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

Crystallization and preliminary X-ray analysis of a native human tRNA synthetase whose allelic variants are associated with Charcot–Marie–Tooth disease

Wei Xie, Paul Schimmel and Xiang-Lei Yang*

Departments of Molecular Biology and Chemistry, The Skaggs Institute for Chemical Biology, The Scripps Research Institute, BCC-379, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Correspondence e-mail: xlyang@scripps.edu

Received 23 September 2006
Accepted 3 November 2006

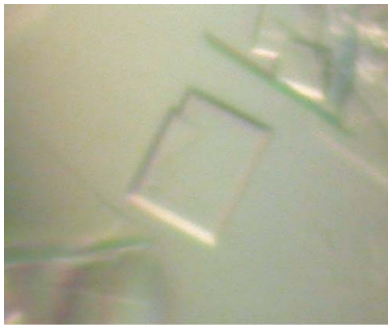
Glycyl-tRNA synthetase (GlyRS) is one of a group of enzymes that catalyze the synthesis of aminoacyl-tRNAs for translation. Mutations of human and mouse GlyRSs are causally associated with Charcot–Marie–Tooth disease, the most common genetic disorder of the peripheral nervous system. As the first step towards a structure–function analysis of this disease, native human GlyRS was expressed, purified and crystallized. The crystal belonged to space group $P4_32_12$ or its enantiomorphic space group $P4_12_12$, with unit-cell parameters $a = b = 91.74$, $c = 247.18$ Å, and diffracted X-rays to 3.0 Å resolution. The asymmetric unit contained one GlyRS molecule and had a solvent content of 69%.

1. Introduction

Aminoacyl-tRNA synthetases (ARSs) are ancient enzymes presumptively involved in the establishment of the genetic code. They catalyze the first step of protein biosynthesis; that is, the ligation of each of the 20 amino acids onto the 3′ ends of their cognate tRNAs, which bear the specific trinucleotide anticodons of the code. Each amino acid has a specific synthetase, giving a total of 20 ARSs in most organisms. In general, the 20 ARSs are evenly separated into two classes according to their sequences and structural homology (Webster *et al.*, 1984; Ludmerer & Schimmel, 1987; Eriani *et al.*, 1990; Cusack *et al.*, 1990).

Our understanding of the functions of ARSs in higher eukaryotes has greatly expanded over the past few years. Novel biological activities have been discovered for several ARSs and for auxiliary factors associated with ARSs (Wakasugi & Schimmel, 1999; Levine *et al.*, 2003; Sampath *et al.*, 2004; Yang *et al.*, 2004; Park *et al.*, 2005; Jordanova *et al.*, 2003; Lee *et al.*, 2006). Among others, the class II glycyl-tRNA synthetase (GlyRS) is connected with specific human diseases. Antibodies against GlyRS were detected in the sera of patients suffering from polymyositis or dermatomyositis and interstitial lung diseases (Freist *et al.*, 1996). More significantly, mutations in both human and mouse GlyRSs are causally associated with one type of Charcot–Marie–Tooth disease. This inherited disorder of the peripheral nervous system (PNS) is the most common genetic disease of the PNS, affecting approximately 1 in 2500 people (Skre, 1974).

Human GlyRS, along with other eukaryotic and archeal GlyRSs, forms α_2 homodimers. In contrast, some bacterial GlyRSs (*e.g. Escherichia coli*) form $\alpha_2\beta_2$ heterotetramers, while others (*e.g. Thermus thermophilus*) are eukaryote-like α_2 homodimers. The two solved structures of GlyRS are from bacterial sources: *T. thermophilus* GlyRS as an α_2 homodimer (Logan *et al.*, 1995) and the α-chain of the $\alpha_2\beta_2$ *Thermotoga maritima* GlyRS (PDB code 1j5w). Human GlyRS shares less than 30% sequence identity to *T. thermophilus* GlyRS, which contains an N-terminal characteristic class II catalytic domain (with three conserved sequence motifs built around a seven-strand mostly antiparallel β-sheet framework) and a C-terminal anticodon-recognition domain. Compared with *T. thermophilus* GlyRS, human GlyRS has extensions at both the N- and C-termini and three insertions (of various sizes) into the catalytic domain (Fig. 1). To date, at least six different CMT-associated mutant alleles of the GlyRS gene (GARS) have been identified in the human population (Antonellis *et al.*, 2003; Sivakumar *et al.*, 2005; Del Bo *et*



