Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray analysis of native and selenomethionine class I tagatose-1,6-bisphosphate aldolase from *Streptococcus pyogenes*

Brigitte Liotard* and Jurgen Sygusch

Département de Biochimie, Université de Montréal, CP 6128, Succursale Centre-Ville, Montréal H3C 3J7, Canada

Correspondence e-mail: brigitte.liotard@umontreal.ca

Received 11 October 2003
Accepted 10 December 2003

Tagatose-1,6-bisphosphate aldolase (EC 4.1.2.40) is situated at the branching of the tagatose-6-phosphate and Embden–Meyerhof–Parnas (glycolysis) metabolic pathways, where it catalyzes the reversible cleavage of tagatose-1,6-bisphosphate to dihydroxyacetone phosphate and glyceraldehyde 3-phosphate. The recombinant protein from *Streptococcus pyogenes* was overexpressed in *Escherichia coli* in its native and selenomethionine-derivative forms and purified using ion-exchange and hydrophobic interaction chromatography. Orthorhombic crystals suitable for structural analysis were obtained by the hanging-drop vapour-diffusion method for both isoforms. The crystals belong to space group $P2_12_12_1$, with unit-cell parameters a = 63.7, b = 108.1, c = 238.7 Å for the native form and a = 64.1, b = 108.3, c = 239.8 Å for the selenomethionine derivative. The asymmetric unit contains four protomers, corresponding to a crystal volume per protein weight (V_M) of 2.8 Å^3 Da^{-1} and a solvent content of 56% by volume.

1. Introduction

Aldolases are functionally important enzymes of sugar catabolism in living organisms. They are significant because of their role in essential metabolic pathways such as gluconeogenesis, glycolysis and the pentose-phosphate pathway, which is involved in the biosynthesis of nucleic acids. Because of their ability to control the stereochemistry of carbon–carbon bond formation, they are used as an alternative to conventional chemical methods in biotransformations and synthetic organic chemistry (Takayama *et al.*, 1997; Wong & Whitesides, 1994), especially in the synthesis of novel antibiotics (Wagner *et al.*, 1995; Barbas *et al.*, 1997).

Tagatose-1,6-bisphosphate (TBP) aldolase is an inducible enzyme that exhibits greatest affinity for D-tagatose-1,6-bisphosphate. However, the enzyme can also use as a substrate the bisphosphorylated D-hexose stereoisomers sorbose-P_2, psicose-P_2 and fructose-P_2; in the presence of glyceraldehyde 3-phosphate (G3P) and dihydroxyacetone phosphate (DHAP), condensation produces a mixture of the four D-hexoses (Bissett & Anderson, 1980). The four sugar substrates differ in stereochemistry with respect to their hydroxyl group at carbons 3 and 4 of the hexose. The specificity of the enzyme is therefore expected to be directed at hexose C atoms 1, 2, 5 and 6. In contrast, fructose-1,6-bisphosphate (FBP) aldolases, which are constitutively expressed, catalyze a highly stereospecific reaction.

Two classes of aldolases (I and II), which differ in their catalytic mechanism, are found in nature. Class I aldolases, which are mainly present in higher organisms, employ a lysine residue to form a covalent bond with their substrate by Schiff-base formation, while class II aldolases, which are found in lower organisms, use a divalent cation (typically Zn^{2+}) to polarize the substrate carbonyl to initiate catalysis (Rutter, 1964). The substrate specificity of TBP aldolases and of FBP aldolases is identical in both classes.

Three-dimensional structures of class I FBP aldolases determined from a number of different organisms are consistent with an $(\alpha/\beta)_8$ polypeptide fold in aldolase tetramer subunits (Sygusch *et al.*, 1987; Hester *et al.*, 1991; Chudzik *et al.*, 2000; Kim *et al.*, 1998; Dalby *et al.*, 2001). More recently, structures of class I aldolases have been determined displaying pentameric (Lorentzen *et al.*, 2003) or decameric (Thorell *et al.*, 2002) oligomerization. In case of class II FBP aldolases, which are generally dimers, two structures have been solved (Blom *et al.*, 1996; Cooper *et al.*, 1996; Izard & Sygusch, 2004), each subunit exhibiting homologous polypeptide folding. For TBP aldolase, only a single structure has been reported: a class II tetrameric enzyme from *Escherichia coli* (Hall *et al.*, 2002) the subunits of which also displayed an $(\alpha/\beta)_8$ polypeptide fold. The structural evidence suggests that the active sites of TBP and FBP class II aldolases differ in the region where G3P is thought to bind, in which amino acids show reduced side-chain volume. Randomization in G3P orien-

tation can result in the formation of stereoisomers at the C3—C4 bond of hexose-P_2 and would explain the lack of specificity of TBP aldolase.

