Acta Crystallographica Section D
Biological Crystallography
ISSN 1399-0047

Structural insight into glucose dehydrogenase from the thermoacidophilic archaeon *Thermoplasma volcanium*

Yoshitaka Kanoh,[a] Seiichiroh Uehara,[a] Hideyuki Iwata,[b] Kazunari Yoneda,[c] Toshihisa Ohshima[d] and Haruhiko Sakuraba[a]*

[a]Department of Applied Biological Science, Faculty of Agriculture, Kagawa University, 2393 Ikenobe, Miki-cho, Kita-gun, Kagawa 761-0795, Japan, [b]Thermostable Enzyme Laboratory, 5-5-2 Minatojima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan, [c]Department of Bioscience, School of Agriculture, Tokai University, Aso, Kumamoto 869-1404, Japan, and [d]Department of Biomedical Engineering, Osaka Institute of Technology, 5-16-1 Ohmiya, Asahi-ku, Osaka 535-8585, Japan

Correspondence e-mail: sakuraba@ag.kagawa-u.ac.jp

Received 22 October 2013
Accepted 1 February 2014

PDB references: tvGlcDH, 3wic; T277F mutant, complex with NADP, 3wid; T277F mutant, complex with NAADP and D-glucose, 3wie

Glucose dehydrogenase from the thermoacidophilic archaeon *Thermoplasma volcanium* (tvGlcDH) is highly active towards D-glucose and D-galactose, but does not utilize aldopentoses such as D-xylose as substrates. In the present study, the crystal structures of substrate/cofactor-free tvGlcDH and of a tvGlcDH T277F mutant in a binary complex with NADP and in a ternary complex with D-glucose and nicotinic acid adenine dinucleotide phosphate, an NADP analogue, were determined at resolutions of 2.6, 2.25 and 2.33 Å, respectively. The overall structure of each monomer showed notable similarity to that of the enzyme from *Sulfolobus solfataricus* (ssGlcDH-1), which accepts a broad range of C5 and C6 sugars as substrates. However, the amino-acid residues of tvGlcDH involved in substrate binding markedly differed from those of ssGlcDH-1. Structural comparison revealed that a decreased number of interactions between the C3-hydroxyl group of the sugar and the enzyme are likely to be responsible for the lack of reactivity of tvGlcDH towards D-xylose.

1. Introduction

Glucose dehydrogenase (GlcDH) catalyzes the oxidation of glucose to gluconate using NAD or NADP as a cofactor. To date, two types of GlcDH have been identified. The first type is classified into the short-chain dehydrogenase reductase (SDR) family. SDR-type GlcDHs are only found in prokaryotes and insects (Jörnvall *et al.*, 1981, 1995). The active form of SDRs is either a tetramer or a dimer, and each subunit typically consists of about 250 amino-acid residues (Jörnvall *et al.*, 1981). The second type of GlcDH occurs in archaea and belongs to the medium-chain dehydrogenase reductase (MDR) family. MDR family enzymes possess somewhat larger subunits (about 360 amino-acid residues) than SDRs and contain an active-site Zn^{2+} ion (Edwards *et al.*, 1996; Jörnvall *et al.*, 1987). The archaeal MDR GlcDH is known to be the first enzyme in a nonphosphorylated variant of the Entner–Doudoroff pathway which is utilized as the central sugar catabolic pathway by several thermoacidophilic archaea (Dandekar *et al.*, 1999), most notably *Sulfolobus solfataricus* (De Rosa *et al.*, 1984) and *Thermoplasma acidophilum* (Danson, 1989). GlcDH from *S. solfataricus* (ssGlcDH-1) shows activity toward a broad range of substrates, including D-glucose, D-galactose, D-xylose and L-arabinose (Lamble *et al.*, 2003). Moreover, because metabolites of both L-glucose and D-galactose can utilized by the same enzymes in the pathway, it was proposed that the modified Entner–Doudoroff pathway in *S. solfataricus* is promiscuous for the metabolism of D-glucose and D-galactose (Lamble *et al.*, 2003). Recently, a second GlcDH (ssGlcDH-2) was identified in *S. solfataricus* (Haferkamp *et al.*, 2011). Because ssGlcDH-2 is absolutely

specific for D-glucose, it is thought to be the major player in glucose catabolism *via* the modified Entner–Doudoroff pathway in *S. solfataricus*, whereas ssGlcDH-1 may have a dominant role in galactose degradation. Some MDR GlcDHs from other sources reportedly show substrate specificities that differ from those of ssGlcDH-1 and ssGlcDH-2. For example, a *Picrophilus torridus* enzyme also exhibits activity towards D-glucose and D-galactose, but does not utilize aldopentoses such as D-xylose, L-arabinose and D-ribose (Angelov *et al.*, 2005). On the other hand, a *S. tokodaii* enzyme (ST1074) shows weak activity towards D-galactose but utilizes D-xylose to a greater extent (Ohshima *et al.*, 2003). To date, the crystal structures of MDR GlcDHs from *T. acidophilum* (John *et al.*, 1994), *S. solfataricus* (ssGlcDH-1; Milburn *et al.*, 2006) and *Haloferax mediterranei* (Baker *et al.*, 2009; Britton *et al.*, 2006) have been reported. Extensive analysis of these structures has resolved the substrate-binding sites and led to the elucidation of the catalytic mechanism of these enzymes. However, little is known about the mechanism underlying the differences in substrate recognition among MDR GlcDHs because the structure of a substrate-bound MDR GlcDH that shows narrow substrate specificity has not been determined.

In the present study, we succeeded in resolving the crystal structure of the MDR GlcDH from the thermoacidophilic archaeon *T. volcanium* (tvGlcDH) in complex with D-glucose and nicotinic acid adenine dinucleotide phosphate (NAADP), an NADP analogue. The amino-acid sequence of tvGlcDH showed 34 and 41% identity to those of ssGlcDH-1 and ssGlcDH-2, respectively (Fig. 1), and the enzyme showed reactivity towards a strict range of substrates (D-glucose and D-galactose), as is the case for the *P. torridus* enzyme. A detailed structural analysis of tvGlcDH could shed light on the substrate-recognition mechanism of this enzyme, information that would provide further insight into the catalytic activity of MDR GlcDHs. We therefore compared the active-site architecture of tvGlcDH with that of ssGlcDH-1. We then evaluated the amino-acid residues that could be responsible for the difference in substrate specificity between tvGlcDH and ssGlcDH-1. This is the first description of the structure of a substrate-bound MDR GlcDH showing narrow substrate specificity.

2. Materials and methods

2.1. DNA manipulation

An expression vector (pTVGlcDH) encoding tvGlcDH without a His tag (ORF TVN1019) was a kind gift from the Thermostable Enzyme Laboratory, Kobe, Japan. An expression vector encoding C-terminally His-tagged tvGlcDH was constructed by amplifying the tvGlcDH gene fragment by PCR. The oligonucleotide primers used for the amplification were 5′-ATA**CATATG**TCCACTATAAACGCCATAGTAACTGATG-3′, which contains a unique *Nde*I restriction site (bold) overlapping the 5′ initiation codon, and 5′-GAA**CTCGAG**CGGCCATTTTATCACCGTCTTTATTTCA-3′, which contains a unique *Xho*I restriction site (bold) proximal to the 3′ end of the open reading frame. The expression vector pTVGlcDH served as the template. The amplified fragment was digested with *Nde*I and *Xho*I, and ligated with the

