

RESEARCH ARTICLE

Open Access

Unusual domain architecture of aminoacyl tRNA synthetases and their paralogs from *Leishmania major*

V S Gowri^{1,2}, Indira Ghosh², Amit Sharma³ and Rentala Madhubala^{1*}

Abstract

Background: *Leishmania major*, a protozoan parasite, is the causative agent of cutaneous leishmaniasis. Due to the development of resistance against the currently available anti-leishmanial drugs, there is a growing need for specific inhibitors and novel drug targets. In this regards, aminoacyl tRNA synthetases, the linchpins of protein synthesis, have received recent attention among the kinetoplastid research community. This is the first comprehensive survey of the aminoacyl tRNA synthetases, their paralogs and other associated proteins from *L. major*.

Results: A total of 26 aminoacyl tRNA synthetases were identified using various computational and bioinformatics tools. Phylogenetic analysis and domain architectures of the *L. major* aminoacyl tRNA synthetases suggest a probable archaeal/eukaryotic origin. Presence of additional domains or N- or C-terminal extensions in 11 aminoacyl tRNA synthetases from *L. major* suggests possibilities such as additional tRNA binding or oligomerization or editing activity. Five freestanding editing domains were identified in *L. major*. Domain assignment revealed a novel asparagine tRNA synthetase paralog, asparagine synthetase A which has been so far reported from prokaryotes and archaea.

Conclusions: A comprehensive bioinformatic analysis revealed 26 aminoacyl tRNA synthetases and five freestanding editing domains in *L. major*. Identification of two EMAP (endothelial monocyte-activating polypeptide) II-like proteins similar to human EMAP II-like proteins suggests their participation in multisynthetase complex formation. While the phylogeny of tRNA synthetases suggests a probable archaeal/eukaryotic origin, phylogeny of asparagine synthetase A strongly suggests a bacterial origin. The unique features identified in this work provide rationale for designing inhibitors against parasite aminoacyl tRNA synthetases and their paralogs.

Keywords: Aminoacyl tRNA synthetases, Paralog, Editing domains

Background

Aminoacyl tRNA synthetases (aaRSs) are indispensable components of protein synthesis. They covalently append amino acids to their cognate tRNA. Most organisms possess separate tRNA synthetases for each of the 20 standard amino acids. There are two classes of aminoacyl tRNA synthetases each comprising of ~10 tRNA synthetase enzymes - Class I tRNA synthetases contain the classical Rossmann-nucleotide binding catalytic domain with two highly conserved 'HIGH' and 'KMSKS' catalytic motifs which are critical for their function [1]. Class II enzymes

contain a central antiparallel β -sheet flanked by α -helices. Despite these structural and sequence differences, both tRNA synthetases catalyse the same two step reaction. The first step involves activation of the amino acid by ATP to form aminoacyl adenylate. The second step is the attachment of the aminoacyl group to the cognate tRNA. While Class I attaches aminoacyl group to the 2'-hydroxyl group of tRNA, Class II synthetases attach them to 3'-hydroxyl group of tRNA. The 3-D structure and the specific anticodon in the tRNA determine the specificity of tRNA synthetases. Most eukaryotes carry two genes for each of the 20 standard aminoacyl tRNA synthetases (cytosolic and mitochondrial). However, trypanosomatids carry only a single copy per amino acid except for Asp, Trp and Lys [2,3].

* Correspondence: madhubala@mail.jnu.ac.in

¹School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India

Full list of author information is available at the end of the article

Gene knock out studies of the Trypanosomal histidyl tRNA synthetase showed a complete arrest of growth in the bloodstream forms of the parasite suggesting an essential role in cell survival [4]. Mammalian methionyl tRNA synthetase provides a cytosolic anchoring site for Aminoacyl tRNA synthetase Interaction Multifunctional Protein-3 (AIMP-3/p18), a potent tumor suppressor in addition to its essential role in initiating translation [5]. Participation of aminoacyl tRNA synthetases in cell apoptosis, rRNA synthesis, RNA trafficking, multisynthetase enzyme complex formation supplement their essential role in protein synthesis [6,7]. As an inevitable component of protein synthesis, aminoacyl tRNA synthetases have been important antibacterial drug targets. An important example of aaRS inhibitor is provided by the antibiotic mupirocin which selectively inactivates bacterial Isoleucyl-tRNA synthetase (IleRS) [8,9]. Four distinct aminoacyl tRNA synthetases in *Anabaena sp.* PCC 7120 contain a novel CAAD domain that bears putative transmembrane helices. Domain deletion studies indicate its essential role in membrane anchoring and a purely structural role and do not alter the catalytic properties of the enzyme [10]. Eukaryotic tRNA synthetases, unlike their prokaryotic counterparts carry additional domains or extensions in their N- or C-terminal regions which mediate protein-protein interaction or involved in tRNA binding. A comprehensive computational analysis of the aminoacyl tRNA synthetases of *Plasmodium falciparum* reveals novel domain architectures. Phylogenetic analyses of several Pf aaRSs reconcile their evolutionary link to plants and bacteria [11]. Recently, the expression and localization profiles of the cis- and trans- aaRS editing domains of *P. falciparum* showed an uneven distribution of 8 aaRS editing domains in the different cellular compartments [12].

Leishmaniasis is one of the deadly diseases caused by the different species of *Leishmania*. Increasing resistance to presently available anti-leishmanial drugs poses a need for identification of novel drug targets as well as specific inhibitors for treating leishmaniasis. In this regard, aminoacyl tRNA synthetases, the versatile players of protein translation machinery have received attention in kinetoplastid research community [13]. Very recently, crystal structure of a methionyl-tRNA synthetase [14] and a novel pseudodimeric structure of a tyrosyl tRNA synthetase from *Leishmania major* have been solved [15]. Substantial differences between the human tRNA synthetases and the *L. major* tRNA synthetase homologue promise a rationale for designing inhibitors to selectively target the parasite enzyme. A comprehensive bioinformatic analysis employing the profile-based hidden markov model (HMM) has identified aaRSs and aaRS related proteins from *L. major*. The sequence features and novel domain architectures of aaRSs from *L. major* were analyzed using a combination

of BLAST and HMM search tools. Domain assignment revealed a novel asparagine tRNA synthetase (AsnRS) paralog Asparagine synthetase A (AsnA) which has been so far reported from prokaryotes and archaea and has been shown to be absent in eukaryotes. We for the first time report the phylogeny and structural analysis of a eukaryotic AsnA from *L. major*.

Results and discussions

A total of 26 aminoacyl tRNA synthetases (11 Class I; 14 Class II; 1 non-standard) were identified in *L. major* (Table 1) using Hidden Markov Models (HMMs). Like other trypanosomatids [2,3], *L. major* also has a single copy of the tRNA synthetases except for Asp, Lys, Trp as well as Pro. The presence of the synthetase and anticodon binding domains were confirmed using the Conserved Domains Database (CDD) domain assignments from NCBI. Based on the generic domain architecture, 25 *L. major* sequences identified using the HMM searches could be certified as authentic aaRSs (Table 1). Among the tRNA synthetase related proteins, LmjF.16.1130 and LmjF.22.0470 contain only an RNA binding domain/Myf domain. However, BLAST sequence search against PDB database identified human EMAP II-like sequences (E-value: $2e-21$; 37%) as the top hit suggesting their sequence relationship with the EMAP II-like sequences such as P43 from human, Arc1p from yeast, Trbp111 from *A. aeolicus* etc. Both LmjF.16.1130 and LmjF.22.0470 also contain a modified heptapeptide motif that has been shown to be essential for the cytokine activity in the human EMAP II-like protein [16]. The presence of 'ELR' motif at the N-terminus has also been shown to be potent promoters of angiogenesis [16]. Aminoacyl tRNA synthetase sequences of Cys (LmjF.12.0250), Asn (LmjF.34.2340), Lys (LmjF.15.0230) and Tyr (LmjF.14.1370) [15] possess an "ELR" motif at the N-terminus. LmjF.26.0830 contains only the Class II synthetase catalytic core with all the three active site motifs conserved. BLAST search against PDB database identified the *E. coli* Asparagine synthetase A structure as the single hit with a reliable statistical value (E-value: $8e-111$). LmjF.26.0830 shares 58% sequence identity with the *E. coli* Asparagine synthetase A.

