*Acta Cryst.* (1998). D**54**, 438–440

# Crystallization and preliminary X-ray diffraction studies of *E. coli* porphobilinogen synthase and its heavy-atom derivatives

L. Shimoni-Livny, H. L. Carrell, T. Wagner, A. Kaufman Katz, C. Afshar, L. W. Mitchell,† M. Volin, E. K. Jaffe and J. P. Glusker* *at Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, PA 19111, USA. E-mail: jp_glusker@fccc.edu*

(*Received 2 May 1997; accepted 6 August 1997*)

**Abstract**

Porphobilinogen synthase (PBGS) catalyzes the condensation of two identical substrate molecules, 5-aminolevulinic acid (ALA), in an asymmetric manner to form porphobilinogen. *E. coli* PBGS is an homooctameric enzyme. The number of active sites is not clear, but each subunit binds one $Zn^{II}$ ion and one $Mg^{II}$ ion. Diffraction-quality crystals of native *E. coli* PBGS have been obtained, and unit-cell dimensions ($a = 130.8$, $c = 144.0$ Å) are reported. These crystals diffract to about 3.0 Å resolution.

## 1. Introduction

The metalloprotein porphobilinogen synthase (PBGS) catalyzes the asymmetric condensation of two molecules of 5-aminolevulinic acid (ALA) to form porphobilinogen (PBG), the common substituted pyrrole precursor of porphyrins, corrins, chlorins and other tetrapyrroles (Shemin & Russell, 1953). *E. coli* PBGS is an homooctameric (280 kDa) enzyme with either four or eight active sites per molecule (Mitchell *et al.*, 1995; Senior *et al.*, 1996; Cheung *et al.*, 1997). Characterization of *E. coli* PBGS has shown it to be similar to bovine PBGS in that there are eight $Zn^{II}$ ions per octamer (Jaffe *et al.*, 1994). One major difference between *E. coli* and mammalian PBGS is that *E. coli* PBGS also has eight equivalent $Mg^{II}$ binding sites per octamer (Mitchell & Jaffe, 1993; Jaffe *et al.*, 1995) in addition to the eight $Zn^{II}$ present in both proteins. There is 41% sequence identity between *E. coli* PBGS and mammalian PBGS. We are studying the structure of *E. coli* PBGS in order to determine its modes of folding and catalytic activity, and to determine the number of active sites (four or eight) per molecule.

## 2. Materials and methods

*E. coli* PBGS, expressed from the construct RP523(pLM1228), was purified according to the procedure of Mitchell *et al.* (1995). For reasons which may be related to the physiology of the host *E. coli* strain RP523, the *E. coli* PBGS expressed in this construct crystallizes much more readily than that expressed in the original construct TB1(pCR261) (Roessner *et al.*, 1995), and therefore is the form that we describe here. The host RP523 is a hemin permeable $hemB^-$ strain of *E. coli* (Humanoff *et al.*, 1989). The plasmid pLM1228 contains a 1400 bp fragment excised from the *Eco*RI and *Bam*HI sites of pCR261 and inserted into the multiple cloning site of pUC119. The expression of PBGS by RP523(pLM1228) is constitutive. Mass spectral analysis of an AspN protease digest of protein purified from RP523(pLM1228) reveals all the expected peptides of wild-type *E. coli* PBGS (Jaffe, personal communication). The recombinant enzyme is fully active and requires the presence of its substrate, ALA, for crystallization. This substrate is then stoichiometrically converted to product. Purified enzyme in a buffer solution composed of 0.1 *M* potassium phosphate (pH 7.0), 10 m*M* 2-mercaptoethanol, 10 µm $Zn^{II}$ and 1 m*M* $Mg^{II}$, was dialyzed against 50 m*M* TES [*N*-tris-(hydroxymethyl)methyl-2-aminoethanesulfonic acid] (TES–KOH, pH 7.0), 10 µ*M* 2-mercaptoethanol, 10 µ*M* $Zn^{II}$ and 1 m*M* $Mg^{II}$ to prevent the formation of phosphate crystals.

Crystals of the *E. coli* enzyme from RP523(pLM1228) were obtained by the vapor-diffusion method using hanging drops (McPherson, 1982). Equilibration of drops containing approximately 3.2 mg $ml^{-1}$ of protein solution, plus 0.44 *M* magnesium sulfate solution in 50 m*M* TES–KOH (pH 8.2), against 1.1 *M* magnesium sulfate in 50 m*M* TES–KOH (pH 8.2), and 2 m*M* ALA resulted in crystals which appeared in approximately 24 h at room temperature in the dark.