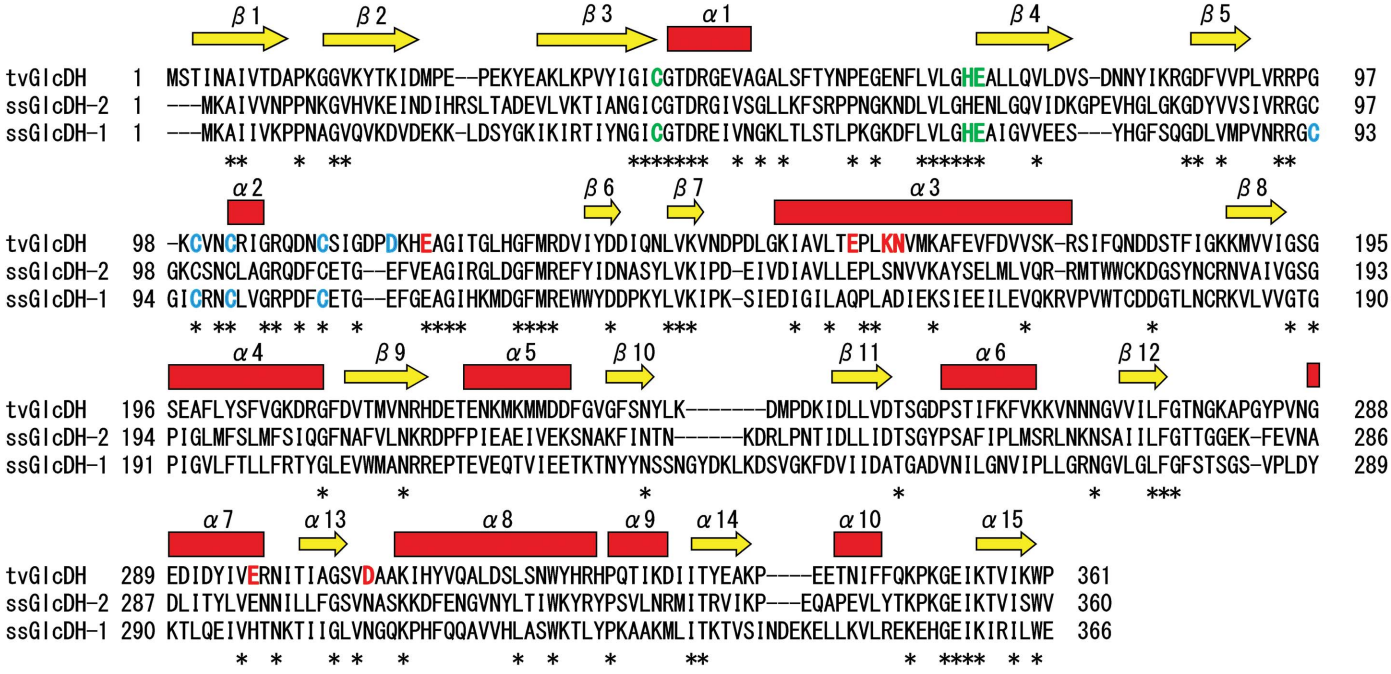


Figure 1
Amino-acid sequence alignment of tvGlcDH, ssGlcDH-2 and ssGlcDH-1. Sequences were aligned using *ClustalW* (Thompson *et al.*, 1994). Asterisks indicate conserved residues. The residues involved in glucose binding in tvGlcDH are shown in red. In the sequences of tvGlcDH and ssGlcDH-1, the residues involved in the binding of catalytic Zn and structural Zn are shown in green and blue, respectively. The secondary-structural assignments for tvGlcDH are shown above the alignment

Table 1
Primer sequences used for site-directed mutagenesis.

The mutations are shown in bold.

Mutation	Sequence
T277F	5′-GGTGTAGTGATACTGTTTGGA**TTC**AACGGAAAAGCCCCGGGATAT-3′
	5′-ATATCCCGGGGCTTTTCCGTT**GAA**TCCAAACAGTATCACTACACC-3′
V93N	5′-TTTCGTAGTACCTCTC**AAC**AGGAGGCCGGGCAAG-3′
	5′-CTTGCCCGGCCTCCT**GTT**GAGAGGTACTACGAAA-3′
E156Q	5′-CCTGATCTTGGAAAGATAGCCGTACTTACT**CAG**CCCCTTAAAAACGTT-3′
	5′-AACGTTTTTAAGGGG**CTG**AGTAAGTACGGCTATCTTTCCAAGATCAGG-3′
K159A	5′-CCGTACTTACTGAGCCCCTT**GCA**AACGTTATGAAAGCGTTCG-3′
	5′-CGAACGCTTTCATAACGTT**TGC**AAGGGGCTCAGTAAGTACGG-3′
N160D	5′-CGTACTTACTGAGCCCCTTAAA**GAC**GTTATGAAAGCGT-3′
	5′-ACGCTTTCATAAC**GTC**TTTAAGGGGCTCAGTAAGTACG-3′
E296H	5′-CGAGGACATTGATTATATCGTG**CAT**AGGAATATAACGATAGCTGGAT-3′
	5′-ATCCAGCTATCGTTATATTCCT**ATG**CACGATATAATCAATGTCCTCG-3′
D306N	5′-GATAGCTGGATCAGTG**AAC**GCCGCGAAGATACA-3′
	5′-TGTATCTTCGCGGC**GTT**CACTGATCCAGCTATC-3′
K159A/N160D	5′-ATAGCCGTACTTACTGAGCCCCTT**GCAGAC**GTTATGAAAGCGTTCG-3′
	5′-CGAACGCTTTCATAAC**GTCTGC**AAGGGGCTCAGTAAGTACGGCTAT-3′
E156Q/K159A/N160D	5′-CCTGATCTTGGAAAGATAGCCGTACTTACT**CAG**CCCCTT**GCAGAC**GTT-3′
	5′-AAC**GTCTGC**AAGGGG**CTG**AGTAAGTACGGCTATCTTTCCAAGATCAGG-3′

expression vector pET-21a (Novagen, Madison, Wisconsin, USA) previously linearized using *Nde*I and *Xho*I, yielding pETGlcDH.

Site-directed mutagenesis was accomplished using a QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, California, USA) according to the manufacturer's instructions. The expression vector (pETGlcDH) for His-tagged tvGlcDH served as the template. The mutagenic primers for T277F, V93N, E156Q, K159A, N160D, E296H, D306N and K159A/N160D mutants are listed in Table 1, and appropriate combinations of the mutations were employed for the construction of V93N/N160D, V93N/K159A, V93N/D306N, V93N/K159A/N160D and V93N/E156Q/K159A/N160D/D306N/E296H mutants.

2.2. Protein expression and purification

Untagged tvGlcDH expressed in *Escherichia coli* and purified by cation-exchange and gel-filtration column chromatography steps was a gift from the Thermostable Enzyme Laboratory, Kobe, Japan. To prepare His-tagged tvGlcDH, *E. coli* strain BL21(DE3) CodonPlus RIPL (Agilent Technologies) was transformed with pETGlcDH, after which the transformants were cultivated at 310 K in 1 l Luria–Bertani medium containing 100 µg ml^{-1} ampicillin until the optical density at 600 nm reached 0.6. Expression was then induced by adding 0.5 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) to the medium and cultivation was continued for an additional 20 h at 303 K. The cells were harvested by centrifugation, suspended in 50 m*M* potassium phosphate buffer pH 7.0 (buffer *A*) and disrupted by ultrasonication, after which the cell debris was removed by centrifugation (15 000*g* for 10 min). The resulting supernatant, which served as the crude extract, was heated at 323 K for 15 min and the denatured proteins were removed by centrifugation (15 000*g* for 10 min). The supernatant from this step was loaded onto a nickel-charged Chelating Sepharose Fast Flow column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) equilibrated with buffer *A*. The column was then washed with the same buffer, and the enzyme was eluted in stepwise fashion with 0.1, 0.3, 0.5 and 0.7 *M* imidazole in the buffer. The active fractions in the 0.5 and 0.7 *M* imidazole effluents were pooled, concentrated and loaded onto a Superdex 200 26/60 column (GE Healthcare) equilibrated with buffer *A* containing 0.2 *M* NaCl. The eluted active fractions were collected and dialyzed against buffer *A* and used for biochemical experiments. Expression and purification of the T277F and V93N mutants were performed using the same method as described for His-tagged tvGlcDH. To prepare the enzyme for crystallization, 10 m*M* Tris–HCl buffer pH 7.5 containing 0.2 *M* NaCl was used as the equilibration buffer in the final gel-filtration step. The eluted enzyme solution was desalted and concentrated by ultrafiltration (Amicon Ultra 30K NMWL) for crystallization trials.

