Enantioselective oxidation of galactitol-1-phosphate by galactitol-1-phosphate 5-dehydrogenase from *Escherichia coli*

Rocío Benavente,[a,♯] María Esteban-Torres,[b,♯] Gert-Wieland Kohring,[c] Álvaro Cortés-Cabrera,[d] Pedro A. Sánchez-Murcia,[d] Federico Gago,[d] Iván Acebrón,[a] Blanca De Las Rivas,[b] Rosario Muñoz,[b] José M. Mancheño[a]*

[a]Department of Crystallography and Structural Biology, Institute of Physical Chemistry Rocasolano, CSIC, Serrano 119, E-28006 Madrid, Spain

[b]Laboratory of Bacterial Biotechnology, Institute of Food Science and Technology and Nutrition (ICTAN), CSIC, Juan de la Cierva 3, E-28006 Madrid, Spain

[c]Microbiology, Saarland University, Campus, Geb. A1.5, D-66123 Saarbruecken, Germany

[d]Department of Biomedical Sciences, School of Medicine and Health Sciences, University of Alcalá, E-28871 Alcalá de Henares, Madrid, Spain

♯ These authors contributed equally to this work

*Correspondence e-mail: xjosemi@iqfr.csic.es

Keywords: crystal structure; galactitol-1-phosphate 5-dehydrogenase; metal-binding; enantioselectivity

Abbreviations used: GPDH, recombinant galactitol-1-phosphate 5-dehydrogenase

Abstract

Galactitol-1-phosphate 5-dehydrogenase (GPDH) is a polyol dehydrogenase that belongs to the medium chain dehydrogenase/reductase superfamily (MDR). It catalyses the Zn^{2+} and NAD^{+}-dependent stereoselective dehydrogenation of L-galactitol-1-phosphate to D-tagatose-6-phosphate. Here, we report three crystal structures of GPDH from *Escherichia coli*: the open state of GPDH with Zn^{2+} in the catalytic site and also those of the closed state in complex with the polyols Tris and glycerol, respectively. The closed state of GPDH reveals no cofactor bound, which is at variance with the conformational transition of the prototypical mammalian liver alcohol dehydrogenase. The main intersubunit-contacting interface within the GPDH homo-dimer presents a large internal cavity that probably facilitates the relative movement between the subunits. The substrate analogue glycerol bound within the active site mimics the catalytically relevant backbone of galactitol-1-phosphate. The glycerol binding mode reveals for the first time in the polyol dehydrogenases a penta-coordinated zinc ion in complex with a polyol and also a strong hydrogen bond between the primary hydroxyl group and the conserved acidic residue Glu144, an interaction originally proposed more than thirty years ago that supports a catalytic role for this residue.

1. Introduction

The superfamily of medium chain dehydrogenase/reductase enzymes (MDR) catalyse the reversible oxidation of primary and secondary alcohols into aldehydes and ketones, respectively, using NAD(P) as a cofactor (Persson *et al*., 2008). These enzymes comprise a group of different Zn^{2+}-dependent alcohol dehydrogenases (EC 1.1.1.1; ADHs), historically typified by yeast and liver ADHs (Riveros-Rosas *et al*., 2003). The polyol dehydrogenases (PDHs) form a distinct family that includes sorbitol dehydrogenase (EC 1.1.1.14; SDH) (Jeffery & Jörnvall, 1988), xylitol dehydrogenase (EC 1.1.1.9) (Persson *et al*., 1994) and galactitol-1-phosphate 5-dehydrogenase (EC 1.1.1.251; GPDH) (Nobelmann & Lengeler, 1995), among many others.

Currently, the structures of many MDR enzymes have been solved (Eklund & Ramaswamy, 2008). Basically, all members share a common subunit architecture that is constructed from two domains, a catalytic domain and a nucleotide-binding domain. Between them there is a deep cleft that hosts the catalytically requisite Zn^{2+} ion and the cofactor. These protein subunits associate into homodimeric assemblies in mammalian ADHs, or homotetramers in fungal, bacterial and yeast ADHs (Eklund & Ramaswamy, 2008), and also in PDHs (Klimacek *et al*., 2007). Together with the catalytic zinc ion, most dehydrogenases from this superfamily have another zinc, which is also located in the catalytic domain, particularly within the so-called lobe loop, which is defined as the structural zinc (Auld & Bergman, 2008) and whose functional role, if any, is not yet clear (Raj *et al.,* 2014). Interestingly, functional tetrameric PDHs lack this structural zinc, hence they contain only one metal per subunit (Jeffery & Jörnvall, 1988).

Whereas the coordination of structural zinc ions with four cysteine residues is, with some exceptions (Guy *et al.,* 2003; Pennacchio *et al.,* 2009), essentially conserved (Auld & Bergman, 2008; (Raj *et al.,* 2014), the coordination of the catalytic zinc ions is much less homogeneous and is currently under intense investigation concerning both the participating ligands and the potential changes occurring along the catalytic cycle that would affect the coordination number of the zinc ion (tetra- and penta-coordinated) or the zinc-ligand bond distances (Ryde, 1995; Baker *et al.,* 2009). Obviously, the elucidation of these latter dynamic aspects is relevant to the wider MDR superfamily since they are intimately connected to the structural basis of catalysis by these enzymes. In this regard, uncertainties

remain about the functional roles of the acidic residues Glu60 (using GPDH numbering) and Glu144, both conserved among the PDHs (Eklund *et al.,* 1985; Jeffery & Jörnvall, 1988), which are very close to the catalytic zinc. Whereas in this latter family of enzymes Glu60 (Glu68 in horse liver ADH; hlADH) is a Zn^{2+} ligand (Klimacev *et al.,* 2007), Glu144 (Cys174 in hlADH) is linked to the zinc ion through a water molecule (also conserved in the family) and thus belongs to the second coordination sphere. In the case of hlADH, Glu68, which does not coordinate the metal when in the absence of substrates (the ligands to the zinc are Cys46, His67, Cys174 and a water molecule), has been suggested to play different roles, namely, a stabilising role in the active site (Al-Karadaghi *et al.*, 1994), modulator of the electrostatic field in this environment (Ganzhorn & Plapp, 1988) and also as a transient ligand to the zinc along the catalytic cycle, permitting the efficient product release (Ryde, 1995). In other cases, however, substitution of this residue by alanine or aspartic acid has minor effects in catalysis (Kleifeld *et al*., 2003).

Regarding Glu144, Eklund *et al.* (1985) originally proposed that this residue would be involved in substrate binding in PDHs by specifically forming a strong hydrogen bond with the primary hydroxyl group adjacent to the secondary hydroxyl group undergoing reaction. Despite no direct structural evidences exist supporting this idea, the crystal structure of human SDH in complex with NADH and the inhibitor CP-166,572 (PDB entry: 1pl6) is consistent with this hypothesis, revealing a 2.4 Å H-bond between the side chain of Glu144 (Glu155 in SDH) and the primary hydroxyl group of the inhibitor (Pauly *et al.,* 2003). Moreover, kinetic isotope effect studies with xylitol dehydrogenase from the yeast *Galactocandida mastotermitis* identified a role for this residue as a modulator of general base catalysis, not being essential for hydrogen transfer (Klimacek *et al.,* 2007).

Conversely, catalysis by hlADH is intimately linked to conformational flexibility and in fact this protein was one of the first enzymes for which different conformational states were reported for apo- and holo-forms (Brändén, 1965). Nevertheless, despite much is known about these conformational changes (Plapp, 2010), the connections between protein dynamics and catalysis are not well understood. Basically, the main conformational changes in hlADH that have been identified during catalysis are: a global conformational change affecting the relative orientation of catalytic and cofactor-binding domains that occurs upon binding of the coenzyme, and loop rearrangements around residues 292-299 of the coenzyme binding domain. Other minor changes affecting the active site result from the exchange of zinc ligands,

perhaps involving penta-coordinated zinc intermediates (Dworschack & Plapp, 1977; Makinen & Yim, 1981). The predominant scenario explaining these conformational changes is that of the induced fit, namely, the enzyme changes conformation upon binding of substrates (Plapp, 2010) and, particularly for hlADH, it has been shown that binding of the complete coenzyme is required for the transition between the open state (apo-form) and the closed state (holo-form) (Eklund & Ramaswamy, 2008). Nevertheless, it is not clear that this transition is required for catalysis because the catalytic residues occupy similar positions in both forms; in this regard, only minor changes have been observed between the apo- and holo-forms in human SDH (Pauly *et al*., 2003).

The crystal structures of GPDH that we report in this work offer novel details for the PDH family on several enzymatic issues that may be relevant for the wider MDR superfamily. Thus, the structure of the apo-form reveals that, in contrast to the currently known crystal structures of PDHs, GPDH has a structural Zn^{2+} per subunit in addition to the catalytic metal ion. Furthermore, the structure of the closed state obtained in complex with the substrate analogue glycerol reveals key structural features: for the first time the Zn^{2+} is observed penta-coordinated with a polyol that is structurally homologous to the catalytically relevant part of the substrate; secondly, also at variance with PDHs, Glu60 is not a ligand to the Zn^{2+}; and, thirdly, this complex reveals the originally predicted strong hydrogen bond between the primary alcohol and Glu144 while the secondary alcohol is directly coordinating the Zn^{2+}. Finally, it is remarkable that the closed state of GPDH is observed with substrate analogues within the active site but in the absence of bound cofactor.

2. Materials and methods

2.1. DNA manipulations

The expression vector pT77-GPDH coding for the wild-type galactitol-1-phosphate 5-dehydrogenase from *Escherichia coli* (K12) (Esteban-Torres *et al*., 2012) was used as template for the preparation of single-point variants affecting residues located in the active site region of GPDH (Ser40, Arg44, His51, Tyr106 and Arg112) by site-directed mutagenesis PCR. The mutagenic primers used (forward and reverse, respectively) were: S40A-fw (5′-GCTTATGTGGTGCCGATTTACC-3′) and S40A-rv (5′-GGTAAATC GGCACCACATAAGC-3′) for the Ser40Ala variant; R44T-fw (5′-CGATTTACCCACCATATTTAAA AATGG-3′) and R44T-rv (CCATTTTTAAATATGGTGGGTAAATCG-3′) for the Arg44Thr variant; H51A-fw (5′-GGTGCAGCTTATTATCC-3′) and H51A-rv (5′-GGATAATAAGCTGCACC-3′) for the His51Ala variant, Y106A-fw (5′- GTGCGCAAAAGCTGATTTTATTGG-3′) and Y106A-rv (CCAATA AAATCAGCTTTTGCGCAC-3′) for the Tyr106Ala variant, and finally, R112A-fw (5′-GGCTCAGCGCGTGATGGTGGATTTG-3′) and R112A-rv (5′-AATCCACCATCACGCGCTGAGCC-3′) for the Arg112Ala variant. The mutated GPDH genes were sequenced to verify the expected nucleotide changes.

2.2. Protein Expression and Purification

Wild-type and single-point mutants of GPDH were produced and purified essentially as previously described (Esteban-Torres *et al*., 2012). Briefly, cells carrying the corresponding recombinant plasmid were grown at 37 ºC in LB media containing ampicillin (100 μg/ml) and induced by adding 0.4 mM isopropyl-β-D-thiogalactopyranoside. After induction, the cells were grown at 22 ºC during 20 h and collected by centrifugation using a Beckman Coulter J-25 Avanti centrifuge (7500 x g for 15 min at 4 ºC). Cells were resuspended in 20 ml of 20 mM Tris-HCl, pH 8.0 containing 100 mM NaCl per liter of cell culture. Crude extracts were prepared by French Press lysis of cell suspensions. The lysate was centrifuged at 17,400 x g for 40 min at 4 ºC using a Beckman Coulter J-25 Avanti centrifuge. The supernatant was filtered through a 0.22 μm filter (Millipore) and subsequently loaded onto a HisTrap-FF

column (GE Healthcare) previously equilibrated in binding buffer (20 mM Tris-HCl, pH 8.0 containing 100 mM NaCl and 10 mM imidazole). The recombinant proteins were eluted with a linear gradient of imidazole (from 10 mM to 500 mM) with an ÄKTA Prime Plus. Fractions containing the target protein were pooled and concentrated by ultrafiltration. The protein (2 mL) was then loaded onto an ion exchange HiTrap Q HP column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 8.0 containing 10 mM NaCl. The elution of the proteins was carried out with a linear gradient of sodium chloride (from 10 mM to 500 mM). As before, fractions containing the target protein were pooled and concentrated by ultrafiltration. The final purified material was stored at -80º C until use.

