Journal of Structural Biology 171 (2010) 238–243

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Structure Report

The structure of tryptophanyl-tRNA synthetase from *Giardia lamblia* reveals divergence from eukaryotic homologs

Tracy L. Arakaki [a,b], Megan Carter [a,c], Alberto J. Napuli [a,c], Christophe L.M.J. Verlinde [a,b], Erkang Fan [a,b], Frank Zucker [a,b], Frederick S. Buckner [a,c], Wesley C. Van Voorhis [a,c], Wim G.J. Hol [a,b], Ethan A. Merritt [a,b,*]

[a] *Medical Structural Genomics of Pathogenic Protozoa*[1], *USA*
[b] *Department of Biochemistry, University of Washington, Seattle, WA 98195, USA*
[c] *Department of Medicine, University of Washington, Seattle, WA 98195, USA*

ARTICLE INFO

Article history:
Received 16 February 2010
Received in revised form 8 April 2010
Accepted 26 April 2010
Available online 8 May 2010

Keywords:
Aminoacyl-tRNA synthetase
Protozoa
Structural genomics

ABSTRACT

The 2.1 Å crystal structure of tryptophanyl-tRNA synthetase (TrpRS) from the diplomonad *Giardia lamblia* reveals that the N-terminus of this class I aminoacyl-tRNA synthetase forms a 16-residue α-helix. This helix replaces a β-hairpin that is required by human TrpRS for normal activity and has been inferred to play a similar role in all eukaryotic TrpRS. The primary sequences of TrpRS homologs from several basal eukaryotes including *Giardia* lack a set of three residues observed to stabilize interactions with this β-hairpin in the human TrpRS. Thus the present structure suggests that the activation reaction mechanism of TrpRS from the basal eukaryote *G. lamblia* differs from that of higher eukaryotes. Furthermore, the protein as observed in the crystal forms an $(\alpha_2)_2$ homotetramer. The canonical dimer interface observed in all previous structures of tryptophanyl-tRNA synthetases is maintained, but in addition each N-terminal α-helix reciprocally interlocks with the equivalent helix from a second dimer to form a dimer of dimers. Although we have no evidence for tetramer formation *in vivo*, modeling indicates that the crystallographically observed tetrameric structure would be compatible with the tRNA binding mode used by dimeric TrpRS and TyrRS.

1. Introduction

The intestinal parasite *Giardia lamblia* (*Giardia intestinalis*) is a eukaryote from the class Diplomonadida. The biology of these organisms presents puzzles that were deepened by analysis of the completed *G. lamblia* genome (Morrison et al., 2007). For example, diplomonads possess two transcriptionally active diploid nuclei, but lack both mitochondria and peroxisomes. These odd features are reflected in the set of aminoacyl-tRNA synthetases (aaRS) found in the *G. lamblia* genome. Eukaryotes in general use two different aaRS for each amino acid, one mitochondrial and one cytosolic. Consistent with its lack of mitochondria, the *G. lamblia* genome contains only a single aaRS for each amino acid. Many of these are more closely related to archaeal homologs than to aaRS sequences from higher eukaryotics, but the sequence of the tryptophanyl-tRNA synthetase (TrpRS) from *G. lamblia* is close to that seen for cytosolic homologs in higher eukaryotes. TrpRS is a class I aminoacyl-tRNA synthetase, characterized by an α-helical anticodon binding domain and a separate canonical Rossman-fold catalytic domain. Class I aaRS contain two characteristic sequence motifs, HIGH and KMSKS, near the active site. The latter motif is KMSAS in *G. lamblia* TrpRS, as it is in mammals and many other eukaryotic TrpRS. The *G. lamblia* TrpRS sequence contains a 49-residue insertion in the catalytic domain relative to homologous sequences in higher eukaryotes. Such insertions are typical for *G. lamblia* genomic sequences (Morrison et al., 2007). Disregarding this insertion, the *Giardia* core domain sequence is ~50% identical to that of mammalian cytosolic TrpRS, a slightly higher identity than it shares with the core domain in archaeal sequences. All TrpRS structures reported to date are α_2 homodimers. A single cognate $tRNA^{Trp}$ molecule spans both monomers on binding. The tRNA anticodon loop binds to the α-helical domain of one monomer, while the tRNA acceptor arm extends into the active site in the catalytic domain of the other monomer (Shen et al., 2006; Yang et al., 2006). This mode of tRNA recognition and binding to an α_2 dimer is shared by bacterial, eukaryotic, and archaeal TrpRS and TyrRS (Yaremchuk et al., 2002; Tsunoda et al., 2007). Eukaryotic and archaeal TrpRS sequences also contain a weakly conserved N-terminal extension from the core domain consisting of up to several hundred additional residues. Schimmel and coworkers have structurally and biochemically characterized the 154 residue extension in human TrpRS as being comprised of several functionally

