PROTEINS: Structure, Function, and Genetics 49:285–288 (2002)

STRUCTURE NOTE

Crystal Structure of MTH169, a Crucial Component of Phosphoribosylformylglycinamidine Synthetase

Renu Batra,[1] Dinesh Christendat,[2] Aled Edwards,[2] Cheryl Arrowsmith,[2] and Liang Tong[1]*
[1]*Department of Biological Sciences, Columbia University, New York, New York*
[2]*Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada*

Introduction Phosphoribosylformylglycinamidine (FGAM) synthetase (EC 6.3.5.3) catalyses the reaction 5′-phosphoribosylformylglycinamide (FGAR) + glutamine + ATP ↔ FGAM + glutamate + ADP + P_i, the fourth step in the de novo purine biosynthetic pathway. In eukaryotes and many bacterial systems (including *Escherichia coli* and *Salmonella typhimurium*), the FGAM synthetase is encoded by a large protein with an N-terminal ATPase domain and a C-terminal glutamine-binding domain.[1] In archaeal and other bacterial systems, however, FGAM synthetase is encoded by separate genes, making it a multisubunit (rather than multidomain) enzyme. For example, in *Bacillus subtilis*, the purL protein is homologous to the ATPase domain, whereas the purQ protein is homologous to the glutamine-binding domain of the single-chain FGAM synthetases.[2]

The *purL* and *purQ* genes are part of the *pur* operon in *B. subtilis*, which encodes 11 of the 12 enzymes in the purine biosynthetic pathway.[2] The genetic studies also identified an open reading frame (ORF) of 84 amino acids in this operon, now known as *purS*,[3] which is conserved in a large group of Gram-positive bacteria and methanogenic archaea (Fig. 1).Recent studies showed that disruption of the *purS* gene in *B. subtilis* resulted in a purine-auxotrophic phenotype, due to defective FGAM synthetase activity.[3] Therefore, the purS protein appears to be required for the function of the purL and purQ subunits of the FGAM synthetase, but the molecular mechanism for the functional role of purS is currently not known.

We recently initiated a prototype structural genomics effort that focused on non-membrane proteins in the proteome of the thermophilic archaeon *Methanobacterium thermoautotrophicum* (Mth).[4] One of the proteins selected for this study is MTH169, which shares 29% amino acid sequence identity with that of purS in *B. subtilis*. The Mth proteome also contains proteins that are homologous to the purL and purQ subunits of *B. subtilis*. MTH169 is, therefore, the likely purS ortholog in Mth. We will use the names purS and MTH169 interchangeably here. In addition, for ease of discussion, all purS proteins are numbered according to the sequence of MTH169 (Fig. 1).

The crystal structure of MTH169 (purS) has been determined at 2.56 Å resolution (Table I)and deposited at the Protein Data Bank (entry 1GTD). This 84-residue protein forms a tetramer, each subunit of which contains well-defined secondary structure elements, including a three-stranded anti-parallel β-sheet (β1 through β3) and two α-helices (αA and αB) [Fig. 2(A)]. Helix αB covers part of one face of the β-sheet, forming the hydrophobic core of the structure. Residues in this core are primarily on strand β1 and helix αB [Fig. 2(A)], and are generally conserved among the family of purS proteins (Fig. 1). On the other hand, residues on the other face of the β-sheet are weakly-conserved and are mostly hydrophilic or charged in nature.

Helix αA extends away from the β-sheet structure [Fig. 2(A)]. Such a conformation is probably unstable for the molecule in the monomeric state, but these residues are stabilized by quaternary interactions in the tetramer of MTH169. There are two molecules of MTH169 in the crystallographic asymmetric unit, which form a non-crystallographic dimer [Fig. 2(B)]. The dimer is situated near a crystallographic twofold symmetry axis, and this produces a tetramer of MTH169 in the crystal [Fig. 2(C)]. Within each monomer, about 1,100 $Å^2$ and 650 $Å^2$ of the surface area is buried at the dimer and tetramer interface, respectively, suggesting that the dimer interface is more extensive than the tetramer interface.

Structural searches against the Protein Data Bank, with the program Dali,[5] showed that there are other structures with similar backbone folds, but none of them are identical to MTH169. All the structures identified by Dali have a four-stranded anti-parallel β-sheet, most frequently with the extra strand inserted between αA and β2 and located next to β2 of MTH169. Moreover, the αA

Grant sponsor: NIGMS; Grant number: P50 GM62413.

Renu Batra and Dinesh Christendat contributed equally to this work.

Renu Batra's present address is Institut für Biologische Informationsverarbeitung, Forschungszentrum Julich, D-52425 Julich, Germany.