As the key players in protein translation, most organisms require 20 standard aminoacyl tRNA synthetases for protein synthesis. However, indirect routes of GlntRNA^{Gln} and AsntRNA^{Asn} synthesis also exist in many organisms which either completely lack the respective tRNA synthetases or lack them in some specific organelles such as mitochondria [17]. Kinetoplastid (*Trypanosoma* and *Leishmania*) Seryl tRNA synthetases (SerRS) show a close functional and evolutionary relationship to the metazoan SerRS which is supported by the presence of a metazoan-trypanosomatid specific sequence insertion in SerRS [18]. The kinetoplastid SerRS also show high

Table 1 List of all the aminoacyl tRNA synthetases and their associated proteins, aaRS paralogs and editing domains with their CDD domain assignments and subcellular localization

	Gene name	Length (aa)	HMM based function assignment	E-value#	Swissprot function assignment	Subcellular location (PSORT II)	CDD based domain assignment	
CLASS I aaRS	LmjF.12.0250	784	Cysteinyl-tRNA synthetase	1.00E-191	Cysteinyl-tRNA synthetase	Cytosol	CysRS_Core; anticodon_la_like_superfamily	
	LmjF.15.1440	570	Glutaminyl-tRNA synthetase	8.70E-208	Glutaminyl-tRNA synthetase	Cytosol	GlnRS_core; GlnRS_core; tRNA_synt_lc_C_superfamily	
	LmjF.30.3240	594	Glutaminyl-tRNA synthetase	7.00E-079	Glutaminyl-tRNA synthetase	Cytosol	GlnRS_core; GlnRS_core; tRNA_synt_lc_C_superfamily	
	LmjF.27.1310	692	Arginyl-tRNA synthetase	2.00E-151	Arginyl-tRNA synthetase	Cytosol	Arg_tRNA_synt_N; ArgRS_Core; anticodon_la_like_superfamily	
	LmjF.36.5620	1009	Isoleucyl-tRNA synthetase	0	Isoleucyl-tRNA synthetase	Cytosol	IleRS_Core; Anticodon_la_like_superfamily	
	LmjF.13.1100	1075	Leucyl-tRNA synthetase	3.70E-140	Leucyl-tRNA synthetase	Cytosol	IleRS_Core; Anticodon_la_like_superfamily	
	LmjF.21.0810*	747	Methionyl-tRNA synthetase	6.00E-145	Methionyl-tRNA synthetase	Cytosol	MetRS_Core_superfamily; Anticodon_la_Met	
	LmjF.23.0300	412	Tryptophanyl-tRNA synthetase	1.90E-062	Tryptophanyl-tRNA synthetase	Cytosol	TrpRS_core	
	LmjF.29.0060	480	Tryptophanyl-tRNA synthetase	1.70E-095	Tryptophanyl-tRNA synthetase	Mitochondrial	TrpRS_core	
	LmjF.14.1370*	682	Tyrosyl-tRNA synthetase	1.10E-051	Tyrosyl-tRNA synthetase	Endoplasmic reticulum	TyrRS_core; TyrRS_core	
	LmjF.30.3130	967	Valyl-tRNA synthetase	0	Valyl-tRNA synthetase	Cytosol	ValRS_core; anticodon_la_Val; Val_tRNA_synt_C_superfamily	
	CLASS II aaRS	LmjF.22.1540	962	Alanyl-tRNA synthetase	3.10E-262	Alanyl-tRNA synthetase	Cytosol	AlaRS_core; tRNA_SAD
		LmjF.30.0460	550	Aspartyl-tRNA synthetase	3.90E-106	Aspartyl-tRNA synthetase	Cytosol	AspRS_cyto_N; AsxRS_Core
LmjF.21.0895		641	Aspartyl-tRNA synthetase	5.70E-097	Aspartyl-tRNA synthetase	Mitochondrial	AspRS_cyto_N; ClassII_aaRS_like_core_superfamily; ClassII_aaRS_like_core_superfamily	
LmjF.34.2340		890	Asparaginyl-tRNA synthetase	2.40E-144	Asparaginyl-tRNA synthetase	Cytosol	tRNA_synt_lc_R1_superfamily; AsnRS_cyto_like_N; AsxRS_core	
LmjF.36.3840		628	Glycyl-tRNA synthetase	5.10E-056	Glycyl-tRNA synthetase	Cytosol	GlyRS_core; HGTP_anticodon	
LmjF.30.0130		586	Lysyl-tRNA synthetase	8.20E-184	Lysyl-tRNA synthetase	Cytosol	LysRS_N; LysRS_Core	
LmjF.15.0230		536	Lysyl-tRNA synthetase	9.30E-148	Lysyl-tRNA synthetase	Cytosol	LysRS_N; LysRS_Core	
LmjF.19.1040		633	Phenylalanyl-tRNA synthetase	3.20E-053	Phenylalanyl-tRNA synthetase	Cytosol	B3_4_superfamily; B5_superfamily; PheRS_beta_core	
LmjF.32.0870		499	Phenylalanyl-tRNA synthetase	5.90E-052	Phenylalanyl-tRNA synthetase	Cytosol	PheRS_alpha_core	
LmjF.18.1210		731	Prolyl-tRNA synthetase	1.30E-188	Prolyl-tRNA synthetase	Cytosol	Ybak_like_superfamily; ProRS_arch_euk; HGTP_anticodon_superfamily	
LmjF.18.1220		731	Prolyl-tRNA synthetase	1.20E-190	Prolyl-tRNA synthetase	Cytosol	Ybak_like_superfamily; ProRS_arch_euk; HGTP_anticodon_superfamily	
LmjF.11.0100		474	Seryl-tRNA synthetase	4.40E-135	Seryl-tRNA synthetase	Cytosol	Seryl_tRNA_N_superfamily; SerRS_Core	
LmjF.30.0630		473	Histidyl-tRNA synthetase	2.30E-103	Histidyl-tRNA synthetase	Cytosol	HisRS_Core; HisRS_anticodon	

Table 1 List of all the aminoacyl tRNA synthetases and their associated proteins, aaRS paralogs and editing domains with their CDD domain assignments and subcellular localization (Continued)

	LmjF.35.1410	787	Threonyl-tRNA synthetase	3.20E-252	Threonyl-tRNA synthetase	Cytosol	TGS_superfamily; tRNA_SAD; ThrRS_Core; HGTP_anticodon
Non-canonical	LmjF.09.0950	595	O-Phosphoseryl-tRNA Selenium transferase	9.10E-127	SLA/LP autoantigen-like protein	Nuclear	Selenium_SpcS
aaRS associated proteins	LmjF.22.0470	426	EMAP-II like	1.20E-023	Hypothetical, Conserved	Extracellular	tRNA_binding_domain_superfamily
	LmjF.16.1130	180	EMAP-II like	1.20E-015	Tyrosyl/methionyl-tRNA synthetase	Cytosol	tRNAbindingdomain_superfamily
AsnRS paralog	LmjF.26.0830	353	Asparagine Synthetase A	2.00E-002	Aspartate, ammonia ligase	Mitochondrial	ClassII_aaRS_like_core_superfamily
<i>Trans</i> Editing domains	LmjF.15.0690	491	AlaX	3.90E-028	Hypothetical, conserved	Cytosol	tRNA_SAD_superfamily
	LmjF.03.0710	253	YbaK	2.20E-002	Hypothetical, conserved	Cytosol	Ybak_like_superfamily
	LmjF.21.0910	275	YbaK	7.20E-003	Hypothetical, conserved	Cytosol	Ybak_like_superfamily
	LmjF.36.2730	152	Dtda	1.00E-026	D-tyrosine deacylase	Cytosol	Dtyr_deacylase
	LmjF.34.3360*	211	Dtda	5.40E-020	Hypothetical, conserved	Cytosol	Dtyr_deacylase_superfamily

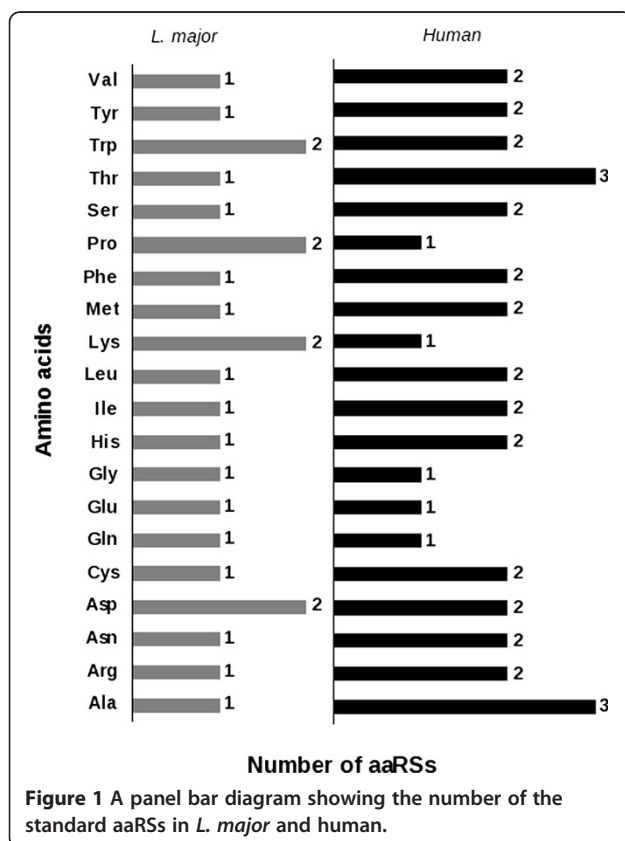
* indicates *L. major* sequences whose crystal structures are available * E-value (Expectation value) is an indication of the significance of a hit to the HMM Model queried. This gives a more quantitative measure of statistical significance. The lower the E-value, the better is the significance of the hit to the query HMM.

affinity for tRNA^{Sec} [18]. Proteins containing non-standard aminoacids such as Sec (Selenocysteine) have been reported from trypanosomatids [19]. The presence of Selenocysteine incorporation is further supported by the presence of a selenophosphosynthetase (LmjF.36.5410), the first enzyme in the selenocysteine tRNA synthesis as well as a selenocysteine specific elongation factor (LmjF.34.2840). LmjF.09.0950, a o-phosphoseryl tRNA (sec) selenium transferase, the third enzyme in the SectRNA^{Sec} synthesis with SepSecS-like domains was also identified using the Hidden Markov Model searches.