The preparation of the binary complex [GPDH:Zn^{2+}] was done as follows: first, wild type GPDH (10 mg/ml) in Tris buffer, pH 8.0 (20 mM Tris-HCl, pH 8.0, 100 mM NaCl) was buffer exchanged to zinc-containing buffer (20 mM Tris-HCl, pH 9.0, 100 mM NaCl, 0.5 mM zinc chloride), with a HiTrap desalting column (GE Healthcare). The eluted protein was concentrated up to 12 mg/ml for crystallization trials or used as starting material for preparation of the tertiary complex [GPDH:Zn^{2+}:NAD^{+}]. In this case, NADH was added to the protein solution up to a final concentration of 1.8 mM and incubated overnight at 4º C. The mixture was then concentrated and loaded onto a HiTrap desalting column (GE Healthcare) equilibrated in Tris buffer, pH 8.0. The eluted protein was concentrated to 12 mg/ml for crystallization trials. UV-VIS absorbance scans of this protein stock solution showed a maximum at 340 nm indicating the presence of NADH.

2.3. Synthesis of galactitol-1-phosphate and activity assays

L-Galactitol-1-phosphate (LG1P) and D-galactitol-1-phosphate (DG1P) were synthesized and verified by Elexopharm, Saarbruecken, Germany (www.elexopharm.de). All enzymatic tests were performed in duplicate. Catalytic constants for the oxidation of LG1P were determined in 100 mM Tris-HCl buffer pH 9.0, containing 0.5 mM $ZnCl_2$, 1.8 mM NAD^{+} and concentrations of GPDH variants between 4 and 50 µg/ml, respectively. The reaction was started by addition of the substrate and the change of extinction was followed at 340 nm using an Ultrospec 2100 pro photometer from GE Healthcare. For inhibition experiments 1 or 3 mM DG1P were added to the buffer prior to reaction start with LG1P.

Protein concentrations were determined by the method of Pierce (BCA™ Protein Assay) with bovine serum albumin as the standard (Smith *et al.*, 1985).

2.4. Crystallization

Whereas GPDH incubated with Zn^{2+} crystallized in the same experimental conditions as those described previously (Esteban-Torres *et al*., 2012), conditions for crystallization of GPDH incubated with Zn^{2+} and NAD(H) were determined *de novo* by the sitting-drop vapour diffusion method with commercial screens from Hampton Research (Riverside, California, USA) and Qiagen in Innovaplate SD-2 96-well plates set up using a Nanodrop Innovadyne robot. Each drop contained 250 nl of protein (12 mg/ml) in Tris–HCl buffer (20 mM Tris–HCl pH 8.0 containing 0.1 M NaCl) and 250 nl of reservoir solution. Drops were equilibrated against 65 µl reservoir solution. Initial crystals were observed in 20% (*w/v*) PEG 6000, 0.1 M Tris-HCl, pH 8.0, 0.2 M $CaCl_2$, and also in 15% (*w/v*) PEG 4000, 0.1 M Tris-HCl, pH 8.5, 0.2 M magnesium chloride. After scaling up and optimization of the crystallization conditions using hanging drops in 24-well VDX plates high quality diffracting crystals where prepared in 16% (*w/v*) PEG 4000, 0.2 M magnesium chloride, 0.1 M Tris-HCl, pH 8.5 (2:1 protein:precipitant volume ratio; total volume 3 µl).

2.5. Data collection, processing and structure solution

Crystals suitable for X-ray analysis were transferred to an optimized cryoprotectant solution (reservoir solution plus 20% (*v/v*) glycerol) for ~5-10 secs (2 minutes for the crystal measured on beamline ID23-1) and then cryocooled at 100 K in the cold nitrogen-gas stream. Diffraction data from crystals of GPDH preincubated with Zn^{2+} were recorded on BL13-XALOC beamline at the ALBA Synchrotron (Barcelona, Spain) at the zinc peak (9.667 keV; 1.28245 Å wavelength) as determined experimentally from the corresponding fluorescence energy scan. A Pilatus 6M detector (Area Detector Systems Corp.) was used with a crystal-to-detector distance of 268.14 mm, and a total of 360 images were collected with a 0.5º oscillation angle. Conversely, two datasets from crystals prepared with GPDH preincubated with Zn^{2+} and NADH were collected on beamlines ID29 and ID23-1 at the ESRF (Grenoble, France) using wavelengths of 1.28149 and 0.96885 Å, respectively. Whereas the crystal measured on

beamline ID29 was incubated for ~5-10 s in the cryoprotectant solution, the one measured in ID23-1 was incubated for ~2 minutes. In both cases, the detector was a Pilatus 6M. A total of 3600 images were collected on beamline ID29, with a 0.10º oscillation angle and a crystal-to-detector distance of 317.48 mm. The dataset from beamline ID23-1 consists of 2400 images collected with a 0.15º oscillation angle and a crystal-to-detector distance of 188.84 mm. Diffraction images were processed with the *XDS* (Kabsch, 2010) program package. Space group examination was done with *POINTLESS* (Evans, 2011) and intensities reduction with *AIMLESS* (Evans, 2011) from the *CCP4* suite of programs (Winn *et al*., 2011). The first crystal form (GPDH incubated only with Zn^{2+}) belonged to the monoclinic space group $P2_1$, with unit cell parameters a = 43.31, b = 76.91, c = 108.65 Å, β = 95.51º, and is herein defined as crystal form I. Conversely, crystals from GPDH preincubated with Zn^{2+} and NADH were also monoclinic (space group $P2_1$), with unit cell parameters a = 65.31, b = 78.81, c = 68.24 Å, β = 94.75º, and are defined as crystal form II. Data statistics are summarized in Table 1.

2.6. Structure determination and refinement

The crystal structure was determined by molecular replacement using the program *Phaser* (McCoy et al., 2007). The atomic coordinates of GPDH (PDB code: 4a2c) were used as a search model. Model rebuilding was done manually using *Coot* (Emsley *et al*., 2010) and the refinement was performed with *phenix.refine* (Afonine *et al.,* 2012) in *Phenix* (Adams *et al*., 2010). Refinement steps included *xyz* refinement, TLS, individual atomic displacement parameters (ADPs), addition of ligands, and automatic addition of water molecules using default parameters. Stereochemical validation was carried out using the program MOLPROBITY (Chen *et al*., 2010). Refinement statistics are shown in Table 1. The final model from crystal form I (model 1) has an R-factor of 19.3 % and an R_{free} of 25.7 %, and included 692 amino acid residues, 5 zinc ions and 276 solvent molecules. Model 2 (the one obtained from crystal form II measured at ID29) has an R-factor of 20.2 % and an R_{free} of 25.5 %, and included 692 amino acid residues, 5 zinc ions, 2 Tris molecules and 471 solvent molecules, and model 3 (obtained from crystal form II measured at ID23-1) has an R-factor of 18.5 %, R_{free} of 22.9 %, and included 692 amino acid residues, 5 zinc ions, 2 glycerol molecules and 654 solvent molecules. Ribbon diagrams were prepared using PyMOL (DeLano, 2008).

2.7. Computational methods

To account for substrate and cofactor recognition the three-dimensional coordinates of the enzyme structure in the apo-form were used to build a model of the Michaelis complex. Initial positions for the NAD^+ cofactor and the catalytic zinc ion were obtained from the best-fit superimposition of this apo form onto *Neurospora crassa* L-arabinitol-4-dehydrogenase (PDB code 3m6i) using the structural alignment tool implemented in PyMOL. The location of the cofactor was compared to that found in other related dehydrogenases from the same MDR superfamily and it was established that the key ligand-protein interactions were present. The geometry of the resulting complex was optimized by following an energy minimization protocol consisting of 2000 steps of steepest descent followed by 3000 steps of conjugate gradient. The AMBER 10 force field and published parameters for NAD^+ (Ryde, 1995) were employed. To place the LG1P and DG1P molecules inside the active site of the modeled enzyme, we used our in-house CRDOCK docking program (Cortés *et al*., 2012) and default parameters. The binding poses obtained for each molecule were ranked according to the calculated interaction energies but also taking into account the feasibility of the enzymatic reaction on the basis of the distances separating C5 of LG1P from C4 of NAD^+ molecule ($\leq$ 4.0 Å) and the O5 and O6 atoms from the zinc ion ($\leq$ 2.5 Å). By following this procedure we were able to select the most likely binding orientations for both substrates within the enzyme active site.

2.8. PDB accession codes

Atomic coordinates and structure factor amplitudes have been deposited in the Protein data Bank and are available under the accession codes 4ueo (open state of GPDH with Zn^{2+} in the catalytic site), 4uej (closed state of GPDH with glycerol in the active center) and 4uek (closed state of GPDH with Tris in the active center).

3. Results and discussion

3.1. Overall structure of the GPDH monomer and some comments of the two crystal forms

Two different monoclinic crystals of GPDH have been prepared in this work. Crystal form I is herein defined as the one corresponding to crystals of GPDH that have not been incubated with NADH; it contains in the catalytic metal-binding site either nickel scavenged during purification (Esteban-Torres *et al*., 2012) or externally added zinc (this work; PDB entry: 4ueo). Conversely, crystal form II corresponds to crystals of GPDH that has been incubated with Zn^{2+} and NADH [this work; PDB entries: 4uek (model 2) and 4uej (model 3)]. Despite these latter incubation conditions permitted the reconstruction of the holoenzyme in solution as revealed spectroscopically (not shown), no electron density assignable to the cofactor is observed. The comparison of the structure of GPDH from crystal form II (either model 2 or 3) with that from crystal form I (model 1) reveals a global conformational change in GPDH, similar to the classical one undergone by horse liver ADH (hlADH) upon cofactor-binding (Plapp, 2010). This conformational change mainly involves a relative orientation of the two subdomains of the protein subunits (see below) and therefore it does not affect the description of the overall structure of GPDH subunits. For this, we have arbitrarily chosen molecule A from crystal form II (subunit conformation of models 2 and 3 are essentially indistinguishable) as reference.

The GPDH monomer comprises 346 amino acid residues, arranged in two distinct subdomains, as typically observed in members of the medium-chain dehydrogenase/reductase (MDR) superfamily (Eklund & Ramaswamy, 2008): a discontinuous, large catalytic domain (residues 1-148 and 286-346) and a smaller cofactor-binding domain (residues 149-285) (**Figure 1**). The subdomains are separated by a deep cleft where the active site is located and therefore where the NAD(H) cofactor, the catalytic zinc and the substrates are accommodated (Eklund & Ramaswamy, 2008). The cofactor-binding domain is a canonical Rossmann fold, characteristic of NAD(H)-binding proteins (Bellamacina, 1996; Fan & Plapp, 1999), which is made up of a six-stranded parallel β-sheet surrounded by α-helices.

Conversely, the core of the catalytic domain is formed by a seven-stranded mixed β-sheet [five antiparallel (β3, β4, β5, β6 and β7) and two parallel (β15, β16) β strands], which is flanked by a two-stranded antiparallel β-sheet (β1, β2) and three α-helices (α1, 3_{10}-3 and α10). The main contacting regions between these two subdomains arise from the loop that houses the structural zinc-binding site (lobe loop), which interacts with two connecting loops from the other subdomain (loop between strand β14 and helix α9 and the loop between helices 3_{10}-3 and α4). As also observed in other members of the MDR

superfamily, a distinct kink is present within the 3_{10}-3 helix preceding the position of the functionally relevant Glu144 (see below).

