Structural characterization of the thermostable Bradyrhizobium japonicum D-sorbitol dehydrogenase

Fredslund, Folmer; Otten, Harm; Gemperlein, Sabrina; Poulsen, Jens-Christian Navarro; Carius, Yvonne; Kohring, Gert-Wieland; Lo Leggio, Leila

Published in:
Acta Crystallographica Section F

DOI:
10.1107/S2053230X16016927

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Document license:
Other

Citation for published version (APA):
Fredslund, F., Otten, H., Gemperlein, S., Poulsen, J-C. N., Carius, Y., Kohring, G-W., & Lo Leggio, L. (2016). Structural characterization of the thermostable *Bradyrhizobium japonicum* D-sorbitol dehydrogenase. *Acta Crystallographica Section F*, *72*(11), 846-852. https://doi.org/10.1107/S2053230X16016927

electronic reprint

STRUCTURAL BIOLOGY
COMMUNICATIONS

ISSN: 2053-230X
journals.iucr.org/f

Structural characterization of the thermostable *Bradyrhizobium japonicum* D-sorbitol dehydrogenase

Folmer Fredslund, Harm Otten, Sabrina Gemperlein, Jens-Christian N. Poulsen, Yvonne Carius, Gert-Wieland Kohring and Leila Lo Leggio

Acta Cryst. (2016). F**72**, 846–852

STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Structural characterization of the thermostable *Bradyrhizobium japonicum* D-sorbitol dehydrogenase

Folmer Fredslund,[a]‡ Harm Otten,[a]‡§ Sabrina Gemperlein,[b] Jens-Christian N. Poulsen,[a] Yvonne Carius,[c] Gert-Wieland Kohring[b] and Leila Lo Leggio[a]*

[a]Department of Chemistry, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark, [b]Microbiology, Saarland University, Campus Building A1.5, Saarbrücken, D-66123 Saarland, Germany, and [c]Department of Structural Biology, ZHMB, Saarland University, Building 60, D-66421 Homburg, Germany. *Correspondence e-mail: leila@chem.ku.dk

Received 3 May 2016
Accepted 21 October 2016

‡ These authors contributed equally.
§ Present address: Novozymes A/S, Kroghoejsvej 36, DK-2880 Bagsvaerd, Denmark.

Keywords: sorbitol dehydrogenase; thermostability; glucitol.

PDB reference: D-sorbitol dehydrogenase, 5jo9

Supporting information: this article has supporting information at journals.iucr.org/f

Bradyrhizobium japonicum sorbitol dehydrogenase is NADH-dependent and is active at elevated temperatures. The best substrate is D-glucitol (a synonym for D-sorbitol), although L-glucitol is also accepted, giving it particular potential in industrial applications. Crystallization led to a hexagonal crystal form, with crystals diffracting to 2.9 Å resolution. In attempts to phase the data, a molecular-replacement solution based upon PDB entry 4nbu (33% identical in sequence to the target) was found. The solution contained one molecule in the asymmetric unit, but a tetramer similar to that found in other short-chain dehydrogenases, including the search model, could be reconstructed by applying crystallographic symmetry operations. The active site contains electron density consistent with D-glucitol and phosphate, but there was not clear evidence for the binding of NADH. In a search for the features that determine the thermostability of the enzyme, the T_m for the orthologue from *Rhodobacter sphaeroides*, for which the structure was already known, was also determined, and this enzyme proved to be considerably less thermostable. A continuous β-sheet is formed between two monomers in the tetramer of the *B. japonicum* enzyme, a feature not generally shared by short-chain dehydrogenases, and which may contribute to thermostability, as may an increased Pro/Gly ratio.

1. Introduction

Rare sugars have been proposed to have a broad potential for application, but low availability often hinders systematic studies (Beerens *et al.*, 2012). They are known as sugar replacements (Lu *et al.*, 2008; Baek *et al.*, 2010), as therapeutics in cancer proliferation, as anti-inflammatory agents and as immunosuppressants (Lim & Oh, 2011). The so-called Izumoring (Granström *et al.*, 2004) shows in detail that all polyols and rare sugars can be produced from natural sources by the use of oxidoreductases, epimerases and isomerases, but the corresponding enzymes often lack sufficient turnover numbers and product yields, generating a constant demand for better catalysts.

Among the rare sugars, D-sorbose is expected to be a building block for industrially interesting compounds in addition to its use as a low-calorie sweetener (Huwig *et al.*, 1996). The subject of this study, *Bradyrhizobium japonicum* D-sorbitol dehydrogenase (*Bj*SDH), was investigated within a project to find stable enzyme variants, specifically for use in

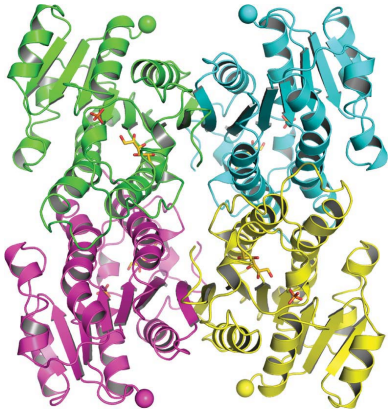


Table 1
Macromolecule-production information.

