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

Protein preparation, crystallization and preliminary crystallographic studies of *Bacillus subtilis* glycinamide ribonucleotide transformylase

Yu-He Liang, Xiang-Yu Liu, Juan Wang and Lan-Fen Li*

The National Laboratory of Protein Engineering and Plant Genetic Engineering, College of Life Sciences, Peking University, Beijing 100871, People's Republic of China

Correspondence e-mail: lilf@pku.edu.cn

Received 26 March 2009
Accepted 1 June 2009

Glycinamide ribonucleotide transformylase (GART) catalyzes the transfer of a formyl group from formyl tetrahydrofolate (FTHF) to glycinamide ribonucleotide (GAR), which is an essential step in the *de novo* synthesis pathway of purines. In *Bacillus subtilis*, GART is encoded by the gene *purN*. In order to study the structure and function of *B. subtilis* GART, the *purN* gene was amplified, cloned into an expression vector and expressed in soluble form in *Escherichia coli*. The protein was purified to homogeneity and crystals suitable for X-ray data collection were obtained. These crystals diffracted to 2.5 Å resolution and belonged to space group $P3_121$, with unit-cell parameters $a = b = 95.5$, $c = 64.0$ Å.

1. Introduction

The *de novo* purine-biosynthesis pathway contains ten enzymatic steps leading from 5-phosphoribosyl pyrophosphate (PRPP) to inosine monophosphate (IMP). IMP is then converted to adenosine monophosphate (AMP) and guanosine monophosphate (GMP). In *Bacillus subtilis*, the genes encoding the enzymes catalyzing the ten steps of the biosynthetic pathway are located in a 12-gene cluster and are organized into three groups of overlapping genes [*purEKB*, *purC(orf)QLF* and *purMNH(J)*] followed by the last singly encoding *purD* gene (Ebbole & Zalkin, 1987; Rappu *et al.*, 1999). The 12-gene *pur* operon has been cloned and sequenced and initial studies of its expression have been reported (Ebbole & Zalkin, 1989). The gene *purN* encodes the folate-dependent glycinamide ribonucleotide transformylase (GART), which catalyzes the transfer of a formyl group from formyl tetrahydrofolate (FTHF) to glycinamide ribonucleotide (GAR), resulting in formyl-glycinamide ribonucleotide (FGAR). This reaction is the third step of the PRPP to IMP biosynthesis pathway (Warren & Buchanan, 1957).

The crystal structures of *Escherichia coli* GART (Almassy *et al.*, 1992; Chen *et al.*, 1992; Klein *et al.*, 1995; Greasley *et al.*, 1999, 2001), *Mycobacterium tuberculosis* GART (Zhang *et al.*, 2009) and the human GART domain (Varney *et al.*, 1997; Zhang *et al.*, 2002, 2003; Dahms *et al.*, 2005) have been determined. The catalytic mechanisms of GART have been studied in some detail (Caperelli & Giroux, 1997; Shim & Benkovic, 1998, 1999; Morikis *et al.*, 2001; Su *et al.*, 1998; Zhang *et al.*, 2002; Manieri *et al.*, 2007). GART is found as a single protein in prokaryotes, while in most eukaryotes it constitutes the C-terminal part of a monomeric but trifunctional protein with PurD–PurM–PurN activities (Almassy *et al.*, 1992). GART is well conserved among different species: *B. subtilis* GART shows 31% (56/178) and 38% (68/181) sequence identity to the *E. coli* enzyme and the C-terminal part of the human enzyme, respectively. GART has been a known target for structure-based drug design in the anticancer therapy field for about 20 years (Taylor *et al.*, 1985; Zhang *et al.*, 2003; Deng *et al.*, 2008). Studies of the *B. subtilis* GART structure and comparisons with other known structures would provide further information on the atomic mechanism of formyl transfer in the synthesis of DNA-precursor purines, which will assist in the structure-based design of antitumour agents.