*Correspondence to: Liang Tong, Department of Biological Sciences, Columbia University, New York, NY 10027. E-mail: tong@como.bio.columbia.edu

Received 28 May 2002; Accepted 30 May 2002

TABLE I. Summary of Crystallographic Information (PDB ID: IGTD)

Space group	$P4_322$
Unit cell parameters (a, c) (Å)	53.8, 142.7
Maximum resolution (Å)	2.56
Number of observations	55,012
R_{merge} (%)[a]	7.0
Number of reflections	12,678
Resolution range for refinement	20–2.56
Completeness (%)	90
R factor (%)[b]	23.5
Free R factor (%)	27.8
rms deviation in bond lengths (Å)	0.007
rms deviation in bond angles (°)	1.2

[a] $R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h \sum_i I_{hi}$.
[b] $R = \sum_h |F_h^o - F_h^c| / \sum_h F_h^o$

helix in these structures runs along the β-sheet and helps cover the hydrophobic core. The conformation for helix αA in MTH169 is not observed in any of these other structures. Interestingly, this folding motif of a four-stranded anti-parallel β-sheet with two helices on one face is also observed to mediate protein oligomerization, producing dimers, trimers, and tetramers for a number of proteins.

A major part of the dimer interface in MTH169 is composed of hydrogen-bonding interactions between strand β2 of one monomer with the same strand in the other monomer. This extends the 3-stranded β-sheet in the monomer to a 6-stranded, highly-twisted anti-parallel β-sheet in the dimer [Fig. 2(B)]. In addition, helix αA and the αA-β2 loop in one monomer are packed against helix αB in the other monomer [Fig. 2(B)], with conserved, hydrophobic residues in this interface (Fig. 1).

All four helices are located on the convex face of the β-sheet in the dimer [Fig. 2(B)]. Helix αA together with the β1-αA and αB-β3 loops form a ring-like structure on the surface of this sheet [Fig. 2(D)], and residues in this structure are mostly conserved and hydrophobic among the purS proteins (Fig. 1). The tetramer of MTH169 is formed by the direct contact of the ring-like structures from the two dimers [Fig. 2(C)]. This contact leaves an open space in the center of the tetramer, with a volume of about 1,100 Å^3 [Fig. 2(C)]. Two charged side chains from each monomer in MTH169 (Fig. 1), Glu19 and Asp41, are pointed towards this pocket. The concave face of the β-sheet in the dimer remains exposed to the solvent in the tetramer [Fig. 2(C)].

Roughly 7,000 Å^2 of the surface areas of the four monomers are buried by the formation of the tetramer, suggesting that this tetramer may be a stable entity. This structural observation was confirmed by gel filtration studies, which demonstrated that MTH169 is a tetramer in solution (unpublished results). The conserved nature of the dimer and tetramer interfaces among the purS proteins (Fig. 1) suggests that the other purS orthologs may also exist as tetramers. This is supported by our mutagenesis studies showing that single-site mutations in the dimer or tetramer interface did not disrupt the tetramer (data not shown), demonstrating the stability of this oligomer.

Our structural analysis suggests that purS is likely a protein-protein interaction module that helps bring the purL and purQ subunits together. Overall, the molecular surface of MTH169 is highly negative electrostatically [Fig. 2(E)], a feature that may be conserved among the purS proteins. All these proteins have calculated isoelectric points between 4.5 and 6, and will therefore be negatively charged at physiological pH. This surface electro-negativity may be functionally important for these proteins. Despite this overall negatively charged surface, there are conserved hydrophobic surface patches on the tetramer as well, near the dimer and tetramer interfaces. It is likely that these surface patches may have important roles in the function of purS, for example the recruitment of the purL and purQ proteins.

Materials and Methods. The MTH169 gene was amplified by PCR using Mth genomic DNA as template and cloned into the pET15b vector (Novagen, Madison, WI). Recombinant protein was expressed with a hexa-histidine fusion tag at the N-terminus in *E. coli* BL21 (DE3). The cells were lysed in a buffer containing 50 mM Hepes (pH 7.5), 500 mM NaCl, 5% (v/v) glycerol, 1 mM DTT, 1 mM PMSF, and 0.5 mM benzamidine. The soluble recombinant protein was bound to Ni^{+2}-affinity resin and eluted in a buffer containing 50 mM Hepes, pH 7.5, 500 mM NaCl, 5% (v/v) glycerol, and 250 mM imidazole. The purified protein was dialyzed extensively against a buffer containing 10 mM Hepes and 500 mM NaCl, concentrated to 20 mg/mL, and stored at 4°C.

The selenomethionine(Se-Met)-labeled protein was expressed in the methionine auxotroph *E. coli* strain B834 (DE3) (Novagen) and purified under the same conditions as the native protein, except that 5 mM β-mercaptoethanol was included in all buffers.

