Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary X-ray diffraction data from selenomethionine glycinamide ribonucleotide synthetase

Todd M. Weaver,† Weiru Wang and Steven E. Ealick*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853, USA

† Present address: Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455, USA.

Correspondence e-mail: see3@cornell.edu

Received 13 April 1998
Accepted 2 July 1998

In this study, the overexpression, purification and crystallization of selenomethionine (SeMet) incorporated glycinamide ribonucleotide synthetase (GAR-syn) from *Escherichia coli* are reported. The overexpression of SeMet GAR-syn was placed under the control of the isopropylthio-β-galactoside (IPTG) inducible T7 RNA-polymerase system. The newly developed construct contained a removable histidine tag on the amino terminus of GAR-syn, which allowed rapid purification using metal-chelate chromatography techniques. The SeMet GAR-syn crystals were grown by hanging-drop vapor diffusion and belong to the space group $P2_12_12_1$ with unit-cell parameters a = 56.2, b = 62.4 and c = 129.8 Å and a single monomer in the asymmetric unit. The crystals diffract to 1.6 Å resolution and have led to the determination of multiple-wavelength anomalous diffraction phases to 2.2 Å resolution.

1. Introduction

Glycinamide ribonucleotide synthetase (GAR-syn) catalyzes the reversible conversion of phosphoribosylamine (PRA), MgATP and glycine to glycinamide ribonucleotide (GAR), MgADP and phosphate. This reaction is the second of eleven catalytic steps in the *de novo* purine biosynthetic pathway. GAR-syn belongs to a class of enzymes that form amide bonds through ATP activation of a carboxyl group to form ADP and a phosphoanhydride intermediate, with subsequent condensation to generate the amide and phosphate. Other well studied structurally characterized enzymes that also belong to this family include D-alanine–D-alanine ligase (Fan *et al.*, 1994), glutathione synthetase (Yamaguchi *et al.*, 1993) and biotin carboxylase (Waldrop *et al.*, 1994).

The first step in the purine biosynthetic pathway catalyzed by phosphoribosylpyrophosphate amidotransferase (PRPP-AT) produces PRA, which is unstable, having a half-life of 5 s under physiological conditions (Mueller *et al.*, 1994). A recent kinetic study (Rudolph & Stubbe, 1995) suggested that the transfer of this unstable intermediate from PRPP-AT to GAR-syn takes place by a process defined as substrate channeling, which would be facilitated by specific protein–protein interactions. The X-ray structure of PRPP-AT has been previously determined (Kim *et al.*, 1996). Therefore, the GAR-syn structure would open the possibility of addressing channeling from the structural point of view.

Recent advances in crystallography have proven that the multiwavelength anomalous diffraction (MAD) method is a powerful approach for phasing of macromolecular crystal structures (Hendrickson, 1991), and selenomethionine incorporation has been shown to be generally useful for MAD phasing (Hendrickson *et al.*, 1990). GAR-syn contains 14 methionine residues out of a total of 429 amino-acid residues, and selenomethionine incorporation thus provides an attractive alternative to heavy-atom phasing methods. In this paper, we describe the overexpression, crystallization and preliminary X-ray diffraction data for SeMet GAR-syn.

2. Materials and methods

2.1. Re-engineering of the gene

In order to obtain large amounts of highly purified GAR-syn in a rapid fashion, a removable histidine tag containing a thrombin cleavage site was engineered onto the amino terminus and the resulting construct was placed under the control of the IPTG-inducible T7 RNA-polymerase expression system (Studier *et al.*, 1990). The source of the *pur*D gene was derived from plasmid pJS187 (Shen *et al.*, 1990). The full-length *pur*D gene was amplified by the polymerase chain reaction (PCR) using the following primers: 5′ primer, 5′-GGAATTCATGAAAGTATTAGTGATT-GG-3′; 3′ primer, 5′-ACTAAGCTTTTAGT-TCTGCTCGCGTTCGA-3′. Both primers above contain the coding sequence for the *pur*D gene along with either an *Eco*RI restriction site for the 5′-primer or an *Hin*dIII restriction site for the 3′-primer (underlined). The recombinant gene was cloned into pET28a.1 between the *Eco*RI and *Hin*dIII restriction sites. This construct was used to

Table 1
X-ray statistics for SeMet GAR-syn data collected at CHESS.

