Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Purification, crystallization and preliminary X-ray analysis of inositol dehydrogenase (IDH) from *Bacillus subtilis*

K. E. Van Straaten, A. Hoffort, D. R. J. Palmer and D. A. R. Sanders*

Department of Chemistry, University of Saskatchewan, 110 Science Place, Saskatoon, Saskatchewan S7N 5C9, Canada

Correspondence e-mail: david.sanders@usask.ca

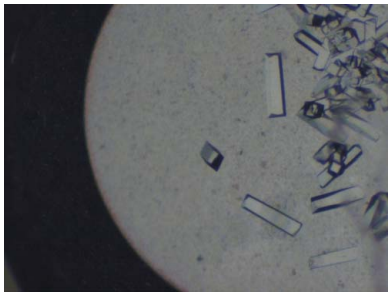
Received 30 October 2007
Accepted 4 January 2008

Inositol dehydrogenase (IDH) is an enzyme that catalyses the NAD^+-dependent oxidation of *myo*-inositol to *scyllo*-inosose. The enzyme has been purified to homogeneity by means of Ni^{2+}-affinity chromatography and was crystallized in both native and selenomethionine (SeMet) labelled forms using the microbatch method. SAD X-ray diffraction data were collected to 2.0 Å resolution from a SeMet-labelled crystal at the Advanced Photon Source (APS) and a MAD data set was collected to 1.75 Å resolution at the Canadian Light Source (CLS); this is the first reported anomalous diffraction experiment from the CLS. The crystals belong to space group *I*222 and contain one molecule per asymmetric unit.

1. Introduction

Inositol dehydrogenase (IDH) from *Bacillus subtilis* is an NAD^+-dependent enzyme that catalyses the oxidation of the axial hydroxyl group of *myo*-inositol to form *scyllo*-inosose (Fig. 1). IDH is the first enzyme in the catabolic pathway of *myo*-inositol, which is a primary carbon source for soil bacteria. The kinetics of IDH have been studied in detail with natural substrates as well as with synthesized 4-*O*-alkyl- or acyl-substituted derivatives of inositol and glycosyl inositols (Daniellou *et al.*, 2006; Daniellou & Palmer, 2006). These studies revealed that IDH has a broad substrate spectrum while remaining highly stereoselective (Daniellou *et al.*, 2005; Daniellou & Palmer, 2006). IDH is able to oxidize the monosaccharides α-D-glucose and α-D-xylose, but not β-D-glucose, D-mannose or D-galactose (Ramaley *et al.*, 1979). IDH also oxidizes the α-(1,6)-linked disaccharides melibiose and isomaltose. Only 1-L-4-*O*-substituted *myo*-inositol derivatives are substrates of IDH; the 1-D-enantiomer is neither a substrate nor an inhibitor. Furthermore, *scyllo*-inositol, the equatorial stereoisomer of *myo*-inositol, is neither a substrate nor an inhibitor of IDH (Ramaley *et al.*, 1979). These observations indicate that an axial hydroxyl group is a requirement for the substrate and that the active site of IDH can selectively discriminate between structural variations in the substrate.

IDH is a member of the GFOR/IDH/MocA family, a group of homologous dehydrogenases, and has a molecular weight of 39.17 kDa. The enzyme is active as a tetramer in solution and has an apparent molecular weight of 160 kDa (Ramaley *et al.*, 1979). Recently, an IDH model has been generated (Daniellou *et al.*, 2007) based on the crystal structure of glucose-fructose oxidoreductase (GFOR) from *Zymomonas mobilis* (PDB code 1ofg; Kingston *et al.*, 1996; Daniellou *et al.*, 2007). Sequence alignment with IDH homologues and homology modelling of IDH allowed us to predict the active-site residues important for binding and catalysis. His176 and Asp172 are proposed to act as the catalytic dyad, in which His176 is proposed to be the acid/base catalyst. H176A and D172N mutants



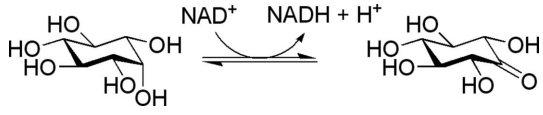


Figure 1
Reaction catalyzed by inositol dehydrogenase (IDH).

showed a marked loss in activity (Daniellou *et al.*, 2007). However, knowledge of their actual role in catalysis requires determination of the precise role of the active-site residues as provided by high-resolution crystal structures with and without substrates or inhibitors.