Comparison of the number of aminoacyl tRNA synthetases (20 standard aminoacids) of the human with *L. major* (Figure 1) shows a disparity in the number of aaRS for all the aminoacids except for Gly, Glu and Gln where a single copy is present in both human and *Leishmania*. Two copies one each for cytoplasm and mitochondria of AspRS and TrpRS are present in both human and *L. major*. While human possess a single copy of LysRS and ProRS, *L. major* has two copies of these predicted to be in the cytoplasm. One of the LysRS (LmjF.15.0230) has an "ELR" motif at the N-terminus. The two copies of ProRS from *L. major* are identical copies probably a product of gene duplication. Humans possess the maximum number of alanyl and threonyl tRNA synthetases (3 copies each) compared to *L. major* which has a single

copy of each of them. Non-canonical roles of tRNA synthetases require their presence in diverse cellular compartments. Hence, prediction of subcellular localization of the *LmaaRSs* was done using PSORT-II. 80% of the tRNA synthetases are cytosolic and 10% of them are present in mitochondria according to PSORT II predictions (Table 1). Nuclear localization was predicted for the non-standard o-phosphoseryl tRNA(sec) selenium transferase (LmjF.09.0950). Mitochondrial localization was predicted for 3 proteins corresponding to an AspRS, AsnA (AsnRS paralog) and TrpRS (Table 1). Many proteins that are imported into mitochondrion have targeting signals typically at the N-terminus [20] or C-terminus or protein internal. However, numerous mitochondrial proteins have been shown to be lacking these signals including those proteins that have been shown to be imported into mitochondria in trypanosomes. Examples include the glutamyl and glutaminyl tRNA synthetases from *L. tarentolae* [21] and *T. brucei* [22]. Although, the glutaminyl tRNA synthetases are shown to be absent in their mitochondria, the glutaminyl tRNA synthetase activity has been shown experimentally in both these organisms. Hence, it is possible that the single copy tRNA synthetases are probably transported to mitochondria during translation although PSORT II is unable to predict the possibility of mitochondrial localization for the single copy ones. Trans splicing of a leader sequence to the 5' end of the mRNA is a common phenomenon among human and protozoa. This results in alternative splicing in these organisms resulting in proteins with different properties such as gain or loss of targeting signals. Such a mapping of the 5' splice sites using the splice leader trapping method in *T. brucei* resulted in the discovery of nearly 2500 alternative splice events in a stage-regulated manner [23]. The splice sites data for *L. major* at the trypdb server suggests an alternate start site as a result of trans splicing in the promastigote stages for several tRNA synthetases including the single copy tRNA synthetases such as the valyl, isoleucyl, leucyl, glutamyl tRNA synthetases [Additional file 1: Table S1].

The tRNA synthetases and other associated proteins such as the EMAP II and the editing domains although present in all the trypanosomatids, show interesting differences between the *Leishmania* and *Trypanosoma* [Additional file 2: Table S2]. There are two AspRS in all the *Leishmania Spp*, *T. brucei brucei* and *T. brucei gambiense*. However, *T. congolense* and *T. vivax* have only a single copy of AspRS. Furthermore, all the *Leishmania spp* and *Trypanosoma spp* carry a single HisRS except *T. congolense*. *T. cruzi Non-Esmeraldo strain* and *T. congolense* carry three TrpRS whereas all other trypanosomatids carry two TrpRS. *T. cruzi Non-Esmeraldo strain* lacks a MetRS and GlnRS. Moreover, *T. cruzi Non-Esmeraldo strain* contains only an alpha chain of PheRS



and *T. cruzi* Esmeraldo strain lacks an AsnRS and ArgRS. Although several tRNA synthetases are syntenic and conserved in *Leishmania*, the protein expression is regulated at different stages in different *Leishmania* species. For example, SerRS (LmjF.11.0100), LysRS (LmjF.15.230), glutamyl (LmjF.30.3240), AsnRS (LmjF.34.2340), GlyRS (LmjF.36.3840), IleRS (LmjF.36.5620) are regulated predominantly in the promastigotes in *L. major* [24].

Domain architecture of tRNA synthetases from *L. major*

All the standard aminoacyl tRNA synthetases contain the synthetase core domain as well as the anticodon binding domain in some order. Hence, the presence of these generic domains helps in distinguishing the aminoacyl tRNA synthetases from the aminoacyl tRNA synthetase associated proteins (aaRS associated proteins) which refers to EMAP II-like proteins containing only the RNA binding domains. In addition to these generic domains, some of the aminoacyl tRNA synthetases also possess additional domains or extensions tethered to either N- or C-terminus which might be involved in RNA binding or oligomerization. Presence of editing domains ensures the fidelity of protein translation in some tRNA synthetases

by hydrolysing the tRNA aminoacylated with non-cognate amino acid [25]. Thus, in addition to the aaRS, the editing domains (both cis- and trans-) are also novel drug targets. The domain architecture of all the aaRSs (including the non-standard o-phosphoseryl tRNA(sec) selenium transferase) with all the additional domains or motifs, *trans* editing domains, aaRS paralogs and other aaRS associated proteins of *L. major* is shown in Figure 2a and b.

Alanyl and threonyl tRNA synthetases often possess a secondary associated domain (tRNA_SAD) containing a HxxxH motif which is typical of a metal dependent hydrolyses [26]. Alanyl (LmjF.22.1540) and threonyl tRNA synthetases (LmjF.35.1410) of *L. major* contain this domain. The presence of a tRNA_SAD domain with a conserved HxxxH motif suggests a functionally important hydrolytic activity (Figure 2b). *Lm*ThrRS also contains a TGS domain tethered N-terminus to the tRNA_SAD (Figure 2b). Based on its occurrence in other regulatory proteins, this domain is proposed to bind ligands (most likely nucleotides) [27]. Hence, the TGS domain in *Lm*ThrRS probably has a regulatory role. In addition to an editing domain tRNA_SAD, *Lm*AlaRS has a C-terminal extension (DHHA1 domain). Crystal structure and

