Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Characterization of *Escherichia coli* OtsA, a trehalose-6-phosphate synthase from glycosyltransferase family 20

Robert P. Gibson, Ruth M. Lloyd, Simon J. Charnock and Gideon J. Davies*

Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO10 5DD, England

Correspondence e-mail: davies@ysbl.york.ac.uk

Received 5 November 2001
Accepted 29 November 2001

The *Ots* gene cluster of *Escherichia coli* encodes the synthetic apparatus for the formation of α,α-1,1-trehalose, a non-reducing glucose disaccharide. The *otsA* gene encodes a trehalose-6-phosphate synthase, a glycosyltransferase which catalyses the synthesis of α,α-1,1-trehalose-6-phosphate from glucose-6-phosphate using a UDP-glucose donor. It has been classified into glycosyltransferase family GT-20 based upon amino-acid sequence similarities. The *otsA* gene has been cloned and recombinant protein overexpressed using a pET-based system in *E. coli* BL21 cells. The recombinant protein (MW $\simeq$ 54.7 kDa) is active and has been crystallized in two forms suitable for X-ray diffraction analysis. The first is orthorhombic, $P2_12_12_1$, with unit-cell parameters a = 104.1, b = 127.8, c = 179.9 Å. Data for this form have been collected to 3.0 Å resolution at the CLRC Daresbury Synchrotron Radiation Source. The second form has unit-cell parameters $a = b$ = 141.9, c = 317.8 Å and displays the apparent space group $P4_2$. These crystals diffract beyond 2 Å resolution, but display merohedral twinning.

1. Introduction

Glycosyltransferases (EC 2.4.1.*x*) catalyse the formation of glycosidic linkages between an activated sugar donor and the acceptor; the latter is typically another sugar, a lipid or a protein. The donor sugar may be activated by either a nucleoside diphosphate, nucleoside monophosphate, lipid phosphate or merely a phosphate leaving group. To date, 55 sequence-based families of glycosyltransferases have been identified (Campbell *et al.*, 1997; Coutinho & Henrissat, 1999; http://afmb.cnrs-mrs.fr/~pedro/CAZY/gtf.html). Both inverting transferases (which invert the stereochemistry of the sugar donor) and retaining enzymes (which likewise retain the anomeric configuration of the donor) are known. To date, three-dimensional structures have only been reported for seven enzymes from inverting glycosyltransferase families: GT-1 (Mulichak *et al.*, 2001), GT-2 (Charnock & Davies, 1999), GT-7 (Gastinel *et al.*, 1999), GT-13 (Ünligil *et al.*, 2000), GT-28 (Ha *et al.*, 2000), GT-43 (Pedersen *et al.*, 2000) and the bacteriophage T4 DNA-glucosyltransferase (Vrielink *et al.*, 1994) (reviewed by Tarbouriech *et al.*, 2001; Ünligil & Rini, 2000). Only two enzymes of the retaining category have been reported, these being from families GT-8 (Persson *et al.*, 2001) and GT-6 (Gastinel *et al.*, 2001). The mechanism of the configuration-retaining enzymes remains obscure, with a variety of conflicting mechanisms proposed (reviewed by Davies, 2001).

Trehalose (α-D-glucopyranosyl-α-D-glucopyranoside) is a non-reducing disaccharide synthesized ubiquitously throughout nature by both prokaryotes and eukaryotes. Its production is frequently associated with physiological stress, such as osmotic shock; trehalose forms a glass-like matrix structure upon dehydration, which has unique stabilizing properties (reviewed by Singer & Lindquist, 1998). Indeed, trehalose production is associated with so-called 'resurrection plants' and other organisms that withstand almost complete desiccation (Singer & Lindquist, 1998). *E. coli* produces trehalose *via* a two-step pathway involving the enzymes OtsA and OtsB (Fig. 1). The first step in the pathway is catalysed by OtsA, a trehalose-6-phosphate (T-6-P) synthase (EC 2.4.1.15). This enzyme catalyses the production of T-6-P from UDP-glucose and glucose-6-phosphate, with a net retention of stereochemistry at C1 of the donor sugar. OtsB, a T-6-P phosphatase, performs the second step in the pathway, removing the 6-phosphate group and resulting in the production of trehalose. Sequence similarity has led to the identification of T-6-P synthases in other species, as well as alternate redundant pathways for its production and subsequent utilization. It has been found that not only does trehalose act as a structural agent, but also that the phosphate precursor is potentially part of a