In this work, the *B. subtilis purN* gene was cloned, the protein was purified to homogeneity and crystals suitable for further X-ray diffraction studies were obtained.

2. Materials and methods

2.1. Cloning and expression

Genomic *B. subtilis* DNA was used as a template to amplify the *purN* gene (GeneID 936045) by the polymerase chain reaction (PCR). The primers for PCR were 5′-CAC CAT GAA AAA GTT TGC GGT ATT-3′ and 5′-TCA TGC CTT TTC ACC TCT GTT-3′. An entry clone for the Gateway cloning system was created by inserting the PCR product into a pENTR/D-TOPO vector using pENTR Directional TOPO Cloning Kits (Invitrogen, USA). Positive clones carrying the *purN* gene were identified by colony PCR. The target gene was then subcloned into a modified pET vector, pET21-DEST, with an N-terminal His_5 tag *via* an LR recombination reaction (Ding *et al.*, 2002). The LR product was transformed into DH5α (Invitrogen) competent *E. coli* cells. The positive clones were again confirmed by colony PCR and transformed into *E. coli* BL21 (DE3) strain (Invitrogen) for protein expression. The transformed BL21 (DE3) cells were grown aerobically in Luria–Bertani (LB) medium containing 100 µg ml^{-1} ampicillin at 310 K until an OD_{600} of between 0.6 and 0.8 was reached. The cells were then induced with 1 m*M* isopropyl β-D-1-thiogalactopyranoside and grown for a further 4 h. The cells were then harvested by centrifugation at 3000*g* for 10 min at 277 K.

2.2. Protein purification

The cell pellets were resuspended in ice-cold lysis buffer (20 m*M* Tris–HCl, 500 m*M* NaCl pH 7.5) containing 1 m*M* phenylmethylsulfonyl fluoride and disrupted by sonication. The crude cell extract was centrifuged at 31 000*g* for 1 h at 277 K. The supernatant was loaded onto a 5 ml HiTrap Ni^{2+}-chelating column (GE Healthcare, USA) which was equilibrated with lysis buffer. The protein was eluted with a linear gradient of 50–500 m*M* imidazole in lysis buffer. The fractions containing the target protein were pooled and concentrated to about 10 mg ml^{-1} using an ultrafiltration device with a 10 kDa cutoff (Millipore, USA). The protein concentration was measured using a Bio-Rad protein-assay kit (Bio-Rad Laboratories, USA) based on the Bradford method. The concentrated protein was then loaded onto a Superdex 75 XK 16/60 gel-filtration column (GE Healthcare, USA) previously equilibrated with elution buffer (20 m*M* Tris–HCl, 100 m*M* NaCl pH 7.5). Gel filtration was carried out at 277 K in elution buffer and the purified protein was concentrated to 30 mg ml^{-1} for crystallization trials. The purity of the protein was checked by SDS–PAGE.

2.3. Crystallization

Crystallization was performed by the hanging-drop vapour-diffusion method at 296 K. Each hanging drop, prepared by mixing 1 µl reservoir solution and 1 µl protein solution (30 mg ml^{-1}, containing 20 m*M* Tris–HCl, 100 m*M* NaCl pH 7.5), was equilibrated against 0.5 ml reservoir solution. Initial crystallization conditions were screened using several commercially available crystallization screening kits (Hampton Research, USA). Crystals were observed after only 2 d in several conditions and crystals suitable for diffraction experiments were obtained using a reservoir solution containing 0.2 *M* lithium sulfate, 0.1 *M* bis-tris pH 5.5, 25%(*w*/*v*) polyethylene glycol (PEG) 3350. The crystals reached their maximum dimensions of 1.0 × 0.2 × 0.1 mm within one week (Fig. 1).