Initial conditions for crystallization of MTH169 were identified using commercial screening kits (Hampton Research). The final crystallization condition consists of 32% (v/v) MPD, 0.1 M sodium citrate (pH 5.6), and 0.2 M ammonium acetate, and the protein concentration is at 10 mg/ml. Crystals appear after approximately 24 h and can be flash-frozen in liquid nitrogen directly in the mother liquor.

X-ray diffraction data to 2.56 Å resolution were collected at Beamline 19-ID at the Advanced Photon Source (APS). A multi-wavelength anomalous diffraction (MAD) data set was collected.[6] However, there was extensive decay in the crystal during the data collection, even though it was kept at 100K. As a result, only the data set at the peak wavelength, which was collected first, was used for the structure determination. The diffraction images were processed with the HKL package.[7] The data processing statistics are summarized in Table I.

The positions of the Se atoms were determined based on the anomalous differences with the program SHELX.[8] After single-wavelength anomalous diffraction (SAD) phasing and solvent flattening, helices and β-strands could be observed in the electron density map. A twofold noncrystal-

lographic symmetry (NCS) axis was located based on the electron density, whose parameters were then refined based on the phase information.[9] After twofold NCS averaging, with a locally written program (Tong, unpublished data), the electron density map can be interpreted based on the amino acid sequence. The atomic model of the protein was built into the electron density map with the program O.[10] The structure refinement was carried out with the program CNS.[11] The refinement statistics are summarized in Table I.

```
                          1   5   10   15   20   25   30   35   40   45   50
S.S.                        |--β1--|        |----αA-----|        |----β2-----|
M. thermoautotrophicum    MKFMVEVRIRLKKGMLNPRAATIERALALLGY.EVEDTDTTDVITFTMDED
B. subtilis                MYK-K-YVS--ESV-D-QGSAVQH--HSMT-N--Q-VRIGKYMEL-I.-K
M. tuberculosis            MAR-V-HVMP-AEI-D-QGQA-VG--GR--HLGIS-VRQGKRFELEV-DT
Staphylococcus aureus      MKTI-LH-T-QPQV-DTQGQ-LT--VHD---AQ-N-IRVGK-LYM-V--V
Aquifex aeolicus           MKL-R-I-MP-E-L-D-QGRAV-EM-KEN-FN.-SNVRVGK-VELEVS--
Lactococcus lactis         MTK-R-YVAY-ASI-D-Q-QA-KA-THKM--Q--S-LNVGKFPD-DF-AE
M. jannaschii              MYKAT-I-K----V----GR--Q---NF--FNN-KEVQ-YKM-DII-EGE
Thermotoga maritima      MPL-KFAIDVQYRSNVRD-RGE----V-REEKGLP-KKLRLGKS-HLEVEAE

                           55   60   65   70   75   80   84
S.S.                         |------αB-----|        |--β3---|
M. thermoautotrophicum     SLEAVEREVEDMCQRLLCNPVIHDYDVSINEMEG
B. subtilis                -DRDLDVL-KE--EK--A-T--E--RYEVE-VVAQ
M. tuberculosis            VDDTTLA...EIAES--A-T--E-WTI-RDPQ
Staphylococcus aureus      -D-K-HNIITTLSEK-FA-T--EE-SYKVLDD-KENA
Aquifex aeolicus           TD......LRKLVEKY-I--L-E--EIEELSQKS
Lactococcus lactis         AE.LAKEKATEIANE--A--NMET-K-E-LTEQSEGVK K
M. jannaschii              NE-K-KE-V-E--KK--A-------EIKVEKI-
Thermotoga maritima        NK-KAYEI-KKA-EE--V---VEE-E-REL
```

Fig. 1. Alignment of representative purS sequences. Residues in the hydrophobic core of the monomer, dimer, and tetramer are shown in green, purple, and cyan, respectively. A total of 26 PurS homologs can be identified based on PSI-Blast searches of the current sequence database.

During the structure refinement, it was recognized that there is a stretch of electron density in the $2F_o - F_c$ as well as the NCS averaged electron density map that is not explained by the atomic model for MTH169. The density is located near a crystallographic twofold axis, and its features suggest that it probably correspond to a peptide segment of three residues. The only possible interpretation for this density is the N-terminal His tag, which was not removed for crystallization. The direction of the peptide segment in this density was not clear, so both a His-Ser-Ser and a Ser-Ser-His model were built into this density. The two models were then refined independently, which showed that the His-Ser-Ser model produced lower *R* factors. This model would bridge the two 6-stranded β-sheets of two dimers into a 14-stranded fully antiparallel β-sheet across this crystallographic two-fold axis.