al., 2006) and one in the mouse (Seburn *et al.*, 2006). The mutations are spread out over the structure of the homologous *T. thermophilus* GlyRS (Fig. 1). Except for D500N, which resides within the third insertion of the catalytic domain, all mutations are contained in the conserved regions shared by human and *T. thermophilus* GlyRS (Mazauric *et al.*, 1998). Using the existing structural information on GlyRS, no explanation of the connection to CMT disease has been obtained.

Thus, in order to establish a structural framework for investigating the association of GlyRS with CMT disease, we sought to determine the crystal structure of human GlyRS and the GlyRSs encoded by the CMT mutant alleles. As the first step, we here describe the expression, purification, crystallization and preliminary X-ray diffraction analysis of native human GlyRS.

2. Protein expression and purification

The original human GlyRS gene *GARS* (accession No. BAA06338), including the mitochondrial-specific coding region, was cloned into a pET-21a(+) vector (EMD Biosciences, Inc., Novagen, Madison, WI, USA) by Anthony Antonellis at the National Human Genome Research Institute of the NIH. This clone was used as a template to create plasmid pCMM171, in which the mitochondrial-specific region (54 amino acids) was removed from the N-terminus. The new construct represents the cytosolic human GlyRS (685 amino acids) linked with a C-terminal coding sequence for a dipeptide spacer (Leu-Glu) immediately preceding a His_6 tag.

A 2 l culture of Luria–Bertani broth containing 100 mg ml^{-1} ampicillin was inoculated with a 5 ml overnight culture of *E. coli* BL21(DE3)/pCMM171 and grown at 303 K to an OD_{600} of 0.5–0.6. Human GlyRS expression was then induced by the addition of 1 m*M* isopropyl β-D-thiogalactopyranoside (IPTG). The *E. coli* cells were grown for a further 12 h before being harvested by centrifugation at 5000*g* for 20 min. The cell pellets were washed with a buffer containing 40 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl and resuspended in 25 ml Ni–NTA buffer *A* (40 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 10 m*M* imidazole, 1 m*M* PMSF). The cells were disrupted by three passes through a French press. The lysate was centrifuged at 20 000*g* for 50 min to remove cell debris. The supernatant was filtered through a 0.22 µm syringe filter and applied onto an Ni–NTA (Qiagen, Valencia, CA, USA) column previously equilibrated with Ni–NTA buffer *A*. The column was washed with 20 column volumes of washing buffer (40 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 15 m*M* imidazole). Using an imidazole gradient, the protein was eluted from the column at approximately 100–210 m*M* imidazole. The GlyRS-containing fractions were pooled, concentrated and applied onto a MonoQ (HR 16/10) column (Amersham Bioscience, Piscataway, NJ, USA) equilibrated with MonoQ buffer *A* (20 m*M* Tris–HCl pH 7.8, 20 m*M* NaCl and 1 m*M* DTT). The column was eluted with buffer containing an NaCl gradient and the protein eluted at a concentration of 200–270 m*M* NaCl. Glycerol [25%(*v*/*v*) final concentration] was added to the eluted protein before it was flash-frozen and stored at 193 K. The protein remained fully active for aminoacylation after it was stored.