2.3. Determination of enzyme activity and protein levels

The enzyme activity was assayed spectrophotometrically using a Shimadzu UV-mini 1240 spectrophotometer equipped with a thermostat. The standard reaction mixture consisted of 60 m*M* potassium phosphate buffer pH 6.5 containing 100 m*M* D-glucose, 2 m*M* NADP, 2 m*M* magnesium acetate and the enzyme in a final volume of 1.0 ml. After warming the reaction mixture by incubation for 2 min at 328 K without the enzyme, reaction was started by addition of the enzyme. The appearance of NADPH was monitored from the increase of absorbance at 340 nm (extinction coefficient ε = 6.22 mM^{-1} cm^{-1}). The protein concentration was determined using the Bradford method; bovine serum albumin served as the standard (Bradford, 1976).

2.4. Polyacrylamide gel electrophoresis and molecular-mass determination

SDS–PAGE (12.5% acrylamide slab gel, 1 mm thick) was carried out using the procedure of Laemmli (1970), after which the protein band was stained with Coomassie Brilliant Blue R-250. The molecular mass of the His-tagged tvGlcDH was determined using a Superose 6 10/300 GL column (GE Healthcare) with 50 m*M* potassium phosphate buffer pH 7.0 containing 0.2 *M* NaCl as the elution buffer. Gel-filtration standards (Bio-Rad, Hercules, California, USA) were used as molecular-mass standards. The subunit molecular mass was determined by SDS–PAGE using eight marker proteins (6–175 kDa; New England Biolabs).

2.5. Stability, pH optimum and kinetic parameters

To determine the effect of temperature on its stability, His-tagged tvGlcDH was incubated for 10 min at different temperatures in 50 m*M* potassium phosphate buffer pH 7.0. The residual activity was determined using the standard assay method. To determine the effect of pH on its stability, the enzyme was incubated for 10 min in buffer at different pH values (333 K) and the remaining activity was again determined using the standard assay method. The buffers (250 m*M*) used for these assays were citrate (pH 5.0–6.0), potassium phosphate (pH 6.0–8.5) and glycine–NaOH (pH 8.5–10.5). The same potassium phosphate buffer was also used to determine the optimal pH for enzyme activity at 328 K. To determine the kinetic parameters, the initial velocity was examined by varying the concentration of one substrate while keeping the concentrations of the other substrates constant, as described previously (Ohshima *et al.*, 1978).

2.6. Crystallization and data collection

For crystallization of the substrate/cofactor-free enzyme, purified untagged tvGlcDH in 10 m*M* Tris–HCl pH 7.5 was concentrated to 10 mg ml^{-1} by ultrafiltration (Amicon Ultra 30K NMWL). Crystallization was then performed using the sitting-drop vapour-diffusion method, in which 1 µl protein solution was mixed with an equal volume of reservoir solution [30% 1,2-propanediol, 20% polyethylene glycol 400 (PEG 400), 0.1 *M* 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid buffer (HEPES) pH 7.5]. Initially, a cluster of rod-shaped crystals were grown from the solution, but because these crystals were difficult to reproduce, they were crushed and suspended in the same solution and a new drop was inoculated with microseeds from the slurry. After several cycles of seeding, diffraction-quality crystals were obtained after two weeks at 293 K. Crystals of NADP-bound or NAADP-bound T277F mutant enzyme were grown in sitting drops composed of 1 µl enzyme solution (30 mg ml^{-1}) containing 1 m*M* NADP or NAADP mixed with an equal volume of the aforementioned reservoir solution. Diffraction-quality crystals appeared after one week at 293 K and were highly reproducible. Crystals of NAADP/D-glucose-bound T277F mutant enzyme were prepared by soaking NAADP-bound T277F mutant crystals for 24 h in reservoir solution containing 1.8 *M* D-glucose.

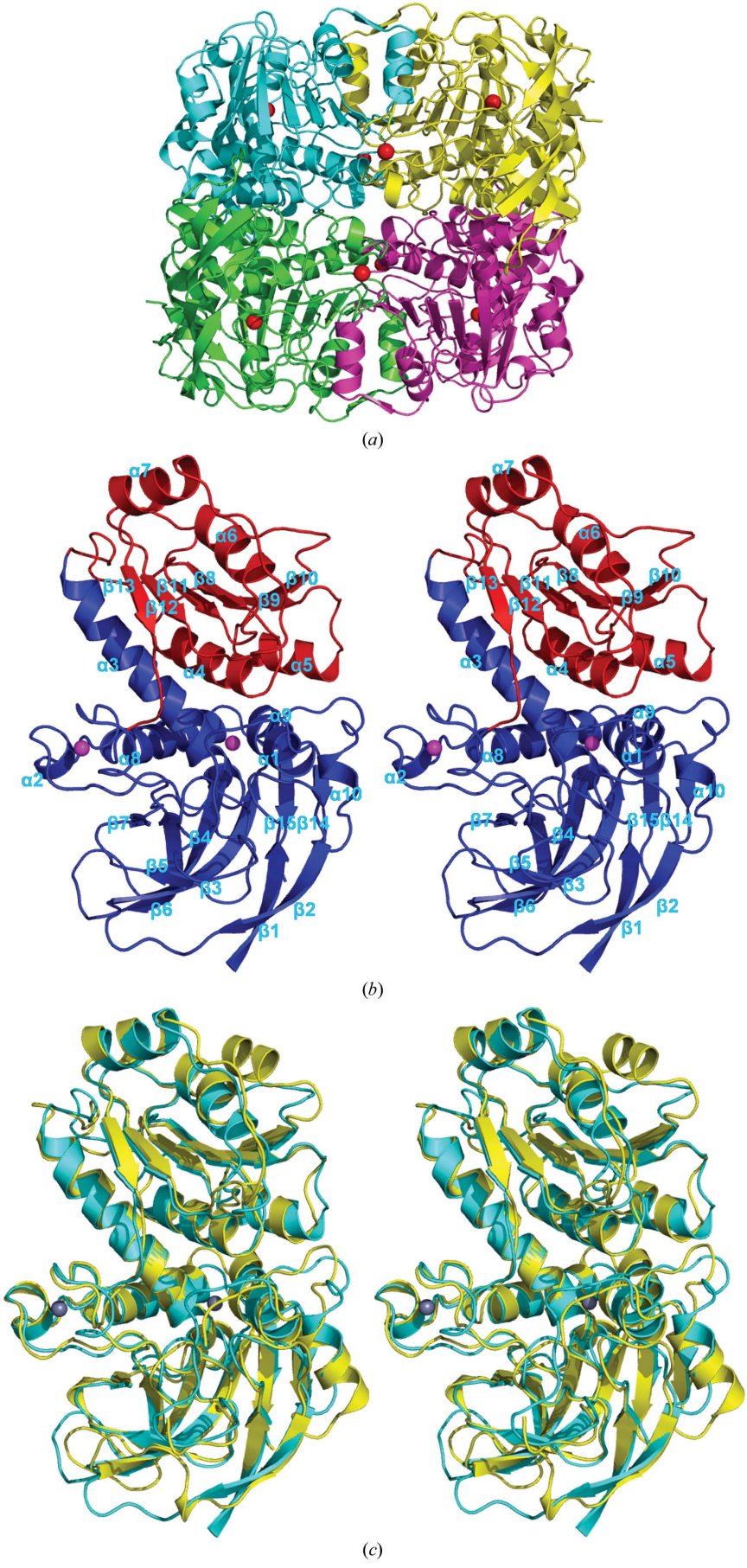
Diffraction data for substrate/cofactor-free enzyme crystals were collected using a Saturn A200 CCD detector system on beamline BL26B1 at SPring-8, Harima, Japan and those for NADP-bound T277F mutant and NAADP/D-glucose-bound T277F mutant crystals were collected using an ADSC CCD detector system on the AR-NW12 and AR-NE3A beamlines at the Photon Factory, Tsukuba, Japan. All diffraction measurements were carried out on crystals cryoprotected with Paratone-N mixed with paraffin oil (1:1) (Hampton Research, Aliso Viejo, California, USA) and cooled to 100 K in a stream of nitrogen gas. The data were processed using the *HKL*-2000 package (Otwinowski & Minor, 1997).

