The crystal structure of galactitol-1-phosphate 5-dehydrogenase from *Escherichia coli* K12 provides insights into its anomalous behavior on IMAC processes

María Esteban-Torres[a], Yanaisis Álvarez[b,c], Iván Acebrón[b], Blanca de las Rivas[a], Rosario Muñoz[a], Gert-Wieland Kohring[d], Ana María Roa[e], Mónica Sobrino[e], José M. Mancheño[b,] *

[a]Laboratorio de Biotecnología Bacteriana, Instituto Ciencia y Tecnología de Alimentos y Nutrición (ICTAN), CSIC, Juan de la Cierva 3, 28006 Madrid, Spain

[b]Grupo de Cristalografía y Biología Estructural, Instituto de Química Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain

[c]Centro de Estudios Avanzados de Cuba, CITMA, Carretera de S. Antonio km 2, 17100 Habana, Cuba

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

[e]Screening & Compound Profiling, CIB, GlaxoSmithKline, Santiago Grisolía 4, 28760 Tres Cantos, Madrid, Spain

*Corresponding author: José M. Mancheño

Instituto de Química Física Rocasolano (CSIC)

Serrano 119, 28006 Madrid, Spain

E-mail: xjosemi@iqfr.csic.es

Phone: +34-917459547 Fax: +34-915642431.

Abbreviations: GPDH, endogenous galactitol-1-phosphate 5-dehydrogenase; rGPDH, recombinant, untagged GPDH; IMAC, immobilized metal affinity chromatography; NTA, nitrilotriacetic acid

Running title: Anomalous IMAC behavior of GPDH

ABSTRACT

Endogenous galactitol–1–phosphate 5–dehydrogenase (GPDH) (EC 1.1.1.251) from *Escherichia coli* spontaneously interacts with Ni^{2+}–NTA matrices becoming a potential contaminant for recombinant, target His–tagged proteins. Purified recombinant, untagged GPDH (rGPDH) converted galactitol into tagatose, and D–tagatose–6–phosphate into galactitol–1–phosphate, in a Zn^{2+}– and NAD(H)–dependent manner and readily crystallized what has permitted to solve its crystal structure. In contrast, N–terminally His–tagged GPDH was marginally stable and readily aggregated. The structure of rGPDH revealed metal–binding sites characteristic from the medium–chain dehydrogenase/reductase protein superfamily which may explain its ability to interact with immobilized metals. The structure also provides clues on the harmful effects of the N–terminal His–tag.

Keywords

Crystal structure; dehydrogenase; D–tagatose–6–phosphate; galactitol; metal binding–site; Zn–metalloenzyme.

Highlights

rGPDH from *E. coli* spontaneously interact with NTA matrices.

rGPDH is a homogeneous, stable dimer in solution.

The presence of an N–terminal His_6–tag in GPDH has adverse effects in its stability.

rGPDH is an NAD(H)–dependent zinc metalloenzyme.

The crystal structure of rGPDH has been solved up to 1.87 Å resolution.

INTRODUCTION

Currently immobilised metal affinity chromatography (IMAC) is one of the most widely used techniques for the purification of recombinant proteins [1,2] due to the standardization of methods for the efficient production of His–tagged proteins [3] and to the fact that the presence of His–tags in recombinant proteins is not normally harmful to their structures [4]. Despite in most cases IMAC leads to high protein purity upon single–step chromatographic purification [1], it is now recognized that other proteins can be co–purified with the target, His–tagged protein. In *E. coli* these "contaminants" have been classified in four groups: i) proteins with metal–binding sites, ii) proteins with exposed histidine clusters, iii) proteins that bind to the overexpressed His–tagged protein, and iv) proteins with affinity for agarose–based matrixes [1,5]. Notably, the presence of these accompanying proteins is not predictable nor universal since many of them are stress–responsive proteins what suggest that experimental variables such as culture conditions, media composition and also the particular bacterial strain, may have important effects on their production.

We show that endogenous GPDH and rGPDH from *Escherichia coli* spontaneously interact with NTA matrices what has permitted its efficient purification and further characterization. The enzyme catalyzed the oxidation of galactitol producing the rare sugar tagatose, and the reduction of D–tagatose–6–phosphate, which rendered galactitol–1–phosphate. Pure rGPDH could be crystallized and its structure solved in contrast to the recombinant, N–terminally His_6–tagged protein (His_6–GPDH), which showed marginal stability in solution and was therefore reluctant to crystallize, revealing a harmful effect of the tag on the GPDH structure.

GPDH is a member of the medium–chain dehydrogenase/reductase (MDR) protein superfamily [6] which is composed of homodimeric or homotetrameric proteins that contain two tetrahedrally coordinated zinc ions per subunit, one catalytic at the active site, and one structural [7]. The structure of rGPDH reveals a structural zinc coordinated by four cysteine residues and a catalytic metal–binding site, occupied by a Ni^{2+} ion, which is coordinated by three protein ligands

(C38, H59, E60) and a water molecule. The presence of this last ion within the catalytic metal–binding site indicates that this center may be involved in the interaction with the Ni–NTA matrix.

MATERIALS AND METHODS

Identification of galactitol–1–phosphate 5–dehydrogenase was done by mass spectrometry techniques from tryptic and chymotryptic peptides obtained from in–gel digestions of protein bands from the SDS–PAGE. MALDI–MS and MS/MS data were combined through the BioTools 3.0 program (Bruker Daltonik) to search the nonredundant protein databases NCBInr and SwissProt using the Mascot software (Matrix Science, London, UK) [8].