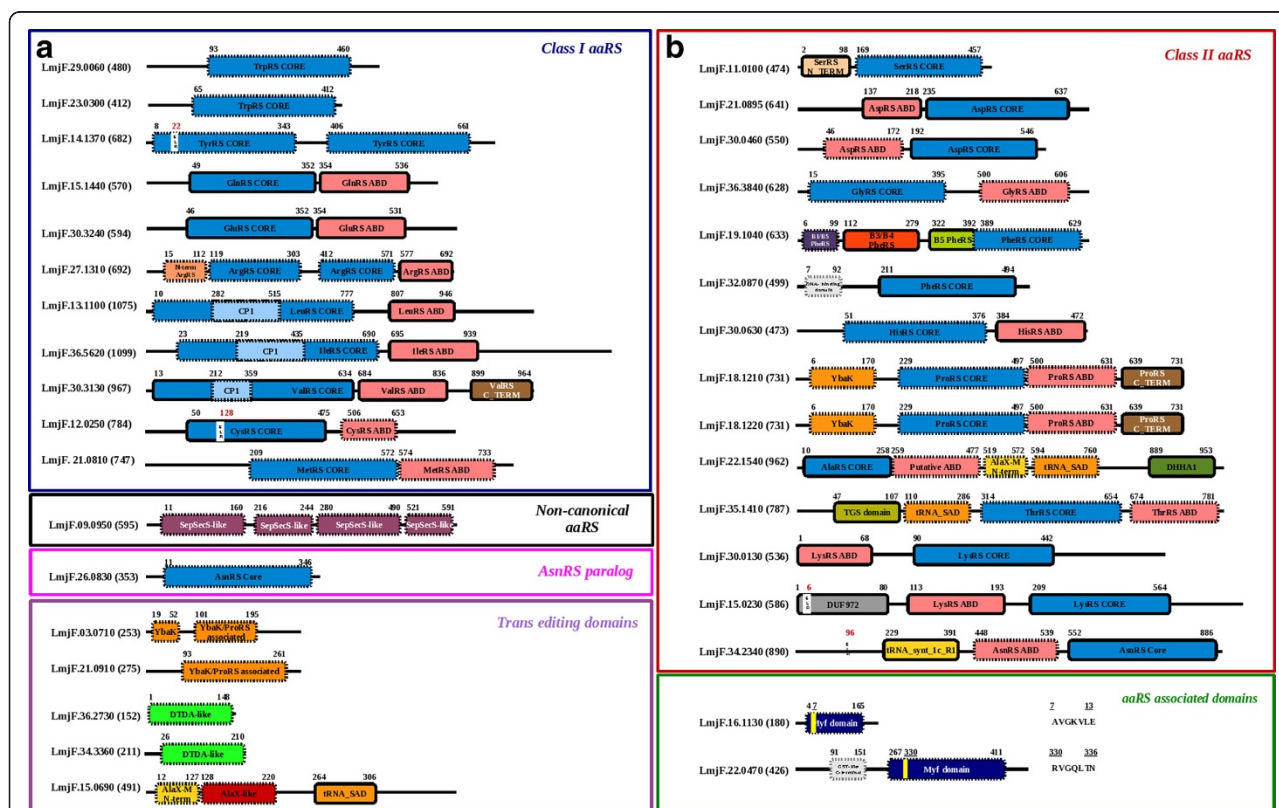


Figure 2 Domain architecture of (a) Class I and (b) Class II *Lmaar*Ss, aaRS associated proteins, and *trans* editing domains. The box line type for domain boundaries correspond to the domain prediction procedure (solid line for PFAM HMM predictions; dotted line for SUPERFAMILY based HMM predictions). The beginning of ELR motif is denoted by the residue number in red color font. The beginning of heptapeptide motif is denoted by an underlined residue number. The grey boxes corresponds to lower prediction accuracy (Evalue: greater than 1E-05).

functional analysis of this C-Ala extension in *A. aeolicus* AlaRS shows that it promotes cooperative binding of the aminoacylation and editing domain to tRNA^{Ala} [28].

While the C-terminal extension of *LmAlaRS* might be involved in oligomerization, N- or C-terminal extensions of SerRS, LeuRS and LysRS have been shown to provide additional tRNA binding to these synthetases [29-31]. The N-terminal extension of *LmLeuRS* (LmjF.13.1100) is present as insertion in the editing domain denoted as CP1 (Connective Polypeptide) (Figure 2a). Only three LeuRS editing domains have been structurally characterized till date [31,32]. The CP1 of *E. coli* LeuRS lacks this insertion and hence lacks the editing activity as an isolated CP1 domain [33]. Recent crystallographic and biochemical evidences reconcile this observation [31]. *LmLeuRS* has an N-terminal extension of approximately 35 residues long. Secondary structure prediction of this insertion using PSIPRED server [29] suggests that this N-terminal extension has a helix of ~15 residues long (Additional file 3: Figure S1). Further, sequence comparison of the editing domain of *LmleuRS* with that of human, *E. coli*, *A. aeolicus* and *G. lamblia* suggest the T-rich region, GTG motif and the conserved Asp essential for function are all conserved. Sequence-based phylogeny suggests a close evolutionary relationship of *LmLeuRSCP1* to *GLeuRSCP1* which has been verified to possess fully functional editing domain in isolation [31]. The anti-fungal drug (AN2690) binding residues of *C. albicans* LeuRS are also highly conserved in *LmLeuRS* editing domain. Although *LmLeuRS* is ~1100 residues long, the presence of a probable functional editing domain in isolation proves it to be a novel drug target and encourages experimental verification for its drug binding abilities.

L. major encodes two cytosolic LysRS (LmjF.15.0230 and LmjF.30.0130) (Table 1). One of the *LmLysRS* (LmjF.15.0230) has an N-terminal extension (DUF972) similar to the mammalian LysRS (Figure 2b). The N-terminal extension of mammalian LysRS has been shown to participate in non-specific tRNA binding [26]. Deletion of this N-terminal extension has been shown to reduce the tRNA binding affinity by 100-fold and hence decreases the aminoacylation of tRNA^{Lys} by 3-fold in mammals [29]. Based on the sequence homology of the *LmLysRS* (LmjF.15.0230) to the mammalian LysRS, the N-terminal extension in LmjF.15.0230 can be expected to participate in a non-specific tRNA binding and could probably play a role in amino acylation activity of this LysRS. LmjF.15.0230 also contains an 'ELR' motif at the N-terminal extension. However, chemokine activity of this 'ELR' motif in *LmLysRS* requires experimental verification.

Stand-alone deacylase/trans editing domains in *L. major*

While the pairing of the correct amino acid to their cognate tRNA is done by the aminoacyl tRNA synthetases,

faithful translation in protein synthesis is ensured by the presence of editing domains(ED) either tethered to the aaRS (*cis* editing domains) or as free standing editing domains (*trans* editing domains). There are 8 *cis* editing domains tethered to AlaRS, ThrRS, PheRS, LeuRS, IleRS, ValRS, ProRS (both the copies) and 5 *trans* editing domains which includes the AlaX (AlaRS ED), two YbaK-like (ProRS ED) and two D-tyrosyl deacylases (DTDA) in *L. major*.

The Second Associated domain (tRNA_SAD) of AlaRS/ThrRS are generally tethered to the synthetase core. In *L. major*, in addition to the tethered editing domains in the AlaRS (LmjF.22.1540) and ThrRS (LmjF.35.1410), a freestanding tRNA_SAD (LmjF.15.0690) domain (Figure 2a) was also found. There are two types of standalone AlaX domains: AlaX-M, AlaX-S. Both the domains differ in their metal coordination types (Zn coordination). In AlaX-M, in addition to coordination with the Cysteine residues, there is coordination with a water molecule. However, in AlaX-S, the metal ion is coordinated only by cysteines [34,35]. The standalone domain has all the four cysteines conserved. Sequence based phylogeny suggests that *L. major* tRNA_SAD standalone domain is closer to AlaX-M family (Additional file 4: Figure S2).

In addition to the tethered YbaK/ProX deacylase domains in the two *LmProRS* copies, two freestanding YbaK domains (LmjF.03.0710 and LmjF.21.0910) are also found (Figure 2a). Sequence based phylogeny of both tethered and standalone deacylase domains in *L. major* with the available crystal structures of YbaK/ProX domains suggest that the tethered deacylase domains are closer in terms of their amino acid sequence to ProX type which specifically deacylate the misacylated tRNA^{Pro} with Alanine [36,37] while the trans deacylase domains are closer to YbaK-like which deacylate the misacylated tRNA^{Pro} with Cysteine [37,38] (Additional file 5: Figure S3). The *trans* editing domain LmjF.03.0710 lacks the active site lysine which is shown to be critical in *H. influenza* YbaK domain by mutagenesis (K46 in PDB: 1DBX) [38-40]. Thus, only one of the *trans* editing domain of *LmProRS* (LmjF.21.0910) which has the active site lysine conserved might be functional.

Detection of D-amino acids in the form of free amino-acids, peptides and proteins in various living organisms from bacteria to human challenges the current concept of protein synthesis [41,42]. D-Tyr tRNA^{Tyr} deacylases (DTDA), a new class of tRNA dependent hydrolases provide a novel checkpoint by recycling the misaminoacylated D-Tyr tRNA^{Tyr}. There are two DTDA in *L. major*. One of the *LmDTDA2* (LmjF.34.3360) has a crystal structure solved using the Structural Genomics of Pathogenic Protozoa Consortium (SGPP) (PDB: 1TC5). This sequence is closer to the human, mouse DTDA2 homologue. The