regulatory system controlling the inhibition of hexokinase (Silva-Udawatta & Cannon, 2001).

OtsA has been classified into family GT-20 in the Henrissat classification (Campbell *et al.*, 1997; Coutinho & Henrissat, 1999). OtsA from *E. coli* is a 54.7 kDa protein whose sequence similarity to other GT-20 enzymes is approximately 40% irrespective of the organism. Here, we present the cloning, expression, purification and crystallization of recombinant OtsA from *E. coli* (strain W3110) along with a preliminary X-ray diffraction analysis of two distinct crystal forms.

2. Materials and methods

2.1. Cloning, expression and purification

The *otsA* gene from *E. coli* strain W3110 was amplified by the polymerase chain reaction using Vent DNA polymerase (New England Biolabs Inc., Beverly, MA, USA) and the oligonucleotide primers CATATgAgTCgTTTAgTCgT and CTCgAgCgCAAgCTTTggAAAg with a thermal program comprising 22 cycles at 368 K for 1 min, 329 K for 1 min and 347 K for 2 min, following two initial cycles with an annealing temperature of 313 K and a further two cycles with an annealing temperature of 315 K. The 1.44 kbp PCR product contained engineered *Nde*I and *Xho*I restriction endonuclease sites at the 5′ and 3′ termini, respectively, and the original stop codon was removed to place *otsA* in frame with the C-terminal polyhistidine purification peptide of the pET22b (Novagen) expression vector. The amplified gene product was ligated into pCR-Blunt (Invitrogen) and sub-cloned into pET22b on a partial *Nde*I/*Xho*I restriction fragment, generating pECTS1. *E. coli* BL21(DE3), transformed with pECTS1, was cultured on 2%(*w*/*v*) LB agar supplemented with 50 mg ml^{-1} ampicillin for 16 h at 310 K. A single colony was used to inoculate 1 l LB liquid growth medium, supplemented with 50 mg ml^{-1} ampicillin and then incubated at 310 K with shaking at 120 rev min^{-1}. When the OD_{550} reached 0.45, 1 m*M* isopropyl-thio-β-D-galactoside was introduced to initiate overexpression. The culture was then incubated at 303 K for a further 12 h, after which the stationary phase had been reached. The bacterial culture was harvested by centrifugation and resuspended in sonication buffer consisting of 50 m*M* Na HEPES buffer pH 7.0 containing 200 m*M* NaCl and 10 m*M* imidazole. The cells were sonicated and debris was removed by centrifugation at 30 000*g*, 277 K for 45 min and the supernatant loaded onto a 1.6 × 5 cm chelating Sepharose column (Pharmacia Biotech), charged with 100 m*M* sodium acetate buffer pH 5.3 containing 12 mg ml^{-1} $NiSO_4$ and equilibrated with the sonication buffer. A stepwise elution was performed with imidazole concentrations of 10, 50, 150, 250 and 500 m*M* which yielded ∼90% pure recombinant OtsA in the 250 m*M* fraction as judged by SDS–PAGE (data not shown). The protein preparation was buffer exchanged into 50 m*M* Na HEPES buffer pH 7.0 containing 200 m*M* NaCl using a Sephadex DP-10 G-25M gel-filtration column (Pharmacia Biotech) and loaded in 2 ml volumes onto a HiLoad16/60 Superdex 200 gel-filtration column (Pharmacia Biotech). The purified protein was finally buffer exchanged into 5 m*M* Na HEPES buffer pH 7.0 containing 60 m*M* $MgCl_2$ by repeated concentration using a 3.5 ml centrifugation membrane (Amicon). Confirmation of the predicted molecular mass, 54.7 kDa, was demonstrated using MALDI-TOF MS analysis (data not shown).