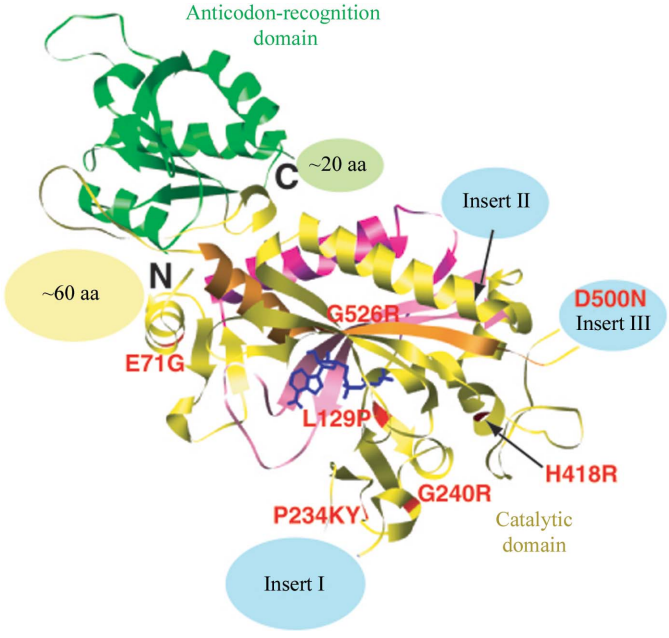


Figure 1
Mapping of the human and mouse CMT disease-causing mutations on the homologous *T. thermophilus* GlyRS structure. Compared with *T. thermophilus* GlyRS, human GlyRS has extensions at both the N- and C-termini and three insertions of various sizes into the catalytic domain. The three conserved sequence motifs in the catalytic domain are colored magenta, pink and orange, respectively.

3. Crystallization

Human GlyRS was buffer-exchanged into 10 m*M* HEPES pH 7.0 with 20 m*M* NaCl before being concentrating to 8 mg ml^{-1} using a Centricon centrifugal filter (Millipore, Billerica, MA, USA). High-throughput crystallization screens were set up with a Mosquito crystallization robot (TTP Labtech, Royston, England) using the sitting-drop vapor-diffusion method. The commercial screens Index (Hampton Research, Aliso Viejo, CA, USA) and Structure Screen (Molecular Dimensions Ltd, Apopka, FL, USA), as well as in-house PEG-based and ammonium sulfate-based screens, were used. 150 nl protein sample was mixed with an equal volume of reservoir solution in each well of 96-well plates at room temperature. Crystallization was monitored at regular intervals with the CrystalPro imaging system (TriTek, Sumerduck, VA, USA). Crystallization hits were obtained from the screens and subsequently optimized with the Opti-Salts suite (Qiagen, Valencia, CA, USA). After optimization, plate-like crystals (0.1 × 0.1 × 0.02 mm; Fig. 2) were obtained within 2 d by vapor diffusion of 2 µl hanging drops (1 µl protein and 1 µl reservoir solution) against 1 ml reservoir solution consisting of 10% PEG 6000,

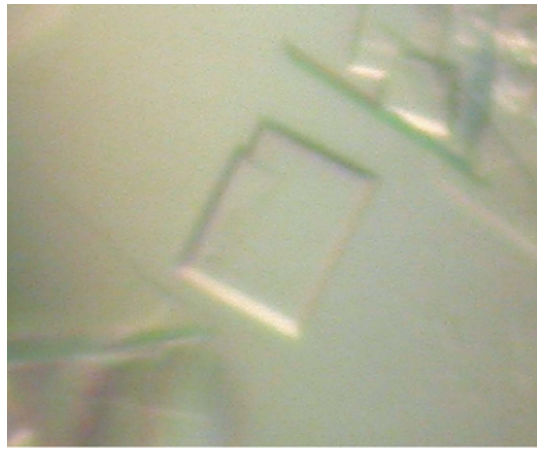


Figure 2
Crystals of human GlyRS. The dimensions of the largest crystal are approximately 0.1 × 0.1 × 0.02 mm.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

X-ray source	SSRL beamline 11-1
Wavelength (Å)	0.97945
Space group	$P4_32_12$ or $P4_12_12$
Unit-cell parameters (Å)	$a = b = 91.74$, $c = 247.18$
Resolution (Å)	50–3.0 (3.12–3.0)
Unique reflections	22076
Completeness (%)	96.9 (78.0)
Redundancy	4.8 (2.8)
$R_{\rm merge}$† (%)	8.9 (34.7)
$\langle I/\sigma(I)\rangle$	13.2 (3.0)

† $R_{\rm merge} = \sum |I - \langle I\rangle| / \sum I$, where I is the intensity of an individual reflection and $\langle I\rangle$ is the average intensity over symmetry equivalents.