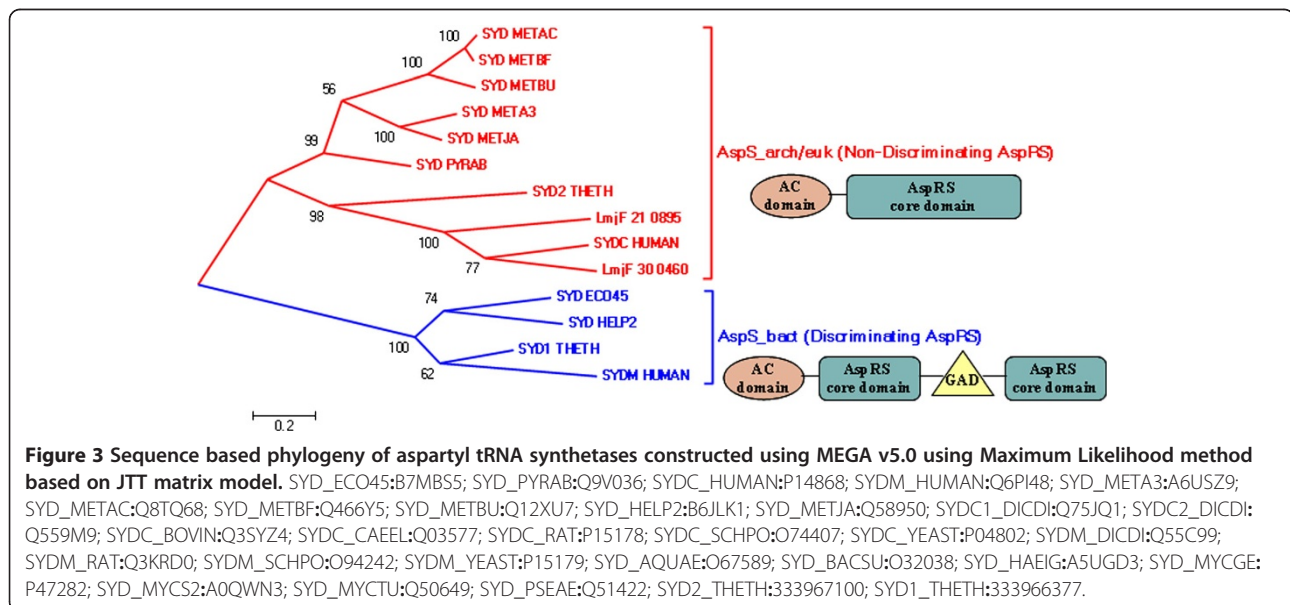
other *LmDTDA1* homologue (LmjF.36.2730) is closer to the *PfDTDA* homologue (Additional file 6: Figure S4) whose structure has been solved recently (3KO5). Like other deacylases/editing domains of aaRSs, DTDA has been shown to be a novel drug target in *P. falciparum* [43]. While structural comparison of the *LmDTDA2* and *PfDTDA1* shows major differences in the length of a specific loop, 3-D structural modeling of *LmDTDA1* homologue (LmjF.36.2730) based on the *PfDTDA* template structure and a comparative structural analysis of *LmDTDA1* with a perspective of inhibitor design is commendable.

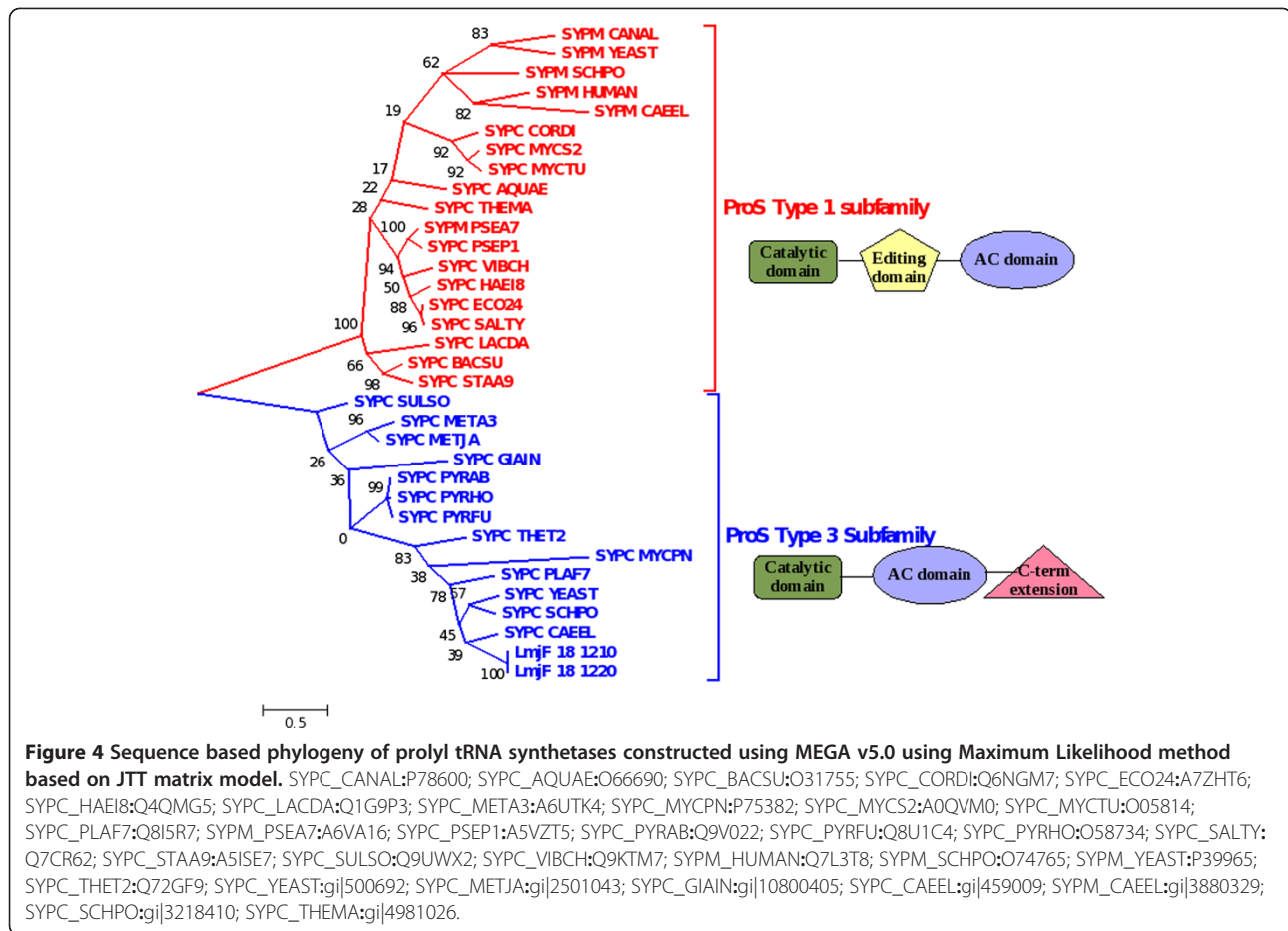
Novel aminoacyl tRNA synthetases from *L. major*

L. major, like other trypanosomatids encodes two AspRS enzymes (one cytosolic and a mitochondrial enzyme). Conserved Domains Database (CDD) based domain architecture suggest that the cytosolic AspRS (LmjF.30.0460) has a non-discriminating catalytic core (AsxRS which can charge Asp/Asn) while the mitochondrial copy (LmjF.21.0895) contains the canonical AspRS catalytic core (Table 1). Sequence-based phylogeny clearly suggests that *Leishmania major* encodes two Eukaryotic or Archaeal type AspRS (Non-discriminating) (Figure 3). This is further confirmed by the PFAM domain assignments. While the Bacterial AspRS contain a GAD domain inserted within the catalytic core which could probably function as an editing domain, the Archaeal/Eukaryotic AspRS lack this GAD domain and hence belong to the non-discriminating type AspRS. An Asp/Asn synthetase domain (AsxRS) can acylate either Asp or Asn in a non-discriminating manner [44,45]. Generally, a non-discriminating type AspRS is involved in the indirect pathway of Asparagine tRNA synthesis [44,45]. The

indirect pathway, in addition to a non-discriminating type AspRS requires GatCAB complex which is a multiprotein complex involved in transamidation of AsptRNA^{Asn}. However, *L. major* has only GatA (LmjF.16.1360) and a very distant homolog of GatC (LmjF.18.0110). This distant homolog of GatC in *L. major* is closely related to GatF of yeast [46]. Yet GatB is absent. In *P. falciparum* only GatA and GatB subunits have been reported [11]. While, the indirect pathway of tRNA(Gln) and tRNA(Asn) charging requires either an GatCAB or GatDE in bacteria and archaea respectively. It has been reported earlier that yeast requires a GatFAB for transamidation [46,47]. The GatF subunit belongs to DUF726 (Domain of unidentified function) family of PFAM. Our database searches showed that *P. falciparum* also has a homolog of GatF (PFL0295c). It is possible that *Leishmania* and *Plasmodium* have a GatFAB instead of GatCAB. However, this requires validation.

There are two cytosolic ProRS enzymes both of which are annotated as bifunctional enzymes from *L. major* (LmjF.18.1210; LmjF.18.1220). Conserved Domains Database (CDD) based domain assignments suggest that they have a YbaK domain in addition to the catalytic core and anticodon binding domains (Table 1). YbaK domain has been suggested to hydrolyse misacylated tRNA^{Pro}; essentially editing function [40]. Sequence based phylogeny of ProRS from all domains of life suggest that the *LmProRS* cluster with other cytosolic eukaryotic and archaeal ProRS all of which belong to the ProS type 3 subfamily domain architecture (according to PFAM domain assignment); while the bacteria and other eukaryotic mitochondrial enzymes have a ProS type 1 subfamily domain architecture (Figure 4). ProRS have been shown to be capable of aminoacylating both tRNA^{Pro} and tRNA^{Cys} with their





respective cognate aminoacids in archaea lacking CysRS. Sequence based clustering of *Lm*ProRS with the bifunctional archaeal ProRS enzymes suggest that the bifunctional *Lm*ProRS enzymes are probably capable of charging both tRNA^{Pro} and tRNA^{Cys} with their respective aminoacids. The presence of a ProX type cis editing domain at the N-terminus of both *Leishmania* ProRS which specifically hydrolyses the misacylated tRNA^{Pro} with Alanine (Additional file 5: Figure S3) further confirms the bifunctional ability of *Lm*ProRS proteins.