Two bound metal ions per GPDH subunit are identified in both crystal forms, which are here unambiguously identified as zinc. This observation agrees with the fact that most NAD(H)-dependent dehydrogenases from the MDR superfamily have two zinc ions per subunit (Auld & Bergman, 2008), namely one catalytic zinc within the active site and another one located in the lobe loop of the catalytic subdomain, but it is at variance with the known structures of functional, tetrameric PDH enzymes, which lack the structural zinc (Jeffery & Jörnvall, 1988). In contrast to this, in some bacterial MDR alcohol dehydrogenases this zinc-containing loop has been shown to be important for tetramer formation (Banfield *et al*., 2001; Esposito *et al*., 2002; Guy *et al*., 2003). Thus, GPDH is a peculiar PDH insofar as it is dimeric and contains two zincs per subunit. In fact, we found that bacterial, dimeric MDR enzymes are reductases but they lack zinc ions (Edwards *et al.,* 1996; Shimomura *et al.,* 2003; Sulzenbacher *et al*., 2004).

As expected, matching a GPDH subunit against the PDB using the DALI server (Holm & Sander, 1995) revealed that the closest structural homologs belong to ADHs and PDHs enzymes from the MDR superfamily, in particular, a putative zinc-binding dehydrogenase from *Sinorhizobium meliloti* (PDB entry: 4ejm; rmsd 2.0 Å for 332 Cα-atoms), L-threonine dehydrogenase from *Pyrococcus horikoshii* (Ishikawa *et al*., 2007) (PDB entry: 2dfv; rmsd 1.9 Å for 331 Cα-atoms), ketose reductase (sorbitol dehydrogenase) from silverleaf whitefly (Banfield *et al.*, 2001) (PDB entry: 1e3j; rmsd 1.9 Å for 331 Cα-atoms), L-threonine 3-dehydrogenase from *Thermus thermophilus* (PDB entry: 2dq4; rmsd 2.0 Å for 336 Cα-atoms), L-threonine dehydrogenase from *Thermococcus kodakaraensis* (Bowyer *et al.,* 2009) (PDB entry: 3gfb; rmsd 2.0 Å for 334 Cα-atoms), human sorbitol dehydrogenase (hSDH) (Pauly *et al.,* 2003) (PDB entry: 1pl8; rmsd 2.1 Å for 335 Cα-atoms), and also many other ADHs with rmsd values around 2 Å. The result reveals a remarkably high degree of conservation of the architecture of the protein subunits despite the low level of sequence identity among them (less than 27%) (Figure 1).

3.2. Quaternary structure and the built-up of an internal cavity

In contrast to bacterial ADHs and PDH enzymes, which are tetrameric assemblies, GPDH is a dimer both in solution and in the crystal (Esteban-Torres *et al*., 2012). The association between GPDH subunits occurs through their cofactor-binding subdomains. The main contacting interface buries approximately 3450 $Å^2$ per dimer and results from the contacts between the β12-β13 connecting loop, strands β13 and β14 and helix α8 from each subunit (**Figure 2**). It is remarkable that, upon dimer formation, these structural elements, together with helix α7 and strand β12 from each subunit, delimit an internal, large solvent inaccessible cavity (310 $Å^3$ in volume as estimated with the CASTp server [Dundas *et al*., 2006]) **(Figures 3a and 3b)** lined exclusively with hydrophobic residues (Val243, Val247, Leu256, Leu258, Leu266, Leu268, Phe273, Ile276, Val283) **(Figure 3c)**. The overall shape of this cavity is that of a rectangular prism: one of its sides is surrounded by the two-stranded, intersubunit, antiparallel β-sheet formed by strands β13, whereas the opposite side is defined by strands β12 and β14 from each subunit. One pair of perpendicular and opposed faces of the cavity would be formed by helix α8 from one subunit and the β12-β13 connecting loop from the other one, whereas the other opposed pair of faces would be defined by helices α7. As far as we know, similar cavities have not been described previously for any other enzyme from the MDR superfamily, irrespective of the highly conserved protein architecture of these proteins. On the contrary, all the structural homologues found for GPDH have tightly packed hydrophobic cores in the equivalent regions.

As expected by the crucial location at the intersubunit-contacting interface, there exist structural elements of GPDH that probably contribute to stabilize the dimeric assembly. In particular, helix α8 (residues 270-278) is an important intersubunit secondary structure element, together with strands β12 and β14. This helix participates in several intersubunit contacts: (i) a cation-π interaction between Arg278 and Trp287, (ii) a salt bridge between Glu280 and His151 (2.3 Å), and (iii) a hydrophobic cluster formed by Phe273 and Leu277, on the one hand, and residues Val243, Leu258 and Leu262 from the contacting subunit, on the other hand. We believe that, in contrast to these intersubunit-stabilizing interactions, the presence of a large, internal cavity at the intersubunit interface should be considered *a priori* a destabilizing factor, essentially due to the inefficient packing of hydrophobic side chains. In fact, it is well

known that cavity-creating mutations within hydrophobic cores in proteins are generally destabilizing (Eriksson *et al.,* 1992; Jackson *et al*., 1993; Buckle *et al.,* 1996).

Although the electron density map does not reveal any ordered water molecule within the cavity, in agreement with the low probability of finding a single water molecule in a small hydrophobic cavity (1 in 20,000) (Wolfenden & Radzicka, 1994), the existence of water molecules in large hydrophobic cavities in proteins is contentious (Matthews *et al*., 1995) perhaps because of the assumption that lack of density in a crystallographic electron density map means absence of matter (Yu *et al*., 1999). Although these maps did not usually reveal any atom whose average position is ill defined, a large body of studies on hydrogen exchange shows that solvent can penetrate deepest into the protein structure (Eglander & Kallenbach, 1983). For this reason, we think it is reasonable to propose the existence of mobile solvent within the observed cavity of GPDH, which may play a role in protein dynamics. In this regard, in this work we have identified two distinct conformational states in GPDH (open and closed states, respectively), which are similar to those originally identified in hlADH (Plapp, 2010). Comparison between these two conformational states reveals displacements of side chains facing the cavity, which most probably occur in synchrony with the global conformational change of GPDH. Considering these observations, we claim that the cavity, filled with mobile, internal molecules would be a structural outcome that integrates both flexibility to the main intersubunit interface allowing significant conformational changes to occur, and stability that would preserve the dimeric assembly.

3.3. A global conformational change in GPDH

As indicated above, we have crystallized the binary complex GPDH incubated with Zn^{2+} (crystal form I) and also made crystallization trials with the tertiary complex [GPDH:Zn^{2+}:NADH], which rendered new monoclinic crystals (crystal form II). Despite the fact that the phased-model difference Fourier map of GPDH calculated with data from these latter crystals did not reveal any electron density assignable to the NADH cofactor, this latter structure (either model 2 or 3), when compared to model 1, reveals the existence of a global conformational change in GPDH similar to that reported for hlADH occurring upon binding of the complete cofactor (Plapp, 2010). Hence, two distinct conformational states for GPDH are identified: an open state derived from the crystal form I (equivalent to the open state of

hlADH obtained with the apo-enzyme), and a closed state, derived from crystal form II (equivalent to the closed state of hlADH, obtained with the holo-enzyme). The conformational change in hlADH can be described as a rigid body rotation between the catalytic and cofactor-binding subdomains (Plapp, 2010), together with some rearrangements in the 292-299 loop involved in cofactor binding (β12-β13 connecting loop in GPDH).

Similarly to hlADH, the observed conformational change in GPDH involves a relative motion of the catalytic and cofactor-binding subdomains, with helices α9 and 3_{10}-3 acting as hinges. This movement leads to a closure of the cleft between subdomains where the active site is located. In particular, the N-terminal end of helix α1 and the α10-β16 connecting loop, both from the catalytic subdomain, and helix α5 and connecting loops between strand β8 and helix α4, and between strands β12 and β13, from the cofactor-binding subdomain, approach the center of the cleft, narrowing its overall width.

When analyzed at the dimeric assembly level, the reorientation of the GPDH subdomains of both subunits results in an in-plane, antiparallel displacement of one β13 strand relative to the other, and also a change in the relative position of helix α8 from one subunit with respect to the β12-β13 connecting loop from the other subunit. Conversely, strands β12 and β14 from the extended intersubunit β-sheet remain stationary, probably contributing to the integrity of the dimer as the conformational change occurs, similarly to helices α8. In fact, several intersubunit interactions are identified at both sides of strands β12 and β14: polar interactions between Glu280 with His151 and with carbonyl oxygen of Trp287, and stacking interactions between the aromatic rings of His253 and Phe99 side chains. Interestingly, the imidazole ring of His151 is highly oriented since it is bridged between Glu280 from the other subunit and Glu298.

Normal mode analyses of GPDH with the web server ElNémo (Suhre & Sanejouand, 2004) reveals that a single normal mode (second normal mode for a single subunit and first normal mode for the dimer) reproduces a conformational transition between an open and a closed conformation of the enzyme in both directions (open → closed and closed → open), in which the crystal structures of GPDH would be snapshots lying along the pathway connecting such conformations.

As indicated above, residues facing the internal cavity of GPDH participate in the global conformational change. The observed antiparallel displacement of strands β13 involves an in-plane

translation of the Leu266 and Leu268 side chains within the cavity, whereas the movements affecting to helices α8 and the β12-β13 connecting loops involve the displacements of Leu262, Phe273 and Ile276 side chains. As above indicated, these internal displacements of side chains are probably facilitated by the presence of the internal cavity, where solvent molecules in protein dynamics may play a role.

Curiously, supporting evidence for a role of the cavity as a facilitator for the intersubunit conformational change may come also from hlADH. In this case, the observed conformational change has no effect on the dimeric assembly of hlADH in the sense that intersubunit contacting surfaces are not affected by the reorientation of protein subdomains. Here, the remarkable observation is that, contrary to GPDH, hlADH has no internal cavity but a highly packed hydrophobic core. This hints at a correlation between the presence of an internal cavity (lack of a tightly packed hydrophobic core) and the existence of local conformational changes at the dimer interface. Obviously, these observations raise numerous questions as to which are the relationships between the specific catalytic mechanisms of the different ADHs and the diverse conformational degrees of freedom provided by the intersubunit dynamics.

Finally, it is noteworthy that the closed state of GPDH is observed with no cofactor bound but containing Tris or glycerol within the active site. This may suggest either that the binding of these molecules induces such a conformational transition in the absence of cofactor or the existence of a conformational equilibrium between open and close states in the absence of ligands, which in our studies would probably be selected by the different crystallization conditions. Although, as far as we know, this would be a novelty for any MDR enzyme where the scenario is dominated by the induced fit concept (Plapp, 2010), the conformational selection scenario is a well-known phenomenon in enzymatic systems (Changeux, 2013). In fact, it has been proposed that both situations are not mutually incompatible but limiting cases of a more general framework that integrates both phenomena (Hammes *et al.,* 2009).

3.4 Metal ligands

Prior characterization of GPDH revealed the necessity of Zn^{2+} and NAD(H) for oxidation of galactitol and for reduction of D-tagatose-6-phosphate (Esteban-Torres *et al*., 2012). The metal-binding sites of GPDH that we reported previously revealed two metals bound per protein subunit: one Zn^{2+} bound in the structural metal-binding site, and another metal located in the catalytic site that we proposed

to be a nickel scavenged from the purification media (Esteban-Torres *et al*., 2012). Now, new crystals of GPDH incubated with Zn^{2+} have permitted the unambiguous identification of the two metals as Zn^{2+}, as revealed by X-ray energy scans and also from anomalous difference maps calculated with data recorded at the Zn peak (9.667 keV) (**Figure 4**).

The structural zinc ion, in all the structures reported here, is located in the lobe loop of the catalytic domain (residues 89-103) where it is coordinated by four cysteine residues (Cys89, Cys92, Cys95, and Cys103), with a tetrahedral geometry (**Figure 4a**). In this type of coordination where Zn^{2+} forms four strong bonds with cysteine residues the metal does not normally dissociate in biological conditions (Harding *et al.*, 2010), usually playing a role in the stabilization of the tertiary structure. This structural metal is also observed in the crystal structures of the close homologs L-threonine dehydrogenases from *Pyrococcus horikoshii* (Ishikawa *et al.,* 2007) and from *Thermus thermophilus* (PDB entry: 2dfv) or whitefly ketose reductase (Banfield *et al.,* 2001), but neither in hSDH (Pauly *et al.,* 2003), sheep liver sorbitol dehydrogenase (Yennawar *et al.,* 2011) nor in L-threonine dehydrogenase from *Thermococcus kodakaraensis* (Bowyer *et al.,* 2009). This variability in the structural metal binding-site correlates well with the variation in the total zinc content observed in general for other more distant ADHs (Bogin *et al.,* 1997; Korkhin *et al.,* 1998; Sulzenbacher *et al.,* 2004).