The gene *gatD* from *Escherichia coli* (K12) coding galactitol–1–phosphate 5–dehydrogenase was PCR–amplified using the primers F–his_gatD (5´– GGTGAAAACCTGTATTTCCAGG GCATGAAATCAGTGGTGAATGAT) and R–his_gatD (5´– ATCGATAAGCTTAGTTAGCTA TTATCAGGGAATGAGCAACACTTT, for producing His_6–GPDH, and F–r_gatD (5´– TAACTTTAAGAAGGAGATATACATATGAAATCAGTGGTGAATGAT) and R–his_gatD for producing rGPDH. Both protein variants were prepared essentially as described [9]. Purification involved an IMAC step on HisTrap FF Ni–affinity column (GE Healthcare) and an ion exchange on HiTrap Q HP column (GE Healthcare). Purity was checked by MALDI–TOF–MS [10].

Km determinations for galactitol oxidation were done from 1 mM to 100 mM galactitol in 100 mM Tris–HCl buffer pH 9.0, containing 0.5 mM $ZnCl_2$ and 1.8 mM NAD^+ with rGPDH concentrations of 200 µg/mL and 100 µg/mL, respectively. Km determinations for D–tagatose–6–phosphate reduction were done from 0.1 mM to 2 mM D–tagatose–6–phosphate (Sigma–Aldrich) in 100 mM Bis–Tris buffer pH 6.5, containing 0.5 mM $ZnCl_2$ and 0.3 mM NADH with rGPDH concentrations of 5 µg/mL and 2.5 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.

The oligomeric state of His_6–GPDH and rGPDH in solution was studied by analytical gel–filtration and analytical ultracentrifugation techniques as described previously [11]. Thermal stability studies were done by differential scanning fluorimetry assays [12] with the LightCycler 480 Real Time PCR System (Roche). Analysis of the data was done with the LightCycler Protein Melting Software (Roche).

Crystallization of GPDH (and rGPDH) was done using the sparse matrix method [13] essentially as described [14]. Optimization of the crystallization conditions to hanging drops containing 2 μl of the protein solution (9.3 mg/ml in 20 mM Tris–HCl, pH 8.0 with 0.1 M NaCl, 0.04% (w/v) sodium azide) and 1 μl of reservoir solution (24% (w/v) PEG 2000 MME, 0.1 M Bis–Tris, pH 6.5) rendered high quality diffraction crystals. rGPDH crystals for X–ray analysis were transferred to an optimized cryoprotectant solution (reservoir solution plus 15% (v/v) glycerol) for ~5 secs and then cryocooled at –173 °C in the cold nitrogen–gas stream. Diffraction data were recorded on a Q210 ADSC CCD detector (Area Detector Systems Corp.) at beamline ID14–1 at the European Synchrotron Radiation Facility (ESRF) (Grenoble, France). The images were processed and scaled using MOSFLM [15] and SCALA from the CCP4 program suite [16]. Intensities were converted to structure–factor amplitudes using TRUNCATE also from the CCP4 suite [16].

The crystal structure of rGPDH was determined by the molecular replacement method with PHASER from the PHENIX program [17] using the atomic coordinates of threonine–3–dehydrogenase from *Thermus thermophilus* (PDB code 2DQ4) as search model. The program CHAINSAW [18] from CCP4 [16] was used to prepare the search model. The estimated Matthews coefficient was 2.40 $Å^3$ Da^{-1}, corresponding to 49% solvent content and two rGPDH molecules within the asymmetric unit. Refinement was performed with PHENIX [17]. Non–crystallographic symmetry restraints were initially applied, but these were removed before the final refinement cycles. Isotropic individual temperature factors were refined, with the TLS parameters added in the final stages of refinement. After several further rounds of restrained and TLS refinement and manual correction using Coot [19] the structural model was finally refined to an *R*–factor of 21.3%

and an R_{free} of 27.3%. Validation of the structure was performed with MOLPROBITY [20]. Data processing and structure refinement statistics are shown in Table 1.

RESULTS AND DISCUSSION

Endogenous GPDH interacts NTA-matrixes

Initial small-scale expression trials aimed at producing the unrelated protein esterase Q88Y25_Lacpl from the lactic acid bacteria *Lactobacillus plantarum* WCFS1 in *E. coli* BL21 (DE3) cells revealed poor levels of expression, which were subsequently improved by the co-overexpression of the enzyme with molecular chaperones GroES/GroEL from *E. coli* [21]. In these conditions, a minor but significant peak was observed in the eluate from the HisTrap FF column. The imidazole concentration required for elution of the protein was ~ 25 mM, indicating a weakly bound protein [5]. The mass value of *m/z* obtained by MALDI-TOF mass spectrometry (37,362 Da) was lower than the expected one for the complete protein (38,628 Da), therefore suggesting either that the expressed protein was not the expected one or that it was a truncated form of the putative esterase.

In-gel trypsin and chymotrypsin digestions and subsequent mass spectrometry analyses unambiguously identified the protein as the galactitol-1-phosphate 5-dehydrogenase from *E. coli* (UniProt code: P0A9S3). The sequence coverage was 32 and 60% for the cleavage with trypsin and chymotrypsin, respectively. Some peptides containing Asn residues were identified as the deamidated species (chymotryptic peptides comprising residues 47–57, 47-61 and 47-64), and a few methionine oxidations were also observed (chymotryptic peptides 211–227, 278–290 and 327–343; tryptic peptides 203–225 and 280–304).