Source organism	*B. japonicum* USDA110
Expression vector	pET-24
Expression host	*E. coli*
Complete amino-acid sequence of the construct produced	HHHHHHMARELEGKVAAVTGAASGIGLASAEAML-AAGARVVMVDRDEAALKALCNKHGDTVIPLVV-DLLDPEDCATLLPRVLEKACQLDILHANAGTY-VGGDLVDADTMAIDRMLNLNVNVVMKNVHDVL-PHMIERRTGDIIVTSSLAAHFPTPWEPVYASS-KWAINCFVQTVRRQVFKHGIRVGSISPGPVVS-ALLADWPPEKLKEARDSGSLLEASDVAEVVMF-MLTRPRGMTIRDVLMLPTNFDL

the construction of an electro-enzymatic flow-cell device for the production of rare sugars with electrochemical cofactor regeneration (Wang *et al.*, 2012). *Bj*SDH has been shown to preferentially catalyse the oxidation of D-glucitol to D-fructose, but additionally to catalyse the oxidation of L-glucitol to D-sorbose, and offers interesting application potential with respect to its apparent K_m value and its temperature stability (Gauer *et al.*, 2014).

*Bj*SDH is part of the superfamily of Zn-independent short-chain dehydrogenases (SDRs; Kallberg *et al.*, 2010), to which about 25% of all dehydrogenases belong. This enzyme belongs to the classic subfamily of SDRs, as it includes the cofactor-binding motif TG*xxx*[AG]*x*G and the active-site residues Y*xxx*K (Kavanagh *et al.*, 2008). There are many structurally determined homologues of the sorbitol dehydrogenases, but the closest in sequence to *Bj*SDH are only 33–35% identical in sequence. Most of them exist in a tetrameric form in solution (Philippsen *et al.*, 2005). The pyridine nucleotides NAD(H) and NADP(H) always act as soluble cofactors, *i.e.* they are not covalently bound to the enzyme (Wang *et al.*, 2012).

2. Materials and methods

2.1. Macromolecule production

*Bj*SDH had previously been cloned from *B. japonicum* USDA110 chromosomal DNA and expressed in *Escherichia coli*. Briefly, the gene was cloned into pET-24a(+) from chromosomal DNA and an N-terminal His_6 tag was included (Table 1). The gene construct was expressed in *E. coli* BL21-Gold (DE3) and after an initial step of heat treatment at 50°C (Gauer *et al.*, 2014) the protein was purified for structural studies by metal-affinity chromatography.

Rhodobacter sphaeroides sorbitol dehydrogenase (*Rs*SDH; Philippsen *et al.*, 2005) was produced for thermostability studies in a similar way. A His-tagged construct, N-His_6-*Rs*SDH, containing the corresponding gene in the expression vector pET-24a(+) (Novagen), was expressed in *E. coli* BL21-Gold (DE3). The crude extract was applied onto a HiTrap FF Crude column (GE Healthcare, column volume 5 ml) equilibrated with binding buffer (20 m*M* NaH_2PO_4, 2 m*M* NaCl, 20 m*M* imidazole pH 7.4). After washing the column with two column volumes of binding buffer to eliminate unbound material, His-tagged dehydrogenase was eluted with elution buffer (20 m*M* NaH_2PO_4, 500 m*M* NaCl, 500 m*M* imidazole pH 7.2). Fractions containing the enzyme were pooled and loaded onto a Desalting XK 26/20 G25 column (GE Healthcare), which was equilibrated with 20 m*M* Tris–HCl buffer pH 7. The enzyme was eluted using the same buffer.

Table 2
Data collection, processing and refinement.

Values in parentheses are for the outer shell.

Data collection	
Diffraction source	ID29-1, ESRF
Wavelength (Å)	0.8999
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	375.13
Rotation range per image (°)	0.5
Total rotation range (°)	100
Exposure time per image (s)	0.2
Space group	$P6_222$
a, *b*, *c* (Å)	100.92, 100.92, 88.77
α, β, γ (°)	90, 90, 120
Mosaicity (°)	0.334
Resolution range (Å)	43.87–2.90 (3.07–2.90)
Total No. of reflections	64896
No. of unique reflections	6287
Completeness (%)	99.5 (98.0)
Multiplicity	10.32
$\langle I/\sigma(I)\rangle$	12.4 (1.14)
$CC_{1/2}$	99.6 (46.3)
R_{meas} (%)	24.7 (248.5)
Overall *B* factor from Wilson plot (Å^2)	71.82
Refinement	
Final R_{cryst} (%)	27.00
Final R_{free} (%)	29.51
No. of non-H atoms	
Total	1817
Protein	1800
Ligand	17
R.m.s. deviations	
Bonds (Å)	0.004
Angles (°)	0.765
Average *B* factors (Å^2)	
Overall	68.05
Protein	68.10
Ligand	63.11
Ramachandran plot	
Favoured regions (%)	88.6
Additionally allowed (%)	9.3
Outliers (%)	2.1