Although the structure of GPDH determined from crystals of form I (model 1) reveals a Zn^{2+} in each catalytic site, the low metal occupancy deduced from *B*-factor analysis (Zn^{2+} overall *B*-factor: 125 $Å^2$ vs ligands overall *B*-factor: 63 $Å^2$) makes the geometry of metal coordination not as reliable as those determined from GPDH models 2 or 3 in which full metal occupancy is observed (model 2: Zn^{2+} overall *B*-factor: 27 $Å^2$ vs ligands overall *B*-factor: 18 $Å^2$; model 3: Zn^{2+} overall *B*-factor: 20 $Å^2$ vs ligands overall *B*-factor: 25 $Å^2$). Nonetheless, it can be easily ascertained in model 1 that the metal is coordinated by Cys38, His59 and two water molecules (**Figure 4b**).

Unexpectedly, a different metal coordination is observed for each of the GPDH models from crystal form II (models 2 and 3), which results from the presence of different polyols in the active site that are ligands to the zinc: whereas in model 2 a Tris molecule is identified in the active site, a glycerol is observed in model 3. Most probably, these complexes resulted from the different incubation times of GPDH crystals in the cryoprotectant solution. Thus, short times (~5-10 secs) yielded GPDH crystals with

Tris in the active site and longer ones (2-3 minutes) led to the replacement of the Tris molecules by glycerols. Since this latter molecule is structurally analogous to the catalytically relevant region of galactitol-1-phosphate, the complex may reveal interactions that most probably occur between GPDH and LG1P.

In model 2 of GPDH the catalytic zinc is coordinated by residues Cys38, His59, a water molecule and hydroxyl OH1 from the Tris molecule (**Figure 4c**). In model 3, however, the zinc is penta-coordinated by Cys38, His59, a water molecule and OH1 and OH2 hydroxyls of glycerol. The shape of the coordination is a distorted tetragonal pyramid (**Figure 4d**), typical for 5-coordination [54]. In this case, the most distorted angle is observed between atoms Nε2/Zn^{2+}/O(OH1), most probably due to the formation of an additional interaction: a strong hydrogen bond to Oε2 atom from Glu144 (distance 2.4 Å). On the contrary, hydroxyl group OH2 (equivalent to the hydroxyl group of LG1P undergoing reaction) exclusively interacts with the metal.

We believe that these latter structural results are remarkable in several aspects: firstly, the existence of the interaction between OH1 and the Glu144 side chain shows that this residue is an essential element of PDH-specific substrate recognition and catalysis. This idea was anticipated almost thirty years ago by Eklund *et al*. (1985) for a model of sorbitol dehydrogenase and is here validated for the first time with a polyol molecule within the active site. Secondly, the structure of bound glycerol supports the binding mode of sorbitol by human sorbitol dehydrogenase, as proposed by Pauly *et al*. (2003) in that C1 and C2 oxygen atoms of sorbitol would coordinate the zinc leading to a penta-coordinated intermediate. Thirdly, this complex with glycerol is, as far as we know, the first structure revealing a Zn^{2+} in the PDHs that is penta-coordinated with a polyol that is a substrate analogue that participates in metal coordination. In truth, the crystal structure of human sorbitol dehydrogenase complexed with NADH and the inhibitor CP-166,572 (Pauly *et al*., 2003) (PDB entry: 1pl6) also reveals a penta-coordinated Zn^{2+}, although this inhibitor is not a polyol. Furthermore, among the ADHs the only structures presenting a penta-coordinated zinc are the binary complex between hlADH and 1,10-phenanthroline (Boiwe & Brändén, 1977), which is not an alcohol, and two other crystal structures (*H. sapiens* ADH3; PDB entries: 1m5c and 1teh), in which one Zn^{2+} coordination is established with a water molecule at distances around 3 Å

that markedly depart from the average metal-water atom distance derived from the Cambridge Structural Database (2.09 ± 0.05) Å (Harding *et al.,* 2010).

3.5. Further analysis of the GPDH active site

The structures of GPDH with Tris or glycerol within the active site reveal features that may be relevant to the wider MDR superfamily, in particular, aspects related to the couple of acidic residues close to the zinc center (Glu60 and Glu144 in GPDH). The acidic residue Glu144 is conserved among the PDHs (Jeffery & Jörnvall, 1988) and is not a ligand to the Zn^{2+}, but coordinates the metal via a water molecule whose conserved character in PDHs [human (Pauly *et al*., 2003) and sheep (Yennawar *et al*., 2011) sorbitol dehydrogenases and whitefly ketose reductase (Banfield *et al*., 2001)] suggests it may play a role in catalysis (Klimacek *et al* 2007). In GPDH, this solvent molecule also interacts with the accompanying acidic residue Glu60, which similarly to Glu144 does not coordinate the Zn^{2+}. The role of Glu60 in catalysis in MDR enzymes in general, and in PDHs in particular, is not clear although it has been proposed that it may coordinate the Zn^{2+} along the catalytic cycle (Ryde, 1995). In agreement with this, Glu60 coordinates the Zn^{2+} in sheep sorbitol dehydrogenase (Yennawar *et al*., 2011) and whitefly ketose reductase (Banfield *et al*., 2001). On the other hand, in human sorbitol dehydrogenase it is a ligand to the Zn^{2+} when the enzyme is unliganded (PDB entry: 1pl7) or with NAD^+ bound (PDB entry: 1pl8), but it is not when the enzyme is complexed with NADH and the inhibitor CP-166,572 (PDB entry: 1pl6). All together, this variability in Glu60 metal coordination is consistent with a dynamic zinc coordination ligand exchange during catalysis (Baker *et al.,* 2009; Plapp, 2010). In this regard, the results of recent, thorough analyses of Zn^{2+} coordination in ADHs (including PDHs) from the MDR superfamily (Auld & Bergman, 2008; Raj *et al*., 2014) are also consistent with this variability.

Another residue strictly conserved in all GPDH homologues, and strongly conserved in the MDR superfamily, situated close to the zinc center is Asp41. No specific role has been assigned to this amino acid in PDHs, although it has been well studied in yeast ADH (Ganzhorn & Plapp, 1988). In GPDH, the carboxyl group of Asp41 interacts with the guanidinium moiety of Arg44, and is at hydrogen bonding distance to two internal, highly ordered water molecules that participate in a dense network of H-bonds **(Figure 5),** and also to the hydroxyl group of Ser40. One of these latter water molecules, strongly

conserved in all structurally characterized PDHs, forms a hydrogen bond to His59. Therefore, it can be thought that Asp41, Arg44 and the latter water molecule may contribute to stabilize the geometry of the active site. To analyse the roles of these residues in deeper detail we have prepared several point mutants.

Removal of the negative charge of Asp41 in the Asp41Asn mutant leads to a variant enzyme with unaltered K_m for LG1P (**Table 2**) but with a dramatically decreased k_{cat} for LG1P oxidation (from 8800 to 250 s^{-1}). Since this mutation involves the amidation of the carboxylate, it seems reasonable to propose that the ionic character of the side chain is key for its catalytic role. To analyse this aspect in deeper detail, the ionic interaction with Arg44 was removed in the Arg44Thr variant. The effects on the kinetic constants determined for this variant were similar to the Asp41Asn variant (**Table 2**), revealing the potential relevance of both residues in catalysis. Nonetheless, since Arg44 also interacts with Ser40, which is equivalent to Ser48 in hlADH that facilitates the deprotonation of the alcohol in this enzyme (Plapp, 2010), the observed effects in GPDH should not be exclusively ascribed to a potential destabilization of the active center induced by the loss of the ionic interaction. In fact, our *in silico* analyses of complexes between GPDH and LG1P (or DG1P) strongly support the existence of direct interactions between the guanidinium moiety of Arg44 and both molecules. This residue would not participate in the initial substrate-binding steps but would assist, together with other residues, in the build-up of the correct, near-attack conformation of the substrate (see below), which can only be attained with LG1P.

The experimental results obtained with the second variant affecting to this residue (Arg44Tyr), where the introduced side chain has a similar length as the parental one, may also support such a role. Now, we observed a smaller effect in turnover number (decreasing from 8800 to 1700 s^{-1}) but a remarkable five-fold reduction in K_m for L1GP (**Table 2**). This suggests a direct interaction between the introduced side chain and the ground state of the substrate that would mainly affect to the initial LG1P-binding steps. The less efficient LG1P oxidation would be consistent with this stronger substrate binding not leading to a catalytically productive conformation as efficiently as with the wild-type enzyme.

As a whole, we believe that our results are consistent with the participation of Asp41 (together with the conserved water molecule) in the correct positioning of the His59 side chain; hence its substitution by Asn could induce structural rearrangements of the ligands to the zinc that would affect hydride transfer, accounting for the low k_{cat} value observed. Conversely, the ionic, buttressing interaction

with Arg44 would fix this side chain such that it will be properly oriented towards the bound substrate (or DG1P). As a result of the established interactions, the catalytically productive complex can only be achieved with LG1P (see below).

3.6. Biochemical studies

Despite the fact that GPDH has been classically described as a dehydrogenase specific for the oxidation of galactitol-1-phosphate, no direct experimental evidence has been reported showing this specificity. Although our previous results indicating a 500-fold increase in specific activity of GPDH for reduction of D-tagatose-6-phosphate (DT6P) [K_m = (1.00 ± 0.05) mM] *versus* galactitol oxidation [K_m = (26 ± 4) mM] supported this annotation (Esteban-Torres *et al.,* 2012), a direct characterization of the Zn^{2+}- and NAD^+-dependent oxidation of galactitol-1-phosphate by GPDH is required. To this end, we synthesized both L-galactitol-1-phosphate (LG1P) and D-galactitol-1-phosphate (DG1P). Our results demonstrate that GPDH exhibits a strict stereospecificity against LG1P, which together with our previous results on DT6P reduction, unambiguously define this enzyme as a stereospecific polyol dehydrogenase. Comparison of k_{cat} and K_m values for LG1P oxidation and DT6P reduction reactions reveals a higher turnover number for DT6P versus LG1P, but a K_m for this latter substrate that makes the final specificity constant (k_{cat}/K_m) three-fold higher for LG1P oxidation than for DT6P reduction.

3.7. NAD^+ recognition

Considering our failed attempts to obtain a well-defined electron density map for the cofactor bound to GPDH, we carried out *in silico* studies of the binding mode for NAD^+ to the closed state of the enzyme (**Figure 6**). In the conservative binding mode proposed, the adenine ring is located in a hydrophobic pocket lined by residues Ile167, Ile193, Ala237, Val239 and Thr242, whereas the attached ribose establishes hydrogen-bonding interactions with the side chains of Asp192 and Lys197 through its O2' and O3' atoms, respectively. This result agrees with the reported structural basis of the specificity for NAD(H) over NADP(H), which resides in the presence of an Asp residue (Asp192 in GPDH), that interacts with the ribose hydroxyl groups (Baker *et al*., 1992). The diphosphate bridge, in turn, is placed at the dipole plus end of helix α4 and directly interacts with residues Gly39, Thr171, Ile172, and Met338.

The nicotinamide riboside, on the other hand, has the pyridine ring surrounded by hydrophobic residues Val148, Ile172 and Val259; the amide moiety interacting with the carbonyl groups of Val259 and Ser286 and the NH of Met288; and the ribose O4' hydroxyl engaged in a hydrogen bond both the carbonyl oxygen of Ala237 and the NH of Thr261.