Data set	λ (Å)	Resolution (Å)	Total number of reflections	Number of unique reflections	I/σ	Overall completeness (%)	Completeness (highest shell) (%)	R_{sym}† (%)
Monochromatic	0.9190	1.6	314135	57571	6.0	97.2	88.0	4.4 (17.4)‡
MAD-λ_1	0.97941	2.2	139716	22702	22.0	97.5	80.9	6.5 (9.3)
MAD-λ_2	0.979104	2.2	138670	22800	22.0	97.7	81.7	7.5 (11.2)
MAD-λ_3	0.967642	2.2	141733	22820	22.0	97.8	83.9	6.7 (10.1)

† $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,i} - \langle I_{hkl}\rangle| / \sum_{hkl}\sum_i I_{hkl}$. ‡ Numbers in parentheses correspond to the highest resolution shell, which is 2.25–2.20 Å for MAD data and 1.69–1.60 Å for monochromatic data.

transform the methionine auxotroph B834(DE3) obtained from Novagen.

2.2. Large-scale growth and purification

A single colony was transferred to 5.0 ml of Luria broth containing 30 mg ml^{-1} of kanamycin. This culture was allowed to grow for 15 h, after which 3.0 ml was harvested *via* micro-centrifugation in a 1.5 ml Eppendorf tube. The cell pellet was resuspended with M9 minimal media supplemented with 40 mg l^{-1} SeMet and recentrifuged. This process was repeated three successive times. The re-suspended culture was used to inoculate 1.4 l of M9 media supplemented with 40 mg l^{-1} of SeMet contained in 2.8 l Fernbach flasks. The cells were grown at 310 K to an OD_{600} of between 0.6 and 1.0, after which they were induced with 1 m*M* IPTG. In conjunction with IPTG, an additional aliquot of 20 mg l^{-1} of SeMet was added to each flask to guarantee full incorporation of SeMet within GAR-syn. The cells were induced overnight for at least 12 h, after which time they were harvested by centrifugation at 10000*g* for 30 min. Cell pellets were either stored at 193 K or utilized for purification of SeMet GAR-syn directly.

Frozen cell pellets were thawed into buffer *A* (50 m*M* Na/K phosphate pH 7.8, 300 m*M* NaCl, 10% glycerol). The cells were lysed by sonication in four 2 min time intervals, while the temperature was maintained at 277 K using a dry ice/2-propanol bath. All subsequent steps were performed at 277 K unless otherwise noted. The cellular debris was removed by centrifugation at 17000*g* for 40 min. Nucleic acids were removed by adjusting the supernatant to 1.0% in protamine sulfate, equilibrating for an additional 20 min and re-centrifuging at 17000*g* for 20 min. A total of 10 ml of Ni^{2+}-nitriloacetic acid was then added directly to the supernatant from the preceding step, and this slurry was stirred for 30 min. After centrifugation of the slurry, the supernatant was removed and the resin poured into a 5.0 × 10 cm column. The column was washed until the absorbance at 280 nm was less than 0.03.

Next, a step gradient of increasing imidazole concentration was performed, starting from 5 m*M* and increasing to a final concentration of 400 m*M*. The elution steps were as follows: (i) 5 m*M*, (ii) 10 m*M*, (iii) 40 m*M* and (iv) 400 m*M* imidazole. Essentially, all contaminants were removed during the initial wash steps performed at 5 m*M* and 10 m*M* imidazole, while purified SeMet GAR-syn was eluted during the 40 m*M* and 400 m*M* imidazole steps. After each elution, 5 m*M* ethylenediaminetetraacetic acid was added to the collection vessel prior to the addition of 10 m*M* dithiothreitol (DTT), which was used to prevent oxidation of the SeMet residues. Purified GAR-syn was concentrated to 20 mg ml^{-1} using an Amicon stir cell fitted with a PM-30 membrane, and then dialyzed against 4 l of 10 m*M* Tris acetate pH 7.5, 10 m*M* DTT. After concentration and dialysis, 1.0 ml aliquots of SeMet GAR-syn were flash frozen with liquid nitrogen and stored at 193 K.