2.2. Crystallization and data collection

The tagged *Bj*SDH was purified to homogeneity as judged by SDS–PAGE. *Bj*SDH was dialyzed three times against 133 m*M* Tris–HCl pH 8.5 and concentrated to 21 mg ml^{-1} using ultrafiltration cellulose membranes (Amicon). Final concentrations of 100 m*M* D-glucitol and 0.4 m*M* NAD^+ were added to the protein solution, which was subsequently spun for 10 min at 9500 RCF to remove higher molecular-weight aggregates. The supernatant was used to set up MRC 2-well sitting-drop trays with an Oryx 8 liquid-handling robot (Douglas Instruments). Initially, 0.3 µl drops (with protein solution:reservoir solution ratios of 1:1 and 3:1) were set up in Index HT (from Hampton Research), PACT and JCSG+ (both from Molecular Dimensions) screens and revealed crystalline material in 1.4 *M* NaH_2PO_4/K_2HPO_4 pH 5.6 (Index HT condition B5). 2 µl sitting drops were then set up manually in

a 1:1 ratio at 10 mg ml^{-1} protein concentration, varying the precipitant concentration. Microcrystals were obtained in 1.8 *M* NaH_2PO_4/K_2HPO_4 pH 5.6 and were used to streak-seed protein solution and 1.4 *M* NaH_2PO_4/K_2HPO_4 pH 5.6 in a 1:1 ratio. Disc-shaped crystals grew within six months. They were cryoprotected with 26.1% glycerol in reservoir solution, mounted in nylon loops and flash-cooled in liquid nitrogen.

Data were collected on ESRF beamline ID29 (de Sanctis *et al.*, 2012) at a wavelength of 0.9 Å and a temperature of 100 K. A total of 250° of data were collected over 500 images. Processing and scaling were performed with *XDS* (Kabsch, 2010) using *XDSAPP* (Krug *et al.*, 2012). The data could be processed in the hexagonal space groups $P6_222$ or $P6_422$. Because of visible radiation damage, only the first 100° of data were used as the final data set, but because of the high symmetry a complete and redundant data set was obtained nonetheless (Table 2). A resolution limit of 2.90 Å was chosen according to the criterion of a $CC_{1/2}$ value (Karplus & Diederichs, 2012) of around 50% in the outer resolution shell. A criterion of $I/\sigma(I) > 2$ in the outer resolution shell would result in a resolution limit of 3.15 Å.

2.3. Structure determination and refinement

Using *BLASTP* to perform a search of the Protein Data Bank with the *Bj*SDH sequence identified the closest relative to be FabG from *Bacillus* sp. strain SG-1 (PDB entry 4nbu; Javidpour *et al.*, 2014), with 33% sequence identity to *Bj*SDH. Using the *A* monomer, a model for molecular replacement was sculpted using *phenix.sculptor* with the *Bj*SDH sequence as input (Bunkóczi & Read, 2011). The model was loaded into *PyMOL* (v.1.8; Schrödinger) together with PDB entries 2jap and 3ai1 (the second and third hits in the *BLASTP* search) and then pruned manually by removing the most variable regions (identified visually as residues 36–48, 191–208 and 238–242; *Bj*SDH numbering). The resulting coordinates were used as a search model in molecular replacement with *phenix.automr* (McCoy *et al.*, 2007), resulting in a solution in