Two heavy-atom derivatives have been obtained to date. That of the uranyl heavy-atom derivative was obtained by soaking crystals of the native *E. coli* enzyme with 0.05 m*M* uranyl nitrate [$UO_2(NO_3)_2$] in a 10 µl drop of magnesium sulfate solution (taken from the bath used to produce the native crystals). The crystals so obtained were soaked and

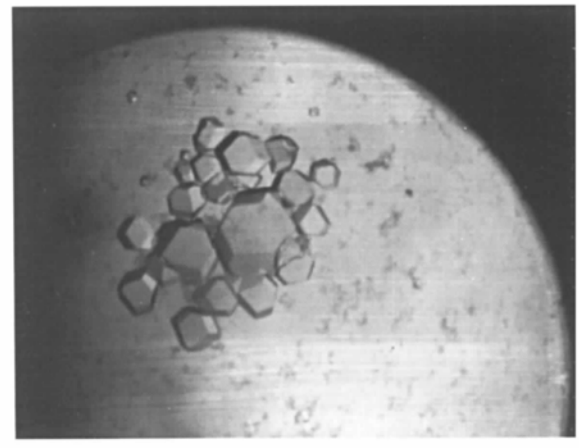

Fig. 1. Crystals of *E. coli* PBGS.

† Current address: Chemistry Department, St Joseph University, 5600 City Avenue, Philadelphia, PA 19131, USA.

Table 1. *Data collection and statistics for native E. coli PBGS and its heavy-atom derivatives*

| | Native | Uranyl derivative | Samarium derivative |
|---|---|---|---|
| Maximum resolution (Å) | 3.0 | 3.4 | 3.4 |
| Number of measured reflections | 229010 | 243199 | 189858 |
| Number of unique reflections | 19618 | 17611 | 12672 |
| Completeness to the maximum resolution shell (%) | 86.3 | 99.8 | 98.1 |
| $R_{merge}$ (%) | 6.9 | 8.4 | 8.8 |
| Soaking concentration (m*M*) | — | 0.05 | 1 |
| Soaking time (h) | — | 5 | 12 |

left in the dark for 5 h, and then back-soaked in a solution of magnesium sulfate in 20% glycerol for 10 min before freezing. A crystal of a samarium heavy-atom derivative was obtained by soaking the native *E. coli* enzyme crystals with 1 m*M* samarium chloride ($SmCl_3 \cdot 6H_2O$) in a 10 µl drop of magnesium sulfate (as described for the uranyl derivative). The crystal was soaked and left in the dark for 12 h, and then back-soaked in a solution of magnesium sulfate in 20% glycerol for 10 min before freezing.

## 3. Results

Colorless crystals of *E. coli* PBGS grew with a multifaceted gem-like shape (see Fig. 1). In general, the crystals are about 0.18 × 0.12 × 0.08 mm in dimensions. X-ray diffraction data for all three crystals (native, uranyl and samarium derivatives) were measured at the beamline X12B at Brookhaven National Laboratory with a MAR Research image-plate system (300 mm plate) using two phases. The crystal-to-detector distance was 275 mm for the native *E. coli* enzyme and 300 mm for the uranyl and samarium derivatives. A 60 s exposure time per frame was employed for all three crystals. Wavelengths used for measuring diffraction data from the native and the uranyl derivative were 1.14 Å, and 1.071 Å for the samarium derivative. The diffraction data are 86.3% complete to 3.0 Å resolution. The diffraction data were processed with the programs *DENZO* and *SCALEPACK* (Otwinowski, 1993). The space group was found to be *I*422, from the observed systematic absences in diffraction data; this space group is the same as that for the bovine PBGS (Carrell *et al.*, 1996), but with different unit-cell parameters ($a$ = 124.8, $c$ = 200.2 for the bovine PBGS). The unit-cell parameters obtained for native *E. coli* PBGS are $a$ = 130.8 and $c$ = 144.0 Å. The total number of reflections measured was 229 010 which gave 19 618 unique reflections to 3.0 Å resolution. The value of $V_m$, the volume per mass unit (Matthews, 1968) is 2.22 $Å^3\ Da^{-1}$ for 32 monomers (four octamers) in the unit cell, compared with 2.80 $Å^3\ Da^{-1}$ for the bovine enzyme.