This paper describes the purification, crystallization and preliminary X-ray analysis of inositol dehydrogenase (IDH) from *B. subtilis*. Furthermore, this is the first time a high-resolution MAD data set has been collected at the Canadian Light Source (CLS), Saskatoon, Canada.

2. Material and methods

2.1. Expression and purification

The expression and purification of IDH was modified from the previously reported protocol (Daniellou *et al.*, 2005) in the following manner. *Escherichia coli* strain Rosetta was used to express IDH. Cells harbouring the IDH-pET28b expression construct for the overexpression of N-terminal His_6-tagged IDH were cultured at 303 K and 250 rev min^{-1} in LB medium containing 50 µg ml^{-1} kanamycin. Expression was induced by the addition of 1 m*M* IPTG at an OD_{600} of 0.3. Cells were grown for an additional 7 h until an OD_{600} of 1.4 was reached. The cells were harvested and resuspended in lysis buffer (buffer *A*) containing 5 m*M* imidazole, 0.50 *M* NaCl, 0.1% NaN_3, 25 m*M* Tris pH 8.2. After addition of 2 m*M* lysozyme, 1 m*M* AEBSF and catalytic amounts of DNAse, the cells were ruptured by sonication and cell debris was removed by centrifugation at 8000*g* for 30 min at 277 K. The IDH-containing supernatant was subjected to a 10% ammonium sulfate cut to remove high-molecular-weight contaminating proteins for 30 min at 277 K. The supernatant was clarified by centrifugation at 17 000*g* for 30 min at 277 K. The supernatant was filtered through 0.22 µm filters (Amicon) and loaded onto a 10 ml Ni-loaded chelating column pre-equilibrated with buffer *A* at 295 K. The column was washed with eight column volumes of buffer *A*. Bound His_6-tagged IDH was eluted with a 15-column-volume linear gradient from 5 to 780 m*M* imidazole in 25 m*M* Tris pH 8.2 containing 0.5 *M* NaCl at a flow rate of 10 ml min^{-1}. His_6-tagged IDH eluted at about 110 m*M* imidazole. Fractions were analyzed by SDS–PAGE. The fractions containing His_6-IDH (molecular weight 39.17 kDa) were pooled and dialyzed against 25 m*M* Tris pH 8.0. His_6-tagged IDH was concentrated to a concentration of 12.0 mg ml^{-1} using a 30K Amicon centrifuge filter device. The amount of protein was determined using UV-absorption spectroscopy at 280 nm with a theoretical molecular extinction coefficient of 42 860 M^{-1} cm^{-1}. A 1 l LB culture produced around 72 mg pure protein.

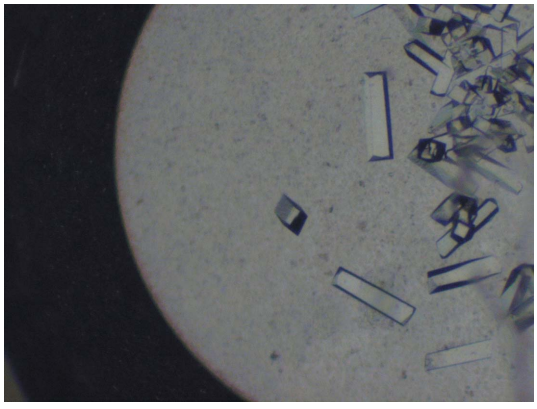


Figure 2
Crystals of inositol dehydrogenase. The crystals have approximate dimensions of 0.15 × 0.15 × 0.3 mm.

A selenomethionine-substituted IDH variant (SeMet-IDH) was prepared by growing *E. coli* Rosetta cells in 2×M9 minimal medium, which was supplemented 15 min before induction with 46 mg l^{-1} L-selenomethionine and a mixture of amino acids known to inhibit methionine biosynthesis (Doublié, 1997; Gerchman *et al.*, 1994). Purification of SeMet-IDH was carried out as described for IDH with 5 m*M* DTT included in the His_6-tagged SeMet-IDH fractions from the Ni-chelating column in order to prevent oxidation of selenomethionine. 320 mg pure SeMet-labelled IDH was obtained from a 2 l M9 SeMet high-yield growth culture (Medicilon Inc.). SeMet-labelled IDH was concentrated to 13.8 mg ml^{-1} in 20 m*M* Tris pH 8.0 and 5 m*M* DTT.

