Molecular & Biochemical Parasitology 177 (2011) 20–28

Contents lists available at ScienceDirect

Molecular & Biochemical Parasitology

Crystal structures of three protozoan homologs of tryptophanyl-tRNA synthetase

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

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

ARTICLE INFO

Article history:
Received 13 July 2010
Received in revised form
27 December 2010
Accepted 5 January 2011
Available online 19 January 2011

Keywords:
Aminoacyl-tRNA synthetase
Cryptosporidium
Trypanosoma brucei
Entamoeba
Giardia

ABSTRACT

Tryptophanyl-tRNA synthetase (TrpRS) is an essential enzyme that is recognizably conserved across all forms of life. It is responsible for activating and attaching tryptophan to a cognate $tRNA^{Trp}$ molecule for use in protein synthesis. In some eukaryotes this original core function has been supplemented or modified through the addition of extra domains or the expression of variant TrpRS isoforms. The three TrpRS structures from pathogenic protozoa described here represent three illustrations of this malleability in eukaryotes. The *Cryptosporidium parvum* genome contains a single TrpRS gene, which codes for an N-terminal domain of uncertain function in addition to the conserved core TrpRS domains. Sequence analysis indicates that this extra domain, conserved among several apicomplexans, is related to the editing domain of some AlaRS and ThrRS. The *C. parvum* enzyme remains fully active in charging $tRNA^{Trp}$ after truncation of this extra domain. The crystal structure of the active, truncated enzyme is presented here at 2.4 Å resolution. The *Trypanosoma brucei* genome contains separate cytosolic and mitochondrial isoforms of TrpRS that have diverged in their respective tRNA recognition domains. The crystal structure of the *T. brucei* cytosolic isoform is presented here at 2.8 Å resolution. The *Entamoeba histolytica* genome contains three sequences that appear to be TrpRS homologs. However one of these, whose structure is presented here at 3.0 Å resolution, has lost the active site motifs characteristic of the Class I aminoacyl-tRNA synthetase catalytic domain while retaining the conserved features of a fully formed $tRNA^{Trp}$ recognition domain. The biological function of this variant *E. histolytica* TrpRS remains unknown, but, on the basis of a completely conserved tRNA recognition region and evidence for ATP but not tryptophan binding, it is tempting to speculate that it may perform an editing function. Together with a previously reported structure of an unusual TrpRS from *Giardia*, these protozoan structures broaden our perspective on the extent of structural variation found in eukaryotic TrpRS homologs.

1. Introduction

The protozoans *Cryptosporidium parvum*, *Trypanosoma brucei*, and *Entamoeba histolytica* are human pathogens responsible for significant public health burden. These organisms are eukaryotes, but they are phylogenetically distant from higher eukaryotes such as their mammalian hosts. There can be substantial differences in the structure and detailed biological role of homologous proteins from protozoans and from humans. The three structures of protozoan tryptophanyl-tRNA synthetase (TrpRS) homologs we describe here were determined as part of a larger effort undertaken by the Medical Structural Genomics of Pathogenic Protozoa (MSGPP) collaboration [1]. We have previously reported the structure of the unique TrpRS from *Giardia lamblia*, which unexpectedly was observed to form a homotetramer [2]. The overall MSGPP goal is to identify and characterize proteins from these disease-causing organisms that may constitute targets for the design and development of new anti-parasitic drugs.

TrpRS is a class I aminoacyl-tRNA synthetase (aaRS). All members of this class are characterized by a Rossmann-fold catalytic domain whose active site is recognizable by the presence of two conserved sequence motifs with consensus sequences HIGH and KMSKS [3,4] (Fig. 1, red boxes). The active site of TrpRS sequences exhibits a third, more weakly conserved, motif (AIDQ in the human sequence) involved in binding ATP [5]. The catalytic domain carries out the two half-reactions necessary to activate and attach

* Corresponding author at: Department of Biochemistry, University of Washington, Mailstop 357742, Seattle, WA 98195, USA. Tel.: +1 206 543 1421; fax: +1 206 685 7002.
E-mail address: merritt@u.washington.edu (E.A. Merritt).
[1] http://msgpp.org.

0166-6851/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.molbiopara.2011.01.003