0.1 *M* HEPES pH 6.5, 0.01 *M* Tris–HCl pH 8.5, 0.5 *M* NaCl and 0.1 *M* magnesium acetate at room temperature. Crystals were soaked for 3–5 min in a freshly made cryoprotective solution containing the components of the reservoir solution plus 20%(*v*/*v*) glycerol. The soaked crystals were mounted on a nylon loop and flash-frozen in liquid nitrogen.

4. Preliminary X-ray analysis

Using the Stanford Automated Mounting system (SAM) at beamline 11-1 of the Stanford Synchrotron Radiation Laboratory (SSRL), pre-frozen crystals were mounted onto the goniometer one by one to screen for diffraction quality. The best crystal diffracted to 3.0 Å resolution (Fig. 3). A complete data set was recorded with an ADSC Quantum-315 CCD detector at a wavelength of 0.979 Å. The crystal was exposed for 10 s and rotated through a 0.3° oscillation per frame at a crystal-to-detector distance of 320 mm. Data were processed with *DENZO* and scaled with *SCALEPACK* from the *HKL*-2000 package (Otwinowski & Minor, 1997). The crystal belonged to space group $P4_32_12$ or the enantiomorphic space group $P4_12_12$, with unit-cell parameters $a = b = 91.74$, $c = 247.18$ Å. The Matthews coefficient was calculated to be 3.32 Å^3 Da^{-1}, indicating one molecule of GlyRS per asymmetric unit and a solvent content of 69.2%. Data-collection statistics are given in Table 1.

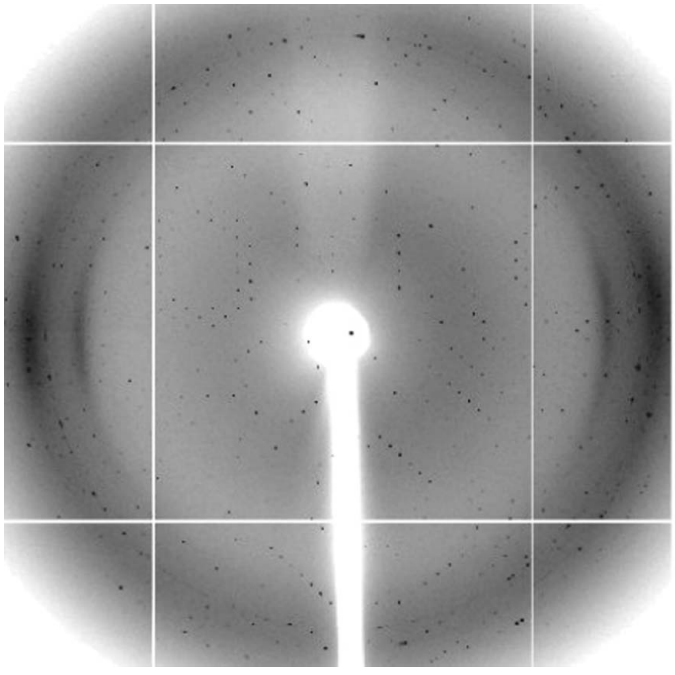


Figure 3
A diffraction image of a human GlyRS crystal that diffracted to 3.0 Å resolution. The crystal was exposed to the X-ray beam (0.979 Å wavelength) for 10 s and rotated through a 0.3° oscillation at a crystal-to-detector distance of 320 mm.

We attempted to solve the structure by molecular replacement. A number of starting models were used for the search. These included the *T. thermophilus* GlyRS structure, the backbone conformation of the same structure with all side chains replaced by alanines and various homology models of human GlyRS generated with *Robetta* (http://www.robetta.org). In addition, various programs (*AMoRe*, *MOLREP* and *Phaser*) from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994) were used. These attempts have not yet produced a successful solution, presumably because of the low sequence identity between human GlyRS and the existing homologous structures. Currently, a *de novo* structure determination is being pursued using crystals of selenomethionine-labeled protein.