Therefore, endogenous (untagged) galactitol-1-phosphate 5-dehydrogenase from *E. coli* spontaneously interacts with HisTrap FF Ni-affinity columns, similarly to other proteins from *E. coli* which exhibit high affinity for divalent cations [5]. In this sense, GPDH shares similarity to

oxydoreductases from the medium-chain dehydrogenase/reductase (MDR) protein superfamily [6] that contain two distinct metal-binding sites namely a structural Zn^{2+}-binding site formed by four Cys residues and a catalytic metal-binding site (Fig 1). In particular, GPDH shows a sequence identity of ~27 % to threonine 3-dehydrogenase from *Thermus thermophilus* (PDB entry: 2DQ4) and also to L-threonine dehydrogenase from *Thermococcus kodakaraensis* (PDB entry: 3GFB) [22]. Protein ligands involved in metal coordination in the latter two dehydrogenases are also present in GPDH (Fig 1), suggesting that they play a similar role in this enzyme, presumably explaining its high affinity for Ni^{2+} ions immobilized in the NTA matrix.

Biochemical and biophysical characterization of recombinant variants of GPDH

Two recombinant variants of the dehydrogenase were produced in *E. coli* BL21 (DE3) cells: His_6-GPDH, which contained an N-terminal His_6-tag followed by a TEV cleavage site, and recombinant GPDH, which is equivalent to the endogenous enzyme.

Both enzymes exhibited dehydrogenase activity, converting galactitol (galactitol-1-phosphate is not commercially available) into tagatose. This reaction was dependent on both Zn^{2+} and NAD^+ (Table 2). The lower but significant activity detected in the absence of Zn^{2+} may be due to Zn^{2+} contaminations in the buffer substances. Since the results obtained with His_6-GPDH were hardly reproducible due to its instability (see below) a subsequent, thorough kinetic analysis was carried out only for the untagged protein.

Galactitol-1-phosphate is expected to be the main substrate for oxidation activity of rGPDH, but this is not commercially available. In comparison, the unphosphorylated galactitol has to be considered as a substrate with low oxidation activity and consequently a low affinity apparent **Km value of 25.8 mM (σ=3.3 mM) was determined** (Fig. 2). Importantly, the reverse reaction of D-tagatose reduction was not detectable. Reduction of D-tagatose-6-phosphate expressed an increase in specific activity of more than 500 fold as compared to galactitol oxidation (Table 2) and an apparent Km of 1 mM, which clearly demonstrates the expected substrate specificity of this enzyme

because galactitol–1–phosphate is the product of this reaction. Among other polyols as substrates only L–sorbitol exhibited oxidation activity and very low activity was found with 1,2–hexanediol (Table 2).

Analytical gel–filtration chromatography and ultracentrifugation assays in conjunction with MALDI–TOF–MS analysis of the intact proteins were used to study the oligomeric state(s) of rGPDH and His_6–GPDH in solution. The first approach revealed that both variants behave as dimeric species (rGPDH: 63 kDa; His_6–GPDH: 72 kDa) and also that His_6–GPDH is metastable in solution, forming large aggregates (Fig 3a).

Conversely, sedimentation velocity analyses revealed that rGPDH (19 and 30 μM) is very homogeneous with an *s* value of 4.5 ± 0.1 S (Fig. 3b). Additionally, the results obtained from the sedimentation equilibrium experiments indicate that the molecular weight of this species (70.7 kDa for the analysis at 14,000 rpm and 77.2 kDa at 18,000 rpm) is consistent with the mass of a dimer (rGPDH monomeric theoretical mass is 37,390 Da; Fig. 3c). This conclusion is definitively supported by the crystal structure of rGPDH which reveals it is a dimeric assembly (see below).

Similar experiments carried out with His_6–GPDH indicated that it is heterogeneous, in agreement with the analytical gel–filtration assays. The sedimentation velocity data indicated that despite the major species corresponded to dimeric His_6–GPDH (*s* value 4.8 ± 0.2 S), this species only represented ~25% out of total of species, whose *s* values were higher and broadly distributed (not shown).

As a whole, these results suggest that the presence of the His_6–tag at the amino–terminus of the enzyme affects negatively to its stability. In agreement with this, protein thermal stability was studied by differential scanning fluorimetry since this method can be applied to samples prone to aggregate [12] due to the small amount and low concentration of protein required. This approach revealed an apparent *T*m value of 63.07 ± 0.06 ºC for rGPDH and 59.96 ± 0.12 ºC for freshly prepared His_6–GPDH. The resulting ΔTm (3.11 ± 0.12 ºC) is within the range observed in other systems [23]. Hence, it is obvious that the His_6–tag has an adverse, destabilizing effect on the

enzyme which we believe it can be explained since the N-terminal segment structurally rationale and in fact the crystal structure of rGPDH provides some clues about this (see below).

Structural basis of the anomalous IMAC behavior of GPDH and a possible explanation for the adverse effect of the N-terminal His_6-tag

The crystal structure of apo rGPDH determined at 1.87 Å resolution revealed that the enzyme is a dimer. This result is notable as deduced from an analysis of the protein assemblies (exploration of the structure folds (CATH)) of the members of the MDR protein superfamily within the PDBe (http://www.ebi.ac.uk/pdbe). This analysis reveals that there exist 158 entries of homooligomers for the MDR group (3.90.180.10), which are distributed as: 110 dimeric (69.6%), 46 tetrameric (29.1 %) and 1 hexameric (1.3%). Since just 7 (PDB entries: 1uuf, 1qor, 1o8c, 1o89, 1wly, 1xa0, and 1iyz) out of 110 dimeric assemblies, and 24 out of 46 tetrameric assemblies come from bacteria, it can be inferred that bacterial MDR members are strongly biased towards tetrameric assemblies. Inspection of these bacterial, dimeric assemblies reveals a similar subunit arrangement, which is the one observed for rGPDH, where the 6-stranded β-sheet of the Rossmann fold of one subunit interacts with the equivalent β-sheet of the accompanying subunit forming an extended 12-stranded β-sheet across the dimer interface to create a large interface between subunits (Fig. 4a). Analysis with PISA (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver) indicates that formation of the rGPDH dimer buried 1,726 $Å^2$ of the 15,156 $Å^2$ accessible surface of each monomer, giving a solvent-accesible surface area of 26,860 $Å^2$ for the dimer. A second point of contact is the lobe loop surrounding the structural Zn^{2+}-binding site. Here, E94 forms a salt bridge with R228 from the other subunit (3.3 Å), and the aromatic ring of F99 stacks against the imidazole ring of H253 and undergoes a hydrophobic interaction with F229 which in turn packs against the aliphatic chain of K97. A hydrogen bond is also observed between Nε2 atom of N102 and the carbonyl oxygen of R278.

