Expression, purification, crystallization and preliminary X-ray analysis of glucose-1-phosphate uridylyltransferase (GalU) from *Erwinia amylovora*

Mirco Toccafondi,^a Michele Ciani^b and Stefano Benini^{a*}

^aLaboratory of Bioorganic Chemistry and Bio-Crystallography (B2CI), Faculty of Science and Technology, Free University of Bolzano, Piazza Università 5, 39100 Bolzano, Italy, and ^bEMBL, Notkestrasse 85, 22607 Hamburg, Germany

Correspondence e-mail: stefano.benini@unibz.it

Received 17 June 2014
Accepted 15 July 2014

Glucose-1-phosphate uridylyltransferase from *Erwinia amylovora* CFPB1430 was expressed as a His-tag fusion protein in *Escherichia coli*. After tag removal, the purified protein was crystallized from 100 mM Tris pH 8.5, 2 M ammonium sulfate, 5% ethylene glycol. Diffraction data sets were collected to a maximum resolution of 2.46 Å using synchrotron radiation. The crystals belonged to the hexagonal space group $P6_2$, with unit-cell parameters $a = 80.67$, $b = 80.67$, $c = 169.18$. The structure was solved by molecular replacement using the structure of the *E. coli* enzyme as a search model.

1. Introduction

Glucose-1-phosphate uridylyltransferase (GalU; EC 2.7.7.9) catalyses the synthesis of UDP-glucose and pyrophosphate from glucose-1-phosphate and UTP. These enzymes are able to catalyse both the forward and reverse reactions and hence are also known as pyrophosphorylases. GalU has a prominent role in sugar metabolism in all living organisms (Kleczkowski *et al.*, 2004; Flores-Díaz *et al.*, 1997). In fact, the forward reaction promotes the formation of UDP-glucose, an activated form of glucose utilized as a glycosyl donor in biosynthetic pathways, while the reverse reaction takes part in the catabolism of complex sugars (Kleczkowski *et al.*, 2004).

In bacteria, GalU is involved in galactose metabolism through the Leloir pathway (Frey, 1996) and in the biosynthesis of exopolysaccharides (EPSs; Chang *et al.*, 1996). EPSs represent the major virulence factors for several Gram-negative bacteria such as *Streptococcus pneumoniae* (Yother, 2011), *Pseudomonas aeruginosa* (Dean & Goldberg, 2002) and *Aeromonas hydrophila* (Vilches *et al.*, 2007), and also for some Gram-positive bacteria such as *Mycobacterium tuberculosis* (Lai *et al.*, 2008).

The Gram-negative plant pathogen *Erwinia amylovora* is the causal agent of fire blight, a necrotic disease affecting several species of the Rosaceae family, including apple and pear trees (Vrancken *et al.*, 2013). The high virulence of *E. amylovora* is related to the production of a biofilm that protects the bacteria from host defences and unfavourable environmental conditions and enables multicellular assembly during host invasion (Koczan *et al.*, 2011). In *E. amylovora*, the two exopolysaccharides levan and amylovan represent the main components of the biofilm (Koczan *et al.*, 2009). *E. amylovora* levan, produced by the enzyme levansucrase, is mainly a mixture of short-chain fructooligosaccharides (Caputi *et al.*, 2013). Amylovan is a branched heteropolymer mainly composed of galactose with some residues of glucose, glucuronic acid and pyruvate (Jumel *et al.*, 1997; Nimtz *et al.*, 1996). Amylovan is one of the most important pathogenicity factors; indeed, *E. amylovora* strains with deficiencies in amylovan biosynthesis showed a loss of pathogenicity (Belle-mann & Geider, 1992). Amylovan biosynthesis requires the proteins encoded by the 12 genes in the *ams* operon and the proteins involved in galactose metabolism (Bugert & Geider, 1995).