2.2. Trehalose-6-phosphate synthase assay

The reaction conditions comprised 50 m*M* Na HEPES buffer pH 7.0 containing 5 m*M* UDP-glucose, 5 m*M* glucose-6-phosphate, 5 m*M* $MgCl_2$, 200 m*M* NaCl and 12 µg ml^{-1} OtsA. The reaction mixture was incubated at 294 K for 60 min before being terminated by the addition of 150 m*M* EDTA. A 15 µl sample diluted 40-fold with water was analysed by anion-exchange HPLC using a Carbopac PA100 column (Dionex; 0.5 × 30 cm). The solvent system comprised 100 m*M* NaOH and the reaction products were eluted using a 0–500 m*M* gradient of NaOAc.

2.3. Crystallization, data collection and processing

Two crystals forms of OtsA have been obtained. Crystals of form 1 were grown by vapour-phase diffusion using the hanging-drop method. 1 µl of recombinant OtsA at a concentration of 15 mg ml^{-1} in 5 m*M* Na HEPES buffer, 60 m*M* $MgCl_2$ pH 7.0 was combined with 1 µl of the reservoir buffer solution after 16 h at 291 K. The tetragonal crystals were produced with a reservoir solution of 0.7 *M* $(NH_4)_2SO_4$, 1.35 *M* Li_2SO_4 and 0.1 *M* trisodium citrate pH 5.4, with the addition of the ligands Mg-UDP and glucose-6-phosphate present at 10 m*M* each. A cryoprotectant for data collection was prepared by supplementing the growth buffer with 25%(*w*/*v*) glucose. Data were collected from a single crystal at the European Synchrotron Radiation Facility (ESRF), Grenoble on beamline ID29 at a temperature of 100 K.

A second, orthorhombic, form was produced in a buffer comprising of 8%(*v*/*v*) monomethylether PEG 550, 8%(*v*/*v*) PEG 20 000, 0.3 *M* NaOAc and 0.1 *M* Na HEPES buffer pH 6.8, with the ligands Mg-UDP and glucose-6-phosphate again present at 10 m*M*. A cryoprotectant solution was produced by the addition of 25%(*v*/*v*) PEG 400 to the growth buffer. Crystals were mounted in rayon-fibre loops and placed in an N_2 stream at 120 K for data collection. Data for the second crystal form were collected on beamline PX9.6 at the CLRC Daresbury SRS. 120° of data were collected as a series of 0.5° oscillation images using an ADSC QUAD-4 CCD as detector, at a wavelength of 0.870 Å. Data were processed and scaled using the *HKL* suite of programs (Otwinowski, 1993; Otwinowski & Minor, 1997).

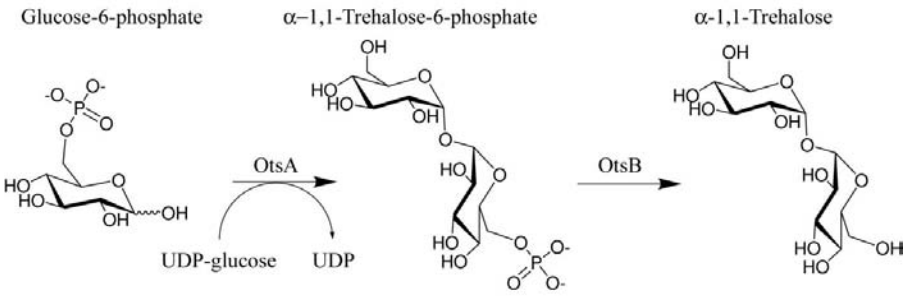

Figure 1
The synthesis of α,α-1,1-trehalose by *E. coli* gene products OtsA and OtsB.