Crystals of the uranyl and samarium derivatives exhibit diffraction and appear to have virtually identical unit-cell dimensions ($a$ = 130.3 and $c$ = 144.1 Å for the uranyl derivative and $a$ = 130.0 and $c$ = 142.3 Å for the samarium derivative) and the same space group, *I*422. The total number of reflections measured for the uranyl derivative was 243 199 which gave 17 611 unique reflections, and gave 99.8% completeness of data to 3.4 Å resolution. The total number of reflections measured for the samarium derivative was 189 858 which gave 12 672 unique reflections (98.1% completeness of data to 3.4 Å resolution). Statistical data for the data collected for the three crystals are summarized in Table 1.

The diffraction data from the native *E. coli* enzyme and the derivatives listed in Table 1 were merged using *CAD* (*CCP*4, Collaborative Computational Project, Number 4, 1994) and then scaled using *SCALEIT* (*CCP*4, Collaborative Computational Project, Number 4, 1994). Difference Patterson maps were solved for the two derivatives, and cross-validated by inspection of the Fourier maps, using *FFT* (*CCP*4, Collaborative Computational Project, Number 4, 1994). The program *MLPHARE* (Otwinowski, 1991) was used to refine the heavy-atom parameters and calculate the phases.

The relationship between the two subunits in the dimer, indicated by a self-rotation function, using *ALMN* (*CCP*4, Collaborative Computational Project, Number 4, 1994), was found to be a non-crystallographic twofold rotation axis perpendicular to the $c$ axis and 22.5° from the $a$ or $b$ axis. This result was also found for the bovine PBGS (Carrell *et al.*, 1996). It is interesting to note that the bovine enzyme is clearly demonstrated to contain only four active sites (Jaffe, 1995). A structure determination of the *E. coli* enzyme is currently in progress, using isomorphous replacement data.

Diffraction data were measured at the National Synchrotron Light Source, Brookhaven National Laboratory, which is supported by the US Department of Energy, Division of Material Sciences and Division of Chemical Sciences. We thank Malcolm Capel and Tom McDonald for the use of their facilities at beamline X12B. Financial support from the DAAD (NATO grant to TW) is gratefully acknowledged. This work was supported by NIH grants CA10925 (JPG), ES03654 (EKJ), CA06927 (ICR), and an appropriation from the Commonwealth of Pennsylvania. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute. This work is part of the dissertation of LS-L for PhD degree at the University of Pennsylvania.

### References

Carrell, H. L., Glusker, J. P., Shimoni, L., Keefe, L. J., Afshar, C., Volin, M. & Jaffe, E. K. (1996). *Acta Cryst.* D**52**, 419–421.
Cheung, K.-M., Spencer, P., Timko, M. P. & Shoolingin-Jordan, P. (1997). *Biochemistry*, **36**, 1148–1156.
Collaborative Computational Project, Number 4. (1994). *Acta Cryst.* D**50**, 760–763.
Humanoff, H., Russell, C. S. & Cosloy, S. D. (1989). *J. Bacteriol.* **170**, 4969–4971.
Jaffe, E. K. (1995). *J. Bioenerget. Biomembr.* **27**, 169–179.
Jaffe, E. K., Volin, M., Myers, C. B. & Abrams, W. R. (1994). *Biochemistry*, **33**, 11554–11562.
Jaffe, E. K., Ali, S., Mitchell, L. W., Taylor, K. M., Volin, M. & Markham, G. D. (1995). *Biochemistry*, **34**, 244–251.
McPherson, A. (1982). *Preparation and Analysis of Protein Crystals.* New York: John Wiley.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mitchell, L. W. & Jaffe, E. K. (1993). *Arch. Biochem. Biophys.* **300**, 169–177.
Mitchell, L. W., Volin, M. & Jaffe, E. K. (1995). *J. Biol. Chem.* **270**, 24054–24059.

Otwinowski, Z. (1991). *Maximum Likelihood Refinement of Heavy-Atom Parameters*. In *Proceedings of the CCP4 Study Weekend*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80–88. Warrington: Daresbury Laboratory.

Otwinowski, Z. (1993). *Oscillation Data Reduction Program*. In *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Roessner, C. A., Spencer, J. B., Ozaki, S., Min, C. H., Atshaves, B. P., Nayar, P., Anousis, N., Stolowich, N. J., Holderman, M. T., Scott, A. I. (1995). *Protein Expr. Purif.* **6**, 155–163.

Senior, N. M., Brocklehurst, K., Cooper, J. B., Wood, S. P., Erskin, P., Shoolingin-Jordan, P. M., Thomas, P. G. & Warren, M. J. (1996). *Biochem. J.* **320**, 401–412.

Shemin, D. & Russell, C. S. (1953). *J. Am. Chem. Soc.* **75**, 4873–4874.