Table 1
Crystallographic parameters and data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P3_121$
Unit-cell parameters (Å)	$a = b = 95.5$, $c = 64.0$
Resolution (Å)	50–2.50 (2.61–2.50)
Data completeness (%)	96.57 (86.0)
No. of observed reflections	76750
No. of unique reflections	11547
$R_{\rm merge}$† (%)	8.18 (38.7)
$\langle I/\sigma(I)\rangle$	7.7 (1.6)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

2.4. Data collection

X-ray diffraction data were collected on a Bruker SMART 6000 CCD detector using Cu $K\alpha$ radiation from a Bruker–Nonius FR591 rotating-anode generator operated at 45 kV and 100 mA. Fine-sliced oscillation images (2000 frames, 0.2°) were collected with an exposure time of 120 s per frame. The crystal-to-detector distance was set to 6 cm. The crystal was flash-cooled and maintained at 100 K in a cold nitrogen stream during data collection without any additional cryoprotectant. The data were processed using the Bruker *PROTEUM* software suite.

3. Results and discussion

The *B. subtilis* GART protein was expressed in a soluble form and without protein aggregation as judged from the gel-filtration elution profile. The protein was expressed with a typical yield of 8 mg of pure protein per litre of cell culture. The protein has an estimated molecular weight of ∼26 kDa and showed greater than 95% purity from SDS–PAGE analysis. This molecular weight matched the theoretical molecular weight of the *B. subtilis* GART protein (21.7 kDa) plus 4 kDa from the His_5 peptide tag (MASMTGGQQMGSSHHHHH-SSGLVPRGSQSTSLYKKAGL) remaining from the cloning procedure.

The best crystal diffracted to 2.5 Å resolution. The structure has been determined by molecular replacement using the human GART

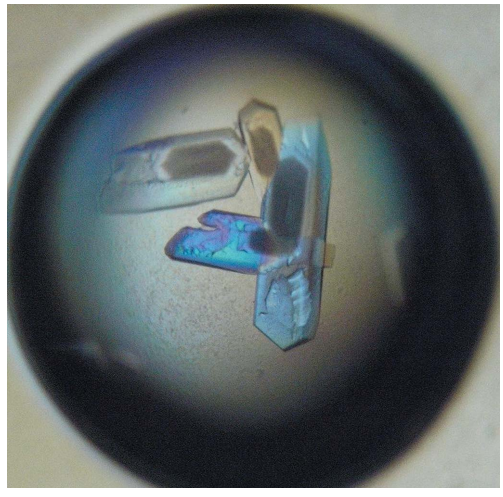

Figure 1
Crystals of *B. subtilis* GART obtained from a reservoir solution containing 0.2 *M* lithium sulfate, 0.1 *M* bis-tris pH 5.5, 25%(*w*/*v*) PEG 3350.

structure (PDB code 1zly; Dahms *et al.*, 2005), which showed 38% sequence identity to the *B. subtilis* GART, as a search model. Molecular replacement was performed with the program *MOLREP* (Vagin & Teplyakov, 1997) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). The $R_{\rm cryst}$ and $R_{\rm free}$ values were 39.7% and 48.1%, respectively, after the first round of rigid-body refinement. There was one molecule per asymmetric unit, corresponding to a solvent content of 63%, with a calculated Matthews coefficient $V_{\rm M}$ of 3.3 Å^3 Da^{-1} (Matthews, 1968). The crystal belonged to space group $P3_121$, with unit-cell parameters $a = b = 95.5$, $c = 64.0$ Å. The crystallographic parameters and data-collection statistics are listed in Table 1. Structure refinement and analysis are in progress.

This work was supported by National Natural Science Foundation of China (NSFC) grant 30600114 to Y-HL. This work was also supported by Peking University's 985 and 211 grants. We are grateful to Dr Ole Kristensen for comments and for reading the manuscript.

References

Almassy, R. J., Janson, C. A., Kan, C. C. & Hostomska, Z. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 6114–6118.