Table 2
Data-collection and refinement statistics for tvGlcDH.

Values in parentheses are for the highest resolution data shell.

	Substrate/cofactor-unbound tvGlcDH (PDB entry 3wic)	NADP-bound T277F mutant (PDB entry 3wid)	NAADP/D-glucose-bound T277F mutant (PDB entry 3wie)
Data collection			
Wavelength (Å)	1.0	1.0	1.0
Space group	$P2_1$	$P2_1$	$P2_1$
Unit-cell parameters			
a (Å)	79.1	79.2	79.9
b (Å)	122.2	90.3	92.3
c (Å)	87.4	120.4	120.5
β (°)	109.9	90.8	91.1
Resolution range (Å)	50–2.60 (2.64–2.60)	50–2.25 (2.29–2.25)	50–2.33 (2.37–2.33)
Total No. of reflections	339658	477199	523098
No. of unique reflections	48068	77124	75199
Multiplicity	7.1 (7.6)	6.2 (5.0)	7.0 (6.4)
Completeness (%)	99.9 (100)	95.4 (83.3)	99.2 (99.4)
$R_{\rm merge}$†	0.080 (0.268)	0.041 (0.233)	0.058 (0.298)
$\langle I/\sigma(I)\rangle$	10.9 (7.8)	32.0 (9.7)	21.8 (8.1)
Refinement			
Resolution range (Å)	50–2.60	50–2.25	50–2.33
$R/R_{\rm free}$‡ (%)	20.5/25.6 (18.1/26.9)	19.7/25.1 (27.9/35.8)	19.1/24.4 (23.2/31.4)
No. of protein atoms	171296	11290	11308
No. of water molecules	234	328	216
No. of ligands			
Zn^{2+} ions	8	4	8
1,2-Propanediol	6	4	—
NAADP	—	—	4
PEG 400	8	8	—
D-Glucose	—	—	4
NADP	—	4	—
B factor (Å^2)			
Protein	35.6	48.3	49.4
Zn^{2+} ion	30.0	35.4	54.3
Water	34.6	44.4	43.3
1,2-Propanediol	44.4	50.1	—
PEG 400	61.8	52.8	—
NADP/NAADP	—	56.2	66.2
D-Glucose	—	—	73.3
R.m.s.d.			
Bond lengths (Å)	0.014	0.017	0.020
Bond angles (°)	1.6	1.9	1.9
Ramachandran statistics (%)			
Favoured	96.2	96.3	96.5
Allowed	3.7	3.7	3.5
Outliers	0.1	0	0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the scaled intensity of the ith observation of reflection hkl, $\langle I(hkl)\rangle$ is the mean value and the summation is over all measurements. ‡ $R_{\rm free}$ calculated with randomly selected reflections (5%).



2.7. Phasing and refinement

The structure of the substrate/cofactor-free enzyme was solved to a resolution of 2.6 Å by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) in the *CCP*4 program suite (Winn *et al.*, 2011); the structure of chain *A* of ssGlcDH-1 (PDB entry 2cd9; Milburn *et al.*, 2006) served as the search model. In the final refined model, R = 0.205 ($R_{\rm free}$ = 0.256). The structure of the NADP-bound T277F mutant was solved to a resolution of 2.25 Å by molecular replacement using *MOLREP* (Vagin & Teplyakov, 2010) with the substrate/cofactor-free enzyme structure as a search model. In the final model, R = 0.198 ($R_{\rm free}$ = 0.249). Based on the model of the NADP-bound T277F mutant, the structure of the NAADP/D-glucose-bound T277F mutant was refined to a resolution of 2.33 Å using difference Fourier maps. In the final model, R = 0.191 ($R_{\rm free}$ = 0.244). In all cases, model building was performed using *Coot* (Emsley *et al.*, 2010) and refinement was carried out using *REFMAC*5 (Murshudov *et al.*, 2011) and *CNS* (Brünger *et al.*, 1998). Water molecules were incorporated using *Coot* (Emsley *et al.*, 2010) and model geometry was analyzed using *RAMPAGE* (Lovell *et al.*, 2003). Data-collection and refinement statistics are listed in Table 2. Molecular-graphics figures were prepared using *PyMOL* (http://www.pymol.org/).

3. Results and discussion

3.1. Gene expression and enzyme purification

After transforming *E. coli* cells with pETGlcDH, an expression vector encoding His-tagged tvGlcDH, the crude extract from the recombinant cells exhibited a high level of GlcDH activity and the enzyme was readily purified from the crude extract of the cells in three steps: heat treatment followed by two column-chromatography steps.

Figure 2
(*a*) Ribbon representation of the tvGlcDH tetramer, with subunits shown in different colours. Zn^{2+} ions are shown as red spheres. (*b*) Overall structure of the tvGlcDH monomer (stereoview). The nucleotide-binding domain (residues 175–307) and catalytic domain (residues 2–174 and 308–361) are shown in red and blue, respectively; Zn^{2+} is shown in magenta. (*c*) Superposition of the structures of tvGlcDH (cyan) and ssGlcDH-1 (yellow) (stereoview).

About 15 mg purified enzyme was obtained from 1 l *E. coli* culture. The purified enzyme gave a single protein band on SDS–PAGE and had a specific activity of about 498 µmol min^{-1} mg^{-1} at 328 K, which was comparable with that observed for *T. acidophilum* GlcDH (458 µmol min^{-1} mg^{-1} at 328 K; Bright *et al.*, 1993), but was higher than those reported for *P. torridus* GlcDH (252 µmol min^{-1} mg^{-1} at 328 K; Angelov *et al.*, 2005), ssGlcDH-1 (110 µmol min^{-1} mg^{-1} at 343 K using NAD as the cofactor) and ssGlcDH-2 (102 µmol min^{-1} mg^{-1} at 343 K using NAD as the cofactor) (Haferkamp *et al.*, 2011).

3.2. Molecular and catalytic properties of tvGlcDH

SDS–PAGE showed the subunit molecular mass of tvGlcDH to be about 42 kDa, which is consistent with the molecular mass (41 311 Da) calculated from the amino-acid sequence. The native molecular mass of about 158 kDa determined by gel filtration suggests that the native enzyme is a homotetramer. Evaluation of the catalytic activity at temperatures ranging from 313 to 353 K revealed the enzyme to be maximally active at around 343 K. The enzyme was also somewhat thermostable, retaining more than 90% of its activity after incubation for 10 min at temperatures of up to 333 K, although a substantial loss of activity was seen at temperatures above 343 K. Evaluation of the catalytic activity at different pH values revealed the enzyme to be maximally active at around pH 6.5. The enzyme lost no activity when incubated at pH values between 6.0 and 9.0 for 10 min at 333 K.