Galactose is the main building block for amylovan production, but it is not available at the infection site. In *E. amylovora*, UDP-galactose-4-epimerase (GalE) is the designated enzyme that provides galactose residues for such a purpose by converting UDP-glucose to

© 2014 International Union of Crystallography
All rights reserved

UDP-galactose. UDP-glucose, as a direct precursor of UDP-galactose, thus represents a critical substrate in amylovoran biosynthesis (Metzger *et al.*, 1994). GalU is the unique enzyme responsible for the formation of UDP-glucose; therefore, it plays a key role in EPS production and it exerts its function upstream from galactose metabolism and the amylovoran-biosynthesis proteins. Despite its deep involvement in *E. amylovora* survival and pathogenesis, *EaGalU* has not yet been characterized in terms of structure and function.

Uridyltransferases are present in eukaryotes and prokaryotes. However, while bacterial GalU isoforms share high sequence identity, they do not seem to have a related eukaryotic counterpart, revealing that the prokaryotic and eukaryotic enzymes are not evolutionarily linked (Mollerach *et al.*, 1998; Daran *et al.*, 1995). This suggests that GalU represents a promising target for the development of new pharmacological compounds against *E. amylovora* infection and host invasion and may also be a common target for other Enterobacteriaceae family members. For such a purpose, we expressed, purified and crystallized *EaGalU* for structural characterization and investigation of the catalytic mechanism as a starting point for a structure-based drug-design approach.

2. Materials and methods

2.1. Cloning, overexpression and purification of *E. amylovora* GalU

The sequence of the *galU* gene from *E. amylovora* CFBP1430 was PCR-amplified using the forward primer 5'-ATTATTCCATGGCTGCCTATAACTCAAAGT-3', which introduced an *NcoI* restriction site 5' to the *galU* gene sequence, and the reverse primer 5'-ATTATTCTCGAGTTACTTTTTGTTGCCGACA-3', which introduced an *XhoI* restriction site at the 3' end. The PCR product was purified using the QIAquick PCR Purification Kit (Qiagen, Germany), double-digested with *NcoI* and *XhoI* restriction enzymes (NEB, USA), gel-purified (Qiagen, Germany) and ligated into the pETM-11 expression vector (Dümmler *et al.*, 2005). The pETM-11 plasmid encodes a six-histidine tag sequence and the *Tobacco etch virus* protease (TEV) cleavage site before the *galU* insertion, resulting in H₆X₁₀ENLYFQGAM-GalU.

The construct was propagated into *Escherichia coli* NovaBlue cells (EMD4 Biosciences, Germany) and purified using a DNA Miniprep kit (Sigma, USA). The plasmid size was checked by agarose gel electrophoresis after single and double digestion and the correctness of the inserted region was verified by sequencing at the Microsynth facilities (Microsynth AG, Balgach, Switzerland).

E. coli BL21(DE3) chemically competent cells (EMD4 Biosciences, Germany) were transformed with the pETM-11(*galU*) construct to express the recombinant protein. An overnight pre-culture was grown in 20 ml 2×YT medium containing kanamycin (30 µg ml⁻¹) at 310 K. Half of the pre-culture (10 ml) was added to 1 l fresh 2×YT medium (dilution 1:100) and grown at 310 K with shaking at 180 rev min⁻¹ until an OD₆₀₀ of 0.6 ± 0.1 was reached. The temperature of the shaker was then decreased to 293 K and the culture was left to equilibrate for 1 h before induction with 0.5 mM IPTG. After 16 h, the cells were harvested by centrifugation at 4500g for 20 min at 277 K and resuspended in cold PBS. After a second centrifugation, the cells were lysed in 20 mM Tris buffer pH 7.4, 150 mM NaCl, 20 mM imidazole with 0.2 mg ml⁻¹ lysozyme and protease-inhibitor cocktail (Sigma–Aldrich) and were disrupted by sonication (Soniprep, MSE, UK) on ice for 2 min using 10 s cycles (15.6 MHz). The lysate was cleared by centrifugation at 18 000g for 20 min at 277 K and filtered with a 0.45 µm cellulose acetate filter.