Based on the reported behaviour of several metalloproteins [5], the most plausible explanation for the intrinsic ability of rGPDH to interact with Ni^{2+}–NTA matrices is the presence of metal–binding sites within the structure. Spherical bunches of electron densities were observed in both the structural and catalytic sites, which were assigned to Zn^{2+} and Ni^{2+} ions, respectively since this resulted in *B*–factors for the metals in better agreement with those of the atoms from the inner–sphere coordination. Average *B*–factors for the atoms of the inner sphere coordination of the Zn^{2+} ions were 20.9 $Å^2$ (chain A) and 23.0 $Å^2$ (chain B), and those for zinc ions are 19.4 $Å^2$ and 21.7 $Å^2$, respectively, whereas the corresponding values for Ni^{2+} ions were 31.2 $Å^2$ (chain A) and 28.7 $Å^2$ (chain B), and 41.6 $Å^2$ (chain A) and 34.1 $Å^2$ (chain B), respectively.

The structural zinc ion is bound within the lobe loop adjacent to the catalytic domain, close to the subunits interface. The Zn^{2+} is tetrahedrally coordinated by C89, C92, C95 and C103 (Fig. 4b). Structural metal–binding site from both subunits are almost perfectly superimposable, with the coordination distances between Zn^{2+} and the ligands varying between 2.2 and 2.3 Å. Despite this loop makes intersubunit interactions (see above) there is no clear correlation between oligomeric state of MDR superfamily members and presence or absence of zinc: whereas the overall trend is that mammalian MDRs are dimeric and contain zinc and bacterial MDRs are tetrameric and lack zinc [24] it is notable that only one (together with rGPDH) out of the 7 dimeric bacterial MDRs (PDB code 1uuf) contains zinc.

Conversely, the active site of rGPDH harbors a nickel ion that is bound by C38, H59, E60 and a water molecule, with E144 being in close proximity also (Fig. 4c). Despite the subtle details of the coordination of the catalytic zinc cannot be inferred from this structure it is now recognized that there exist important differences in zinc ligation along the catalytic mechanism of MDRs [25] and therefore each one of the above potential ligands may play a role in the catalytic mechanism of rGPDH. Similarly to the structural Zn^{2+}–binding site, conformation of protein ligands from the catalytic sites from both subunits is essentially identical, although in this case small differences in coordination distances are observed: for instance, distances between E60 and Ni^{2+} are 3.1 Å and 2.6

Å in subunits A and B, respectively and 3.1 Å and 2.4 Å between Ni^{2+} and the closest water molecule in subunits A and B, respectively.

The presence of a Ni^{2+} ion within the catalytic metal–binding site of rGPDH supports the notion that this particular environment may mediate the interaction with the IMAC matrix. In this sense, it is important to remark that the presence of a Ni^{2+} ion in this metal binding–site does not preclude the interaction since pure rGPDH interacts with the matrix (not shown) and therefore initial Ni^{2+}–depletion conditions are not required. The existence of this interaction may open up the possibility to configure straightforward purification protocols for other members of the MDR superfamily by IMAC.

Obviously, the potential interaction between the IMAC matrix and solvent exposed histidine residues cannot be discarded. In this regard, the crystal structure of rGPDH reveals an additional Ni^{2+} ion coordinated to the imidazole rings of H267 from both subunits, a water molecule and the carboxylate group of E300 from a symmetry related molecule (Fig. S1). Despite this metal binding–site is crystallographic in that it results from the protein crystal packing it cannot be discarded as a latent Ni^{2+} binding–site in solution interacting with the matrix.

The N–terminal end of rGPDH is not flexible but rigid as indicated by: firstly, the N–terminal methionine of rGPDH is perfectly defined in the electron density map with the side chain being buried within the structure (Fig. 4d); secondly, the β1 strand (residues 1–6) is tightly packed between the β2 strand (residues 13–17) and the extended segment formed by residues 117–119; thirdly, a salt bridge is formed between E16 carboxylic group and the N–terminal amino group. It is obvious that incorporation of additional residues at the N–terminal end (as in His_6–GPDH) eliminates this last interaction what may destabilize locally the protein. This fact together with the expected, increased flexibility within this region due to the presence of the complete tag (His_6–tag plus TEV cleavage site) may result in an improper folding of this region making His_6–GPDH a metastable molecular species.

ACKNOWLEDGEMENTS

JMM and RM thank the *Ministerio de Ciencia e Innovación* for their research grants BFU2010–17929, CSD2006–00015 and AGL2011–22745, CSD2007–00063 FUN–C–FOOD and Comunidad de Madrid S2009/AGR–1469, respectively. JMM thanks ESRF (Grenoble, France) for provision of synchrotron radiation facilities.

REFERENCES

1. Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubiek, J., Fabis, R., Labahn, J. and Schäfer, F. (2009). Immobilized–metal affinity chromatography (IMAC): A Review. *Methods Enzymol*. 463, 439–473.