We next assessed the electron-donor specificity for oxidation catalyzed by tvGlcDH. With NADP (2 m*M*) as the electron acceptor, the enzyme acted exclusively on D-glucose and D-galactose (100 m*M* substrate), and the activity towards D-galactose was 58% of that observed towards D-glucose. The activities with D-xylose, L-arabinose, D-ribose, D-fructose and D-mannose were all less than 4% of that with D-glucose. This is indicative of the narrow substrate specificity of tvGlcDH, which is similar to that shown by *P. torridus* GlcDH (Angelov *et al.*, 2005). The enzyme utilized NAD as a cofactor with nearly the same substrate specificity as NADP, but the electron-acceptor activity of NAD (2 m*M*) was only 4% of that observed with NADP (at 100 m*M* D-glucose). TvGlcDH showed typical Michaelis–Menten kinetics for oxidation; the K_m values for NADP and D-glucose were 0.14 and 7.1 m*M*, respectively. The K_m value for D-galactose was 8.7 m*M* when NADP (2 m*M*) was used as a cofactor.

3.3. Overall structure and structural homologue

The structure of substrate/cofactor-free tvGlcDH was determined using molecular replacement and refined at a resolution of 2.6 Å (Table 2). The asymmetric unit consisted of one homotetramer with a solvent content of 50.2%, which corresponded to a Matthews coefficient (Matthews, 1968) of 2.5 $Å^3$ Da^{-1}. The model of the tetramer (Fig. 2*a*) contained the ordered residues 2–361 in each subunit, six 1,2-propanediol molecules, eight PEG 400 molecules, eight Zn atoms and 234 water molecules. Each monomer consisted of two domains: a nucleotide-binding domain (residues 175–307) and a catalytic domain (residues 2–174 and 308–361) (Fig. 2*b*). When the model of the tvGlcDH monomer was sent to the *DALI* server (Holm & Rosenström, 2010) for the identification of proteins with similar structures (as of August 24, 2013), the protein with the highest structural similarity was ssGlcDH-1 (PDB entries 2cd9, 2cdb, 2cda and 2cdc; r.m.s.d. of between 1.5 and 1.6 Å; Milburn *et al.*, 2006), as expected. The arrangement of the four subunits of the enzyme was essentially the same as that of ssGlcDH-1. The tvGlcDH monomer possesses a secondary/tertiary structure resembling that of ssGlcDH-1; the N-terminal segment of the catalytic domain is linked to the nucleotide-binding domain by a large central helix (α3), the α8 helix connects the end of the nucleotide-binding domain to the C-terminal segment of the catalytic domain, and the nucleotide-binding domain consists of a Rossmann fold containing six parallel strands (Fig. 2*c*). In addition, the catalytic Zn-binding site of

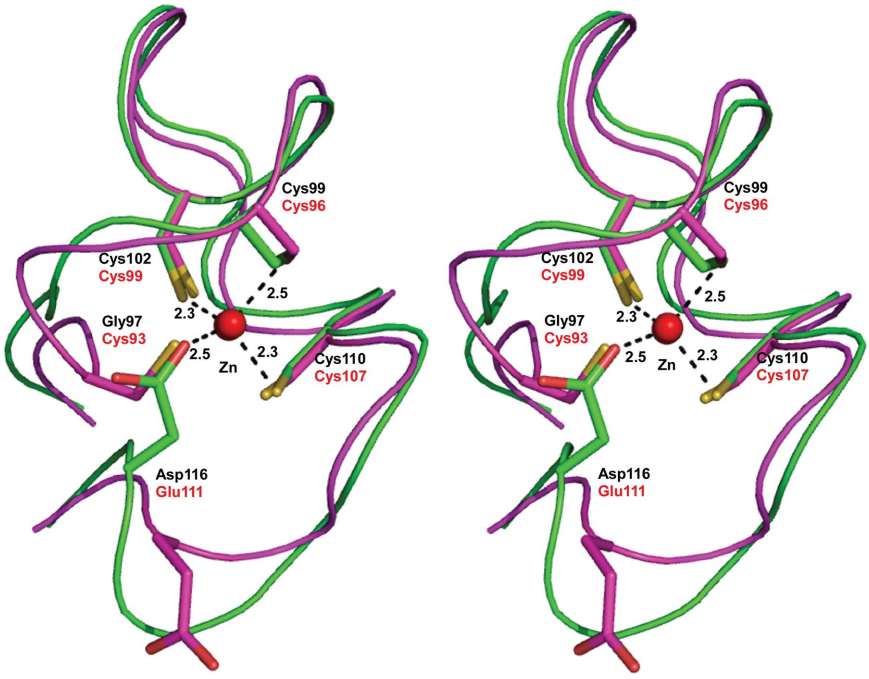


Figure 3
Stereographic close-up of the structural Zn-binding lobe. The structure of tvGlcDH (green and black labels) is superimposed on that of ssGlcDH-1 (magenta and red labels); Zn^{2+} is shown in red. The interactions between Zn^{2+} and the residues in tvGlcDH are shown as dashed lines. Bonding lines are labelled with the corresponding distances (Å). O and S atoms are shown in red and yellow, respectively.

ssGlcDH-1 is well conserved in tvGlcDH. The three residues (Cys39, His66 and Glu67) that coordinate to the catalytic Zn in ssGlcDH-1 are strictly conserved in tvGlcDH as Cys41, His68 and Glu69, respectively (Fig. 1). By contrast, a clear topological difference was found in the conserved structural Zn-binding lobe. In ssGlcDH-1, residues 90–109 coil to form a lobe in which four cysteine residues (93, 96, 99 and 107) coordinate the Zn^{2+} ion. Of these residues, Cys93 is replaced by Gly97 in tvGlcDH, while Cys96, Cys99 and Cys107 are conserved as Cys99, Cys102 and Cys110, respectively (Figs. 1 and 3). Alternatively, the side chain of Asp116, which corresponds to Glu111 in ssGlcDH-1, locates towards the structural Zn and participates in the Zn coordination. To form this unique coordination, the topology of the main chain around Gly97 (residues 94–98) and around Asp116 (residues 113–118) in tvGlcDH must distinctly differ from that in the corresponding regions of ssGlcDH-1 (Figs. 1 and 3).

3.4. NADP-bound structure

In preliminary studies, we co-crystallized the His-tagged tvGlcDH with NADP, determined the initial phases by molecular replacement and refined the structure at a resolution of 2.6 Å (data not shown). Using this approach, however, we did not obtain clear electron density for the NADP molecule, particularly for the nicotinamide ring moiety. This may reflect the relatively low resolution and partial occupancy of this portion. In ssGlcDH-1, NADP reportedly forms several specific interactions with residues in the nucleotide-binding site, one being a stacking interaction between one side of the nicotinamide ring and Phe277, while Phe279 stacks the other side of the nicotinamide ring (Milburn *et al.*, 2006). When we compared the structure of substrate/cofactor-free tvGlcDH with that of NADP-bound ssGlcDH-1 (PDB entry 2cda), we found that Phe279 in ssGlcDH-1 was replaced by Thr277 in tvGlcDH, although Phe277 was conserved as Phe275. We therefore constructed a T277F tvGlcDH mutant and then co-crystallized the mutant with NADP to solve the structure of NADP-bound T277F mutant.