We thank Anthony Antonellis at the National Human Genome Research Institute of the NIH for kindly providing us with the original GlyRS clone, Leslie Nangle and Candace Motta for generating the modified plasmid and helpful discussions, Min Guo for technical assistance and suggestions on the project, and Manal Swairjo and Zhanna Druzina for their help with data collection. This work was supported by grant GM15539 from the National Institutes of Health and by a fellowship from the National Foundation for Cancer Research.

References

Antonellis, A., Ellsworth, R. E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S. Q., Jordanova, A., Kremensky, I., Christodoulou, K., Middleton, L. T., Sivakumar, K., Ionasescu, V., Funalot, B., Vance, J. M., Goldfarb, L. G., Fischbeck, K. H. & Green, E. D. (2003). *Am. J. Hum. Genet.* **72**, 1293–1299.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N. & Leberman, R. (1990). *Nature (London)*, **347**, 249–255.
Del Bo, R., Locatelli, F., Corti, S., Scarlato, M., Ghezzi, S., Prelle, A., Fagiolari, G., Moggio, M., Carpo, M., Bresolin, N. & Comi, G. P. (2006). *Neurology*, **66**, 752–754.
Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990). *Nature (London)*, **347**, 203–206.
Freist, W., Logan, D. T. & Gauss, D. H. (1996). *Biol. Chem. Hoppe-Seyler*, **377**, 343–356.
Jordanova, A., Thomas, F. P., Guergueltcheva, V., Tournev, I., Gondim, F. A., Ishpekova, B., De Vriendt, E., Jacobs, A., Litvinenko, I., Ivanova, N., Buzhov, B., De Jonghe, P., Kremensky, I. & Timmerman, V. (2003). *Am. J. Hum. Genet.* **73**, 1423–1430.
Lee, J. W., Beebe, K., Nangle, L. A., Jang, J., Longo-Guess, C. M., Cook, S. A., Davisson, M. T., Sundberg, J. P., Schimmel, P. & Ackerman, S. L. (2006). *Nature (London)*, **443**, 50–55.
Levine, S. M., Rosen, A. & Casciola-Rosen, L. A. (2003). *Curr. Opin. Rheumatol.* **15**, 708–713.
Logan, D. T., Mazauric, M. H., Kern, D. & Moras, D. (1995). *EMBO J.* **14**, 4156–4167.
Ludmerer, S. W. & Schimmel, P. (1987). *J. Biol. Chem.* **262**, 10801–10806.
Mazauric, M. H., Keith, G., Logan, D., Kreutzer, R., Giegé, R. & Kern, D. (1998). *Eur. J. Biochem.* **251**, 744–757.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Park, S. G., Ewalt, K. L. & Kim, S. (2005). *Trends Biochem. Sci.* **30**, 569–574.
Sampath, P., Mazumder, B., Seshadri, V., Gerber, C. A., Chavatte, L., Kinter, M., Ting, S. M., Dignam, J. D., Kim, S., Driscoll, D. M. & Fox, P. L. (2004). *Cell*, **119**, 195–208.
Seburn, K. L., Nangle, L. A., Cox, G. A., Schimmel, P. & Burgess, R. W. (2006). *Neuron*, **51**, 715–726.
Sivakumar, K., Kyriakides, T., Puls, I., Nicholson, G. A., Funalot, B., Antonellis, A., Sambuughin, N., Christodoulou, K., Beggs, J. L., Zamba-

Papanicolaou, E., Ionasescu, V., Dalakas, M. C., Green, E. D., Fischbeck, K. H. & Goldfarb, L. G. (2005). *Brain*, **128**, 2304–2314.
Skre, H. (1974). *Clin. Genet.* **6**, 98–118.
Wakasugi, K. & Schimmel, P. (1999). *Science*, **284**, 147–151.
Webster, T., Tsai, H., Kula, M., Mackie, G. A. & Schimmel, P. (1984). *Science*, **226**, 1315–1317.
Yang, X. L., Schimmel, P. & Ewalt, K. L. (2004). *Trends Biochem. Sci.* **29**, 250–256.