In contrast, structural information regarding active-site residues implicated in catalysis and specificity with respect to class I TBP aldolases is lacking. Thus, we have initiated structural analysis of the TBP aldolase from *Streptococcus pyogenes* in an attempt to address the mechanistic features of catalysis in class I aldolases.

2. Materials and methods

2.1. Protein expression and purification

S. pyogenes TBP aldolase (class I) was overexpressed as a 326 amino-acid recombinant protein in *E. coli* strain JM109 (Promega), previously transformed by expression vector pKK-223-3, with T7 IPTG-inducible promoter. Selenomethionine (SeMet) labelling of the protein used methionine-biosynthesis inhibition in minimal medium (Doublié, 1997). The two isoforms were purified to homogeneity using a two-step chromatographic protocol. The clarified lysate was applied onto an anion-exchange column (DEAE Sepharose Fast Flow) and eluted with a 50 m*M* step NaCl gradient at pH 8.0. The active fraction was dialyzed overnight against a solution of 1.5 *M* $(NH_4)_2SO_4$, 50 m*M* NaH_2PO_4 pH 7.0 and applied onto a hydrophobic affinity column (Phenyl Sepharose, Amersham Biosciences), which was eluted at a rate of 2 ml min^{-1} with a continuous $(NH_4)_2SO_4$ gradient using an FPLC apparatus. The purified protein was precipitated by dialysis against 90% saturated $(NH_4)_2SO_4$ and stored at 277 K. Dithiothreitol was added to all buffer solutions to a concentration of 5 m*M* during purification of the SeMet derivative in order to prevent selenomethionine oxidation. MALDI analysis of the N-terminus of the SeMet derivative (data not shown) revealed that the expression system had cleaved the first methionine residue in the primary sequence of TBP aldolase (AAK34623), giving 327 residues, and was corroborated by N-terminal sequencing, which showed the first residue to be a threonine.

2.2. Crystallization conditions

Protein crystal-growth conditions were determined by the hanging-drop vapour-diffusion method using Clear Strategy Screens I and II from Molecular Dimensions Ltd. Equal volumes (2 µl) of protein and crystallization buffer were mixed and equilibrated against 500 µl of solution. Prior to screening of crystallization conditions, the protein was dialyzed against 10 m*M* Tris–HCl pH 7.0 (5 m*M* dithiothreitol was also added to the SeMet isoform) and the protein concentration was brought to 5 mg ml^{-1}.

Crystals suitable for diffraction studies grew as thick plates within 5–7 d at 298 K (largest crystal dimensions 0.2 × 0.5 × 0.5 mm) using 9–11% PEG 4000, 0.2 *M* calcium acetate and 0.1 *M* Tris–acetic acid pH 7.5. Crystallization of native and SeMet isoforms yielded identical unit-cell parameters under these conditions.

Native protein crystals could be stored for several months without noticeable degradation, whereas the SeMet isoform crystals degraded within a few weeks, presumably owing to SeMet oxidation.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shell.

		SeMet		
	Native	Inflection	Peak	Remote
Resolution (Å)	2.4 (2.48–2.4)	2.6 (2.69–2.6)	2.6 (2.69–2.6)	2.6 (2.69–2.6)
Wavelength (Å)	0.98	0.9790	0.9785	0.93
Unique reflections	65506	48526	49255	48526
Completeness (%)	97.5 (85)	100 (99)	100 (99)	99.0 (97)
Redundancy	8.5	16.3	16.6	16.4
Average $I/\sigma(I)$	38.9 (6.3)	48.2 (9.3)	49.9 (9.2)	55.9 (10.9)
R_{merge}†	0.090 (0.18)	0.075 (0.21)	0.070 (0.20)	0.062 (0.18)

† $R_{merge} = \sum_{unique\ reflections}(\sum_{i=1}^{N}|I_i - \bar{I}|)/\sum_{unique\ reflections}(\sum_{i=1}^{N} I_i)$, where N represents the number of equivalent reflections and I the measured intensity.