The maximum resolution of diffraction for NADP-bound T277F mutant crystals was enhanced to 2.25 Å, even though the crystals were obtained under the same conditions as used for the NADP-bound His-tagged tvGlcDH crystals. The final model (tetramer) was composed of amino-acid residues 2–361 in each subunit (residues 27–28 in subunit *C* were disordered and were not visible in the electron-density map), four NADP coenzymes, four 1,2-propanediol molecules, eight PEG 400 molecules, four Zn atoms and 328 water molecules. As expected, the nicotinamide ring of NADP in the T277F mutant was tightly held at the active site *via* stacking interactions with Phe275 and Phe277. The electron density corresponding to the NADP bound within the nucleotide-binding site was very clear, enabling us to place the ligand with reasonable accuracy (Fig. 4). Three hydrogen-bonding interactions are made between the nicotinamide amide group (O7N and N7N) and the backbones of Ser304, Phe275 and Asp306. The C2 and C3 hydroxyl groups (O2D and O3D) of the nicotinamide ribose interact with the backbone N and O atoms of Phe277. The nicotinamide phosphate (O2N) interacts with the backbone N atom of Glu197. The C3 hydroxyl group (O3B) of the adenine ribose interacts with the backbone NH groups of Gly195 and Ser194 and the side chain of Asn216. The C2 phosphate group (O2X and O3X) of the adenine ribose interacts with the side chains of His218 and Arg217, while the O atom (O2B) linking the phosphate to the nucleotide forms a hydrogen bond to the side chain of Asn216. N6A of the adenine base forms a hydrogen bond to the side chain of Thr258. N1A and N3A interact with the side chains of Tyr238 and Asn216, respectively. Most of these interactions are conserved in the NADP-bound ssGlcDH-1 structure, although Ser304, Asp306, Glu197, Ser194, His218, Thr258 and Tyr238 in tvGlcDH are replaced by Leu305, Asn307, Ile192, Thr189, Arg213, Ile260 and Ser233,

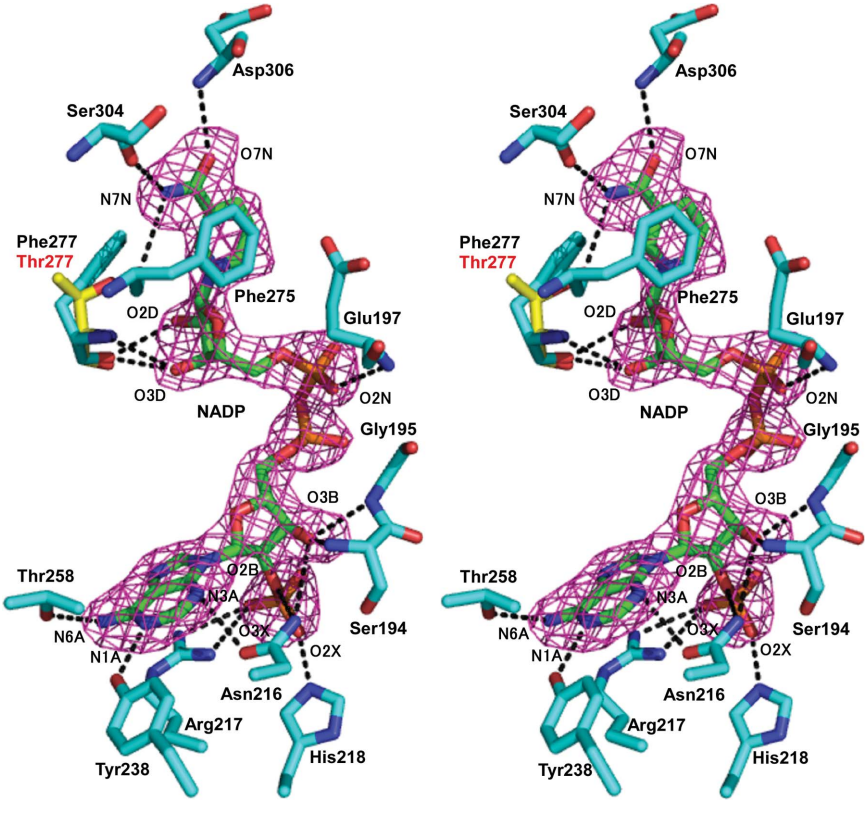


Figure 4
Stereographic close-up of NADP bound to tvGlcDH. The networks of hydrogen bonds are shown as dashed lines. Residues that interact with NADP are shown in cyan. NADP is shown in green. Thr277 in wild-type tvGlcDH is shown in yellow. O and N atoms are shown in red and blue, respectively. The final σ_A-weighted $F_o - F_c$ OMIT electron-density map for NADP is shown at the 2σ level (magenta mesh).

respectively, in ssGlcDH-1. When the structure of substrate/cofactor-free tvGlcDH (wild-type tvGlcDH) was superimposed on the structure of NADP-bound T277F, the side chain, backbone N atom and backbone O atom of Thr277 are situated within hydrogen-bonding distance of the C2 and C3 hydroxyl groups of the nicotinamide ribose, respectively (Fig. 4). In contrast to ssGlcDH-1, in which NADP binding leads to a conformational change within residues 252–260 (Milburn *et al.*, 2006), we detected no substantive difference between the corresponding regions in the structures of wild-type tvGlcDH and NADP-bound T277F.

A remarkable structural change that occurs upon NADP binding to the enzyme is the rotation of the side chain of Glu197 to avoid steric clashes with the nicotinamide phosphate. Within the active-site pocket of the NADP-bound T277F mutant structure, a water molecule (HOH325) was observed in the position corresponding to that of the catalytic Zn in the wild-type tvGlcDH structure. Because the electron-density peak of HOH325 was too weak to be interpreted as a Zn, we modelled a water molecule at this position. However, the possibility that the HOH325 is actually Zn at low occupancy could not been ruled out, as a partially occupied catalytic Zn site was also reported for ssGlcDH-1 (Milburn *et al.*, 2006).

The V_{max} for the T277F mutant was observed to be about 12% of that for the wild-type enzyme [58 µmol min^{-1} mg^{-1} (mutant) *versus* 498 µmol min^{-1} mg^{-1} (wild type) at 328 K], although there was no difference in the K_m values for NADP between the mutant and wild-type enzymes (0.14 m*M* at 328 K). This suggests that the T277F mutation positions the nicotinamide ring in a slightly different position that affects the V_{max}.