The soluble fraction was loaded onto a HisTrap HP 5 ml column (GE Healthcare, Sweden) pre-equilibrated with resuspension buffer (20 mM Tris buffer pH 7.4, 150 mM NaCl, 20 mM imidazole) at a flow rate of 1 ml min⁻¹. The column was washed until the A₂₈₀ reached the baseline to remove contaminants and the bound protein was eluted with 20 mM Tris buffer pH 7.4, 150 mM NaCl, 250 mM imidazole. Buffer was exchanged using a 16/20 HiTrap desalting column (GE Healthcare, Sweden) at a 5 ml min⁻¹ flow rate to remove imidazole and leave the protein in 20 mM Tris pH 7.4, 150 mM NaCl. The N-terminal His tag was cleaved by the addition of solubility-enhanced L56V/S135G TEV protease (Cabrita *et al.*, 2007) at a 1:25 (TEV: GalU) ratio at room temperature overnight. After centrifugation at 4500g for 20 min and filtration with a 0.45 µm filter, the reaction mixture was again loaded onto a His-Trap HP 5 ml column to remove the TEV protease, cleaved His tag and uncleaved protein from the solution. Tag-free GalU was collected as the flowthrough fraction. The GalU sample was concentrated to 5 ml and loaded onto a Sephadex S75 16/60 column (GE Healthcare, Sweden) equilibrated with 20 mM Tris pH 7.5, 150 mM NaCl at a flow rate of 1 ml min⁻¹. All purification steps were carried out at 293 K. The purity of recombinant GalU was confirmed by SDS–PAGE electrophoresis. The protein stability, the prevalent form in solution and the molecular mass were analysed by Thermofluor, DLS and MALDI-TOF, respectively, at the EMBL SPC facilities, Hamburg, Germany.

2.2. Crystallization

The purified protein was concentrated to 5, 10 and 15 mg ml⁻¹ using a Centricon (Millipore Corporation, Massachusetts, USA) in 20 mM Tris pH 7.4, 150 mM NaCl. The protein concentration was determined by direct UV measurement at 280 nm with a NanoVue spectrophotometer (GE Healthcare, Sweden) using an extinction coefficient of 28 670 M⁻¹ cm⁻¹ calculated by *ProtParam* (ExPASy). Crystallization trials were performed using microbatch-under-oil in 96-well MRC plates (Cambridge, England) using volatile oil (MD2-06, Molecular Dimensions) at 293 K. Drops of 1 µl protein solution were mixed with the same volume of conditions from the commercially available crystallization kits PACT, CSS1, CSS2, JCGS, MIDAS and MORPHEUS (Molecular Dimensions). Crystals appeared in 13 different conditions. Selected conditions were optimized using hanging-drop vapour diffusion and adjusting the salt concentration, pH and additives. Drops of 1 µl protein solution (5 mg ml⁻¹) were added to 1 µl precipitant and equilibrated against 1 ml precipitant in 24-well Linbro plates (Hampton Research). Diffracting crystals grew within a week in 100 mM Tris pH 8.5, 2 M ammonium sulfate, 5% ethylene glycol. Crystals were cryoprotected by soaking in a solution consisting of 100 mM Tris pH 8.5, 2.25 M ammonium sulfate, 5% ethylene glycol, 20% glycerol for 1 d. Crystals were scooped out from the cryoprotectant using cryoloops and flash-cooled in liquid nitrogen for storage and subsequent transport to the beamline.

2.3. Data collection and processing

Diffraction data were collected at 100 K using synchrotron radiation on the EMBL P13 beamline at the PETRA III storage ring, DESY, Hamburg, Germany. Data were processed using *XDS* (Kabsch, 2010) and scaled with *SCALA* (Evans, 2006). The structure was solved by molecular replacement using the *BALBES* pipeline (Long *et al.*, 2008). *BALBES* chose as the best model *E. coli* glucose-1-phosphate uridyltransferase (Thoden & Holden, 2007b; PDB entry 2e3d), which has a sequence identity of 89% with *E. amylovora* GalU. Table 1 reports a summary of data collection and processing.