2.3. Crystallization

Small initial crystals of SeMet GAR-syn were obtained using the sparse-matrix protocol (Jancarik & Kim, 1991). Optimized crystallization conditions were obtained using the hanging-drop vapor-diffusion method at 291 K. The protein concentration was 20 mg ml^{-1} and the well contained 0.1 *M* 2-(*N*-morpholino) ethanesulfonic acid (MES) pH 6.3, 0.2 *M* ammonium sulfate and 26% polyethylene glycol 5K monomethylether. The hanging drop was prepared by mixing 2 µl of protein solution containing 10 m*M* DTT and 2 µl of well solution. Crystals appeared within 2 d and reached maximum size over the course of an

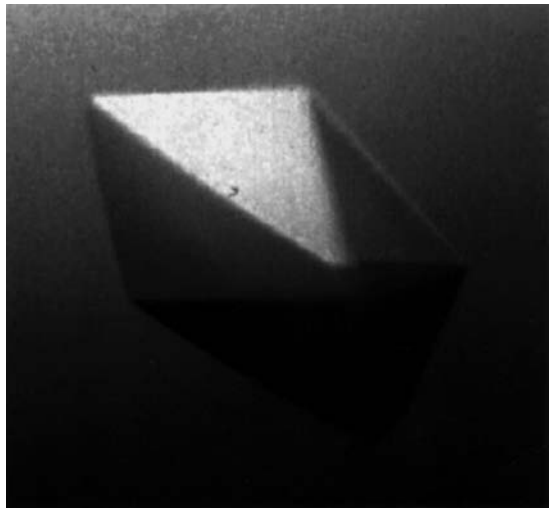

Figure 1
A crystal of SeMet GAR-syn. Crystals appear after 2 d and reach dimensions of 0.2 × 0.3 × 0.6 mm after one week.

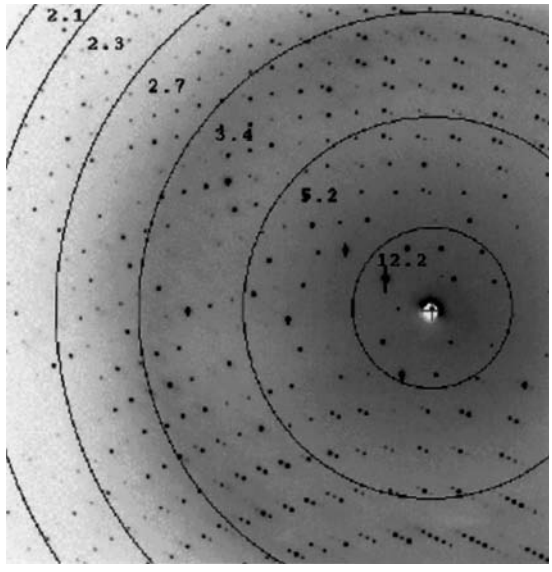

Figure 2
A typical diffraction pattern from SeMet GAR-syn collected on the F1 beamline at CHESS. As is evident in the illustration, the SeMet GAR-syn crystals diffract to beyond 2.1 Å resolution.

Table 2
Se-atom positions, occupancy and *B* factors refined using *MLPHARE* (Otwinowski, 1993).

Atom number	*X*	*Y*	*Z*	Occupancy	Ano. occupancy	*B* factor
1	−0.312	−0.391	−0.739	0.126	−2.407	13.606
2	−0.373	−0.032	−0.199	0.116	−2.126	13.212
3	−0.259	−0.024	−0.795	0.128	−1.919	11.693
4	−0.289	−0.097	−0.759	0.113	−2.203	10.696
5	−0.206	−0.047	−0.811	0.130	−2.055	7.458
6	−0.099	−0.042	−0.964	0.089	−1.873	21.109
7	−0.324	−0.088	−0.783	0.103	−1.914	16.014
8	−0.408	−0.355	−0.911	0.123	−1.578	16.412
9	−0.096	−0.213	−0.949	0.145	−1.403	17.614
10	−0.028	−0.049	−0.954	0.101	−1.410	10.659
11	−0.297	−0.418	−0.458	0.088	−1.632	29.504
12	−0.469	−0.205	−0.787	0.157	−0.732	20.823
13	−0.038	−0.480	−0.487	0.126	−1.651	26.608

additional week. Fig. 1 shows a crystal of SeMet GAR-syn from one of the initial crystallization trials. This crystal reached dimensions of 0.2 × 0.3 × 0.6 mm in about one week.