2.3. Data collection

Crystals were mounted in nylon cryoloops (Hampton Research) after brief immersion in cryoprotectant (mother liquor plus 20% glycerol). Crystals were then flash-cooled in a nitrogen-gas stream at 100 K. Diffraction data were collected at beamline X8-C of the Brookhaven National Laboratory National Synchrotron Light Source, Brookhaven, USA. A single SeMet-derivative crystal was used to collect MAD data at three different wavelengths. The optimal wavelengths for data collection were chosen after recording an X-ray fluorescence spectrum. The wavelengths selected were 0.9790 Å for the inflection point of the fluorescence spectrum (f' minimum), 0.9785 Å for the peak (f'' maximum) and 0.9300 Å for the remote high-energy point. Data were collected with a crystal-to-detector distance of 240 mm, an oscillation range of 0.66° and an exposure time of 60 s per frame.

A preliminary data set was collected at a wavelength of 0.98 Å from a single native crystal that was flash-cooled and cryoprotected in an identical manner; data-collection statistics are shown in Table 1. Fig. 1 shows the diffraction pattern for the native enzyme. Data indexing and integration were performed with the *HKL*2000 package (Otwinowski & Minor, 1997).

3. Results and discussion

Overexpression and purification of the recombinant native enzyme from *S. pyogenes* yielded 30–40 mg per litre of *E. coli* culture, while ∼20 mg were obtained per litre of culture for the SeMet derivative. SDS–PAGE showed TBP aldolase to be >95% pure, with an approximate molecular weight of 37 kDa. Gel filtration of TBP aldolase was consistent with a dimeric quaternary structure when using a mobile phase made up in crystallization buffer.

A BLAST analysis (Altschul *et al.*, 1990) of the TBP aldolase amino-acid sequence against reported TBP aldolases (SWISS-

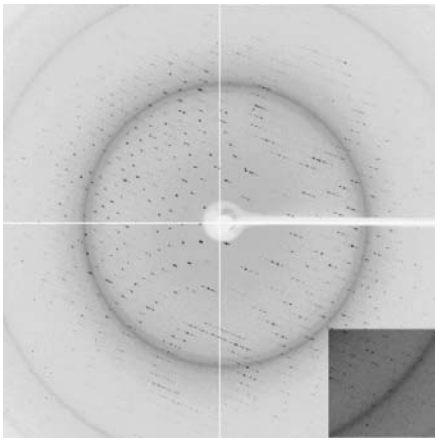

Figure 1
Diffraction pattern obtained from a 0.1 mm thick crystal plate of class I tagatose-1,6-bisphosphate aldolase cloned from *S. pyogenes*. The diffraction experiment was performed at the NSLS synchrotron beamline X8-C and corresponded to a collimator size closed to 0.1 × 0.1 mm to maximally resolve the *c* axis. Exposure time was 60 s. The image in the lower right-hand portion is contrast-enhanced and shows diffraction intensities extending to 2.0 Å resolution.

PROT database) exclusively yielded similarities to aldolase sequences derived from bacterial organisms (74% identity and 87% homology). Secondary-structure prediction using *NNPREDICT* (McClelland & Rumelhart, 1988; Kneller *et al.*, 1990) predicted a protein fold of alternating α-helices and β-sheet motifs interleaved by loops of variable length, which correlates with the $(\alpha/\beta)_8$-barrel folding motif of the aldolase family. A second sequence alignment was performed adding the sequence of rabbit muscle FBP aldolase (P00883). From the sequence alignment (see supplementary data[1]), the predicted secondary structure of TBP aldolase is not inconsistent with the secondary structure observed for FBP aldolase. Active-site residues conserved in FBP aldolase are indicated. A lack of homology with Lys107 in FBP aldolase, associated with the C6 phosphate-binding site (Anai *et al.*, 1973), corroborates the findings in the class II aldolases and suggesting binding-mode differences for G3P in TBP and FBP class I aldolases.

Complete data sets were collected from both isoforms and are shown in Table 1. The resolution limits for data reduction were selected on the basis of completeness and signal-to-noise ratio and set to 2.4 and 2.6 Å for the native and SeMet derivative, respectively. Crystals of both isoforms belong to the orthorhombic system and have space group $P2_12_12_1$. The unit-cell parameters are essentially identical for both the native and SeMet isoforms: $a = 63.7$, $b = 108.1$, $c = 238.7$ Å for the native and $a = 64.1$, $b = 108.3$, $c = 239.8$ Å for the SeMet isoform.

[1] Supplementary data have been deposited in the IUCr electronic archive (Reference PU5032). Services for accessing these data are described at the back of the journal.