3. Results and discussion

Recombinant protein was expressed as 80% total soluble protein by *E. coli* BL21(DE3) cells using the pET-based expression system described above. MALDI-MS gave a value of 54 650.27 ± 50 Da compared with the theoretical mass of 54 675.3 Da. Recombinant OtsA was shown to catalyse the synthesis of trehalose-6-phosphate, using

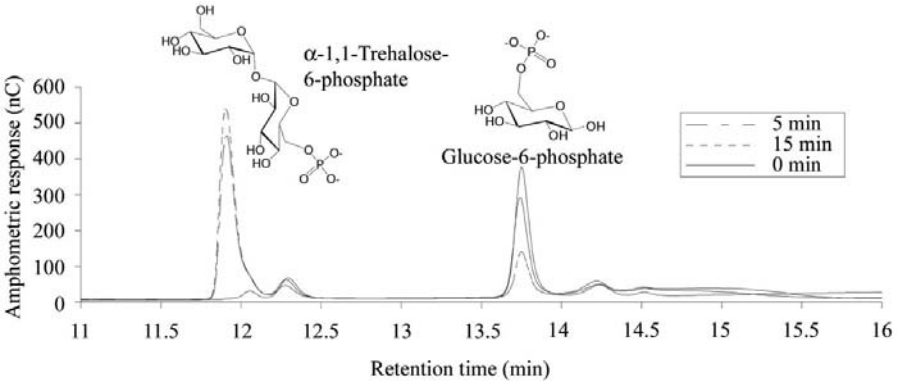

Figure 2
The time-dependent enzymatic synthesis of trehalose-6-phoshate and concomitant depletion of glucose-6-phosphate using a UDP-glucose donor and glucose-6-phosphate as acceptor. The response of the pulsed amperometric ED40 detector (Dionex corporation, Sunnyvale, CA, USA) is plotted as a function of retention time in minutes.

glucose-6-phosphate as the acceptor and either Mg-UDPGlc or Mn-UDPGlc as the sugar donor, by an HPLC-based assay procedure (as described for xylooligosaccharides in Charnock *et al.*, 1998) which demonstrates the time-dependent production of T-6-P with concomitant depletion of the acceptor glucose-6-phosphate (Fig. 2). In order to determine the true kinetic constants and perhaps dissect the ordered binding of substrates, the use of more sophisticated assay measurements, such as the detection of UDP release *via* a linked assay procedure, will be necessary.

Two distinct crystal forms of OtsA were obtained as described above. The first crystal form has unit-cell parameters $a = b = 141.8$, $c = 317.8$ Å, $\alpha = \beta = \gamma = 90^\circ$. Data collected to 2.2 Å on ESRF beamline ID29 reduce in point group 422 ($R_{\rm merge} = 0.10$). Systematic absences are somewhat confusing, with both $l = 4n$ and $l = 2n$ present, but the former more intense than the latter. Analysis of the cumulative intensity distribution revealed a clear merohedral twin with an estimated twinning fraction of 50%. Further confirmation of twinning comes from the analysis of the first and third acentric moments, which gave values of 0.89 and 1.19 compared with theoretical twinned values of 0.94 and 1.18 (non-twinned values 0.866 and 1.34). In order to overcome twinning problems, a screen for an alternative crystal form was undertaken.

The second crystal form of OtsA belongs to space group $P2_12_12_1$ and has unit-cell parameters $a = 104.1$, $b = 127.8$, $c = 179.9$ Å. This crystal form diffracts more weakly than the first crystal form. Preliminary diffraction on beamline ID14-4 reveals diffraction to 2.2 Å and data extending to 3 Å resolution were subsequently collected at the CLRC Daresbury SRS on beamline PX 9.6. Data are 99.9% complete to 3.0 Å with an $R_{\rm merge}$ of 0.070, a multiplicity of 4.9 observations per reflection and a mean $I/\sigma(I)$ of 22. In the outer resolution (3.1–3.0 Å) bin the data are 100.0% complete, with an $R_{\rm merge}$ of 0.39, a mean $I/\sigma(I)$ of 4.6 and a multiplicity of 4.9. The asymmetric unit may contain between from two to five molecules of OtsA, corresponding to 77–44% solvent.