2.4. Data collection and X-ray analysis

The SeMet GAR-syn crystals were flash frozen by transferring to an artificial mother liquor containing the well solution plus 10% glycerol for 2 min and then directly mounting them in loops and placing them on a goniometer head centered in a gaseous stream produced by liquid nitrogen boil off. A complete monochromatic data set to 1.6 Å resolution was collected on the F1 beamline at the Cornell High Energy Synchrotron Source (CHESS) using a Area Detector Systems Corporation Quantum 4 mosaic CCD detector, and a three-wavelength MAD data set was collected to 2.2 Å resolution on the CHESS F2 beamline using a Princeton Scientific Instruments 2k CCD detector. Fig. 2 shows a typical diffraction image with data extending beyond 2.1 Å resolution. The monochromatic data were processed with the *MOSFLM* suite of programs (Leslie, 1992) and the MAD data were processed with *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). A summary of the data analysis is provided in Table 1.

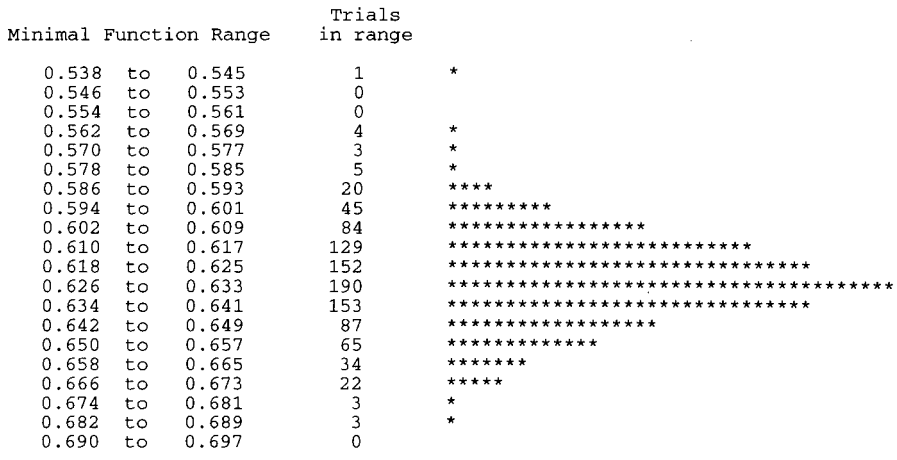

Figure 3
Histogram of the *SnB* results showing the number of trails *versus* minimal function value. (*a*) *SnB* results using F_a values. (*b*) *SnB* results using the anomalous differences from the peak wavelength.

2.5. Location of the Se-atom positions

Se-atom positions were located using direct methods and the program *SnB* (Miller *et al.*, 1994), which implements Hauptman's *Shake-and-Bake* procedure (Miller *et al.*, 1993). First, the anomalous-scattering contributions to the structure-factor magnitudes, F_a, were estimated using the *MADSYS* programs with locally scaled data (Hendrickson, 1985). A total of 3318 F_a values were used to calculate normalized structure-factor magnitudes, *E*, using the program *BAYES* (Blessing, personal communication). A total of 1000 *SnB* trials were performed using 500 *E* values, 5000 triple phase relationships and 20 phase-refinement cycles. Out of 1000 trials, one phase set converged to a solution that revealed nine correct Se-atom positions (Fig. 3*a*). A second *SnB* run was performed using only the peak wavelength, which contains the maximum average anomalous differences. A total of 41412 anomalous differences were converted to *E* values. A total of 1000 *SnB* trials were performed using the 500 largest *E* values, 5000 triple phase relationships and 30 phase-refinement cycles. Three of these trials resulted in solutions revealing 13 of the 14 possible Se-atom positions (Fig. 3*b*). The 13 Se-atom positions are listed in Table 2.