Interestingly, all crystal structures of GPDH reported in this work reveal ordered water molecules interacting with the carbonyl groups of Val259 and Ser286 and additional ones in model 1 (that lacks a polyol in the active center) interacting with the main chain NH of Met288. Hence, the distribution of ordered waters in this region faithfully resembles the predicted position of the amide moiety of the nicotinamide ring bound to GPDH. Moreover, the absence of the latter water molecule interacting with Met288 in models 2 and 3 could be due to the presence of glycerol or Tris in the active center since the distal parts of these molecules protrude towards the predicted position of the nicotinamide ring, making the simultaneous presence of both cofactor and polyol mutually exclusive.

3.8. GDPH substrate recognition and reaction mechanism

The two conformational states of the GPDH subunits may represent structures relevant to the catalytic mechanism of the enzyme namely the apoenzyme (open state) and the ternary complex (closed state), as suggested by the structural similarity to hlADH whose isomerization step has been described in detail (Sekhar & Plapp, 1990; Kovaleva & Plapp, 2005). Therefore we propose the following scenario for GPDH: in the absence of any bound ligand there would exist an equilibrium between the open and the closed states of the enzyme. In both cases, the zinc ion is tetra-coordinated by two protein side chains (Cys38, His59) and two water molecules. The required cofactor NAD^+ would bind the closed state of the enzyme first (conformational selection model), followed by the binding of LG1P. LG1P binds with both its C5 and C6 oxygen atoms coordinated to the zinc, with the concurrent release of a water molecule: zinc ligand exchanges would then yield a penta-coordinated metal ion (model 3). Our *in silico* model of the ternary complex (**Figure 7**) shows that LG1P stacks in part against the nicotinamide ring and this is consistent with an ordered mechanism.

The proton transfer from the C5 hydroxyl group of LG1P required for LG1P oxidation to DT6P demands a base, a role which, for hSDH (Pauly *et al*., 2003) and sheep liver SDH (Klimacek *et al.*, 2007),

has been assigned to the water molecule that in GPDH interacts with Glu60 and Glu144. Based on our structural and functional analyses, we claim that the catalysis by GPDH is consistent with the generally accepted mechanism proposed by Cui *et al*. (2002) [see also (Plapp, 2010) for a review]. In this scenario, the hydroxyl group of Ser40 in GPDH (Ser or Thr residues systematically appear in this position in PDHs) acts as part of the relay system that transfers the proton of the C5 hydroxyl group to the bulk water. The predicted conformation of bound LG1P indicates that upon Zn^{2+} coordination by the substrate, most likely funnelled by the strong hydrogen bond between primary C6 hydroxyl group and the carboxylate of Glu144, the orientation of the secondary C5 hydroxyl group is adequate for such a proton transfer *exclusively* for LG1P. The participation of Ser40 in catalysis can be inferred from the kinetic parameters determined for the variant Ser40Ala: while K_m for LG1P remains essentially unaltered, an important reduction in the turnover number is observed (from 8800 to 1500 s^{-1}) (**Table II**), which is consistent with the assigned role in abstracting the proton from the C5 hydroxyl group of the substrate. Subsequent hydride transfer from C5 to the pyridine ring of the NAD^+ coenzyme, facilitated by the polarization of the C5–O5 bond due to the specific Zn^{2+} coordination to NAD^+, and only possible for LG1P, would lead to cofactor reduction and oxidation of the C5 hydroxyl to the keto group of DT6P.

Although according to the substrate-binding mode many residues contribute to the definition of the complete substrate-binding pocket (Cys38, Ser40, Arg44, His51, His59, Tyr106, Phe108, Ser111, Arg112, Glu144, Trp287, Met288), we observe that two residues, namely Arg44 and Met288, would directly interact with LG1P: Arg44 would interact with C4 hydroxyl group and Met288 would establish hydrophobic interactions with the C4-C5-C6 backbone atoms of LG1P **(Figure 7)**. Hence, a role in substrate orientation (that will ultimately lead to enantiomer discrimination) can be predicted for them. The important role in catalysis of Arg44 could be inferred from the two variants discussed above (Arg44Thr and Arg44Tyr), and that of Met288 was checked with the Met288Arg variant, which revealed a drastic decrease in k_{cat} for LG1P oxidation (similar to that observed for the Arg44Thr variant), with K_m remaining unaltered.

It is worth mentioning that the configuration of the substrate-binding pocket is compatible with DG1P binding **(Figure 7)**. However, the predicted interactions established with Zn^{2+}, Glu144 carboxylate, Arg44 and Met288 would lead to a conformation incompatible with both proton transfer, due to the

improper orientation of the secondary C5 hydroxyl group relative to Ser40, and hydride transfer, since the C5 H atom would be oriented away from the pyridine ring of the NAD^+ coenzyme. Inhibition of LG1P activity by increasing concentrations of DG1P could be demonstrated in biochemical analysis. From Lineweaver-Burke-plots of inhibition kinetics, a competitive inhibition by DG1P can be deduced because the intersection of the plots is close to the reciprocal velocity axis. The K_m for LG1P oxidation is highly affected and drops almost 6-fold by addition of 3mM DG1P while the reaction velocity is only reduced by 20% (Figure S1).

Also, we propose that three residues of the pocket, namely His51, Tyr106 and Arg112, are also important in substrate specificity by playing a bait role for attracting the phosphate group, and probably facilitating product release as well. This proposal is supported by available structural evidence and results from our molecular dynamics simulations (data not shown) revealing that these residues can be oriented towards the solvent (open state) or towards the active site (closed state). The finding that site-directed mutagenesis at these three positions resulted in enzymes displaying decreased turnover numbers but K_m values similar to those of the wild-type enzyme (**Table II**) appears to support this assigned function. Finally, these same studies also revealed that the Arg278 side chain from one subunit might be also engaged in the fixation of the phosphate moiety of the substrate by the other subunit, suggesting the existence of a 'cross-talk' between subunits that is currently under investigation.

Acknowledgements

We thank the ESRF (Grenoble, France) and ALBA synchrotron (Barcelona, Spain) for provision of synchrotron radiation facilities (ID23-1 and ID29 beamlines at the ESRF and XALOC in ALBA). Financial support from the Ministerio de Educación y Ciencia (BFU2010-17929/BMC) and the Factoría de Cristalización (Consolider-Ingenio-2007) (to J.M.M.), and from the European Community through the FP7-NMP-2007-SMALL 1 collaborative project ERUDESP to G.W.K.

References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, T. C., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213-221.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W. Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst*. D**68**, 352-367.

Al-Karadaghi, S., Cedergren-Zeppezauer, E. S. & Hovmoller, S. (1994). *Acta Cryst.* D**50**:793-807.

Auld, D. S. & Bergman, T. (2008). *Cell. Mol. Life Sci.* **65,** 3961-3970.

Baker, P. J., Britton, K. L., Fisher, M., Esclapez, J., Pire, C., Bonete, J., Ferrer, J. & Rice, D. W. (2009). *Proc. Natl. Acad. Sci. USA* **106,** 779-784.

Baker, P. J., Britton, K. L., Rice, D. W., Rob, A. & Stillman, T. J. (1992). *J. Mol. Biol*. **228**, 661-671.

Banfield, M. J., Salvucci, M. E., Baker, E. N. & Smith, C. A. (2001). *J. Mol. Biol.* **306**, 239–250.

Brändén C-I. (1965). *Arch. Biochem. Biophys.* **112**, 215-217.

Bellamacina, C. R. (1996). *FASEB J.* **11**, 1257-1269.

Bogin, O., Peretz, M. & Burstein, Y. (1997). *Protein Sci*. **6**, 450-458.

Boiwe, T. & Brändén C. I. (1977). *Eur. J. Biochem*. **77**, 173-179.

Bowyer, A., Mikolajek, H., Stuart, J. W., Wood, S. P., Jamil, F., Rashid, N., Akhtar, M. & Cooper, J. B. (2009). *J. Struct. Biol.* **168**, 294-304.

Buckle, A. M., Cramer, P. & Fersht, A. R. (1996). *Biochemistry* **35**, 4298-4305.

Changeux, J-P. (2013). *Nat. Rev. Mol. Cell. Biol*. **14**, 819-829.

Chen, V. B., Arendall III, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst*. D**66**, 12-21.

Cortés, A., Klett, J., Dos Santos, H. G., Perona, A., Gil-Redondo, R., Francis, S. M., Priego, E. M., Gago, F. & Morreale, A. (2012). *J. Chem. Inf. Model*. **52**, 2300-2309.

Cui, Q., Elstner, M. & Karplus, M. (2002). *J. Phys. Chem. B* **106**, 2721-2740.

DeLano, W. L. (2008). *The PyMOL Molecular Graphics System.* http://www.pymol.org.

Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J. (2006). *Nucleic Acid Res*. **34**, W116-W118.

Dworschack, R. T. & Plapp, B. V. (1977). *Biochemistry* **16**, 2716-2725.

Edwards, K. J., Barton, J. D., Rossjohn, J., Thorn, J. M., Taylor, G. L. & Ollis, D. L. (1996). *Arch. Biochem. Biophys.* **328**:173-183.

Eklund, H., Horjales, E., Jörnvall, H., Brändén, C. I. & Jeffery, J. (1985). *Biochemistry* **24**, 8005-8012.

Eklund, H., Plapp, B. V., Samama, J. P. & Brändén, C-I. (1982). *J. Biol. Chem*. **257**, 14349-14358.

Eklund, H. & Ramaswamy, S. (2008). *Cell. Mol. Life Sci.* **65,** 3907-3917.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486-501.

Englander, S. W. & Kallenbach, N. R. (1983). *Q. Rev. Biophys.* **16**, 521-655.

Eriksson, A. E., Baase, W. A., Zhang, X. J., Heinz, D. W., Blaber, M., Baldwin, E. P. & Matthews, B. W. (1992). *Science* **255**, 178-83.

Esposito, L., Sica, F., Raia, C. A., Giordano, A., Rossi, M., Mazzarella, L. & Zagari, A. (2002). *J. Mol. Biol.* **318**, 463-477.

Esteban-Torres, M., Álvarez, Y., Acebrón, I., De Las Rivas, B., Muñoz, R., Kohring, G-W, Roa, A. M., Sobrino, M. & Mancheño, J. M. (2012). *FEBS Lett.* **586**, 3127-3133.

Evans, P. R. (2011). *Acta Cryst* D**67**, 282-292.

Fan, F. & Plapp, B. V. (1999). *Arch. Biochim. Biophys*. **367**, 240-249.

Ganzhorn, A. J. & Plapp, B. V. (1988). *J. Biol. Chem.* **263**, 5446-5454.

Guy, J. E., Isupov, M. N. & Littlechild, J. A. (2003). *J. Mol. Biol.* **29**, 1041-1051.

Harding, M. M., Nowicki, M. W. & Walkinshaw, M. D. (2010). *Crystallogr. Rev*. **16**, 247-302.

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

Ishikawa, K., Higashi, N., Nakamura, T., Matsuura, T. & Nakagawa, A. (2007). *J. Mol. Biol.* **366**, 857-867.

Jackson, S. E., Moracci, M., elMasry, N., Johnson, C. M. & Fersht, A. R. (1993). *Biochemistry* **32**, 11259-11269.

Jeffery, J. & Jörnvall, H. (1988). *Adv. Enzymol. Relat. Areas Mol. Biol.* **61**, 47-106.

Kabsch, W. (2010). *Acta Cryst.* D**66**,125-132.

Kleifeld, O., Shi, S. P., Zarivach, R., Eisenstein, M. & Sagi, I. (2003). *Prot. Sci.* **12**, 468-479.

Klimacek, M., Hellmer, H. & Nidetzky, B. (2007). *Biochem. J.* **404,** 421-429.

Korkhin, Y., Kalb(Gilboa), A. J., Peretz, M., Bogin, O., Burstein, Y. & Frolow, F. (1998). *J. Mol. Biol.* **278**, 967–81.

Kovaleva, E. G. & Plapp, B. V. (2005). *Biochemistry* **44**, 12797-12808.

Makinen, M. W. & Yim, M. B. (1981). *Proc. Natl. Acad. Sci. USA* **78**, 6221-6225.