* Corresponding author at: Department of Biochemistry, University of Washington, Seattle, WA 98195, USA. Fax: +1 206 685 7002.
E-mail address: merritt@u.washington.edu (E.A. Merritt).
[1] http://msgpp.org.

1047-8477/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2010.04.010

important components (Yang et al., 2003, 2007). In *Giardia* TrpRS, this extension contains 68 residues (Fig. 1).

We report here a crystal structure of the TrpRS from *Giardia lamblia* determined at 2.1 Å resolution. Unlike all previously characterized TrpRS, the *Giardia* protein as seen in the crystal forms an $(\alpha_2)_2$ homotetramer. An unexpected monomer:monomer interface is formed by reciprocal interdigitation of the N-terminal α-helix containing residues 6–22. The unanticipated N-terminal helix structurally replaces a β-hairpin that has been assumed to be a feature common to eukaryotic TrpRS homologs and has been implicated as being crucial for ATP binding in human TrpRS (Yang et al., 2007). This structural difference between the *G. lamblia* TrpRS and previously studied eukaryotic TrpRS corresponds to a recognizable sequence difference. The *Giardia* sequence lacks a set of hydrophobic residues that stabilize previously observed interactions between the β-hairpin element and the ATP binding site (Fig. 1).

2. Target choice, protein expression and crystallization strategy

The tryptophanyl-tRNA synthetase from *Giardia* was chosen for study as part of a broader effort by the Medical Structural Genomics of Pathogenic Protozoa (MSGPP) collaboration to investigate aminoacyl-tRNA synthetases from parasitic protozoa as possible targets for drug discovery (Fan et al., 2008). The structure reported here corresponds to the sole TrpRS found in the *G. lamblia* genome. The gene encoding the full-length 429-residue *G. lamblia* tryptophanyl-tRNA synthetase (GiardiaDB GL50803_3032) was cloned from strain WB genomic DNA into vector AVA421 (Alexandrov et al., 2004) containing a cleavable N-terminal His6-tag, and expressed in *Escherichia coli* (Mehlin et al., 2006; Arakaki et al., 2006). The protein was purified using a Ni-NTA column and cleaved on the column by protease 3C overnight at 4 °C. The released protein was further purified by gel filtration on a HiLoad Superdex 75 26/60 at a flow rate of 0.5 ml/min and eluted in a single peak corresponding to a dimer. DTT was added to a final concentration of 2 mM prior to loading on the gel filtration column. The dimeric state in solution was later confirmed by HPLC (Shimadzu Prominence), which showed an apparent molecular weight of 105.8 kDa.

Purified protein was screened at the high-throughput crystallization facility at Hauptman-Woodward Institute as previously described (Luft et al., 2003; Arakaki et al., 2008). Initial crystallization condition hits were further optimized to produce crystals used for data collection. Crystals were grown at room temperature by vapor equilibration from sitting drops containing 2 μl protein solution (11 mg/ml protein, 500 mM NaCl, 20 mM HEPES, 2 mM BME, 5% glycerol, 0.025% NaAzide, pH 7.5) mixed with 1 μl crystallization buffer (2.7 M ammonium sulfate, 5 mM DTT, 0.1 M citric acid, pH 5.2).