2.2. Crystallization

Crystals of IDH were grown at 277 K using a modified microbatch method that uses a 50:50 mixture of silicone oil and paraffin oil to cover the drops (D'Arcy *et al.*, 1996, 2004). Initial screening for crystallization conditions was performed using commercial crystallization screens (Qiagen). Equal volumes of protein solution (1.0 µl) and precipitant solution (1.0 µl) were mixed. The crystallization drop was overlaid with a 1:1 mixture of silicone and paraffin oils (Al's Oil; Hampton Research), allowing the slow evaporation of water from the drop. Conditions that yielded positive results shared two common features: all the crystals grew at pH 8.5–9.0 using salt and polyethylene glycol (PEG) 3350 or PEG 4000 as the precipitant. The best-looking and best-diffracting crystals were obtained from conditions E2 [20%(*w*/*v*) PEG 3350, 0.20 *M* potassium fluoride] and G10 [20%(*w*/*v*) PEG 3350, 0.20 *M* potassium sulfate] of The PEGs crystallization screen (Qiagen) and condition H5 [15–30%(*w*/*v*) PEG 4000, 0.20 *M* lithium sulfate, 100 m*M* Tris pH 8.5] of The Classics (Lite) crystallization screen (Qiagen) using the microbatch method. Well defined single rod-shaped crystals were obtained from The PEGs condition G10 (Fig. 2). These crystals appeared about a week after setup. SeMet-labelled IDH crystallized under exactly the same conditions as native IDH. However, the best-diffracting crystals were obtained from condition D2 (0.2 *M* $MgCl_2$, 0.1 *M* HEPES pH 7.5, 30% PEG 400) of the JCSG$^+$ screen (Qiagen). These rod-shaped crystals appeared after 1 d and grew to dimensions of 0.2 × 0.2 × 0.35 mm within 4 d.

2.3. Data collection and processing

For data collection, a crystal was transferred for a few seconds into mother liquor containing 25% ethylene glycol as a cryoprotectant. The crystal was mounted in a cryoloop (Hampton Research) and subsequently flash-frozen in liquid nitrogen. All data collections were performed at 100 K in a cold nitrogen stream.

A native IDH data set was collected to 2.9 Å resolution in-house at the Saskatchewan Structural Science Centre (SSSC) at the University of Saskatchewan on a DX8 Proteum diffractometer equipped with a 4k CCD detector. The total oscillation range covered was 180°, with an oscillation range per image of 1° and an exposure time of 120 s. The images were processed and scaled using *SAINT* and *PROSCALE* (Bruker AXS). The software used for data collection and processing was *PROTEUM* v.1.4.1 (Bruker AXS).

A SAD data set was collected to 2.0 Å resolution on the NE-CAT beamline 24-ID at the APS (Argonne, Illinois, USA) using a MAR CCD detector. Selenomethionine incorporation was confirmed by the appearance of a clear selenium absorption edge in the XAFS experiment, enabling us to determine the peak wavelength (data not

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

		SeMet IDH		
	Native IDH	Peak	Peak	Inflection
Beamline	SSSC	APS	CLS	CLS
Wavelength (Å)	1.54	0.9793	0.9797	0.9798
Space group	*I*222	*I*222	*I*222	
Unit-cell parameters (Å)	a = 51.4, b = 122.5, c = 129.9	a = 51.2, b = 122.4, c = 129.4	a = 52.5, b = 120.4, c = 129.2	
Resolution (Å)	61.0–2.90 (3.03–2.90)	50–2.04 (2.12–2.04)	29.3–1.75 (1.81–1.75)	29.3–1.75 (1.81–1.75)
Observed reflections	64581 (8078)	194112 (25789)	275659 (25099)	274938 (24974)
Unique reflections	9424 (1188)	48289 (3687)	41186 (3964)	41116 (3952)
Completeness (%)	99.7 (100)	96.4 (73.1)	98.7 (96.3)	98.7 (96.2)
Multiplicity	6.8 (6.5)	4.0 (3.1)	6.7 (6.3)	6.7 (6.3)
$I/\sigma(I)$	6.7 (2.4)	27.3 (4.3)	10.5 (3.2)	10.3 (3.5)
$R_{\rm merge}$†	9.6 (27.6)	4.6 (14.5)	7.8 (45.6)	7.8 (42.1)
Phasing statistics				
Se sites (found/all)		5/6	5/6	
FOM after *SHELXD*		0.34	0.35	

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \overline{I(hkl)}|/\sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\overline{I(hkl)}$ is the average intensity over symmetry-equivalent measurements.

shown). A total of 200 images was collected, each of which was exposed for 1 s with 1.0° oscillation at a crystal-to-detector distance of 300 mm. The diffraction images were processed and scaled using *HKL*-2000 (Otwinowski & Minor, 1997).