Matthews, B. W., Morton, A. G. & Dahlquist, F. W. (1995). *Science* **270**, 1847-1849.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. Read, R. (2007). *J. Appl. Cryst.* **40**, 658-674.

Nobelmann, B. & Lengeler, J. W. (1995). *Biochim. Biophys. Acta* **1262**, 69–72.

Pauly, T. A., Ekstrom, J. L., Beebe, D. A., Chrunyk, B., Cunningham, D., Griffor, M., Kamath, A., Lee, S. E., Madura, R., Mcguire, D., Subashi, T., Wasilko, D., Watts, P., Mylari, B. L., Oates, P. J., Adams, P. D. & Rath, V. L. (2003). *Structure* **11**, 1071-1085.

Pennacchio, A., Esposito, L., Zagari, A., Rossi, M. & Raira, C. A. (2009). *Extremophiles* **13**, 751-761.

Persson, B., Hedlund, J. & Jörnvall, H. (2008). *Cell. Mol. Life Sci.* **65**, 3879-3894.

Persson, B., Ziggler Jr, J. S. & Jörnvall, H. (1994). *Eur. J. Biochem.* **226**, 15-22.

Plapp, B. V. (2010). Arch. Biochem. Biophys. **493**, 3-12.

Raj, S. B., Ramaswamy S. & Plapp, B. V. (2014). *Biochemistry* **53**, 5791-5803.

Riveros-Rosas, H., Julián-Sánchez, A., Villalobos-Molina, R., Pardo, J. P. & Piña, E. (2003). *Eur. J. Biochem.* **270**, 3309-3334.

Ryde, U. (1995). *Protein Sci.* **4**, 1124-1132.

Sekhar, V. C. & Plapp, B. V. (1990). *Biochemistry* **29**, 4289-4295.

Shimomura, Y., Kakuta, Y. & Fukuyama, K. (2003). *J. Bacteriol.* **185**, 4211-4218.

Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fujimoto, E. K., Goeke, N. M., Olson, B. J. & Klenk, D. C. (1985). *Anal. Biochem*. **150**, 76-85.

Suhre, K. & Sanejouand, Y-H. (2004). *Nucleic Acid Res.* **32**, W610-W614.

Sulzenbacher, G., Alvarez, K., Van Den Heuvel, R. H., Versluis, C., Spinelli, S., Campanacci, V., Valencia, C., Cambillau, C., Eklund, H. & Tegoni, M. (2004). *J. Mol. Biol.* **342**, 489–502.

Sulzenbacher, G., Roig-Zamboni, V., Pagot, F., Grisel, S., Salomoni, A., Valencia, Ch., Campanacci, V., Vincentelli, R., Tegoni, M., Eklund, H. & Cambillau, Ch. (2004). *Acta Cryst*. D**60**: 1855-1862.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, K. D., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R. & Read R. (2011). *Acta Cryst.* D**67**, 235-242.

Wolfenden, R. & Radzicka, A. (1994). *Science* **265**, 936-937.

Yennawar, H., Moller, M., Gillian, R. & Yennavar, N. (2011). *Acta Cryst*. D**67**:440-446.

Yu, B., Blaber, M., Gronenborn, A. M., Clore, G. M. & Caspar, D. L. D. (1999). *Proc. Natl. Acad. Sci. USA* **96,**103-108.

Table 1
Data-collection and refinement statistics for GPDH structures.

Values in parentheses are for the last shell.

	GPDH:Zn^{2+} (model 1)	GPDH:Zn^{2+}:Tris (model 2)	GPDH:Zn^{2+}:glycerol (model 3)
Data statistics			
Beamline	BL13-XALOC	ID29	ID23-1
Wavelength (Å)	1.2824	1.2815	0.9688
Space group	$P2_1$	$P2_1$	$P2_1$
a, b, c (Å)	43.3, 76.9, 108.6	65.6, 78.9, 68.6	65.3, 78.8, 68.2
α, β, γ (º)	90.0, 95.5, 90.0	90.0, 94.6, 90.0	90.0, 94.7, 90.0
Resolution (Å)	2.0	1.9	1.7
Completeness (%)	97.8 (96.0)	89.7 (54.7)	99.2 (98.7)
No. of observed reflections	160822	299907	483893
No. of unique reflections	47058	49264	70089
$R_{p.i.m.}$† (%)	4.3 (41.1)	5.1 (29.6)	4.0 (25.3)
$\langle I/\sigma(I)\rangle$	8.7 (1.8)	7.9 (2.3)	11.2 (3.2)
Wilson B value (Å^2)	31.3	11.9	15.4
Refinement statistics			
Resolution limits (Å)	44.2 (2.0)	45.2 (1.9)	41.7 (1.7)
Z	2	2	2
No. of reflections used	44566	46624	66417
R_{cryst}‡ (%)	19.5	20.2	18.5
R_{free}§ (%)	25.6	25.8	22.8
No. of atoms			
Protein	5254	5270	5349
Ligands	5	7	7
Water	276	471	654
Average *B* factor (Å^2)			
Protein			
Chain A	50.7	22.8	23.6
Chain B	60.9	24.6	19.5
Ligands			
Zn^{2+}	85.5	27.9	18.6
Tris	---	30.2	---
Glycerol	---	---	31.1
Water oxygen	49.7	32.4	33.2
R.m.s.d., bond lengths (Å)	0.009	0.008	0.007
R.m.s.d., bond angles (º)	1.28	1.16	1.11
Ramachandran Plot			
Favoured (%)	91.7	97.1	96.9
Allowed (%)	6.4	2.9	2.9
Outlier (%)	1.9	0	0.2
PDB code	4ueo	4uek	4uej

†$R_{p.i.m.} = \Sigma_{hkl}\{1/[N(hkl)-1]\}^{1/2} \Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \Sigma_i I_i(hkl)$.
‡$R_{cryst} = \Sigma_{hkl} \big||F_{obs}| - |F_{calc}|\big| / \Sigma_{hkl} |F_{obs}|$. §$R_{free}$ is the cross validation *R* factor for the 5% of the reflections against which the model was not refined.

Table 2
Kinetic parameters for wild-type GPDH and single-point variants

Variant	Substrate	K_m (mM)	k_{cat} (1/s)	k_{cat}/K_m [1/(s· M)]
Wild-type	Galactitol	26 ± 4	37	1400
Wild-type	LG1P	0.16 ± 0.03	8800	$55.0 \cdot 10^6$
Wild-type	DT6P	1.00 ± 0.05	19500	$19.5 \cdot 10^6$
Ser40Ala	LG1P	0.14 ± 0.01	1500	$10.7 \cdot 10^6$
Asp41Asn	LG1P	0.17 ± 0.08	250	$1.5 \cdot 10^6$
Arg44Thr	LG1P	0.15 ± 0.01	170	$1.1 \cdot 10^6$
Arg44Tyr	LG1P	0.033 ± 0.001	1700	$51.5 \cdot 10^6$
His51Ala	LG1P	0.17 ± 0.01	2500	$14.7 \cdot 10^6$
Tyr106Ala	LG1P	0.20 ± 0.02	3000	$15.0 \cdot 10^6$
Arg112Ala	LG1P	0.09 ± 0.01	2700	$30.0 \cdot 10^6$
Met288Arg	LG1P	0.23 ± 0.01	1200	$5.2 \cdot 10^6$

Figure captions

Figure 1. Multiple amino acid sequence alignment of GPDH with the closest structural homologs: zinc-binding dehydrogenase from *Sinorhizobium meliloti* (4ejm), threonine dehydrogenase from *Pyrococcus horikoshii* (2dfv), sorbitol dehydrogenase from silverleaf whitefly (1e3j), threonine 3-dehydrogenase from *Thermus thermophiles* (2dq4), L-threonine dehydrogenase from *Thermococcus kodakaraensis* (3gfb) and human sorbitol dehydrogenase (1pl8). Secondary structural elements of GPDH according to DSSP are indicated as orange arrows (β-strands) and green cylinders (α-helices). In the upper part an overall view of GPDH (represented as ribbon model) is shown indicating the locations of the catalytic and the structural zincs (green spheres) and the relative orientation of the GPDH subdomains.

Figure 2. Main contacting interface between GPDH subunits. A subunit is shown as an orange surface and the accompanying subunit as a ribbon models. The principal structural elements that participate in intersubunit interactions are strands β13 and β14 and helix α8 from both subunits.

Figure 3. Representation of the regular secondary structural elements of GPDH that delimitate an internal cavity situated at the main contacting interface between subunits. (a) Two orthogonal views of such structural elements (helices and strands from one subunit are shown as green *cylinders* and cyan *arrows*, respectively, and as magenta *cylinders* and yellow *arrows* from the other subunit). (b) Close-up view of the cavity showing the side chains of the residues facing its lumen. All residues are hydrophobic. Since the cavity exhibits a two-fold symmetry, only unique residues are indicated for simplicity.

Figure 4. Zinc coordination observed in the different GPDH complexes. (a) Structural zinc in the lobe loop. The structural zinc is shown as a green sphere. The electron density for Zn^{2+} (orange) is derived from an anomalous map (5σ level) prepared from data collected at the zinc peak (9.667 keV; 1.28245 Å wavelength) and that for the ligands is a composite omit map (1σ level) calculated with phenix. (b) Catalytic zinc ligands in GPDH preincubated with the metal (model 1). The zinc (shown as a grey *sphere*) is coordinated by interactions with Cys38, His59 (*stick* representation) and two water molecules (red

spheres). The electron density maps are as in panel (a). (c) Catalytic zinc ligands in GPDH with bound Tris. The ligands are Cys38, His59 a conserved water molecule (red *sphere*) and a hydroxyl group of the Tris molecule. The electron density map is a composite omit map contoured at 1σ level (for clarity the electron density for Tris is shown as an orange mesh). (d) Catalytic zinc ligands in GPDH with bound glycerol. The metal is penta-coordinated by Cys38, His59 a conserved water molecule (red *sphere*) and two hydroxyl groups of the glycerol molecule. A strong hydrogen bond between the one primary –OH group of glycerol and the carboxylate moiety of Glu144 is represented. In all cases, potential hydrogen bonds are represented as dashed lines. Distances are indicated in Å.

Figure 5. The active site of GPDH. Details of the interactions established between the catalytically important residues Ser40, Asp41 and Arg44, and location of the internal, ordered water molecules (these molecules also appear in model 2). The composite omit map around these water molecules (contoured at 1σ level) is also shown. Potential hydrogen bonds are represented as dashed lines, and the ionic interaction between Arg44 and Asp41 as a magenta dotted line. Distances are indicated in Å.

Figure 6. The ternary GPDH-LG1P-NAD+ complex. GPDH residues from the substrate-binding site (see the text) are shown in *stick* representation, as are the substrate (LG1P) and the coenzyme (NAD^+). The guanidinium moiety of the Arg278 side chain from the accompanying subunit (cyan *sticks*) interacts with the phosphate moiety of the substrate. The first coordination sphere of the catalytic zinc (that appears as a purple *sphere*) is completed by Cys38, His59, a conserved water molecule and the C5 and C6 oxygen atoms from LG1P. The C5 H atom to be transferred as hydride to the pyridine of NAD^+ during the catalytic cycle is shown in cyan.

Figure 7. Orientation of LG1P and DG1P in the catalytic center of GPDH. (a) The complex determined for LG1P is compatible with catalysis by GPDH: first, the proton of the C5 hydroxyl group of DG1P is to the side chain of Ser40 that together with the C2D hydroxyl group of the ribose moiety from the NAD^+ cofactor would constitute the proton relay system of GPDH and second, the C5 H atom is located facing the pyridine ring of the NAD^+ coenzyme, satisfactorily oriented for hydride transfer to the C4 atom of the

pyridine moiety. (b) The complex with DG1P is incompatible with catalysis by GPDH. The restrictions imposed by the amino acid side chains configuring the substrate-binding pocket together with the zinc ion impede the formation of a catalytically productive complex (see the text for further details).