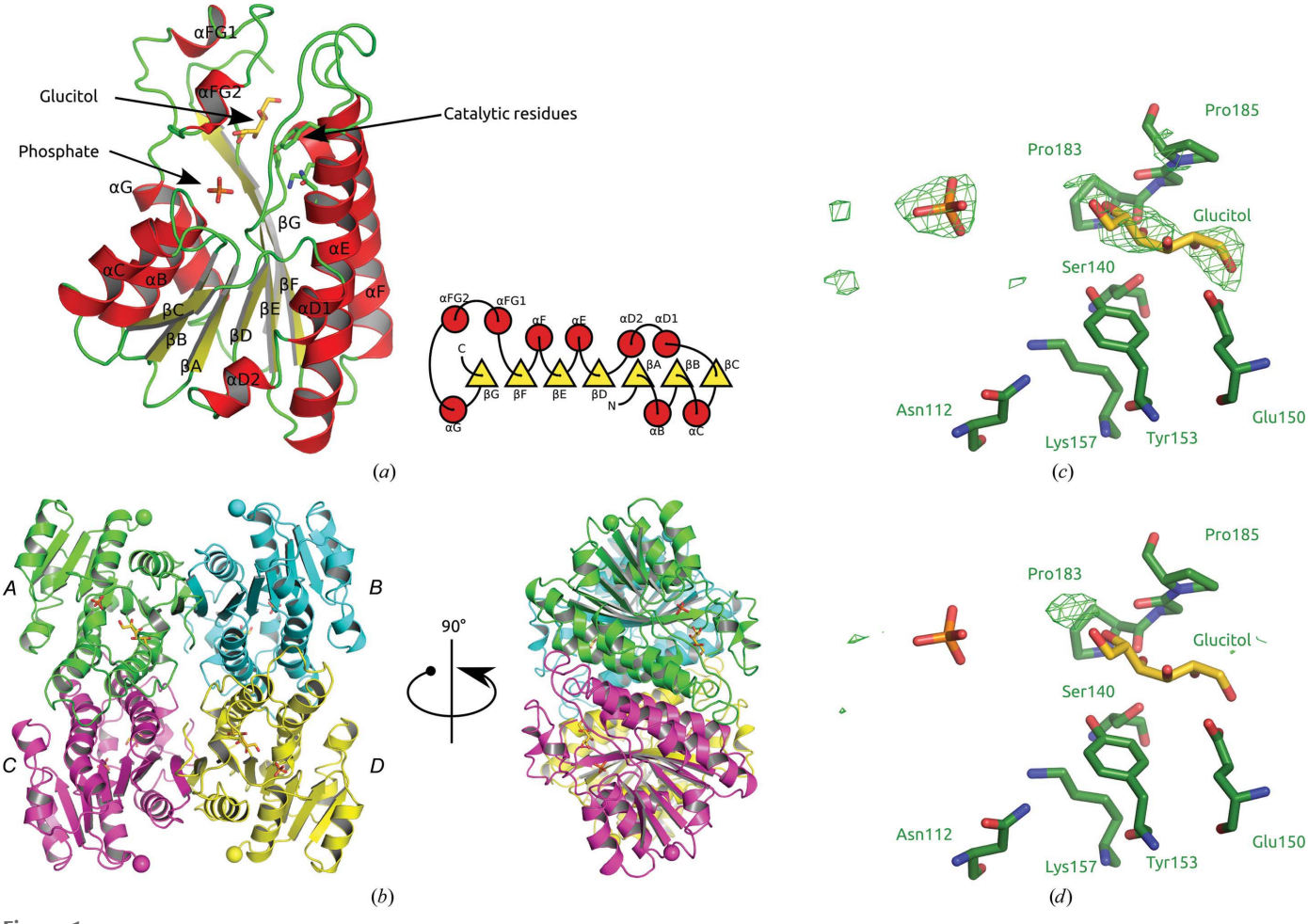
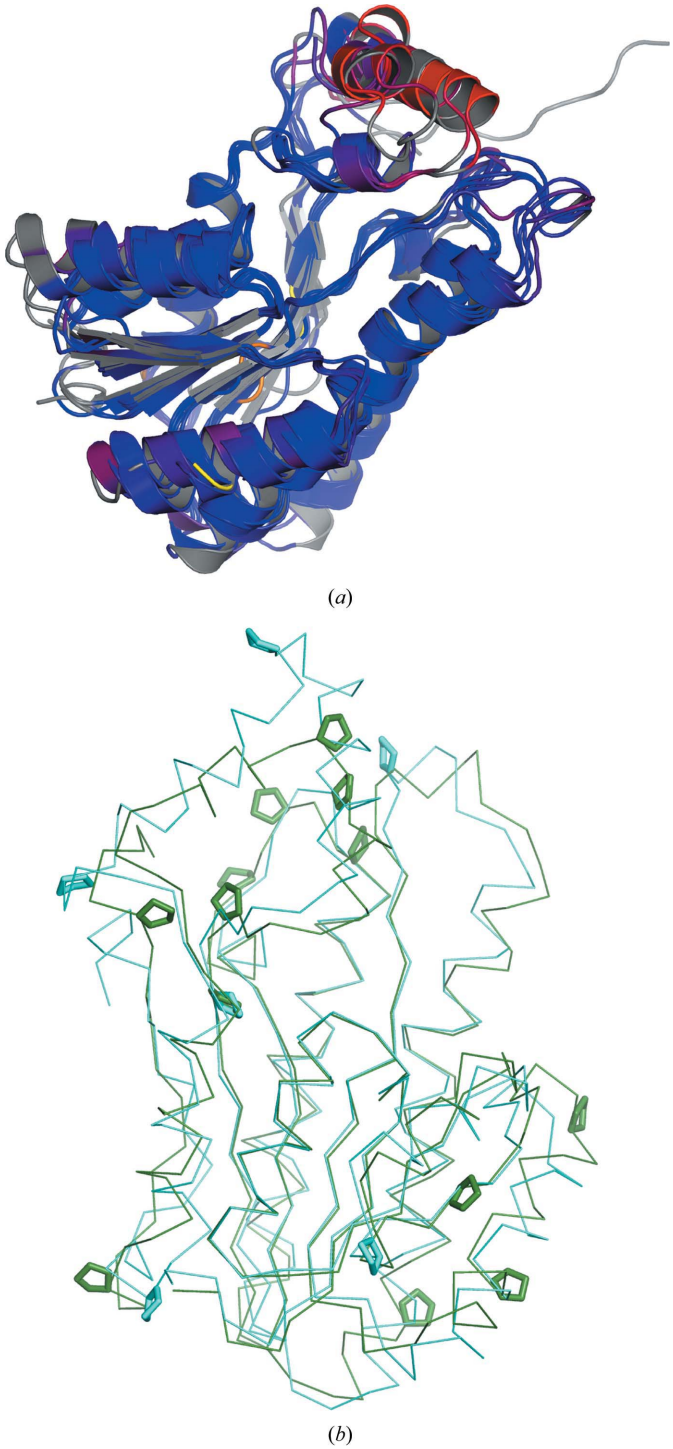