2. Malhotra, A. (2009). Tagging for protein expression. *Methods Enzymol*. 463, 239–258.

3. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. and Stüber, D. (1988). Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. *Biotechnology*. 6, 1321–1325.

4. Carson, M., Johnson, D.H., McDonald, H., Brounillette, Ch. and DeLucas, L.J. (2007). His–tag impact on structure. *Acta Crystallogr. D Biol. Crystallogr*. 63, 295–301.

5. Bolanos–García, V.M. and Davies, O.R. (2006). Structural analysis and classification of native proteins from E. coli commonly co–purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta*. 1760, 1304–1313.

6. Persson, B., Hedlund, J. and Jörnvall, H. (2008). The MDR superfamily. *Cell. Mol. Life Sci*. 65, 3879–3894.

7. Auld, D.S. and Bergman, T. (2008). The role of zinc for alcohol dehydrogenase structure and function. *Cell. Mol. Life Sci*. 65, 3961–3970.

8. Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. (1999). Probability–based protein

identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.

9. Curiel, J.A., de las Rivas, B., Mancheño, J.M. and Muñoz, R. (2011). The pURI family of expression vectors: a versatile set of ligation independent cloning plasmids for producing recombinant His-fusion proteins. *Prot. Express. Purif.* 76, 44–53.

10. Acebrón, I., Curiel, J.A., de las Rivas, B., Muñoz, R. and Mancheño, J.M. (2009). Cloning, production, purification and preliminary crystallographic analysis of a novel glycosidase from the food lactic acid bacterium *Lactobacillus plantarum* CECT 748^{T}. *Prot. Express. Purif.* 68, 177–182.

11. de las Rivas, B., Fox, G.C., Angulo, I., Martínez-Ripoll, M., Rodríguez, H., Muñoz, R. and Mancheño, J.M. (2009). Ornithine transcarbamylase from *Lactobaillus hilgardii*: structural insights into the oligomeric assembly and metal binding. *J. Mol. Biol.* 393, 425–434.

12. Niesen, H., Berglund, H. and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols.* 2, 2212–2221.

13. Jancarik, J. and Kim, S.H. (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J. Appl. Cryst.* 24, 409-411.

14. Angulo, I., Acebrón, I., de las Rivas, B., Muñoz, R., Rodríguez-Crespo, J.I., Menéndez, M., García, P., Tateno, H., Goldstein, I.J., Pérez-Agote, B. and Mancheño, J.M. (2011). High-resolution structural insights on the sugar-recognition and fusion tag properties of a versatile β-trefoil lectin domain from the mushroom Laetiporus sulphureus. *Glycobiology* 21, 1349–1361.

15. Leslie, A.G. (1999). Integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* 55, 1696–1702.

16. The CCP4 suite: programs for protein crystallography. (1994). *Acta Crystallogr. D Biol. Crystallogr.* 50, 760–763.

17. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echoo, 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, J.S., Terwilliger, T.C. and Zwart, P.H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst. D.* 66, 213–221.

18. Stein, N. (2008). CHAINSAW: a program for mutating pdb files used as templates for molecular replacement. *J. Appl. Cryst.* 41, 641–643.

19. Emsley, P. and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Cryst. D* 60, 2126–2132.

20. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X. et al. (2007). MolProbity: all atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 35, W375–W383.

21. Álvarez, Y., Esteban-Torres, M., Acebrón, I., de las Rivas, B., Muñoz, R., Martínez-Ripoll, M. and Mancheño, J.M. (2012). Preliminary X-ray analysis of twinned crystals of the Q88Y25_Lacpl

esterase from *Lactobacillus plantarum* WCFS1. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 67, 1436–1439.

22. A. Bowyer, A., Mikolajek, H., Stuart, J.W., Wood, S.P., Jamil, F., Rashid, N., Akhtar, M. and Cooper, J.B. (2009). Structure and function of the L–threonine dehydrogenase (TkTDH) from the hyperthermophilic archaeon *Thermococcus kodakaraensis*. *J. Struct. Biol.* 168, 294–304.

23. M. Vedadi, M., Niesen, F.H., Allali–Hassani, A., Fedorov, O.Y., Finerty, P.J., Wasney, G.A., Yeung, R., Arrowsmith, Ch., Ball, L.J., Berglund, H., Hui, R., Marsden, B.D., Nordlund, P., Sundstrom, M. and Edwards, A.M. (2006). Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. USA.* 103(43), 15835–15840.

24. Banfield, M.J., Salvucci, M.E., Baker, E.N. and Smith, C.A. (2001). Crystal Structure of the NADP(H)–dependent Ketose Reductase from *Bemisia argentifolii* at 2.3 Å Resolution. *J. Mol. Biol.* 306, 239–250.

25. Baker, P.J., Brittona, K.L., Fisher, M., Esclapez, J., Pire, C., Bonete, M.J., Ferrer, J., and Rice, D.W. (2012). Active site dynamics in the zinc–dependent medium chain alcohol dehydrogenase superfamily. *Proc. Natl. Acad. Sci. USA.* 106(3), 779–784.

26. Vaguine, A.A., Richelle, J. and Wodak, S.J. (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure–factor data and their agreement with the atomic model. *Acta Crystallog. sect. D*, 55, 191–205.

FIGURE LEGENDS

Fig.1. Multiple amino acid sequence alignment of GPDH with threonine 3–dehydrogenase from *Thermus thermophilus* (PDB entry: 2dq4) and with L–threonine dehydrogenase from *Thermococcus kodakaraensis* (PDB entry: 3gfb). Conserved positions are shown in *red* characters; positions conserved in GPDH and either 2dq4 or 3gfb are shown as *white* characters in a blue box. Residues that coordinate the catalytic Zn^{2+} ion are in *green* boxes, and the four cysteine residues that coordinate the structural Zn^{2+} ion are within *cyan* boxes.