The catalytic mechanism of retaining glycosyltransferases remains one of the most important and intriguing questions in glycobiology. Conflicting mechanistic proposals featuring covalent intermediates (by analogy with retaining glycoside hydrolases; Vocadlo *et al.*, 2001), long-lived oxocarbenium ions (as proposed for glycogen phosphorylase) and even exploded S_Ni and substrate participation (Persson *et al.*, 2001) have all been considered (Davies, 2001). Currently, there is no substantial experimental evidence to favour one mechanistic path over any other. The crystal structure of a T-6-P synthase will provide additional structural insight into this mechanism, as well as unlocking the secrets of trehalose synthesis: trehalose is a vastly important sugar, especially in pathogenic organisms such as *Mycobacterium tuberculosis*. To facilitate this, a selenomethionine-derivative enzyme is being prepared for multi-wavelength anomalous dispersion (MAD) analysis in harness with fluorinated substrate analogues to probe the catalytic mechanism.

The authors would like to thank the BBSRC and the Wellcome Trust for funding, the staff of the ESRF and SRS for provision of data-collection facilities and the EU and ESRF for financial assistance. GJD is a Royal Society University Research Fellow.

References

Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. (1997). *Biochem. J.* **326**, 929–942.
Charnock, S. & Davies, G. J. (1999). *Biochemistry*, **38**, 6380–6385.
Charnock, S. J., Spurway, T. D., Xie, H., Beylot, M. H., Virden, R., Warren, R. A., Hazlewood, G. P. & Gilbert, H. J. (1998). *J. Biol. Chem.* **273**, 32187–32199.
Coutinho, P. M. & Henrissat, B. (1999). *Recent Advances in Carbohydrate Engineering*, edited by H. J. Gilbert, G. J. Davies, B. Svensson & B. Henrissat, pp. 3–12. Cambridge: Royal Society of Chemistry.
Davies, G. J. (2001). *Nature Struct. Biol.* **8**, 98–100.
Gastinel, L. N., Bignon, C., Misra, A. K., Hindsgaul, O., Shaper, J. H. & Joziasse, D. H. (2001). *EMBO J.* **20**, 638–649.
Gastinel, L. N., Cambillau, C. & Bourne, Y. (1999). *EMBO J.* **18**, 3546–3557.
Ha, S., Walker, D., Shi, Y. & Walker, S. (2000). *Protein Sci.* **9**, 1045–1052.
Mulichak, A. M., Losey, H. C., Walsh, C. T. & Garavito, R. M. (2001). *Structure*, **9**, 547–557.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A. & Negishi, M. (2000). *J. Biol. Chem.* **275**, 34580–34585.
Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. & Strynadka, N. C. J. (2001). *Nature Struct. Biol.* **8**, 166–175.
Silva-Udawatta, M. N. & Cannon, J. F. (2001). *Mol. Microbiol.* **40**, 1345–1356.
Singer, M. A. & Lindquist, S. (1998). *Trends Biotechnol.* **16**, 460–468.
Tarbouriech, N., Charnock, S. J. & Davies, G. J. (2001). *J. Mol. Biol.* **314**, 665–661.
Ünligil, U. M. & Rini, J. M. (2000). *Curr. Opin. Struct. Biol.* **10**, 510–517.
Ünligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. & Rini, J. M. (2000). *EMBO J.* **19**, 5269–5280.
Vocadlo, D. J., Davies, G. J., Laine, R. & Withers, S. G. (2001). *Nature (London)*, **412**, 835–838.
Vrielink, A., Rüger, W., Driessen, H. P. C. & Freemont, P. S. (1994). *EMBO J.* **13**, 3413–3422.