Figure 1
(*a*) Cartoon representation of the monomer of *Bj*SDH, coloured according to secondary structure (red, helices; yellow, β-strands; green, loop regions). The seven β-strands βA–βG form a parallel β-sheet surrounded by α-helices. A longer insert between βF and αG contains a small helix–turn–helix motif serving as a 'lid' to close the active site. Glucitol, phosphate and the residues involved in catalysis are shown in stick representation. (*b*) Cartoon representation of the tetramer of *Bj*SHD. The seven-stranded β-sheet can be seen as it continues between monomers (teal/green and pink/yellow). (*c*) The active-site catalytic residues with $mF_{obs} - DF_c$ difference OMIT density at +3σ shown in green. The map is from before ligands were included in the model (*d*) The residual $mF_{obs} - DF_c$ electron-density map shown in green at +3σ after the inclusion of glucitol and phosphate in the model.

space group $P6_222$ with a final LLG of 135.4 and a TFZ of 14.3 (a value of >8 indicates a clear solution). After one round of refinement with *phenix.refine* (Afonine *et al.*, 2012) initial R_{work} and R_{free} values of 43 and 48%, respectively, were obtained, further indicating a correct solution. Several rounds of refinement and careful manual model building using *Coot* (Emsley *et al.*, 2010) gave a final model (Fig. 1) including residues 4–242, one molecule of glucitol and one phosphate ion, with a final R_{work} and R_{free} of 27.0 and 29.5%, respectively (see Table 2 for further statistics). Glucitol and one phosphate ion were included in the model based on the difference density at 3σ in a refinement map (shown in Fig. 1*c*). No water molecules have been included in the model owing to the low resolution. Five residues (Cys80, Ile105, Arg131, Leu199 and Ser207) were flagged as Ramachandran outliers in the PDB validation report. The model was checked against a composite OMIT map calculated with *phenix.composite_omit_map* using standard parameters.



(*a*)

(*b*)

Figure 2
(*a*) Comparison of the top *DALI* hits, coloured from low (blue) to high (red) r.m.s.d. compared with the *Bj*SDH structure. Residues with very high r.m.s.d. were removed from the alignment and are coloured grey. *THESEUS* was used to align the structures and *PyMOL* and the `colorbyrmsd.py` community script were used to calculate the r.m.s.d. (*b*) Comparison of proline residues in the structures of *Bj*SDH (green) and *Rs*SDH (PDB entry 1k2w; cyan).

2.4. Structure analysis and representation

The *DALI* server (Holm & Rosenström, 2010) was used to identify the closest structural relatives to the *Bj*SDH structure. Each chain in a PDB file is included in the server database, so the final list was filtered for identical chains to obtain the top five most similar structures. The quaternary assembly of monomers was analysed with *Proteins, Interfaces, Structures and Assemblies* (*PISA*) through *qtPISA* (Krissinel, 2015) from the *CCP*4 suite of programs (Winn *et al.*, 2011). *THESEUS* was used to superpose structures (Theobald & Steindel, 2012). *PyMOL* was used to generate the representations of the analysed structures and *Inkscape* (http://inkscape.org; v.0.48) was used to annotate and arrange the figures as well as to draw the structural diagram.

2.5. Circular-dichroism measurements

A J-815 CD spectrometer (Jasco) equipped with a Jasco CDF-426S/15 temperature-control unit was used to record CD spectra and denaturation curves for *Rs*SDH. Owing to some precipitation, all samples were centrifuged (10 min, 9500 RFC) before use. The clear supernatant was used for CD measurements. Spectra were recorded in water and diluted buffer (*Rs*SDH at 1 m*M* Tris–HCl pH 7.0). Quartz glass far-UV and near-UV CD cuvettes (Hellma) with light paths of 1 mm were used. The denaturation curves were recorded at a single wavelength at 222 nm. The temperature range was 4–60°C. The peak in the second numerical derivation was used to estimate T_m after fitting of the sigmoidal curves in *Excel* (Microsoft).

3. Results and discussion

3.1. Crystal structure

The *Bj*SDH structure folds in the Rossmann fold (Rao & Rossmann, 1973), which is typical for NADH-dependent enzymes and the SDR family (Fig. 1*a*). The monomer of *Bj*SDH is composed of a central parallel β-sheet surrounded by three α-helices on each side. At the top of the β-sheet a small helix–turn–helix motif is located between the two final β-strands βF and βG (see Fig. 1*a*). The conserved catalytic

residues are located at the beginning of αE (Asn112), just after βE (Ser140) and at the beginning of αF (Tyr153 and Lys157). The glucitol is primarily bound by the helix–turn–helix motif, which acts as a lid on top of the structure.

As *Bj*SDH was crystallized with both NAD^+ and glucitol, the active site was expected to include both of these ligands. After carefully modelling the glucitol molecule into the density corresponding to its most likely location (Fig. 1*c*), residual electron density was still visible, some of which could be modelled as a phosphate ion. However, the remaining density does not warrant the modelling of a full NAD molecule. Residual electron density in the active site after the modelling of glucitol and phosphate is shown in Fig. 1(*d*).