Fig.2. Lineweaver–Burk plots for rGPDH from *E. coli* with galactitol (a) and D–tagatose–6–phosphate (b) as substrates. Four different experiments were done for the galactitol oxidation reaction at two different enzyme concentrations: 100 μg/ml (closed symbols) and 200 μg/ml (open symbols). Reduction of D–tagatose–6–phosphate was studied at two enzyme concentrations: 2.5 μg/ml (closed symbols) and 5.0 μg/ml (open symbols).

Fig.3. Analysis of the oligomeric state of rGPDH in solution. (a) Analytical gel–filtration analysis of rGPDH and His_6–GPDH. The elution profiles of both proteins (red, rGPDH; black, His_6–GPDH) are shown together with the eluted positions of some standard proteins. Inset, semilog plot of the molecular mass of standard proteins (kDa) *versus* their Kav values. (b) Sedimentation coefficient distribution c(*s*) corresponding to the sedimentation velocity of purified rGPDH (30 μM). (c)

Sedimentation equilibrium data of rGPDH (19 μM) at 14,000 rpm (open circles) and 18,000 (open squares). The best fit to the data sets (solid line curves) corresponded to an ideal dimeric species. Residuals from the fits are shown in the two lower panels.

Fig.4. Crystal structure of rGPDH. (a) Dimeric assembly of rGPDH within the monoclinic crystals. The 6–stranded β–sheet of the Rossmann fold of one subunit interacts with the equivalent β–sheet of the accompanying subunit forming an extended 12–stranded β–sheet. (b) View of the composite omit map calculated using the program SFCHECK [26] contoured at 1.0 σ around the structural zinc ion and (c) the nickel ion at the active site. The structural zinc ion is tetrahedrally coordinated by C89 (2.3 Å), C92 (2.3 Å), C95 (2.3 Å) and C103 (2.3 Å). (c) The active site nickel ion is coordinated by C38 (2.5 Å), H59 (2.2 Å), E60 (3.1 Å) and a water molecule (3.1 Å). E144 is also in close proximity to the metal (3.5 Å). (d) N-terminal region of rGPDH. The first four residues of the enzyme are defined in the composite omit map (except the flexible K2 side chain) indicating that it is a rigid segment. It is bridged by β2 strand and the loop between residues 117-119. The figures have been prepared with PyMol (http://pymol.sourceforge.net/).

Table(s)

Table 1
Data–collection and refinement statistics.

	rGPDH PDB ID: 4A2C
Data collection	
Beamline	ID14–1 (ESRF)
Wavelength (Å)	0.9334
Space group	$P2_1$
Unit–cell parameters (Å)	a= 43.7, b = 77.1, c = 107.5 $\alpha = 90°$, $\beta = 95.36°$, $\gamma = 90°$
Resolution range (Å)	29.81–1.87 [1.97–1.87][a]
No. of measured reflections	211,151
No. of unique reflections	56,481 [7,753]
Mean I/σ(I)	9.2 [2.5]
Completeness (%)	95.9 [95.9]
Multiplicity	3.7 [3.7]
R_{sym}(%)[b]	8.9 [52.8]
Wilson *B* factor (Å^2)	20.2
Refinement	
Resolution range (Å)	29.81–1.87
R_{work}(%)/R_{free}(%)	21.2/27.3
Molecules and nohydrogen atoms	
Protein molecules	2
Protein	5312
Ni^{2+}	3
Zn^{2+}	2
Water	608
Average B–factors (Å^2)	
Main chain	27.1
Side chain	31.6
Ni^{2+}	37.8
Zn^{2+}	20.6
water	34.7
Rms deviation from ideality	
Bonds (Å)	0.007
Angles (°)	1.127
Ramachandran plot (favored/allowed/generously allowed/disallowed) (%)[c]	97/99.9/0/0.1

[a]Values for the highest resolution shell are given in square brackets.
[b]$R_{sym} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity for the ith measurement of an equivalent reflection with indices h, k, and l and $\langle I(hkl)\rangle$ is the average intensity.
[c]Calculated using MOLPROBITY [20].

Table 2
rGPDH activity at different conditions and with different substrates

Substrate 20 mM	Cation 0.5 mM	NAD^+ 1.8 mM	NADH 0.3 mM	Spec. activity [U/mg]	Activity [%]
Galactitol	Zn^{2+}	+	–	0.060	100
Galactitol	Zn^{2+}	–	–	0.001	–
Galactitol	Mg^{2+}	+	–	0.016	26.7
Galactitol	–	+	–	0.020	33.3
Galactitol	–	–	–	0.001	–
D–Tagatose–6–P	Zn^{2+}	–	+	31.500	52,500.0
D–Tagatose	Zn^{2+}	–	+	0.001	–
L–Sorbitol	Zn^{2+}	+	–	0.090	150
D–Sorbitol	Zn^{2+}	+	–	0.001	–
Mannitol	Zn^{2+}	+	–	0.003	5.0
Ribitol	Zn^{2+}	+	–	0.004	6.7
Xylitol	Zn^{2+}	+	–	0.003	5.0
1,2–Hexanediol	Zn^{2+}	+	–	0.010	16.7