Table 1

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	EMBL P13, PETRA III storage ring
Wavelength (Å)	1.033
Temperature (K)	100
Detector	Dectris PILATUS 6M
Crystal-to-detector distance (mm)	506.52
Rotation range per image (°)	0.05
Total rotation range (°)	90
Exposure time per image (s)	0.1
Space group	<i>P</i> 6 ₂
<i>a</i> , <i>b</i> , <i>c</i> (Å)	80.67, 80.67, 169.18
α , β , γ (°)	90, 90, 120
Mosaicity (°)	0.16
Resolution range (Å)	66.66–2.46 (2.56–2.46)
Total No. of reflections	107891 (7226)
No. of unique reflections	12434 (1284)
Completeness (%)	99.4 (94.9)
Multiplicity	8.7 (5.6)
$\langle I/\sigma(I) \rangle$	12.1 (3.4)
$R_{\text{r.i.m.}}^{\dagger}$	0.105 (0.472)
Overall <i>B</i> factor from Wilson plot (Å ²)	57.06

[†] Estimated by multiplying the conventional R_{merge} value by the factor $[N/(N-1)]^{1/2}$, where *N* is the data multiplicity [$R_{\text{merge}} = 0.099$ (0.428), data multiplicity = 8.7 (5.6)].

3. Results and discussion

Glucose-1-phosphate uridylyltransferase from *E. amylovora* CFPB1430 was cloned and expressed in *E. coli*. The *Nco*I restriction site at the N-terminus of the gene includes an in-frame ATG codon that corresponded to the first GalU codon in our cloning. Consequently, a Ser-to-Gly mutation of the second amino acid was necessary in order to complete the *Nco*I recognition sequence. As a consequence of the cloning strategy, after tag cleavage the N-terminus of the protein in the crystal was GAM₁G-GalU instead of M₁S-GalU (Met1 is underlined). EaGalU was overexpressed in *E. coli*, yielding on average 100 mg purified protein per litre of culture.

EaGalU molecular-mass determination by MALDI-TOF analysis gave a value of 33 249 Da, while the theoretical molecular mass (obtained by *ProtParam* calculation) was 32 887 Da. Size-exclusion chromatography suggested that the protein behaves as a tetramer in solution. The formation of homotetramers was expected, as it is a conserved characteristic of GalU in other organisms such as *E. coli* and *Corynebacterium glutamicum* (Thoden & Holden, 2007a,b).

The Thermofluor stability experiments determined a T_m of around 321 K. Activity tests were performed at 298 K monitoring the decrease of substrate and the increase of product against a timescale using NMR spectroscopy.

The crystals of *E. amylovora* glucose-1-phosphate uridylyltransferase diffracted to 2.46 Å resolution and the space-group determination assigned the crystal to the hexagonal space group *P*6₂ (Table 1). The structure was solved by molecular replacement with $R_{\text{work}} = 35.9\%$ and $R_{\text{free}} = 36.7\%$. The asymmetric unit contains two molecules with a solvent content of 52.09% and a Matthews coefficient of 2.57 Å³ Da^{−1}.

The authors thank Dr L. Caputi and Dr J. Wuerges for helpful discussion. We thank Dr M. Malnoy for providing a sample of

E. amylovora CFBP1430 genomic DNA. We thank Professor Steve Bottomley from Monash University for kindly providing the solubility-enhanced L56V/S135G TEV protease clone. We thank Dr S. Boivin and the Sample Preparation and Characterization (SPC) facility of EMBL-Hamburg for protein biophysical characterization. Plasmid pETM-11 was obtained from the European Molecular Biology Laboratory (EMBL) under a signed Material Transfer Agreement. Thanks are due to Professor G. Guella and R. Ferrazza for the NMR experiments. Data were collected under the European Molecular Biology Laboratory beamtime award No. MX-40. Thanks are due to Dr M. Serio for her useful suggestions throughout the work. This work was financially supported by the Autonomous Province of Bolzano (project: A Structural Genomics Approach for the Study of the Virulence and Pathogenesis of *E. amylovora*) and the Free University of Bolzano (project GAMES: Galactose and Glucuronic Acid Metabolism in *Erwinia* spp.).