Calculation of V_M (Matthews, 1968) yielded a typical value of 2.8 Å^3 Da^{-1} and a solvent content of 56%. These values predict the presence of four protomers per asymmetric unit. On basis of size-exclusion chromatography, this is consistent with two dimers per asymmetric unit.

Knowledge of the tertiary structure and that of catalytic intermediates derived from different substrates should provide insight into the catalytic mechanism of the TBP aldolase and its reduced specificity. Comparison of structural information derived from TBP and FBP aldolases is expected to provide information regarding molecular determinants of the stereospecific condensation/cleavage of carbon–carbon bonds as well as substrate recognition in class I aldolases.

The generous assistance of Miguel St-Jean in data collection and reduction is gratefully acknowledged. The collaboration of Dr Claude Lazure in performing MALDI analysis and N-terminal sequencing was appreciated. This work was supported by a studentship from the Natural Sciences and Engineering Research Council of Canada (NSERC) to BL and a Multi-User Maintenance Grant from NSERC and CIHR (Canada) to JS for synchrotron beamline operation. Work was carried out in part at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the US Department of Energy, Division of Materials Sciences and Division of Chemical Sciences under contract No. DE-AC02-98CH10886.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). *J. Mol. Biol.* **215**, 403–410.
Anai, M., Lai, C. Y. & Horecker, B. L. (1973). *Arch. Biochem. Biophys.* **156**, 712–719.
Barbas, C. F. III, Heine, A., Zhong, G., Hoffmann, T., Gramatikova, S., Bjornestedt, R., List, B., Anderson, J., Stura, E. A., Wilson, I. A. & Lerner, R. A. (1997). *Science*, **278**, 2085–2092.
Bissett, D. L. & Anderson, R. L (1980). *J. Biol. Chem.* **255**, 8750–8755.
Blom, N. S., Tétreault, S. & Sygusch J. (1996). *Nature Struct. Biol.* **10**, 856–862.
Chudzik, D. M., Michels, P. A., de Walque, S. & Hol, W. G. (2000). *J. Mol. Biol.* **300**, 697–707.
Cooper, S. J., Leonard, G. A., McSweeney, S. M., Thompson, A. W., Naismith, J. H., Qamar, S., Plater, A., Berry, A. & Hunter, W. N. (1996). *Structure*, **4**, 1303–1315.
Dalby, A. R., Tolan, D. R. & Littlechild, J. A. (2001). *Acta Cryst.* D**57**, 1526–1533.
Doublié, S. (1997). *Methods Enzymol.* **276**, 523–532.
Hall, D., Bond, C. S., Gordon, A. L., Watt, C. I., Berry, A. & Hunter, W. N. (2002). *J. Biol. Chem.* **277**, 22018–22024.
Hester, G., Brenner-Holzach, O., Rossi, F. A., Struck-Donatz, M., Winterhalter, K. H., Smith J. D. G. & Piontek, K. (1991). *FEBS Lett.* **292**, 237–242.
Izard, T. & Sygusch, J. (2004). Submitted.
Kim, H., Certa, U., Dobeli, H., Jakob, P. & Hol, W. G. (1998). *Biochemistry*, **37**, 4388–4396.
Kneller, D. G., Cohen, F. E. & Langridge, R. (1990). *J. Mol. Biol.* **214**, 171–182.
Lorentzen, E., Pohl, E., Zwart, P., Stark, A., Russell, R. B., Knura, T., Hensel, R. & Sieberg, B. (2003). *J. Biol. Chem.* **278**, 47253–47260.
McClelland, J. L. & Rumelhart, D. E. (1988). *Explorations in Parallel Distributed Processing*, Vol. 3, pp. 318–362. Cambridge, MA, USA: MIT Press.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rutter, W. J. (1964). *Fed. Proc.* **23**, 1248–1257.
Sygusch, J., Beaudry, D. & Allaire, M. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 7846–7850.
Takayama, S., McGarvey, G. J. & Wong, C. H. (1997). *Annu. Rev. Microbiol.* **51**, 285–310.
Thorell, S., Schurmann, M., Sprenger, G. A. & Schneider, G. (2002). *J. Mol. Biol.* **319**, 161–171.
Wagner, J., Lerner, R. A. & Barbas, C. F. III (1995). *Science*, **270**, 1797–1800.
Wong, C.-H. & Whitesides, G. M. (1994). *Enzymes in Synthetic Organic Chemistry*, Vol. 12. New York: Pergamon Press.