificity of glutamate dehydrogenase from *Clostridium symbiosum*. *J. Mol. Biol.* 285, 875-885.
32. Colonna-Cesari, F., Perahia, D., Karplus, M., Eklund, H., Brändén, C.I. & Tapia, O. (1986). Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode. *J. Biol. Chem.* 261, 15273-15280.
33. Skarzynski, T. & Wonacott, A.J. (1988). Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus. J. Mol. Biol.* 203, 1097-1118.
34. Otto, J., Argos, P. & Rossmann, M.G. (1980). Prediction of secondary structural elements in glycerol-3-phosphate dehydrogenase by comparison with other dehydrogenases. *Eur. J. Biochem.* 109, 325-330.
35. Holm, L. & Sander, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* 233, 123-138.
36. Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., Job, D. & Pebay-Peyroula, E. (1997). The crystal structure of plant acetohydroxy acid isomeroreductase complexed with NADPH, two magnesium ions and a herbicidal transition state analog determined at 1.65 Å resolution. *EMBO J.* 16, 3405-3415.
37. Redinbo, M.R., Stewart, L., Kuhn, P., Champoux, J.J. & Hol, W.G.J. (1998). Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* 279, 1504-1513.
38. Armah, D.A. & Mensa-Wilmot, K. (1999). *S*-myristoylation of a glycosylphosphatidylinositol-specific phospholipase C in *Trypanosoma brucei*. *J. Biol. Chem.* 274, 5931-5938.
39. Armah, D.A. & Mensa-Wilmot, K. (1999). Protein *S*-myristoylation in *Leishmania* revealed with a heterologous reporter. *Biochem. Biophys. Res. Commun.* 256, 569-572.
40. Godsel, L.M. & Engman, D.M. (1999). Flagellar protein localization mediated by a calcium-myristoyl/palmitoyl switch mechanism. *EMBO J.* 18, 2057-2065.
41. Stebeck, C.E., Frevert, U., Mommsen, T.P., Vassella, E., Roditi, I. & Pearson, T.W. (1996). Molecular characterization of glycosomal NAD^+-dependent glycerol 3-phosphate dehydrogenase from *Trypanosoma brucei rhodesiense. Mol. Biochem. Parasitol.* 76, 145-158.
42. Fairlamb, A.H. & Bowman, I.B. (1980). Uptake of the trypanocidal drug suramin by bloodstream forms of *Trypanosoma brucei* and its effect on respiration and growth rate *in vivo. Mol. Biochem. Parasitol.* 1, 315-333.
43. Willson, M., Callens, M., Kuntz, D.A., Perie, J. & Opperdoes, F.R. (1993). Synthesis and activity of inhibitors highly specific for the glycolytic enzymes from *Trypanosoma brucei. Mol. Biochem. Parasitol.* 59, 201-210.
44. de la Fortelle, E. & Bricogne, G. (1997). Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. In *Methods in Enzymology.* (Carter, C.W.J. & Sweet, R.M., eds), Vol. 276, pp. 472-494, Academic Press, San Diego.
45. Perrakis, A., Sixma, T.K., Wilson, K.S. & Lamzin, V.S. (1997). wARP: Improvement and extension of crystallographic phases by weighted averaging of multiple refined dummy atomic models. *Acta Crystallogr. D* 53, 448-455.
46. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* 47, 110-119.
47. Brünger, A.T. (1988). X-PLOR Manual, Version 3.1, Yale University, New Haven, CT.
48. Tronrud, D.E., Ten Eyck, L.F. & Matthews, B.W. (1987). An efficient general-purpose least-squares refinement program for macromolecular structures. *Acta Crystallogr. A* 43, 489-501.
49. Brünger, A.T., *et al.*, & Warren, G.L. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* 54, 905-921.
50. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* 24, 946-950.
51. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr. D* 50, 869-873.
52. Esnouf, R.M. (1997). An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* 15, 133-138.
53. Brünger, A.T., Krukowski, A. & Erickson, J.W. (1990). Slow-cooling protocols for crystallographic refinement by simulated annealing. *Acta Crystallogr. A* 46, 585-593.
54. Wallace, A.C., Laskowski, R.A. & Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein–ligand interactions. *Protein Eng.* 8, 127-134.
55. Nicholls, A., Shar, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 282-296.