Figure1

Click here to download high resolution image

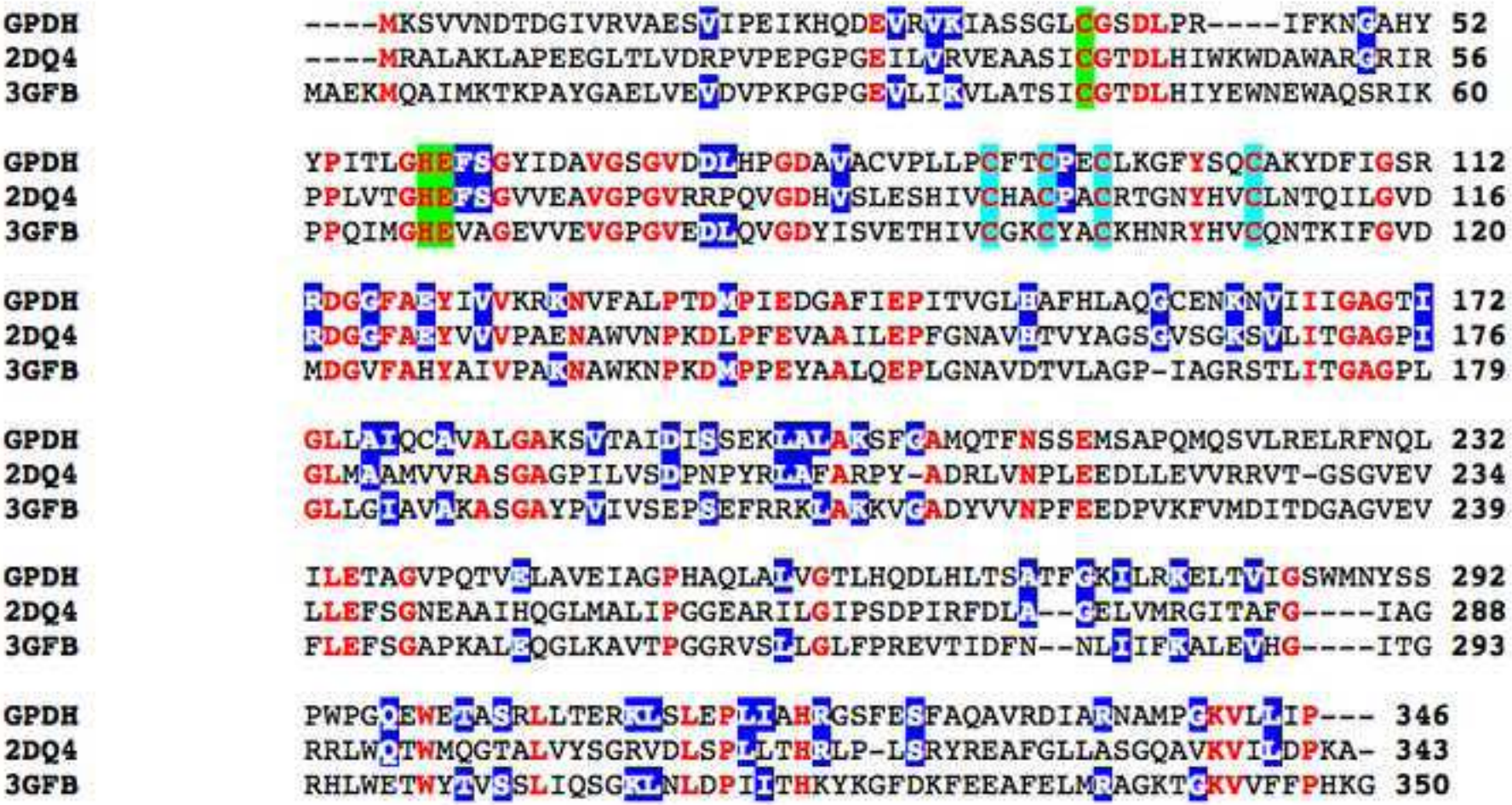

Figure2

Click here to download high resolution image