3. Results and discussion

SeMet GAR-syn was purified to homogeneity in a single step by using Ni-dependent affinity chromatography on an iminodiacetic acid column. Using the newly designed construct and purification scheme, 600 mg of purified SeMet GAR-syn was obtained from 10 g of cells. The histidine tag was resistant to cleavage by the Novagen thrombin cleavage kit and remained in the purified protein. Overloaded SDS–PAGE showed only one very faint band other than the major GAR-syn band, and the estimated purity was better than 97%. The crystallized GAR-syn contains a site mutation, *i.e.*

Pro294→Leu, that was introduced during the PCR reaction. GAR-syn crystals diffract to 1.6 Å resolution at CHESS. X-ray intensity data showed orthorhombic Laue symmetry and systematic absences indicative of space group $P2_12_12_1$. The unit-cell dimensions are a = 56.2, b = 62.4 and c = 129.8 Å. Assuming one molecule per asymmetric unit, the Matthews number is 2.32 Å^3 Da^{-1}, which corresponds to about 50% solvent.

The use of SeMet-incorporated proteins combined with MAD phasing measurements offers an alternative to trial-and-error heavy-atom searching in protein structure determination. Experimental measurements are easily made using tunable synchrotron radiation sources. The availability of techniques using polyhistidine tags and metal-chelating columns allows for rapid protein purification and minimizes problems of oxidation that frequently occur in SeMet proteins. The remaining problem of determining the Se-atom substructure has been greatly facilitated by the availability of direct-methods procedures such as *Shake-and-Bake*. The MAD data collected on GAR-syn at CHESS has been used to determine the Se-atom substructure. To date, 13 of the expected 14 Se sites have been located using *SnB* and phases have been calculated. The MAD-phased map is of excellent quality and a detailed structural analysis is under way. It is hoped that the structure of GAR-syn will provide details of the catalytic mechanism and information about how PRA is channeled between PRPP-AT and GAR-syn.

We would like to thank Ulrike Breitinger for her assistance in setting up the initial crystallization trials. This work is supported by NIH grant GM46733 and is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS), which is supported by the National Science Foundation under award DMR-9311772, using the Macromolecular Diffraction at CHESS (MacCHESS) facility, which is supported by award RR-01646 from the National Institutes of Health. SEE is indebted to the W. M. Keck Foundation and the Lucille P. Markey Charitable Trust.

References

Fan, C., Movews, P. C., Walsh, C. T. & Konx, J. R. (1994). *Science*, **266**, 439–443.
Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.
Hendrickson, W. A. (1991). *Science*, **254**, 51–58.
Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.
Kim, J. H., Karhn, J. M., Tomchick, D. R., Smith, J. L., Zalkin, H. (1996). *J. Biol. Chem.* **271**(26), 15549–15557.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Leslie, A. G. W. (1992). In *Joint CCP4 ESF-EACMB Newsletter on Protein Crystallography*, Vol. 26. Warrington: Daresbury Laboratory.
Miller, R., DeTitta, G. T., Jones, R., Langs, D. A. Weeks, C. M. & Hauptman, H. A. (1993). *Science*, **259**, 1430–1433.
Miller, R., Gallo, S. M., Khalak, H. G. & Weeks, C. W. (1994). *J. Appl. Cryst.* **27**, 613–621.
Mueller, E. J., Meyer, E., Rudolph, J., Davison, V. J. & Stubbe, J. (1994). *Biochemistry*, **33**, 2267–2278.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Rudolph, J. & Stubbe, J. (1995). *Biochemistry*, **34**, 2241–2250.
Shen, Y., Rudolph, J., Stern, M. & Stubbe, J. (1990). *Biochemistry*, **29**, 218–227.
Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). *Methods Enzyymol.* **185**, 60–89.
Waldrop, G. L., Rayment, I. & Holden, H. M. (1994). *Biochemistry*, **33**, 10249–10256.
Yamaguchi, H., Kato, H., Hata, Y., Nishioka, T., Kimura, A., Oda, J. & Katsuube, Y. (1993). *J. Mol. Biol.* **229**, 1083–1100.