References

- Bellemann, P. & Geider, K. (1992). *J. Gen. Microbiol.* **138**, 931–940.
- Bugert, P. & Geider, K. (1995). *Mol. Microbiol.* **15**, 917–933.
- Cabrita, L. D., Gilis, D., Robertson, A. L., Dehouck, Y., Rooman, M. & Bottomley, S. P. (2007). *Protein Sci.* **16**, 2360–2367.
- Caputi, L., Nepogodiev, S. A., Malnoy, M., Rejzek, M., Field, R. A. & Benini, S. (2013). *J. Agric. Food Chem.* **61**, 12265–12273.
- Chang, H.-Y., Lee, J.-H., Deng, W.-L., Fu, T.-F. & Peng, H.-L. (1996). *Microb. Pathog.* **20**, 255–261.
- Daran, J. M., Dallies, N., Thines-Sempoux, D., Paquet, V. & François, J. (1995). *Eur. J. Biochem.* **233**, 520–530.
- Dean, C. R. & Goldberg, J. B. (2002). *FEMS Microbiol. Lett.* **210**, 277–283.
- Dümmler, A., Lawrence, A. M. & de Marco, A. (2005). *Microb. Cell Fact.* **4**, 34.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Flores-Díaz, M., Alape-Girón, A., Persson, B., Pollesello, P., Moos, M., von Eichel-Streiber, C., Thelestam, M. & Florin, I. (1997). *J. Biol. Chem.* **272**, 23784–23791.
- Frey, P. A. (1996). *FASEB J.* **10**, 461–470.
- Jumel, K., Geider, K. & Harding, S. E. (1997). *Int. J. Biol. Macromol.* **20**, 251–258.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kleczkowski, L. A., Geisler, M., Ciereszko, I. & Johansson, H. (2004). *Plant Physiol.* **134**, 912–918.
- Koczan, J. M., Lenneman, B. R., McGrath, M. J. & Sundin, G. W. (2011). *Appl. Environ. Microbiol.* **77**, 7031–7039.
- Koczan, J. M., McGrath, M. J., Zhao, Y. & Sundin, G. W. (2009). *Phytopathology*, **99**, 1237–1244.
- Lai, X., Wu, J., Chen, S., Zhang, X. & Wang, H. (2008). *Protein Expr. Purif.* **61**, 50–56.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* **D64**, 125–132.
- Metzger, M., Bellemann, P., Bugert, P. & Geider, K. (1994). *J. Bacteriol.* **176**, 450–459.
- Mollerach, M., López, R. & García, E. (1998). *J. Exp. Med.* **188**, 2047–2056.
- Nimtz, M., Mort, A., Domke, T., Wray, V., Zhang, Y., Qiu, F., Coplin, D. & Geider, K. (1996). *Carbohydr. Res.* **287**, 59–76.
- Thoden, J. B. & Holden, H. M. (2007a). *Protein Sci.* **16**, 1379–1388.
- Thoden, J. B. & Holden, H. M. (2007b). *Protein Sci.* **16**, 432–440.
- Vilches, S., Canals, R., Wilhelms, M., Saló, M. T., Knirel, Y. A., Vinogradov, E., Merino, S. & Tomás, J. M. (2007). *Microbiology*, **153**, 2393–2404.
- Vrancken, K., Holtappels, M., Schoofs, H., Deckers, T. & Valcke, R. (2013). *Microbiology*, **159**, 823–832.
- Yother, J. (2011). *Annu. Rev. Microbiol.* **65**, 563–581.