3.2. Most similar structures (*DALI*)

To identify the closest structural relatives to *Bj*SDH, the *DALI* server (Holm & Rosenström, 2010) was used, and the top hits were for clavulanic acid dehydrogenase (MacKenzie *et al.*, 2007) from *Streptomyces clavuligerus* (sequence identity 33%), with a *Z*-score range from 34.2 to 33.9 in the chains from PDB entries 2jah and 2jap based on 236 aligned C^α positions out of 245 with an r.m.s.d. of 1.4 Å. The second hit revealed a high similarity to a putative blue fluorescent protein from *Vibrio vulnificus* (Kao *et al.*, 2011; sequence identity 31%), with a *Z*-score range from 32.5 to 31.3 in the chains from PDB entry 3p19 based on 234 aligned C^α positions out of 239 with an r.m.s.d. of 1.7 Å for the best hit. The third hit was a human putative dehydrogenase (PDB entry 1xg5; Structural Genomics Consortium, unpublished work), with a sequence identity of 28%, a *Z*-score from 31.5 to 30.5 and an r.m.s.d. of 2.0 Å based on 235 aligned C^α positions out of 257. A serine dehydrogenase from *E. coli* was the fourth hit (Yamazawa *et al.*, 2011), with a sequence identity of 26% and a *Z*-score between 31.3 and 31.0 over six chains in PDB entry 3asv, giving an r.m.s.d. of 2.1 Å based on 234 aligned C^α positions out of 248. The model used for molecular replacement (Javidpour *et al.*, 2014) is the fifth hit on the *DALI* list (sequence identity 29%), with a *Z*-score range from 30.9 to 30.7 in the chains from PDB entry 4nbu based on 237 aligned C^α positions out of 242 with an r.m.s.d. of 2.2 Å for the best hit.

Evidently the list of closest relatives represents a large sequence diversity of enzymes based on the Rossmann fold that use either NADH or NADPH as the cofactor. Fig. 2(*a*) shows the five most closely related structures, coloured from blue (low r.m.s.d.) to red (high r.m.s.d.), with non-aligned residues coloured grey. As can be seen from this figure, the helix–turn–helix motif in the loop between βF and αG is the most diverse part of the structure, with the rest of the Rossmann fold very well structurally conserved (at the C^α level).

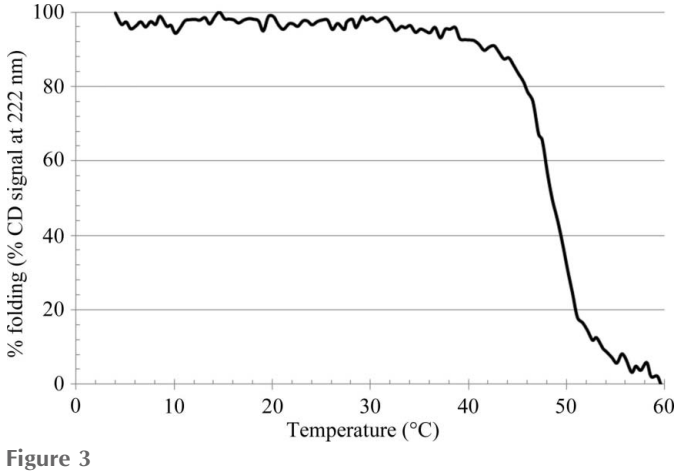


Figure 3
Heat-denaturation curve of *Rs*SDH as followed by CD spectroscopy at 222 nm

3.3. Quaternary structure

Even though the asymmetric unit is limited to one monomer chain, *Bj*SDH is a tetrameric enzyme in the crystal, with the tetramer formed by crystallographic symmetry. This is evident, for example, from *PISA* analysis of the structure (Krissinel, 2015). According to *PISA*, the tetrameric assembly has a calculated free energy of dissociation of 31.2 kcal mol^{-1}, with a total buried surface area of 13 000 $Å^2$ (3233 $Å^2$ per monomer). The tetrameric quaternary organization formed in the crystal (222 point-group symmetry) is the quaternary organization also found in the five best *DALI* hits above. In *Bj*SDH the primary interactions between monomers are hydrogen bonds between the β-sheet from the Rossmann fold (between monomers *A* and *B* and monomers *C* and *D* in Fig. 1*b*), which effectively form one continuous sheet, and interactions between helices in two monomers to form a shared four-helix bundle (between monomers *A* and *C* and monomers *B* and *D* in Fig. 1*b*). The C-terminal part of *Bj*SDH, starting with Pro237 at the end of βG, is tightly coordinated by arginines 168, 169 and 231 from a symmetry-related monomer (in pairs *A*/*B* and *C*/*D*; see Fig. 1*b*), which also coordinate the C-terminal carboxyl group of Leu242.

Previously, characterization of *Bj*SDH by native PAGE and gel filtration showed molecular masses of 77.7 and 83.2 kDa (Gauer *et al.*, 2014). Given the monomer molecular mass of 26.9 kDa, this would suggest a trimer in solution, but taken together with the tetrameric state in the crystal resembling the quaternary structure of sequence-related enzymes, a tetrameric solution structure seems more likely. Some ambiguity in the solution structure was also found for a previously structurally characterized bacterial sorbitol dehydrogenase from *R. spaeroides* (Philippsen *et al.*, 2005), which was reported to be dimeric by sedimentation-centrifugation analysis but forms a tetramer in the crystal.