Caperelli, C. A. & Giroux, E. L. (1997). *Arch. Biochem. Biophys.* **341**, 98–103.

Chen, P., Schulze-Gahmen, U., Stura, E. A., Inglese, J., Johnson, D. L., Marolewski, A., Benkovic, S. J. & Wilson, I. A. (1992). *J. Mol. Biol.* **227**, 283–292.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Dahms, T. E., Sainz, G., Giroux, E. L., Caperelli, C. A. & Smith, J. L. (2005). *Biochemistry*, **44**, 9841–9850.

Deng, Y., Wang, Y., Cherian, C., Hou, Z., Buck, S. A., Matherly, L. H. & Gangjee, A. (2008). *J. Med. Chem.* **51**, 5052–5063.

Ding, H. *et al.* (2002). *Acta Cryst.* D**58**, 2102–2108.

Ebbole, D. J. & Zalkin, H. (1987). *J. Biol. Chem.* **262**, 8274–8287.

Ebbole, D. J. & Zalkin, H. (1989). *J. Bacteriol.* **171**, 2136–2141.

Greasley, S. E., Marsilje, T. H., Cai, H., Baker, S., Benkovic, S. J., Boger, D. L. & Wilson, I. A. (2001). *Biochemistry*, **40**, 13538–13547.

Greasley, S. E., Yamashita, M. M., Cai, H., Benkovic, S. J., Boger, D. L. & Wilson, I. A. (1999). *Biochemistry*, **38**, 16783–16793.

Klein, C., Chen, P., Arevalo, J. H., Stura, E. A., Marolewski, A., Warren, M. S., Benkovic, S. J. & Wilson, I. A. (1995). *J. Mol. Biol.* **249**, 153–175.

Manieri, W., Moore, M. E., Soellner, M. B., Tsang, P. & Caperelli, C. A. (2007). *Biochemistry*, **46**, 156–163.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Morikis, D., Elcock, A. H., Jennings, P. A. & McCammon, J. A. (2001). *Protein Sci.* **10**, 2379–2392.

Rappu, P., Shin, B. S., Zalkin, H. & Mantsala, P. (1999). *J. Bacteriol.* **181**, 3810–3815.

Shim, J. H. & Benkovic, S. J. (1998). *Biochemistry*, **37**, 8776–8782.

Shim, J. H. & Benkovic, S. J. (1999). *Biochemistry*, **38**, 10024–10031.

Su, Y., Yamashita, M. M., Greasley, S. E., Mullen, C. A., Shim, J. H., Jennings, P. A., Benkovic, S. J. & Wilson, I. A. (1998). *J. Mol. Biol.* **281**, 485–499.

Taylor, E. C., Harrington, P. J., Fletcher, S. R., Beardsley, G. P. & Moran, R. G. (1985). *J. Med. Chem.* **28**, 914–921.

Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.

Varney, M. D., Palmer, C. L., Romines, W. H. III, Boritzki, T., Margosiak, S. A., Almassy, R., Janson, C. A., Bartlett, C., Howland, E. J. & Ferre, R. (1997). *J. Med. Chem.* **40**, 2502–2524.

Warren, L. & Buchanan, J. M. (1957). *J. Biol. Chem.* **229**, 613–626.

Zhang, Y., Desharnais, J., Greasley, S. E., Beardsley, G. P., Boger, D. L. & Wilson, I. A. (2002). *Biochemistry*, **41**, 14206–14215.

Zhang, Y., Desharnais, J., Marsilje, T. H., Li, C., Hedrick, M. P., Gooljarsingh, L. T., Tavassoli, A., Benkovic, S. J., Olson, A. J., Boger, D. L. & Wilson, I. A. (2003). *Biochemistry*, **42**, 6043–6056.

Zhang, Z., Caradoc-Davies, T. T., Dickson, J. M., Baker, E. N. & Squire, C. J. (2009). *J. Mol. Biol.* **389**, 722–733.