A two-wavelength MAD data set was collected to 1.75 Å resolution on beamline O8-ID.1 at the Canadian Light Source. A clear selenium absorption edge was observed in the XAFS experiment, enabling us to determine the peak, inflection and remote wavelengths (Fig. 3). Beamline O8-ID.1 is equipped with a MAR225 CCD X-ray detector, which was used to collect the data. The diffraction data consisted of a total of 180 images, each of which was exposed for 1 s with 1.0° oscillation at a crystal-to-detector distance of 180 mm. The images were integrated and scaled using *XDS*/*XSCALE* (Kabsch, 1993). For all experiments, intensities greater than six standard deviations from the mean were rejected from scaling.

3. Results and discussion

A native data set was collected in-house to 2.9 Å resolution. Autoindexing of the diffraction data with *SAINT* showed that the crystals belong to the orthorhombic *I*222 space group, with unit-cell parameters a = 51.4, b = 122.5, c = 129.9 Å, $\alpha = \beta = \gamma = 90°$. The asymmetric unit appears to contain one monomer, giving a Matthews coefficient of 2.60 Å^3 Da^{-1} and a solvent content of 52.8% (Matthews, 1968). The preliminary X-ray crystallographic statistics are summarized in Table 1. Programs from the *CCP*4 suite were used to perform most of the calculations (Collaborative Computational Project, Number 4, 1994).

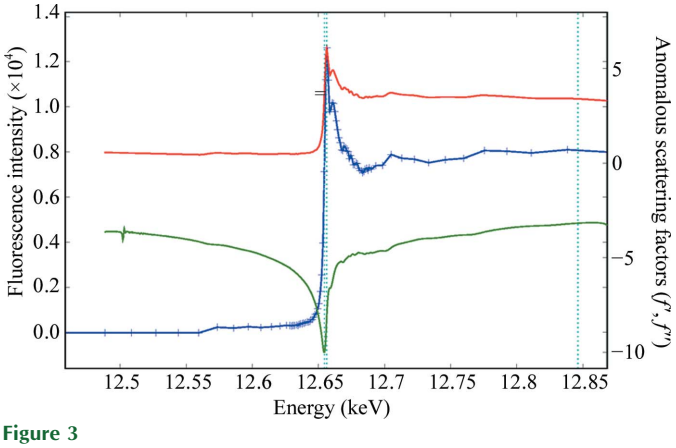


Figure 3
Scattering factors (f' and f'') of SeMet-IDH crystals. The scan was taken on beamline 08ID.1 at the CLS. The energies indicated correspond to the beamline calibration.

3.1. Molecular replacement

For molecular replacement, the programs *MrBUMP* (Keegan & Winn, 2007) and *Phaser* (McCoy *et al.*, 2005) from the *CCP*4 suite were used, with search models that included structures with 15–22% sequence identity to IDH and a molecular model for the *B. subtilis* IDH that had recently been generated (Daniellou *et al.*, 2007) with *MODELLER* (Marti-Renom *et al.*, 2000) using GFOR from *Z. mobilus* as a template (PDB code 1ofg; Kingston *et al.*, 1996; Daniellou *et al.*, 2007). A clear molecular-replacement solution was found using a monomer of PDB entry 2glx [NADP(H)-dependent 1,5-anhydro-D-fructose reductase from *Sinorhizobium morelense*; 22% sequence identity to IDH] as the search model (with the cofactor and water molecules removed). The rotation search, which was performed within the resolution range 20–2.9 Å, yielded one clear top solution with a Z score of 10.04 and an LL gain of 379.2, which confirmed the presence of one monomer per asymmetric unit. The best solution after the translation function had an R factor of 54.7%. Examination of the molecular-replacement solution using the program *Coot* (Emsley & Cowtan, 2004) showed no steric clashes or overlaps with symmetry-related molecules. The solution has a tetrameric arrangement in the asymmetric unit together with symmetry-related molecules, which is physiologically relevant. However, rebuilding and refinement of the model using simulated annealing in *CNS* (Brünger *et al.*, 1998) stalled at high R factors. Reprocessing the data in *C*2 yielded two clear top solutions with Z scores of 10.4 and 10.3 (in contrast to 8.05 for the third solution) and a log-likelihood gain of 679.2. The dimer in the asymmetric unit generated the same tetramer with symmetry-related molecules. Again, rebuilding and refinement of the model using simulated annealing in *CNS* (Brünger *et al.*, 1998) stalled at high R factors. A self-rotation function in *P*1 shows clear 222 point-group symmetry, suggesting space group *I*222. Examination of the self-rotation function did not reveal the presence of any obvious noncrystallographic symmetry or twinning.