Asparagine synthetase A from *L. major* – A novel enzyme specific to Prokaryotes and Archaea

Aminoacyl tRNA synthetase paralogs so far have been reported from prokaryotes [48]. These paralogs while retaining the aaRS catalytic domain with the characteristic motifs, are primarily involved in aminoacid biosynthesis [49,50]. Examples include AsnA, HisZ, lysylation of a specific lysine in EF-P (GenX/PoxA) [49-54] etc. Absence of these paralogs in mammals makes them unique antibacterial drug targets. Biochemical characterization of AsnA and GenX/PoxA from *E. coli* are available [51-55]. One of the *L. major* protein (LmjF.26.0830) has

a AsnRS catalytic core with all the three characteristic class II motifs conserved (Figure 4b). But, it lacks the anticodon binding domain essential for tRNA binding (Table 1). A blast sequence search against PDB database suggests close sequence similarity (~58%) with *E. coli* AsnA. Crystal structure of the *E. coli* protein indeed shows a class II tRNA synthetase core domain structure [55]. The structure based sequence comparison of yeast AspRS catalytic core with *Ec*AsnA shows conservation of structurally and catalytically important residues between the two sequences [55]. While the two substrates (mgATP and aspartic acid) of these two sequences are similar, the reactive carboxyl groups of aspartic acid are different. While in AspRS, the α -carboxyl group of aspartic acid is activated by ATP, β -carboxyl group is activated in AsnA (Figure 5a). In prokaryotes, asparagine is formed by two structurally distinct asparagine synthetases. One is the ammonia utilizing asparagine synthetase referred as AsnA and the other is the glutamine utilizing asparagine synthetase referred as AsnB. Although, AsnB can utilize glutamine or ammonia as the amide donor, glutamine is preferred over ammonia [56,57]. Recently, the crystal structure of AsnA from an

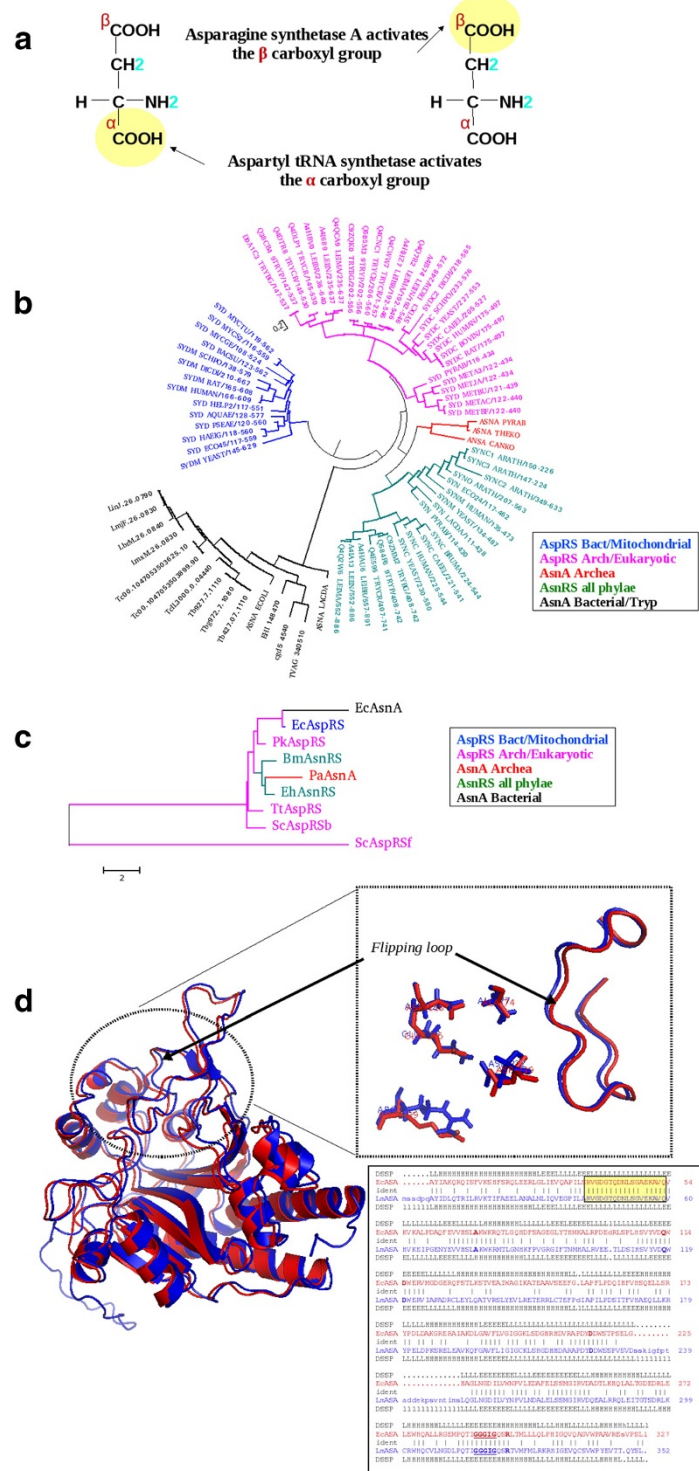


Figure 5 (See legend on next page.)

(See figure on previous page.)

Figure 5 (a) Schematic representation of the difference in the activation of L-Aspartic acid by aspartyl-tRNA synthetase (AspRS) and Asparagine synthetase A (AsnA). (b) Lysyl tRNA synthetase rooted sequence based phylogeny of Asparagine synthetase A with aspartyl and asparagine tRNA synthetases. This tree is constructed using MEGA v5.0 using maximum likelihood method based on JTT matrix model. (c) Structure dependent sequence based phylogeny of the crystal structures of AsnA from *E. coli* (EcAsnA: 11AS); *P. abyssi* (PaAsnA: 3P8Y), AspRS from *T. kodakarensis* (TkAspRS: 3NEM); *T. thermophilus* (TtAspRS: 1EFW); *T. thermophilus* (TtAspRS2: 1IL2); *S. cerevisiae* (ScAspRSf: 1E0V) and AsnRS from *E. histolytica* (EhAsnRS: 3M4P); *B. malayi* (BmAsnRS: 2XGT). The tree is constructed using the structure based sequence alignment generated using MUSTANG structural alignment program. (d) A snapshot of the active site residues in EcASA (PDB: 11AS A chain; Red color) with LmASA structural model (blue color). Structural comparison shows the superposition of the flipping loop and active site residues from both the structures. The pairwise sequence alignment of the corresponding structural superposition with the respective color coding is also shown. The ATP binding Glycine residues are shown in BOLD, underlined fonts. The residues in the flipping loop are shown in the yellow boxes and the active site residues are shown in BOLD font.

archaea (*Pyrococcus abyssi*) with different substrate bound forms including AMP, asparatate, asparagine has guided in decoding the plausible mechanism of asparagine synthesis by the archaeal AsnA enzyme [58,59].

AsnA genes have been reported from prokaryotes and archaea [59,60], while AsnB genes are reported from all three domains of life. Leishmania and trypanosoma albeit being eukaryotes surprisingly possess AsnA (LmjF.26.0830) and AsnB (LmjF.29.1490). Swissprot data [18,19] suggests the presence of AsnA in almost 368 organisms all belonging to prokaryotic origin. In addition to kinetoplastids, blast searches against EupathDB database [61] suggest parasites such as *Trichomonas vaginalis* (TVAG_340510; *E-value*: 8.9E-50); *Entamoeba histolytica* (EHI_148470; *E-value*: 1.4E-73); *Cryptosporidium hominis* (Chro.50501; *E-value*: 6.0E-26); *Cryptosporidium parvum* (cgd5_4540; *E-value*: 2.1E-52) possess a copy of AsnA gene. LmAsnA is predicted to be a mitochondrial copy (Table 1). Among the Class II synthetases, lysine, asparagine and aspartic acid are closely related in their structure and belong to the same subtype (Class 2b) [62]. However, evidences support the evolutionary link between the asparagine synthetase and aspartyl tRNA synthetases as they both recognize aspartic acid and ATP [55]. Thus, a LysRS rooted sequence based phylogeny of AsnA along with AspRS and AsnRS catalytic core from all three domains of life clearly shows that the kinetoplastid and other eukaryotic pathogen AsnA enzymes are of bacterial origin (Figure 5b) while the archaeal AsnA is derived from gene duplication events from the ancestral AspRS as previously mentioned by Blaise and workers [59]. Structure dependent sequence based phylogeny of all the available crystal structures of AspRS, AsnRS with the EcAsnA and PaAsnA enzymes show a similar tree branching with the EcAsnA structurally closer to the bacterial AspRS (Figure 5c). The distinct branching of yeast AspRS free form clearly reflects the conformational rearrangements upon tRNA binding to the yeast AspRS [63].