A dimeric constitution has been reported by analytical ultracentrifugation (Kao *et al.*, 2011) for the blue fluorescent protein from *V. vulnificus*, one of the *DALI* hits above, where mutations in αE and αF could disrupt the dimer, indicating that this shared four-helix bundle was the interface for the observed dimer. For *Bj*SDH the interaction through βG is more pronounced, suggesting that this particular enzyme is a tetramer in solution.

3.4. Thermostability

*Bj*SDH has previously been reported (Gauer *et al.*, 2014) to have a T_m of 62°C as measured by CD at 222 nm. Furthermore, it was shown that the enzyme activity was stable for more than 1 h at 55°C. This thermostability is favourable for its use as an industrial enzyme. In order to have another enzyme for detailed comparison, *Rs*SDH, a previously structurally characterized sorbitol dehydrogenase (Philippsen *et al.*, 2005), was here characterized in terms of melting temperature by CD spectroscopy, as performed for *Bj*SDH in Gauer *et al.* (2014), and was found to have a T_m of 47°C (Fig. 3). Over the years, many structural determinants of thermostability in proteins have been identified, for example an increased number of hydrogen bonds/salt bridges, increased rigidity and increased secondary-structure content, to name a few, forming the overall view that a number of stabilization strategies exist (Petsko, 2001). In order to find the structural source of the added thermostability, the sequences and structures of *Rs*SDH and *Bj*SDH were compared with respect to some of the known thermostabilization features.

We found that the *Bj*SDH sequence contains roughly twice as many proline residues (13 in *Bj*SDH *versus* 6 in *Rs*SDH), but only half the number of glycine residues (15 in *Bj*SDH *versus* 27 in *Rs*SDH), which could explain some of the stability differences. See Fig. 2(*b*) for a comparison of the positions of prolines in *Bj*SDH and *Rs*SDH.

As mentioned in §3.3, in the *Bj*SDH structure the parallel β-sheets of two monomers (in pairs) interact with each other to effectively form one continuous intermolecular β-sheet. The sheets from each monomer are arranged in an antiparallel fashion with respect to each other and form a total of 22 hydrogen bonds. Corresponding interactions are not found in *Rs*SDH; thus, the formation of this continuous sheet in *Bj*SDH might contribute to the tetramer stability. On the other hand, the number of intramolecular hydrogen bonds and salt bridges is comparable between *Bj*SDH and *Rs*SDH.

Since the Pro/Gly content seems to be one of the most likely determinants of thermostability, this factor was also analysed for the top five *DALI* hits, and *Bj*SDH was found to have the highest Pro/Gly ratio of 0.87 among the analyzed proteins, compared with ratios between 0.26 and 0.6 (for the blue fluorescent protein).

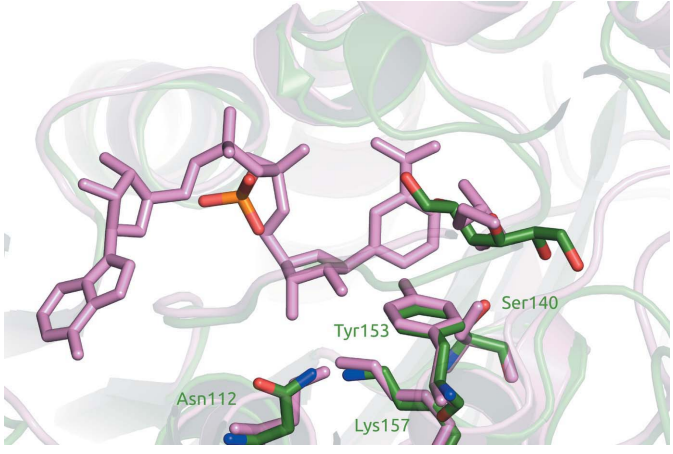


Figure 4
Active site of *Bj*SDH compared with the structure of GatDH (Carius *et al.*, 2010; PDB entry 2wdz; pink).

3.5. Specificity and mechanism

B. japonicum SDH showed a 90% conversion of L-glucitol to D-sorbose, which makes it interesting for the conversion of sugar alcohols (Gauer *et al.*, 2014). To the best of our knowledge, only a few other L-glucitol-acting dehydrogenases have been structurally characterized, and none with comparable properties. The new thermostable *Bj*SDH structure is an important step towards the understanding and engineering of the substrate specificity and thermostability of the sorbitol dehydrogenases.

The amino acids building the catalytic tetrad are conserved throughout the SDR family. A structural comparison of the active site of *Bj*SDH with those of galactitol dehydrogenase (GatDH; PDB entry 2wdz; Carius *et al.*, 2010) and sorbitol dehydrogenase from *R. sphaeroides* (PDB entry 1k2w; Philippsen *et al.*, 2005) shows that the conserved amino acids forming the catalytic tetrad (Asn112, Ser140, Tyr153 and Lys157 in *Bj*SDH) superpose well (Fig. 4). The substrate in *Bj*SDH is stabilized *via* the side chain of Glu150, Ser140 and the carbonyl groups of Pro183 and Gly184, leading to a 90° orientation compared with the pentanediol bound to GatDH (Fig. 4). The most significant differences are in the loop between βE and αF and the helix–turn–helix motif between the βF and αG loop, containing the prolines Pro195 and Pro196, which closes the active site to the environment in this structure. Despite the addition of NAD, only a phosphate group is observable in the density of the *Bj*SDH structure. The position of this phosphate is closest to the phosphate adjacent to the nicotinamide group in the GatDH structure (Fig. 4).