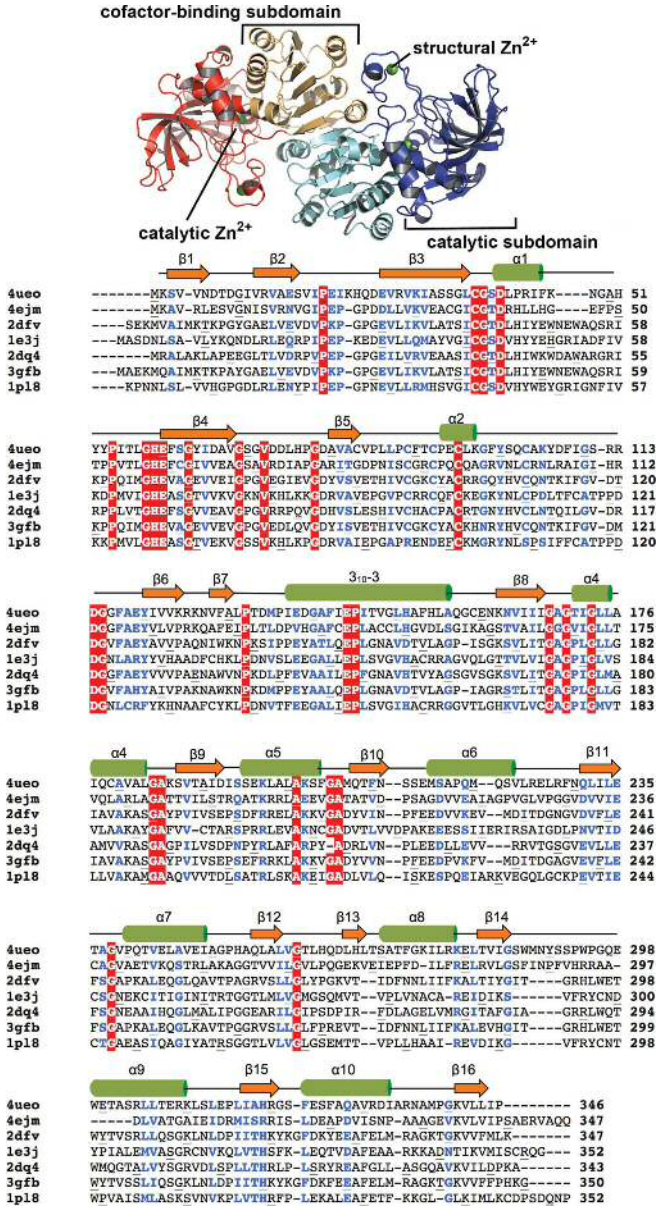
cofactor-binding subdomain
structural Zn2+
catalytic Zn2+
catalytic subdomain
β1 β2 β3 α1
4ueo -------MKSV-VNDTDGIVRVAESVIPEIKHQDEVRVKIASSGLCGSDLPRIFK----NGAH 51
4ejm -------MKAV-RLESVGNISVRNVGIPEP-GPDDLLVKVEACGICGTDRHLLHG----EFPS 50
2dfv ----SEKMVAIMKTKPGYGAELVEVDVPKP-GPGEVLIKVLATSICGTDLHIYEWNEWAQSRI 58
1e3j ---MASDNLSA-VLYKQNDLRLEQRPIPEP-KEDEVLLQMAYVGICGSDVHYYEHGRIADFIV 58
2dq4 -------MRALAKLAPEEGLTLVDRPVPEP-GPGEILVRVEAASICGTDLHIWKWDAWARGRI 55
3gfb ---MAEKMQAIMKTKPAYGAELVEVDVPKP-GPGEVLIKVLATSICGTDLHIYEWNEWAQSRI 59
1pl8 ----KPNNLSL-VVHGPGDLRLENYPIPEP-GPNEVLLRMHSVGICGSDVHYWEYGRIGNFIV 57
β4 β5 α2
4ueo YYPITLGHEFSGYIDAVGSGVDDLHPGDAVACVPLLPCFTCPECLKGFYSQCAKYDFIGS-RR 113
4ejm TPPVTLGHEFCGIVVEACSAVRDIAPGARITGDPNISCGRCPQCQAGRVNLCRNLRAIGI-HR 112
2dfv KPPQIMGHEVAGEVVEIGPGVEGIEVGDYVSVETHIVCGKCYACRRGQYHVCQNTKIFGV-DT 120
1e3j KDPMVIGHEASGTVVKVGKNVKHLKKGDRVAVEPGVPCRRCQFCKEGKYNLCPDLTFCATPPD 121
2dq4 RPPLVTGHEFSGVVEAVGPGVRRPQVGDHVSLESHIVCHACPACRTGNYHVCLNTQILGV-DR 117
3gfb KPPQIMGHEVAGEVVEVGPGVEDLQVGDYISVETHIVCGKCYACKHNRYHVCQNTKIFGV-DM 121
1pl8 KKPMVLGHEASGTVEKVGSSVKHLKPGDRVAIEPGAPRENDEFCKMGRYNLSPSIFFCATPPD 120
β6 β7 310-3 β8 α4
4ueo DGGFAEYIVVKRKNVFALPTDMPIEDGAFIEPITVGLHAFHLAQGCENKNVIIIGAGTIGLLA 176
4ejm DGGFAEYVLVPRKQAFEIGLTLDPVHGAFCEPLACCLHGVDLSGIKAGSTVAILGGGVIGLLT 175
2dfv DGVFAEYAVVPAQNIWKNPKSIPPEYATLQEPLGNAVDTVLAGP-ISGKSVLITGAGPLGLLG 182
1e3j DGNLARYYVHAADFCHKLPDNVSLEEGALLEPLSVGVHACRRAGVQLGTTVLVIGAGPIGLVS 184
2dq4 DGGFAEYVVVPAENAWVNPKDLPFEVAAILEPFGNAVHTVYAGSGVSGKSVLITGAGPIGLMA 180
3gfb DGVFAHYAIVPAKNAWKNPKDMPPEYAALQEPLGNAVDTVLAGP-IAGRSTLITGAGPLGLLG 183
1pl8 DGNLCRFYKHNAAFCYKLPDNVTFEEGALIEPLSVGIHACRRGGVTLGHKVLVCGAGPIGMVT 183
α4 β9 α5 β10 α6 β11
4ueo IQCAVALGAKSVTAIDISSEKLALAKSFGAMQTFN--SSEMSAPQM--QSVLRELRFNQLILE 235
4ejm VQLARLAGATTVILSTRQATKRRDAEEVGATATVD--PSAGDVVEAIAGPVGLVPGGVDVVIE 236
2dfv IAVAKASGAYPVIVSEPSDFRRELAKKVGADYVIN--PFEEDVVKEV--MDITDGNGVDVFLE 241
1e3j VLAAKAYGAFVV-CTARSPRRLEVAKNCGADVTLVVDPAKEEESSIIERIRSAIGDLPNVTID 246
2dq4 AMVVRASGAGPILVSDPNPYRLAFARPY-ADRLVN--PLEEDLLEVV---RRVTGSGVEVLLE 237
3gfb IAVAKASGAYPVIVSEPSEFRRKLAKKVGADYVVN--PFEEDPVKFV--MDITDGAGVEVFLE 242
1pl8 LLVAKAMGAAQVVVTDLSATRLSKAKEIGADLVLQ--ISKESPQEIARKVEGQLGCKPEVTIE 244
α7 β12 β13 α8 β14
4ueo TAGVPQTVELAVEIAGPHAQLALVGTLHQDLHLTSATFGKILRKELTVIGSWMNYSSPWPGQE 298
4ejm CAGVAETVKQSTRLAKAGGTVVILGVLPQGEKVEIEPFD-ILFRELRVLGSFINPFVHRRAA- 297
2dfv FSGAPKALEQGLQAVTPAGRVSLLGLYPGKVT--IDFNNLIIFKALTIYGIT----GRHLWET 298
1e3j CSGNEKCITIGINITRTGGTLMLVGMGSQMVT--VPLVNACA-REIDIKS------VFRYCND 300
2dq4 FSGNEAAIHQGLMALIPGGEARILGIPSDPIR--FDLAGELVMRGITAFGIA----GRRLWQT 294
3gfb FSGAPKALEQGLKAVTPGGRVSLLGLFPREVT--IDFNNLIIFKALEVHGIT----GRHLWET 299
1pl8 CTGAEASIQAGIYATRSGGTLVLVGLGSEMTT--VPLLHAAI-REVDIKG------VFRYCNT 298
α9 β15 α10 β16
4ueo WETASRLLTERKLSLEPLIAHRGS-FESFAQAVRDIARNAMPGKVLLIP-------- 346
4ejm -----DLVATGAIEIDRMISRRIS-LDEAPDVISNP-AAAGEVKVLVIPSAERVAQQ 347
2dfv WYTVSRLLQSGKLNLDPIITHKYKGFDKYEEAFELM-RAGKTGKVVFMLK------- 347
1e3j YPIALEMVASGRCNVKQLVTHSFK-LEQTVDAFEAA-RKKADNTIKVMISCRQG--- 352
2dq4 WMQGTALVYSGRVDLSPLLTHRLP-LSRYREAFGLL-ASGQAVKVILDPKA------ 343
3gfb WYTVSSLIQSGKLNLDPIITHKYKGFDKFEEAFELM-RAGKTGKVVFFPHKG----- 350
1pl8 WPVAISMLASKSVNVKPLVTHRFP-LEKALEAFETF-KKGL-GLKIMLKCDPSDQNP 352