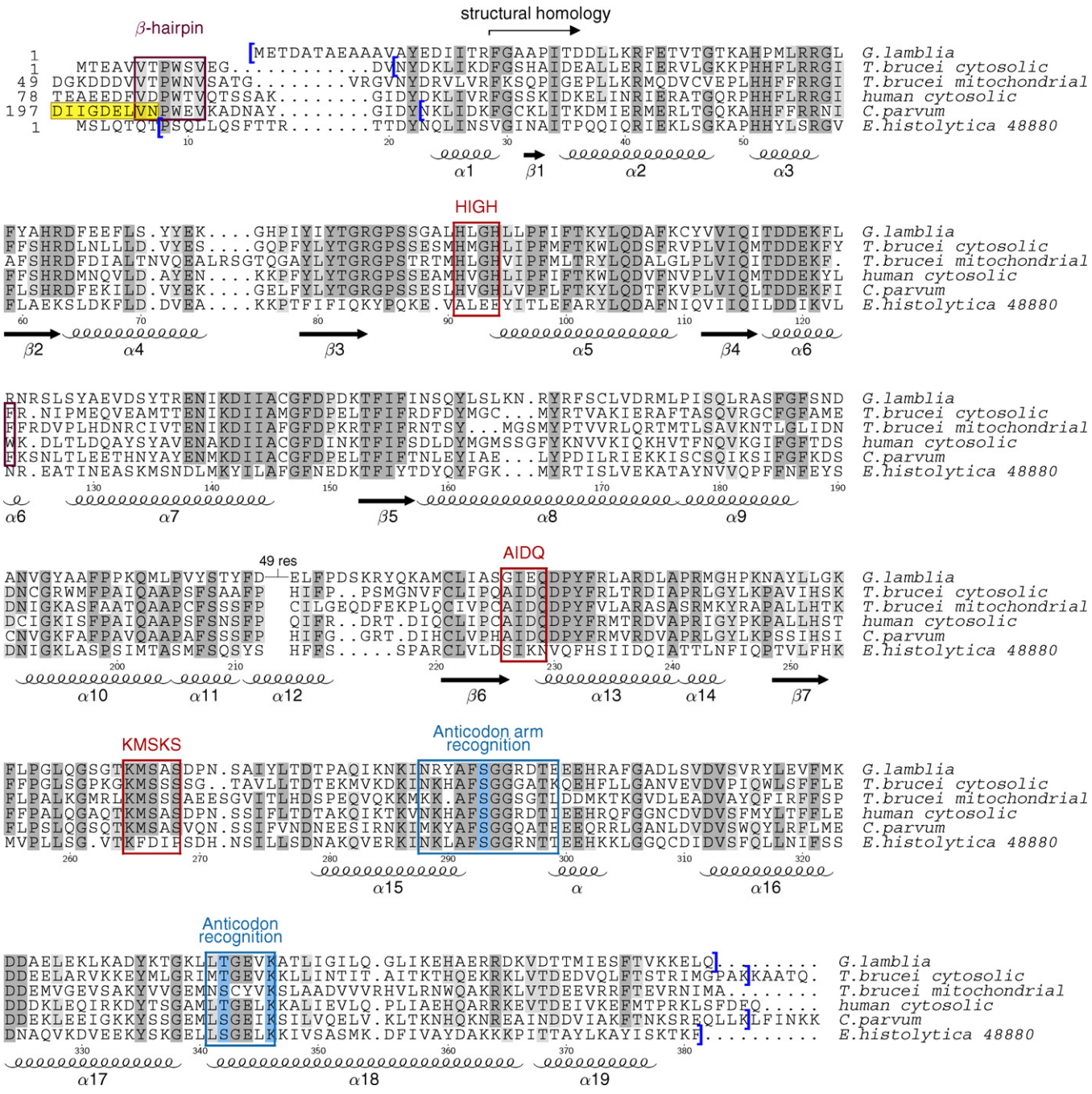

Fig. 1. Sequence alignment of five protozoan TrpRS with human cytosolic TrpRS. Residue numbering is given for the *E. histolytica* sequence. The full-length *T. brucei* mitochondrial, human cytosolic, and *C. parvum* TrpRS contain non-homologous N-terminal extensions whose sequences are not shown here. *C. parvum* residues N-terminal to Pro 206 (yellow highlight) have been replaced by the sequence GPGSM remaining after cleavage of an expression tag from the construct studied here. Sequence motifs associated with the canonical TrpRS active site are boxed in red. These are notably absent in the *E. histolytica* homolog. Regions involved in binding the tRNA anticodon arm as seen in the human TrpRS:tRNA^{Trp} complex (PDB ID: 2dr2; [7]) are boxed in blue. Residues whose sidechains are directly involved in anticodon base recognition in the human TrpRS:tRNA^{Trp} complex are further highlighted in blue. One of these is a strictly conserved lysine. The β-hairpin motif VxxWxV and the paired aromatic residue at the active site present in the human TrpRS and three of the protozoan homologs are boxed in purple. Secondary structure elements are shown for *C. parvum*; they are essentially the same in the human homolog and in the other two protozoan structures described here. The first and last well-ordered residue in each of the respective crystal structures is indicated by blue square brackets. Figure composed using texshade [48]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

L-tryptophan to its cognate tRNA:

$$\mathrm{Trp} + \mathrm{ATP} \Rightarrow \mathrm{Trp}\cdot\mathrm{AMP} + \mathrm{PP}_i$$

$$\mathrm{Trp}\cdot\mathrm{AMP} + \mathrm{tRNA}^{Trp} \Rightarrow \mathrm{AMP} + \mathrm{Trp}\cdot\mathrm{tRNA}^{Trp}$$

Cognate tRNA recognition is carried out by a second, α-helical, domain whose three-dimensional fold is also found in TyrRS. Both TrpRS and TyrRS are homodimeric. Both employ a tRNA binding mode in which the cognate tRNA anticodon arm is recognized and bound by one monomer within the dimer, positioning the acceptor stem of that same tRNA molecule in the active site of the other monomer. Two cognate tRNA molecules can bind simultaneously to one dimer [6–8]. Crystal structures of human cytosolic TrpRS in complex with the cognate tRNA^{Trp} (PDB ID: 2ake, 2dr2; [7]) provide a model for TrpRS:tRNA^{Trp} interaction by eukaryotic homologs including the three protozoan TrpRS examined here. Residues involved in the recognition of tRNA^{Trp} are shown in Fig. 1