If NAD or NADH bound to *Bj*SDH exactly in the way it does to GatDH, the distance between the nicotinamide ring and the substrate as oriented in the described structure would be too short (Fig. 4). As the bound phosphate is slightly shifted in *Bj*SDH compared with the corresponding phosphate of the NAD bound in the compared dehydrogenase structures, this could indicate a slightly different position of the cofactor in *Bj*SDH. A reorientation of the substrate in the presence of the cofactor or flexibility of the NAD might be also possible: the latter has been described, for example, for the sheep liver cytosolic aldehyde dehydrogenase (Moore *et al.*, 1998).

However, the limited resolution of the structure and the lack of bound NAD do not allow us to establish with confidence the structural features that allow this enzyme to accept L-glucitol to some extent as a substrate. A higher resolution structure with cofactor bound would definitely be desirable.

Acknowledgements

We acknowledge Dorthe Boelskifte for help with laboratory work, Jose Antonio Cuesta-Seijo for data collection, DANSCATT for the funding of travel to synchrotrons, Morten J. Bjerrum for use of the circular-dichroism instrument and the

ESRF for synchrotron beamtime. This work was financially supported by the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. NMP3-SL-2008-213487.

References

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.

Baek, S. H., Park, S. J. & Lee, H. G. (2010). *J. Food Sci.* **75**, H49–H53.

Beerens, K., Desmet, T. & Soetaert, W. (2012). *J. Ind. Microbiol. Biotechnol.* **39**, 823–834.

Bunkóczi, G. & Read, R. J. (2011). *Acta Cryst.* D**67**, 303–312.

Carius, Y., Christian, H., Faust, A., Zander, U., Klink, B. U., Kornberger, P., Kohring, G.-W., Giffhorn, F. & Scheidig, A. J. (2010). *J. Biol. Chem.* **285**, 20006–20014.

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

Gauer, S., Wang, Z., Otten, H., Etienne, M., Bjerrum, M. J., Lo Leggio, L., Walcarius, A., Giffhorn, F. & Kohring, G. W. (2014). *Appl. Microbiol. Biotechnol.* **98**, 3023–3032.

Granström, T. B., Takata, G., Tokuda, M. & Izumori, K. (2004). *J. Biosci. Bioeng.* **97**, 89–94.

Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.

Huwig, A., Emmel, S. & Giffhorn, F. (1996). *Carbohydr. Res.* **281**, 183–186.

Javidpour, P., Pereira, J. H., Goh, E.-B., McAndrew, R. P., Ma, S. M., Friedland, G. D., Keasling, J. D., Chhabra, S. R., Adams, P. D. & Beller, H. R. (2014). *Appl. Environ. Microbiol.* **80**, 497–505.

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

Kallberg, Y., Oppermann, U. & Persson, B. (2010). *FEBS J.* **277**, 2375–2386.

Kao, T.-H., Chen, Y., Pai, C.-H., Chang, M.-C. & Wang, A. H.-J. (2011). *J. Struct. Biol.* **174**, 485–493.

Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.

Kavanagh, K. L., Jörnvall, H., Persson, B. & Oppermann, U. (2008). *Cell. Mol. Life Sci.* **65**, 3895–3906.

Krissinel, E. (2015). *Nucleic Acids Res.* **43**, W314–W319.

Krug, M., Weiss, M. S., Heinemann, U. & Mueller, U. (2012). *J. Appl. Cryst.* **45**, 568–572.

Lim, Y.-R. & Oh, D.-K. (2011). *Appl. Microbiol. Biotechnol.* **91**, 229–235.

Lu, Y., Levin, G. V. & Donner, T. W. (2008). *Diabetes Obes. Metab.* **10**, 109–134.

MacKenzie, A. K., Kershaw, N. J., Hernandez, H., Robinson, C. V., Schofield, C. J. & Andersson, I. (2007). *Biochemistry*, **46**, 1523–1533.

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

Moore, S., Baker, H. M., Blythe, T. J., Kitson, K. E., Kitson, T. M. & Baker, E. N. (1998). *Structure*, **6**, 1541–1551.

Petsko, G. A. (2001). *Methods Enzymol.* **334**, 469–478.

Philippsen, A., Schirmer, T., Stein, M. A., Giffhorn, F. & Stetefeld, J. (2005). *Acta Cryst.* D**61**, 374–379.

Rao, S. T. & Rossmann, M. G. (1973). *J. Mol. Biol.* **76**, 241–256.

Sanctis, D. de *et al.* (2012). *J. Synchrotron Rad.* **19**, 455–461.

Theobald, D. L. & Steindel, P. A. (2012). *Bioinformatics*, **28**, 1972–1979.

Wang, Z., Etienne, M., Quilès, F., Kohring, G. W. & Walcarius, A. (2012). *Biosens. Bioelectron.* **32**, 111–117.

Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

Yamazawa, R., Nakajima, Y., Mushiake, K., Yoshimoto, T. & Ito, K. (2011). *J. Biochem.* **149**, 701–712.