3. X-ray diffraction and structure determination

Crystals of *G. lamblia* TrpRS were briefly soaked in paraffin oil and then frozen directly in liquid nitrogen. Data were collected using SSRL beamline 9.1 at an X-ray energy of 12.658 keV. Data were processed and scaled using HKL2000 (Otwinowski and Minor, 1997). The program Molrep (Vagin and Teplyakov, 1997) found an initial molecular replacement solution using a model based on the structure of yeast TrpRS (Malkowski et al., 2007) (PDB accession code 2ip1, 38% overall sequence identity). Initial structure determination and refinement used data to 2.8 Å resolution from a single crystal. Iterative manual model building using Coot (Emsley and Cowtan, 2004) and in Refmac (Murshudov et al., 1997) continued until R and R_{free} reached 0.243 and 0.304, respectively. This initial model was later refined against a higher resolution data set obtained from a second crystal at 2.1 Å resolution to final values $R = 0.199$, $R_{free} = 0.234$. Model quality was validated using Coot and MolProbity (Lovell et al., 2003). Crystallographic data statistics are given in Table 1. The refined model has been deposited in the Protein Data Bank with accession code 3foc.

4. Overall structure

4.1. Monomer

The overall structure of the *Giardia* TrpRS monomer is that of a typical class I aminoacyl-tRNA synthetase. It contains an N-terminal extension (residues 1–64), a Rossman-fold catalytic domain (residues 65–324, 415–429), and an anticodon recognition domain (residues 325–414). The crystallographic asymmetric unit contains two monomers, related by a rotation of 177°. The two monomers were treated independently during crystallographic refinement; their respective C^{α} atoms superimpose with an RMSD of 0.6 Å aligned over 395 residues. Residues 1–219 and 244–420 are well-defined by electron density in monomer A and residues 3–220, 244–344, and 351–429 are well-defined in monomer B. The electron density corresponding to residues of 220–243 in A and 221–243 in B was very weak and these residues, part of a *Giardia*-specific insertion in the catalytic domain, were not modeled. One sulfate was found at the surface of each monomer between the side chains of residues Arg280 and Arg283.

4.2. N-terminal dimer interface

The two monomers in the asymmetric unit form an unusual dimer in which helix α1 in each monomer's N-terminal extension intertwines with helix α1 of the other monomer (Fig. 2). The α1 helices from monomer A monomer B cross at an angle of 70°. Each

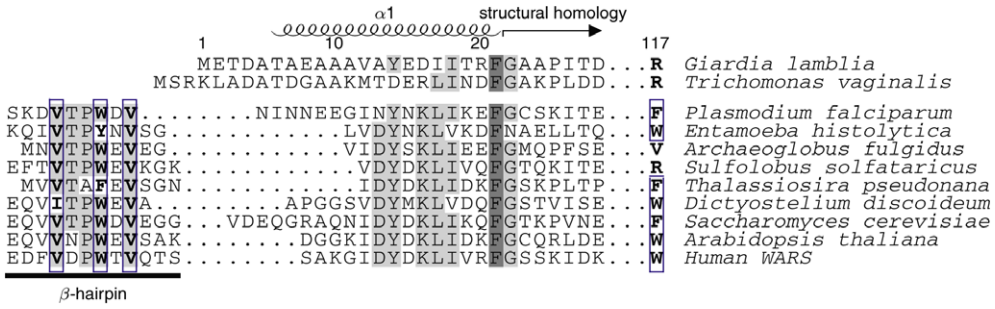

Fig. 1. *Sequence alignment of the G. lamblia TrpRS N-terminus with other eukaryotic and archaeal sequences.* The *Giardia* and human TrpRS are structurally homologous beginning at residue 20. In the *Giardia* structure, residues N-terminal to this point form part of an initial α-helix comprising residues 6–22. In the human TrpRS these residues have been observed instead to form a β-hairpin involved in ATP binding. A triad Val-Trp-Val of hydrophobic residues, shown boxed in the figure, stabilizes association of the hairpin with a Trp residue at the positions equivalent to Arg 117 in the *Giardia* sequence, and with a highly conserved Phe residue equivalent to Phe 301 in the *Giardia* sequence. This hydrophobic triad and the corresponding tryptophan are recognizably present in most eukaryotic TrpRS sequences, but are missing from some protozoan homologs.