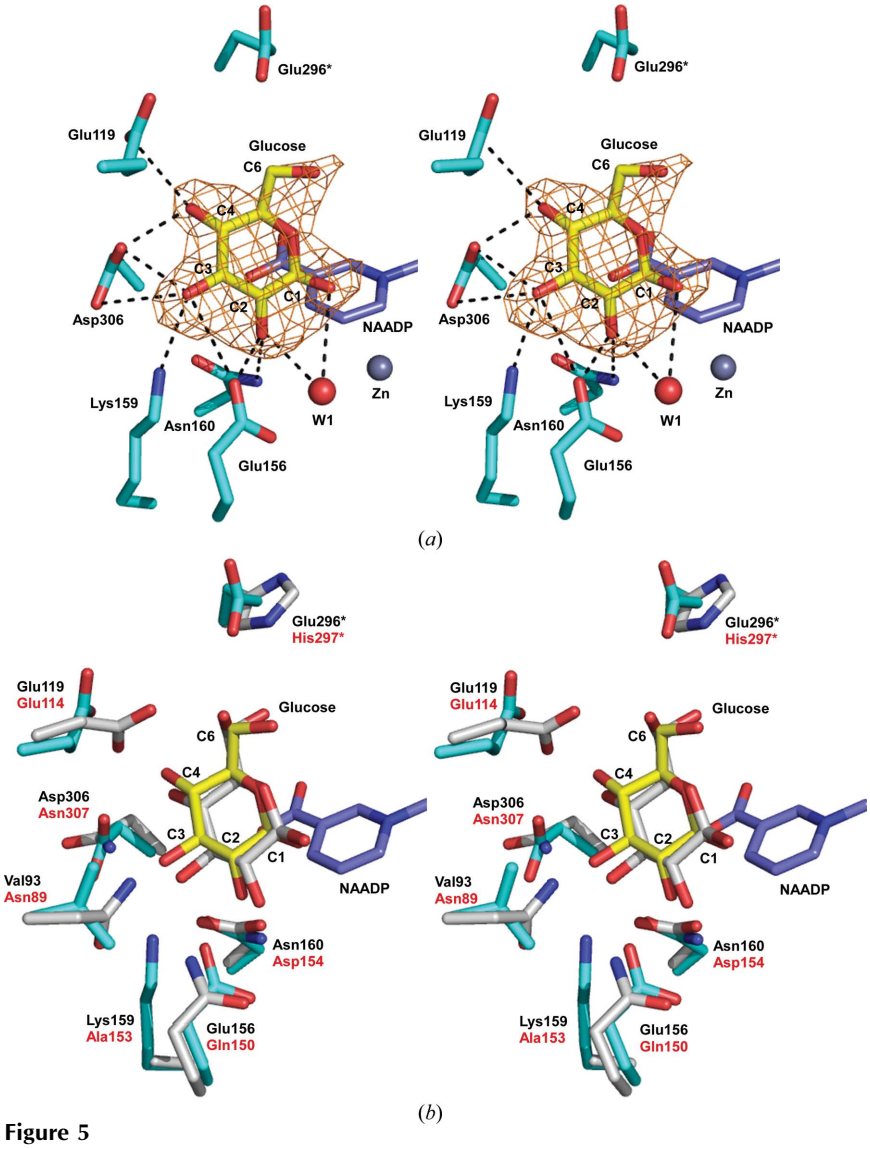


Figure 5
(*a*) Stereographic close-up of the D-glucose-binding site in tvGlcDH. The networks of hydrogen bonds are shown as dashed lines. W1 indicates a Zn-coordinated water molecule. A Zn^{2+} ion is shown as a grey sphere. NAADP and D-glucose are shown in purple and yellow, respectively. Atoms are coloured as in Fig. 4. The final σ_A-weighted $F_o - F_c$ OMIT electron density map for D-glucose is shown at the 2.1σ level (orange mesh). (*b*) Comparison of active-site structures in tvGlcDH (cyan and black labels) and ssGlcDH-1 (grey and red labels). NAADP in tvGlcDH is shown in purple and D-glucose molecules are shown in yellow and grey in tvGlcDH and ssGlcDH-1, respectively.

3.5. Glucose/NAADP-bound structure

Because tvGlcDH exhibits no detectable activity towards NAADP (at 100 m*M* D-glucose), crystals of NAADP-bound T277F mutant were prepared and then soaked in 1.8 *M* D-glucose. The model of the NAADP/D-glucose-bound T277F mutant was refined at a resolution of 2.33 Å. The final model (a tetramer) was composed of amino-acid residues 2–361 in each subunit, four NAADP molecules, four D-glucose molecules, eight Zn atoms and 216 water molecules. The NAADP molecules in all four subunits were positioned/configured in a nearly identical fashion to the NADP molecule from the NADP-bound T277F mutant structure. In the initial electron-density map of the NAADP/D-glucose-bound T277F mutant we observed an extra density within the active-site cavity, and after construction and refinement of the peptide chain a D-glucose molecule could be modelled into this density. Although detailed discussion of the structure of the C6 hydroxyl of D-glucose is difficult owing to poor electron density, the significant electron density at O1, O2, O3 and O4 enables us to deduce a reasonable positioning of the D-glucose molecule (Fig. 5*a*). The bound D-glucose was in the β form, with the C1 hydroxyl in the equatorial configuration. In our model, the C1 hydroxyl forms a hydrogen bond to the Zn-coordinated water molecule (W1). The C2 hydroxyl also forms

a hydrogen bond to W1 and interacts with the side chains of Glu156 and Asn160. The C3 hydroxyl interacts with the side chains of Glu156, Lys159 and Asp306. The C4 hydroxyl forms hydrogen bonds to the side chains of Asp306 and Glu119. Moreover, if the C6 hydroxyl was located towards Glu296*, an additional hydrogen bond would be available between the C6 hydroxyl and the side chain of Glu296* (where an asterisk indicates a residue in the neighbouring subunit; see below). Through these interactions, C1 of the glucose ring is situated in front of the C4 position of the nicotinamide ring, which enables hydride transfer to NADP.

When the NAADP/D-glucose-bound T277F mutant structure was superimposed on the NADP/D-glucose-bound ssGlcDH-1 (T41A mutant) structure (PDB entry 2cdb), we found that the amino-acid residues involved in substrate binding differed significantly between the two enzymes. Among the residues interacting with D-glucose, Gln150, Asp154 and Asn307 in ssGlcDH-1 are replaced by Glu156, Asn160 and Asp306, respectively, in tvGlcDH, although Glu114 in ssGlcDH-1 is conserved as Glu119 in tvGlcDH (Fig. 5*b*). The change from Gln150 to Glu156 does not affect the number of interactions (three) between the residue and the substrate. By contrast, the change from Asp154 to Asn160 reduces the number of the interactions from three to one because the positioning of the Asn160 side chain is too far away to interact with the C3 hydroxyl of the sugar. On the other hand, the change from Asn307 to Asp306 increases the number of interactions from two to three. A marked change from His297* in ssGlcDH-1 to Glu296* in tvGlcDH was recognized, although both are derived from an adjacent monomer in their respective enzymes. SsGlcDH-1 reportedly assembles into a tetramer, with His297* from one monomer contributing to the active site of an adjacent monomer, leading to a narrowing of one end of the substrate-binding cleft (Milburn *et al.*, 2006). In subunit *A* of ssGlucDH-1, the C6 hydroxyl of D-glucose was found in two alternative positions with approximately equal occupancies, each potentially forming just one hydrogen bond to the side chain of His297* or Glu114. In our model, the interaction between the C6 hydroxyl of the sugar and Glu119 (corresponding to Glu114 in ssGlcDH-1) was not observed because the positioning of the Glu119 side chain is too far away to interact with the C6 hydroxyl (Fig. 5*b*). The most striking change is from Asn89 in ssGlcDH-1 to Val93 in tvGlcDH, which results in the loss of a hydrogen bond between the C3 hydroxyl of D-glucose and the side chain of Asn89. Instead, the side chain of Lys159 in tvGlcDH (corresponding to Ala153 in ssGlcDH-1) locates toward the C3 hydroxyl and participates in hydrogen-bonding interactions from the opposite side to Asn89 in ssGlcDH-1 (Fig. 5*b*).