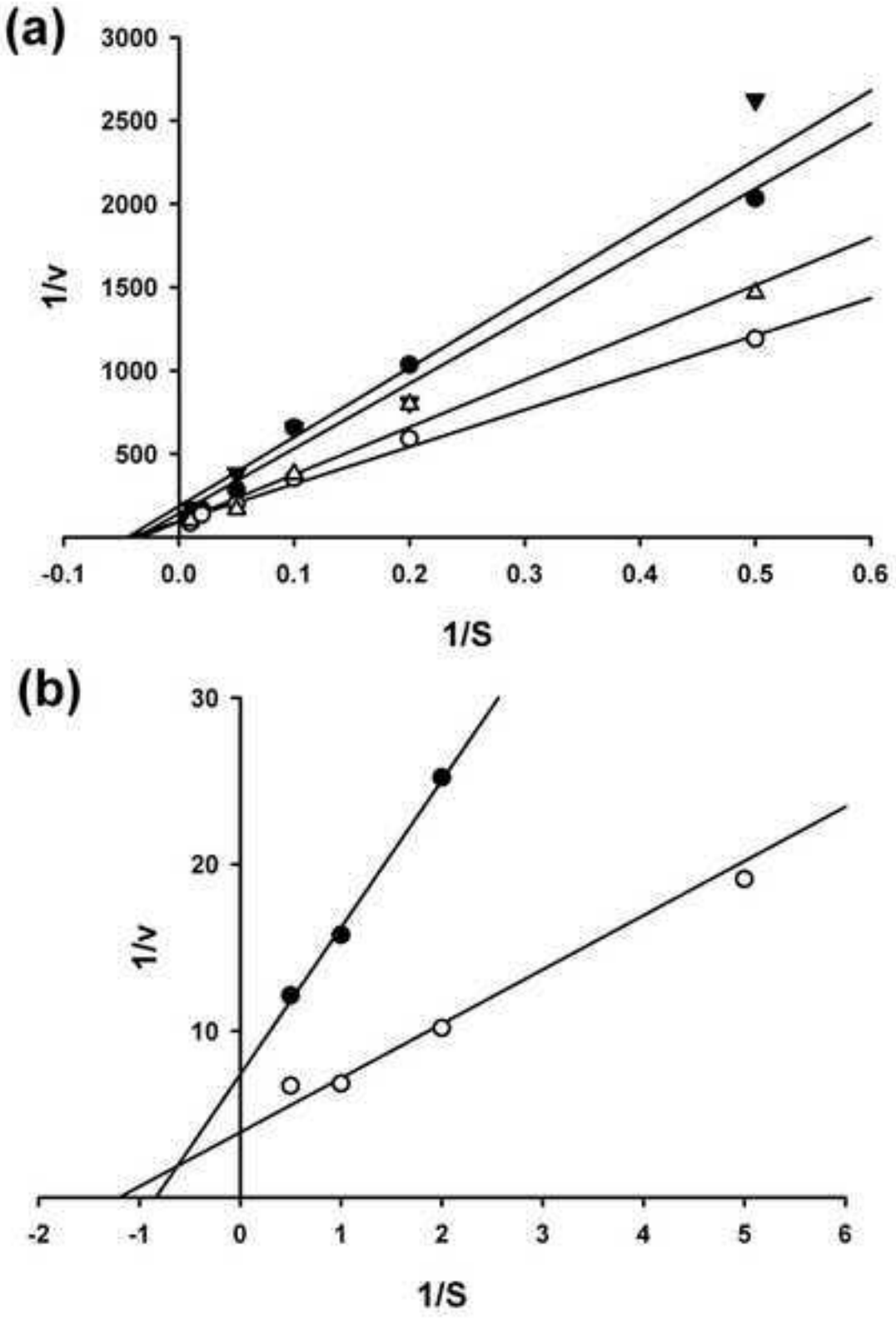

Figure3

Click here to download high resolution image

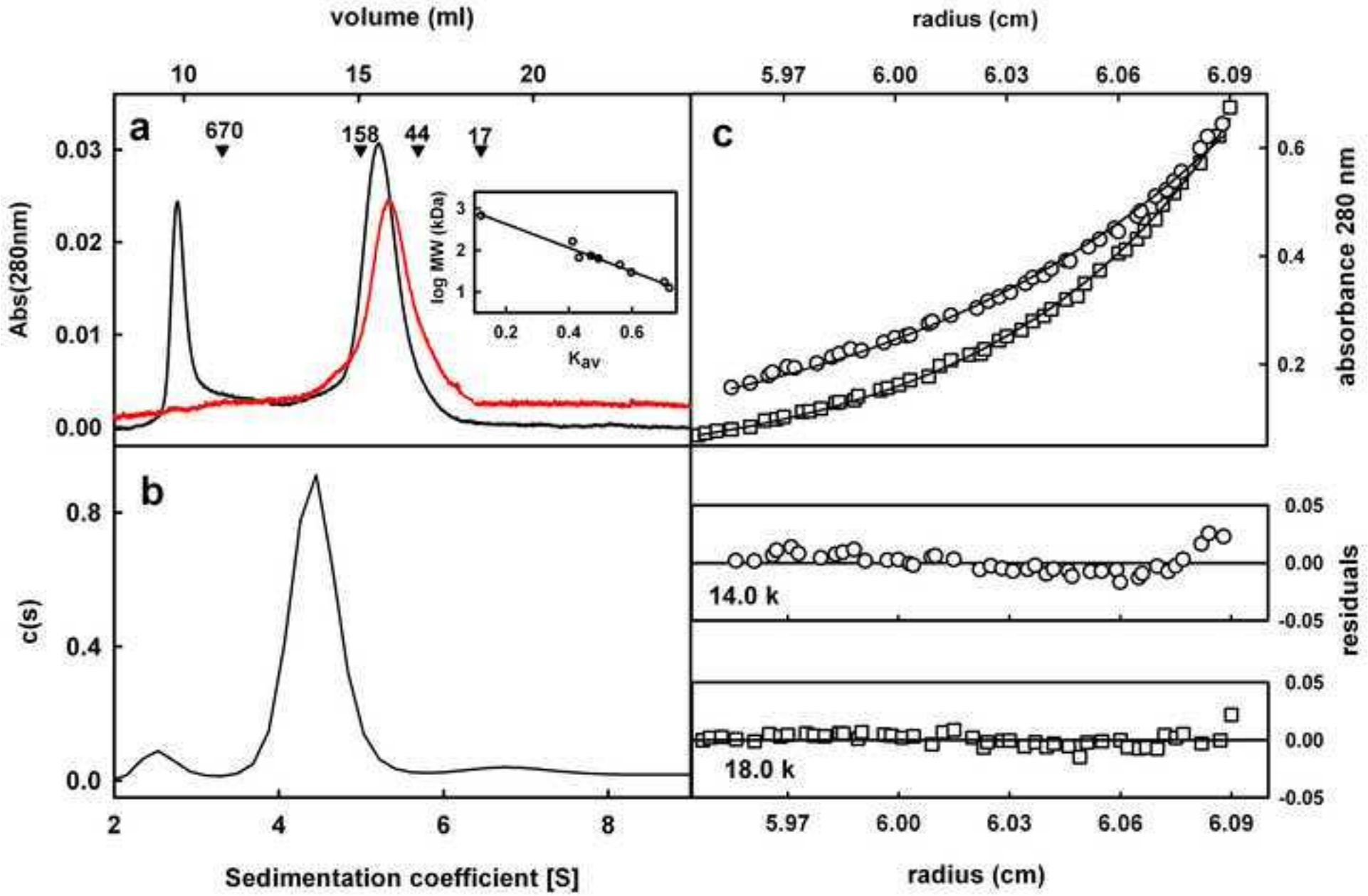

Figure4

Click here to download high resolution image