To date crystal structures of three AsnA enzymes (from *E. coli*, *P.abyssi* and an AsnA peptide structure

from *P. furiosus*) are available. Among the three, amino acid sequences of *L. major* AsnA and EcAsnA closer to each other. Hence, a structural model of LmAsnA was built using EcAsnA (PDB:11AS) as the template using Modeller v 9.0. The model was energy minimized using Amber96 forcefield in gromacs. The quality of the energy minimized model is then verified using PROCHECK available at PDBSUM [64] webserver. Structural comparison of the LmAsnA model with the *E. coli* and *P.abyssi* homologue suggests that the LmAsnA shares 58% sequence identity with the bacterial homologue and 19% sequence identity with the archaeal homologue. The *L. major* model superposes with the *E. coli* and *P. abyssi* structures at 0.5Å and 2.9Å RMSD values respectively. 3-D structural comparison of the LmAsnA model with the EcAsnA (PDB:11AS) shows a complete conservation of catalytic residues, ATP binding Glycine rich region and identical flipping loop lengths that covers the active site at the sequence level and in 3-Dimension (Figure 5d). Based on the sequence and structural similarities between the *E. coli* and the *L. major* enzymes, D222-Q118 pair of LmAsnA can be expected to anchor the beta carboxylate group of the L-aspartic acid. In yeast AspRS, these residues are substituted by a threonine and a glycine respectively and an Asp342 and Q303 which are at structurally different positions anchors the beta carboxylate group of the substrate [55]. While in LmAsnA, EcAsnA and yeast AspRS, an Asp-Gln pair anchors the beta carboxylate group of L-aspartic acid, in archaeal AsnA and AspRS homologues; Aspartic acid is anchored by two arginines [55]. This suggests that the altered substrate specificity and the reaction chemistry between the AspRS and AsnA have been achieved by a few residue substitutions at the active site. The basic difference in the substrate anchoring residues between the archaeal and bacterial/kinetoplastid AsnA enzymes suggests a distinct evolutionary origin between the archaeal and bacterial/kinetoplastid AsnA. These key differences between the archaeal and the kinetoplastid AsnA substrate recognition modes and the absence of AsnA in human make the kinetoplastid enzyme unique drug target for antiparasitic drug design.

Conclusions

Aminoacyl tRNA synthetases are ubiquitous enzymes essential for cell viability. Hence, they have been one of the promising drug targets in antimicrobial infections. Due to increasing resistance to currently available antileishmanial drugs, aminoacyl tRNA synthetases have received attention among the kinetoplastid research community in the recent times. In this study, aminoacyl tRNA synthetases and their associated proteins from *L. major* have been explored for their novel domain architectures and sequence features. Based on the domain architecture, we identified 26 indisputable aminoacyl tRNA synthetases from *L. major*, with a predominant predicted localisation of them in the cell cytosol. Sequence based phylogeny of some specific tRNA synthetases (AspRS and ProRS) confirm their close evolutionary relationship with archaeal/eukaryotic tRNA synthetases. In addition to the appended editing domains and N- or C-terminal extensions which provide additional tRNA binding, we also identified free standing editing domains of AlaRS/ThrRS, two ProRS deacylases and two D-tyrosine deacylases (DTD). Two novel EMAP II-like sequences containing a heptapeptide motif similar to the human EMAP II-like sequences were also identified. The presence of such EMAP II-like sequences suggests the formation of a probable multisynthetase protein complex as seen in the case of human or their probable role in trans-activation of certain aminoacyl tRNA synthetases. Presence of 'ELR' motif in Lys, Asn, Cys and Tyr tRNA synthetases provides clues for their participation in angiogenesis likely. We also highlight the sequence analysis and 3-D structural modelling of a unique enzyme that is completely absent in human, Asparagine synthetase A from *L. major* for the first time. While the aminoacyl tRNA synthetases of *L. major* show archaeal/eukaryotic origin, Asparagine synthetase A of *L. major* shows bacterial origin. The different substrate recognition modes of the bacterial and archaeal enzymes makes them unique and worth exploring.

Methods

Leishmania major (Version 3.1) from TritypDB database [65] is used here. Hidden Markov Models (HMM) [66] were generated using aminoacyl tRNA synthetase sequences and the editing domain sequences (Ybak, DTDA, AlaX) from Swissprot database Release 4.0, 2011 [67,68] for each of the 21 tRNA synthetases (20 standard tRNA synthetases + α -phosphoseryl tRNA(sec) selenium transferase) and the deacylases of ProRS (YbaK), AlaRS (AlaX) and D-tyrosine deacylase (DTDA). The distribution of aaRS sequences from the individual domains of life used for the generation of HMMs is given in the Additional file 7: Table S3. hmmbuild and hmmsearch

options in the suite of HMMER 3.0 package [66] was used for generation and searches using the HMMs respectively. Multiple sequence alignment used for model generation was done using MAFFT multiple sequence alignment tool [69] which employs fast fourier transforms (FFT) for rapid identification of homologous regions. The accuracy of alignments generated by MAFFT has been proved comparable to CLUSTALW and T-coffee progressive alignment methods with the rapid reduction of CPU time [69]. BLAST Webserver at NCBI was used extensively for sequence searches against PDB database [70]. PSORT II [71] was used for subcellular localization prediction analysis. The prediction accuracy for cross validation of yeast sequences is about 57%. PSORT II does not account for multiple localization of protein sequences. PSIPRED (Protein structure prediction server) [72] is used for secondary structural prediction of the leucyl tRNA synthetase of *L. major*. PFAM database [73] from the Sanger Institute, Conserved Domains Database (CDD Server) at NCBI [74] and SUPERFAMILY database (Version 1.75) [75] were used for domain assignments. BLAST searches against PDB database was used for the assignment of deacylase (Connective peptide; CP) domains for Leu, Ile and Val tRNA synthetases.

Phylogenetic analysis

Phylogenetic analysis of the *LmaarSs* was performed combining the set of sequences from the Swissprot/UniprotKB database [67]. Multiple sequence alignment (MSA) of these sequences is generated using CLUSTALW with default parameters [76]. These MSAs were used as seed sequences for phylogenetic tree generation using Jones-Taylor-Thornton (JTT) model [77]. MEGA v5 [78] was used for both analysis and visualization of the phylogenetic trees.

Model building and validation

Comparative structural model of *L. major* Asparagine Synthetase A was built using Modeller v9 [79]. Stereochemical quality of the model was verified using PROCHECK in PDBSUM web resource at EBI [63]. Structural mapping of the active site residues was performed using Pymol [80].

Additional files

Additional file 1: Table S1. Alternative splicing of aaRSs, aaRS paralogs, editing domains and other associated proteins in Promastigote stages in *Leishmania major*.

Additional file 2: Table S2. Distribution of aaRSs, aaRS paralogs, editing domains and other associated proteins in *Leishmania* and *Trypanosoma*. Organisms are Lm: *L. major*; Lbr: *L. braziliensis*; Lta: *L. tarentolae*; Lmx: *L. mexicana*; Lin: *L. infantum*; Tb427: *T. brucei brucei*427; Tb927: *T. brucei brucei* (TREU927); Tbg: *T. brucei gambiense*; TcEs: *T. cruzi* Esmeraldo; TcNEs: *T. cruzi* Non-Esmeraldo; Tcon: *T. congolense*; Tv: *T. vivax*.

Additional file 3: Figure S1. Secondary structure prediction of LmLeuRS showing the first 300 residues which includes the CP1 domain. The 35 residue N-terminal insertion in the CP1 domain is highlighted in yellow. The numerical values corresponds to the confidence levels (ranging from 0–9). H refers to Helix; C refers to Coil and E refers to Sheet.

Additional file 4: Figure S2. Sequence based phylogeny of tRNA_SAD domains (cis/trans) of alanyl tRNA synthetases constructed using MEGA v5.0 using Maximum Likelihood method based on JTT matrix model. Bootstrap values are indicated at the inner nodes.

Additional file 5: Figure S3. Sequence based phylogeny of prolyl tRNA synthetase editing domains (Ybak/ProX) constructed using MEGA v5.0 using Maximum Likelihood method based on JTT matrix model. Bootstrap values indicated at the inner nodes indicate the similarity of tethered editing domains of LmProRS to AlatRNA^{Pro} type editing domains (ProX) and the standalone editing domains of LmProRS to CystRNA^{Pro} type (Ybak).