Figure 1

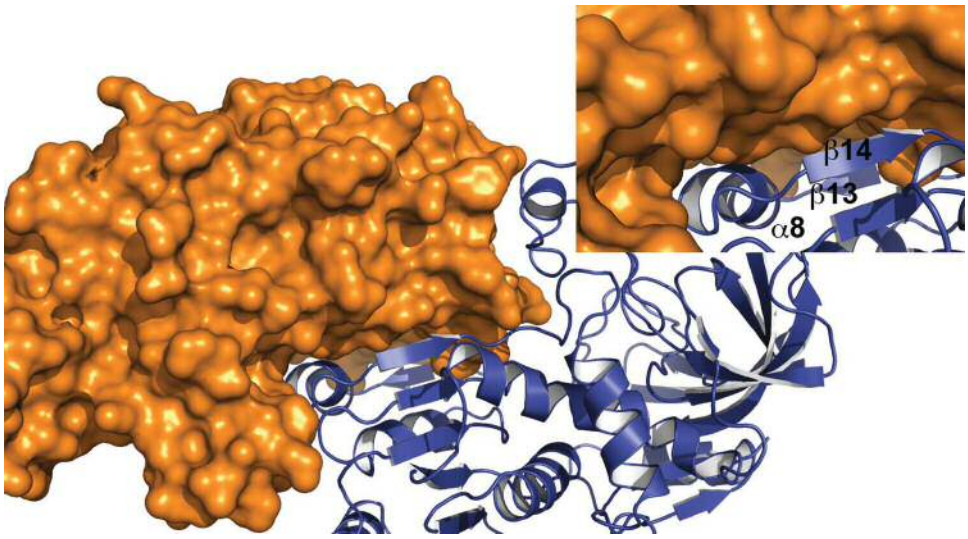

Figure 2

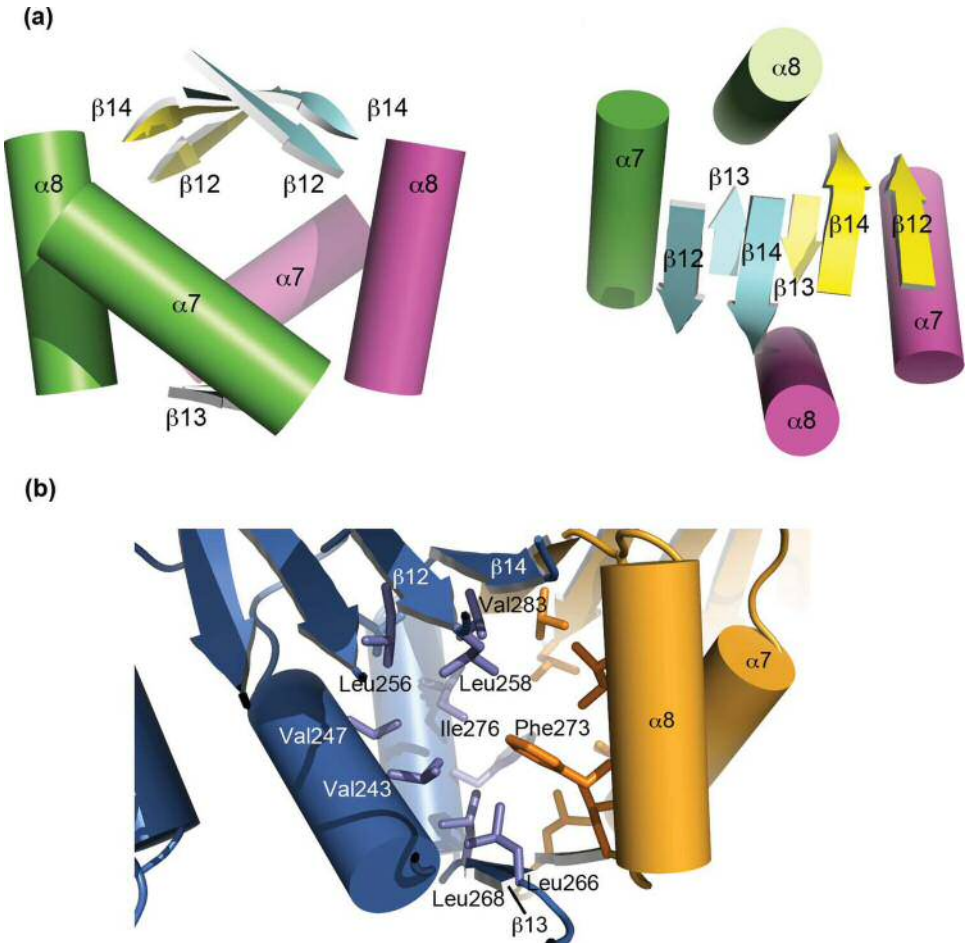
(a)
β14
β14
α8
β12
β12
α8
α7
α7
β13
α8
α7
β13
β12
β14
β14
β13
β12
α7
α8
(b)
β12
β14
Val283
α7
Leu256
Leu258
α8
Val247
Ile276
Phe273
Val243
Leu268
Leu266
β13

Figure 3

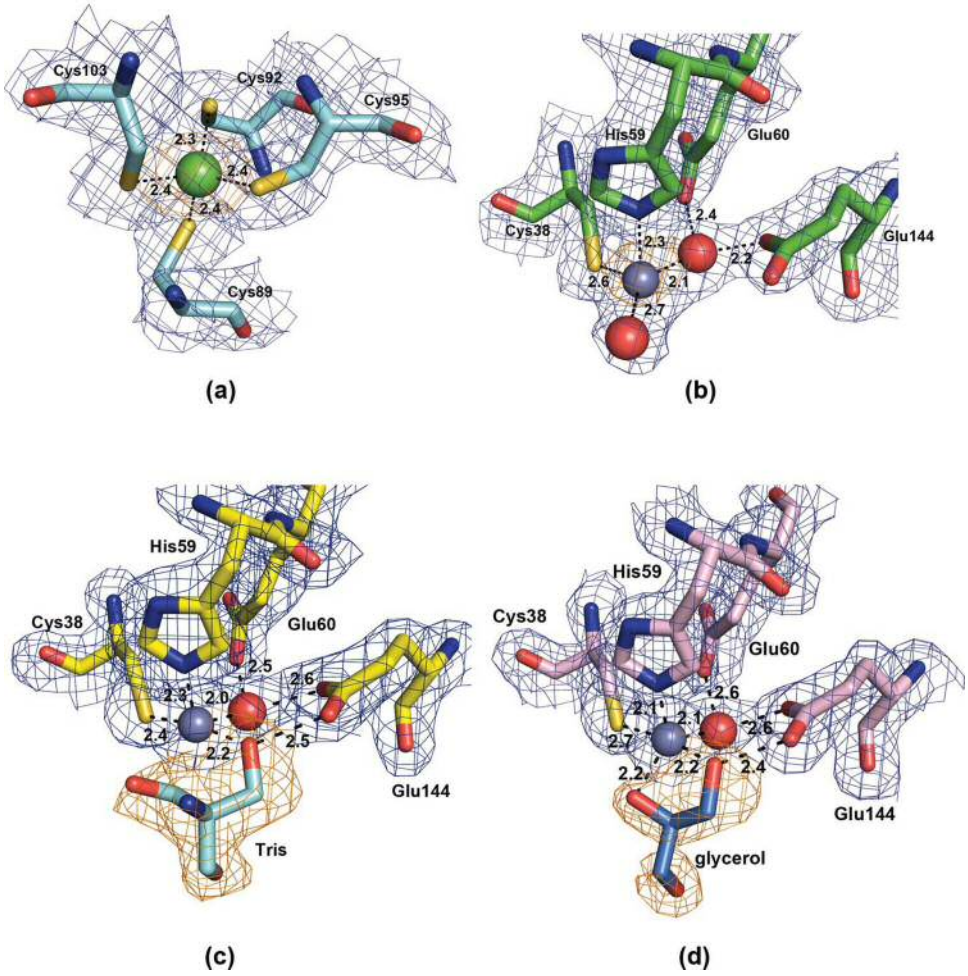
Cys103
Cys92
Cys95
2.3
2.4
2.4
2.4
Cys89
(a)
His59
Glu60
Cys38
2.4
2.3
2.2
Glu144
2.6
2.1
2.7
(b)
His59
Cys38
Glu60
2.5
2.6
2.3
2.0
2.4
2.5
2.2
Glu144
Tris
(c)
His59
Cys38
Glu60
2.6
2.1
2.1
2.6
2.7
2.2
2.2
2.4
Glu144
glycerol
(d)

Figure 4

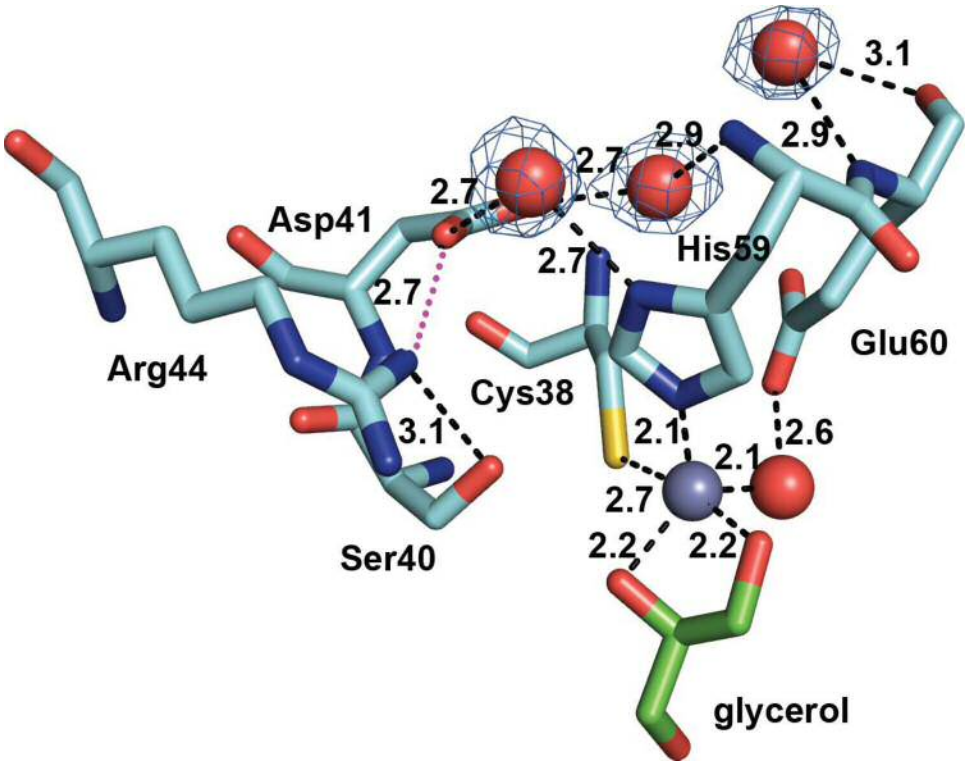
3.1
2.9
2.9
2.7
2.7
Asp41
His59
2.7
2.7
Glu60
Arg44
Cys38
3.1
2.1
2.6
2.1
2.7
2.2
2.2
Ser40
glycerol

Figure 5

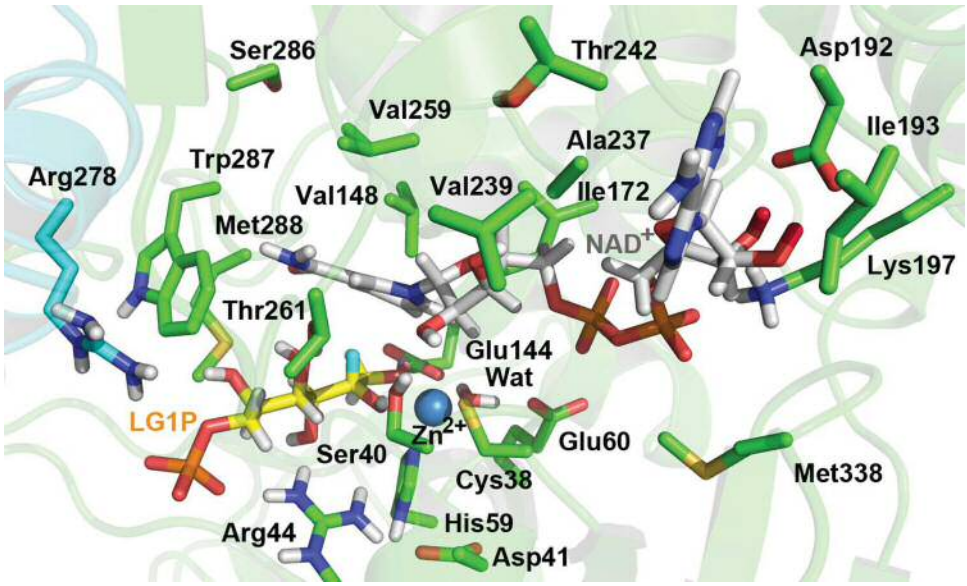
Ser286
Thr242
Asp192
Val259
Ile193
Ala237
Trp287
Arg278
Val148
Val239
Ile172
Met288
NAD+
Lys197
Thr261
Glu144
Wat
LG1P
Zn2+
Glu60
Ser40
Met338
Cys38
His59
Arg44
Asp41

Figure 6

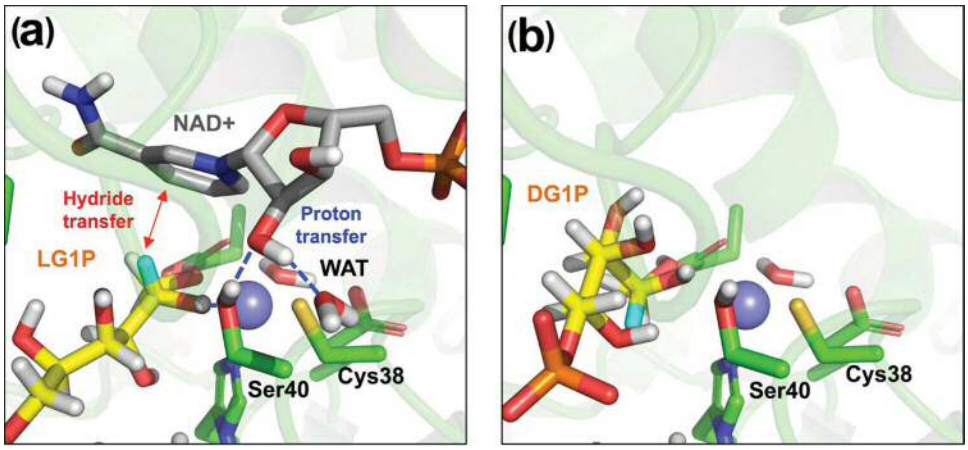
(a)
NAD+
Hydride
transfer
LG1P
Proton
transfer
WAT
Ser40
Cys38
(b)
DG1P
Ser40
Cys38

Figure 7