3.6. Insight into substrate specificity

The crystal structure of ssGlcDH-1 (T41A mutant) in complex with NADP and D-xylose has also been determined (Milburn *et al.*, 2006). Comparison of this structure with that of NADP/D-glucose-bound ssGlcDH-1 revealed that the C2 and C3 hydroxyls of D-xylose are located at positions nearly identical to those of the corresponding hydroxyls of D-glucose in the glucose-bound structure. By contrast, the xylose ring appears to be pushed 0.9 Å away from the NADP nicotinamide ring at the pyranose O atom (Milburn *et al.*, 2006; Fig. 6*a*). Consequently, the C1 and C4 hydroxyls of the xylose are also displaced 0.4 Å away from the position of the equivalent glucose atoms. Xylose, in contrast to glucose, has no C6 hydroxy group and thus cannot form a hydrogen bond to His297* or Glu114. The loss of this hydrogen bond is likely to be the main reason for the different positioning of the two sugars, as all other interactions between glucose and the enzyme are conserved in the xylose-bound structure. This suggests that the anchoring of the C2 and C3 hydroxyls of D-xylose at appropriate positions is important for its utilization as a substrate. By comparing the interactions around the C2 and C3 hydroxyls of the sugar between the substrate-bound ssGlcDH-1 and tvGlcDH structures, we found that the C3 hydroxyl of the sugar is tightly held in position through five surrounding hydrogen bonds in ssGlcDH-1 (Fig. 6*a*), whereas only four interactions were observed at the corresponding position in tvGlcDH (Fig. 6*b*). The number of interactions (three) between the sugar C2 hydroxyl

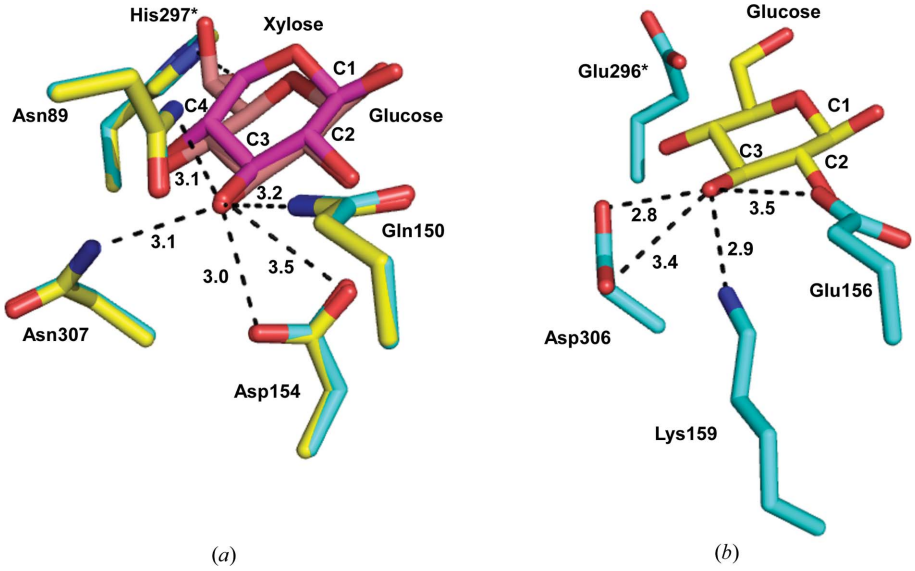


Figure 6
Interactions around the C3 hydroxyls of the sugars. (*a*) The structure of D-glucose-bound ssGlcDH-1 (yellow) is superimposed on that of D-xylose-bound ssGlcDH-1 (cyan). D-Glucose and D-xylose are shown in pink and magenta, respectively. (*b*) The structure of D-glucose-bound tvGlcDH. D-Glucose is shown in yellow. Atoms are coloured as in Fig. 4. The interactions between C3 hydroxyls of the sugars and the surrounding residues are shown as dashed lines. Bonding lines are labelled with the corresponding distances (Å).

and enzyme did not differ between the two enzymes. Given that xylose also cannot form the hydrogen bond to Glu296* in tvGlcDH, these observations suggest that the smaller number of interactions around the C3 hydroxyl of the sugar weakens the affinity of tvGlcDH for D-xylose, making the binding insufficient for catalysis. To test this hypothesis, a V93N mutant of tvGlcDH was constructed in which the number of hydrogen-bonding interactions around the C3 hydroxyl of the sugar might be increased. Contrary to expectation, however, an enhancement of the activity towards D-xylose was not observed. Moreover, the V_{max} value for D-glucose oxidation by the purified V93N mutant was found to be about 71% of the V_{max} value for the wild-type enzyme [356 µmol min^{-1} mg^{-1} (mutant) *versus* 498 µmol min^{-1} mg^{-1} (wild type) at 328 K]. We therefore also constructed E156Q, K159A, N160D, E296H and D306N mutants and compared the activities after heat treatment of the crude extracts. In all cases, improvement of the reactivity towards D-xylose could not be achieved, while the mutations reduced the D-glucose oxidation activity (to less than 19% of that of the wild-type enzyme at 100 m*M* D-glucose). Similar results were obtained for V93N/N160D, V93N/K159A, V93N/D306N, K159A/N160D and V93N/K159A/N160D mutants. Finally, a mutation including all of the substitutions (V93N/E156Q/K159A/N160D/D306N/E296H) completely abolished the activity of the enzyme. These results indicate that all of the substitutions have negative effects on the enzyme reaction. At the present stage, it remains unclear whether or not these substitutions reproduce the interactions that exist in ssGlcDH-1. The structure of the cofactor/D-xylose-bound tvGlcDH mutant may be a useful focus for further investigation.

We are grateful to the staff of the Photon Factory for their assistance with data collection, which was approved by the Photon Factory Program Advisory Committee (Proposal 2013G003). This work was supported in part by Grant-in-Aids for Scientific Research (C) to HS and (A) to TO from the Japan Society for the Promotion of Science.

References

Angelov, A., Fütterer, O., Valerius, O., Braus, G. H. & Liebl, W. (2005). *FEBS J.* **272**, 1054–1062.
Baker, P. J., Britton, K. L., Fisher, M., Esclapez, J., Pire, C., Bonete, M. J., Ferrer, J. & Rice, D. W. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 779–784.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Bright, J. R., Byrom, D., Danson, M. J., Hough, D. W. & Towner, P. (1993). *Eur. J. Biochem.* **211**, 549–554.
Britton, K. L., Baker, P. J., Fisher, M., Ruzheinikov, S., Gilmour, D. J., Bonete, M. J., Ferrer, J., Pire, C., Esclapez, J. & Rice, D. W. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 4846–4851.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Dandekar, T., Schuster, S., Snel, B., Huynen, M. & Bork, P. (1999). *Biochem. J.* **343**, 115–124.
Danson, M. J. (1989). *Can. J. Microbiol.* **35**, 58–64.
De Rosa, M., Gambacorta, A., Nicolaus, B., Giardina, P., Poerio, E. & Buonocore, V. (1984). *Biochem. J.* **224**, 407–414.
Edwards, K. J., Barton, J. D., Rossjohn, J., Thorn, J. M., Taylor, G. L. & Ollis, D. L. (1996). *Arch. Biochem. Biophys.* **328**, 173–183.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Haferkamp, P., Kutschki, S., Treichel, J., Hemeda, H., Sewczyk, K., Hoffmann, D., Zaparty, M. & Siebers, B. (2011). *Biochem. Soc. Trans.* **39**, 77–81.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
John, J., Crennell, S. J., Hough, D. W., Danson, M. J. & Taylor, G. L. (1994). *Structure*, **2**, 385–393.
Jörnvall, H., Persson, M. & Jeffery, J. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 4226–4230.
Jörnvall, H., Persson, B. & Jeffery, J. (1987). *Eur. J. Biochem.* **167**, 195–201.
Jörnvall, H., Persson, B., Krook, M., Atrian, S., Gonzàlez-Duarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Lamble, H. J., Heyer, N. I., Bull, S. D., Hough, D. W. & Danson, M. J. (2003). *J. Biol. Chem.* **278**, 34066–34072.
Lovell, S. C., Davis, I. W., Arendall, W. B. III, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Milburn, C. C., Lamble, H. J., Theodossis, A., Bull, S. D., Hough, D. W., Danson, M. J. & Taylor, G. L. (2006). *J. Biol. Chem.* **281**, 14796–14804.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Ohshima, T., Ito, Y., Sakuraba, H., Goda, S. & Kawarabayasi, Y. (2003). *J. Mol. Catal. B Enzym.* **23**, 281–289.
Ohshima, T., Misono, H. & Soda, K. (1978). *J. Biol. Chem.* **253**, 5719–5725.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.