3.2. SAD experiment

In order to solve the structure of IDH, we intend to obtain experimental phases by the multiple-wavelength anomalous diffraction (MAD) approach using selenomethionine-substituted IDH (Walsh *et al.*, 1999; Doublié, 1997). SeMet IDH crystallizes in the same space group and with the same unit-cell parameters as native IDH. The crystals diffracted to beyond 2 Å resolution at the synchrotron. A 2 Å resolution single-wavelength anomalous diffraction (SAD) data set was collected at the APS. Very recently, we collected the first MAD data set ever to be collected at the Canadian Light Source (Saskatoon, Canada). The statistics of data collection and phasing are summarized in Table 1. A search for Se-atom sites was carried out with *SHELXD* (Sheldrick, 2008) and found five of six Se sites, assuming the presence of six SeMet residues in one molecule. The phases were further improved by density-modification methods using *autoSHARP* (Vonrhein *et al.*, 2007) and *DM* (Collaborative Computational Project, Number 4, 1994). The resulting electron-density map allowed tracing of the main chain of the polypeptides. Model building and structure refinement are now in progress.

We would like to thank Dr Igor Kourinov and Dr K. Rajashankar from NE-CAT beamline 24-ID at the Advanced Photon Source (APS, Argonne) for their assistance at the APS. We wish to thank Dr Pawel Grochulski and Dr Michel Fodje from the Canadian Light Source for their assistance at the CLS. This work was supported by grants from NSERC (DARS, DRJP) and SHRF (Team Grant). The work conducted at APS is based upon research conducted at the Northeastern Collaborative Access Team beamlines of the Advanced Photon Source, supported by award RR-15301 from the National Center for Research Resources at the National Institutes of Health. Use of the Advanced Photon Source is supported by the US Department of Energy, Office of Basic Energy Sciences under contract No.DE-AC02-06CH11357. Additionally, studies were conducted at the Canadian Light Source, which is supported by NSERC, NRC, CIHR and the University of Saskatchewan.

References

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.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
D'Arcy, A., Elmore, C., Stihle, M. & Johnston, J. E. (1996). *J. Cryst. Growth*, **168**, 175–180.
D'Arcy, A., MacSweeney, A. & Haber, A. (2004). *Methods*, **34**, 323–328.
Daniellou, R. & Palmer, D. R. J. (2006). *Carbohydr. Res.* **341**, 2145–2150.
Daniellou, R., Phenix, C. P., Tam, P. H., Laliberte, M. C. & Palmer, D. R. J. (2005). *Org. Biomol. Chem.* **3**, 401–403.
Daniellou, R., Zheng, H. Y., Langill, D. M., Sanders, D. A. R. & Palmer, D. R. J. (2007). *Biochemistry*, **46**, 7469–7477.
Daniellou, R., Zheng, H. Y. & Palmer, D. R. J. (2006). *Can. J. Chem.* **84**, 522–527.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–530.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Gerchman, S. E., Graziano, V. & Ramakrishnan, V. (1994). *Protein Expr. Purif.* **5**, 242–251.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Keegan, R. M. & Winn, M. D. (2007). *Acta Cryst.* D**63**, 447–457.
Kingston, R. L., Scopes, R. K. & Baker, E. N. (1996). *Structure*, **4**, 1413–1428.
McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. (2005). *Acta Cryst.* D**61**, 458–464.
Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. & Sali, A. (2000). *Annu. Rev. Biophys. Biomol. Struct.* **29**, 291–325.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Ramaley, R., Fujita, Y. & Freese, E. (1979). *J. Biol. Chem.* **254**, 7684–7690.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
Walsh, M. A., Evans, G., Sanishvili, R., Dementieva, I. & Joachimiak, A. (1999). *Acta Cryst.* D**55**, 1726–1732.