Acknowledgments We thank Reza Khayat and Hao Wu for help with the data collection, Stephan Ginell and Andrzej Joachimiak for access to the 19-ID beamline at APS, and members of the Ontario Center for Structural Proteomics for help with protein purification and crystallization. This research was supported by a grant to the Northeast Structural Genomics Consortium from the Protein Structure Initiative of the National Institutes of

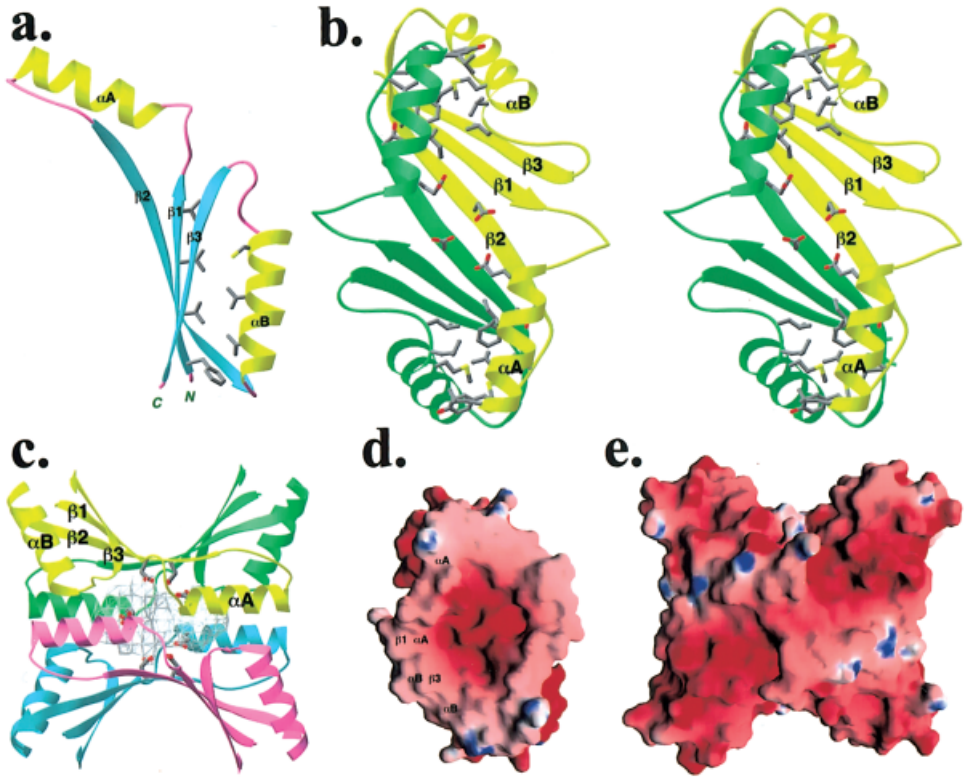

Fig. 2. Structure of MTH169. **a:**. Schematic drawing of the monomer of MTH169. The β-strands are shown in cyan, and the α-helices in yellow. The side chains in the hydrophobic core of the monomer are shown as stick models. **b:**. Stereo diagram of the dimer of MTH169, and the two monomers are colored in yellow and green, respectively. **c:**. The tetramer of MTH169, with the monomers colored yellow, green, cyan, and purple. A cavity in the center of the tetramer is represented by the cage. **d:**. Molecular surface of the dimer of MTH169. The ring-like structure mediates the formation of the tetramer. **e:** Molecular surface of the tetramer of MTH169. a–c were produced with Ribbons,[12] and d and e were produced with Grasp.[13]

Health (P50 GM62413) and the Ontario Research and Development Challenge Fund.

REFERENCES

1. Schendel FJ, Mueller E, Stubbe J, Shiau A, Smith JM. Formylglycinamide ribonucleotide synthetase from *Escherichia coli*: cloning, sequencing, overproduction, isolation, and characterization. Biochemistry 1989;28:2459–2471.
2. Ebbole DJ, Zalkin H. Cloning and characterization of a 12-gene cluster from *Bacillus subtilis* encoding nine enzymes for de novo purine nucleotide synthesis. J Biol Chem 1987;262:8274–8287.
3. Saxild HH, Nygaard P. The yexA gene product is required for phosphoribosylformylglycinamidine synthetase activity in *Bacillus subtilis*. Microbiology 2000;146:807–814.
4. Christendat D, et al. Structural proteomics of an archaeon. Nature Struct Biol 2000;7:903–909.
5. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993;233:123–138.
6. Hendrickson WA. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science 1991;254:51–58.
7. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
8. Sheldrick GM, Schneider TR. SHELXL: High-resolution refinement. Methods Enzymol 1997;277:319–343.
9. Tong L, Replace: a suite of computer programs for molecular-replacement calculations. J Appl Cryst 1993;26:748–751.
10. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Cryst 1991;A47:110–119.
11. Brunger AT, et al. Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Cryst.1998;D54:905–921.
12. Carson M. Ribbon models of macromolecules. J Mol Graph 1987;5:103–106.
13. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991;11:281–296.