Additional file 6: Figure S4. Sequence based phylogeny of D-tyrosine deacylases (Dtdas) from DTDA2_HUMAN (sp|Q96FN9); DTD2_MOUSE (sp|Q8BHA3); DTD2_DANRE (sp|Q68EL2); DTD1_ECOLI (1JKE); DTD2_HINF (1J7G); DTD1_AAQU (2DBO); DTD1_PFAL (3K05); DTD1_BOVIN (sp|Q2T9V8); DTD1_HUMAN (2OKV); GEK1_ARATH (sp|Q9ZPQ3); DTDA_PYRFU (sp|P58852); DTDA_PYRHO (sp|O57774); DTDA_PYRAB (sp|Q9V2R8) constructed using MEGA v5.0 using Maximum Likelihood method based on JTT matrix model. Bootstrap values are indicated at the nodes.

Additional file 7: Table S3. Distribution of aaRS sequences used for generation of HMMs for all the 21 aaRS from the Swissprot database in Bacteria, Archea, Eukaryotes and Virus is given in the table.

Abbreviations

aaRS: Aminoacyl tRNA synthetases; CDD: Conserved domains database; PDB: Protein data bank.

Competing interests

The author declares that they have no competing interests.

Authors' contributions

VSG performed the analysis and wrote the manuscript. RMB, AS and IG analyzed the problem and reviewed the manuscript. All authors read and approved the final manuscript.

Acknowledgements

RMB is a JC Bose National Fellow. VSG is currently a UGC D.S. Kothari Postdoctoral research fellow. This work is supported by DBT-COE postdoctoral research fellowship to VSG.

Author details

¹School of Life Sciences, Jawaharlal Nehru University, New Delhi 110 067, India. ²School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi 110 067, India. ³Structural and Computational biology, International Centre for Genetic Engineering and Biotechnology, New Delhi 110 067, India.

Received: 11 April 2012 Accepted: 30 October 2012

Published: 14 November 2012

References

1. Ibba M, Söll D: Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 2000, **69**:617–650.
2. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, et al: The genome of the African trypanosome *Trypanosoma brucei*. *Science* 2005, **309**:416–422.
3. Charrière F, Helgadóttir S, Horn EK, Söll D, Schneider A: Dual targeting of a single tRNA(Trp) requires two different tryptophanyl-tRNA synthetases in *Trypanosoma brucei*. *Proc Natl Acad Sci USA* 2006, **103**:6847–6852.
4. Merritt EA, Arakaki TL, Gillespie JR, Larson ET, Kelley A, et al: Crystal structures of trypanosomal histidyl-tRNA synthetase illuminate differences between eukaryotic and prokaryotic homologs. *J Mol Biol* 2010, **397**:481–494.
5. Kwon NH, Kang T, Lee JY, Kim HH, Kim HR, et al: Dual role of methionyl-tRNA synthetase in the regulation of translation and tumor suppressor activity of aminoacyl-tRNA synthetase-interacting multifunctional protein-3. *Proc Natl Acad Sci USA* 2011, **108**:19635–19640.
6. Park SG, Choi EC, Kim S: Aminoacyl-tRNA synthetase-interacting multifunctional proteins (AIMPs): a triad for cellular homeostasis. *IUBMB Life* 2010, **62**:296–302.
7. Bhatt TK, Khan S, Dwivedi VP, Banday MM, Sharma A, et al: Malaria parasite tyrosyl-tRNA synthetase secretion triggers pro-inflammatory responses. *Nat Commun* 2011, **2**:530.
8. Hurdle JG, O'Neill AJ, Ingham E, Fishwick C, Chopra I: Analysis of mupirocin resistance and fitness in *Staphylococcus aureus* by molecular genetic and structural modeling techniques. *Antimicrob Agents Chemother* 2004, **48**:4366–4376.
9. Hurdle JG, O'Neill AJ, Chopra I: Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents. *Antimicrob Agents Chemother* 2005, **49**:4821–4833.
10. Olmedo-Verd E, Santamaría-Gómez J, Ochoa De Alda JA, Ribas De Pouplana L, Luque I: Membrane Anchoring of Aminoacyl-tRNA Synthetases by Convergent Acquisition of a Novel Protein Domain. *J Biol Chem* 2011, **286**:41057–41068.
11. Bhatt TK, Kapil C, Khan S, Jairajpuri MA, Sharma V, et al: A genomic glimpse of aminoacyl-tRNA synthetases in malaria parasite *Plasmodium falciparum*. *BMC Genomics* 2009, **10**:644–657.
12. Khan S, Sharma A, Jamwal A, Sharma V, Pole AK, et al: Uneven spread of cis- and trans-editing aminoacyl-tRNA synthetase domains within translational compartments of *P. falciparum*. *Nature Scientific reports* 2011, **1**:1–11.
13. Shibata S, Gillespie JR, Kelley AM, Napuli AJ, Zhang Z, et al: Selective inhibitors of methionyl-tRNA synthetase have potent activity against *Trypanosoma brucei* Infection in Mice. *Antimicrob Agents Chemother* 2011, **55**:1982–1989.
14. Larson ET, Kim JE, Zucker FH, Kelley A, Mueller N, et al: Structure of *Leishmania major* methionyl-tRNA synthetase in complex with intermediate products methionyladenylate and pyrophosphate. *Biochimie* 2011, **93**:570–582.
15. Larson ET, Kim JE, Castaneda LJ, Napuli AJ, Zhang Z, et al: The double-length tyrosyl-tRNA synthetase from the eukaryote *Leishmania major* forms an intrinsically asymmetric pseudo-dimer. *J Mol Biol* 2011, **409**:159–176.
16. Wakasugi K, Schimmel P: Highly differentiated motifs responsible for two cytokine activities of a split human tRNA synthetase. *J Biol Chem* 1999, **274**:23155–23159.
17. Ibba M, Söll D: The renaissance of aminoacyl-tRNA synthesis. *EMBO Rep* 2001, **2**:382–387.
18. Geslain R, Aeby E, Guitart T, Jones TE, Castro De Moura M, et al: *Trypanosoma seryl-tRNA synthetase* is a metazoan-like enzyme with high affinity for tRNA^{Sec}. *J Biol Chem* 2006, **281**:38217–38225.
19. Cassago A, Rodrigues EM, Prieto EL, Gaston KW, Alfonso JD, et al: Identification of *Leishmania* selenoproteins and SECIS element. *Mol Biochem Parasitol* 2006, **149**:128–134.
20. Gurvitz A: Identification of the *Leishmania major* proteins LmjF07.0430, LmjF07.0440, and LmjF27.2440 as components of fatty acid synthase I. *J Biomed Biotechnol* 2009, **2009**:950864–950872.
21. Nabholz C, Hauser R, Schneider A: *Leishmania tarentolae* contains distinct cytosolic and mitochondrial glutamyl-tRNA synthetase activities. *Proc Natl Acad Sci USA* 1997, **94**:7903–7908.
22. Rinehart J, Horn EK, Wei D, Soll D, Schneider A: Non-canonical eukaryotic glutamyl- and glutamyl-tRNA synthetases form mitochondrial aminoacyl-tRNA in *Trypanosoma brucei*. *J Biol Chem* 2004, **279**:1161–1166.
23. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, et al: Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. *PLoS Pathog* 2010, **6**:e1001037.
24. Depledge DP, Evans KJ, Ivens AC, Naveed A, Asher M, Kaye PM, et al: Comparative expression profiling of leishmania: modulation in gene expression between species and in different host genetic backgrounds. *PLoS Negl Trop Dis* 2009, **3**:e476.
25. Lee JW, Beebe K, Nangle LA, Jang J, Longo-Guess CM, et al: Editing-defective tRNA synthetase causes protein misfolding and neurodegeneration. *Nature* 2006, **443**:50–55.

26. Sankaranarayanan R, Dock-Bregeon AC, Romby P, Caillet J, Springer M, et al: **The structure of threonyl-tRNA synthetase-tRNA(Thr) complex enlightens its repressor activity and reveals an essential zinc ion in the active site.** *Cell* 1999, **97**:371–381.
27. Wolf YI, Aravind L, Grishin NV, Koonin EV: **Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events.** *Genome Res* 1999, **9**:689–710.
28. Guo M, Chong YE, Beebe K, Shapiro R, Yang XL, Schimmel P: **The C-Ala domain brings together editing and aminoacylation functions on one tRNA.** *Science* 2009, **325**:744–747.
29. Francin M, Kaminska M, Kerjan P, Mirande M: **The N-terminal domain of mammalian Lysyl-tRNA synthetase is a functional tRNA-binding domain.** *J Biol Chem* 2002, **277**:1762–1769.
30. Bilokapic S, Maier T, Ahel D, Gruic-Sovulj I, Söll D, et al: **Structure of the unusual seryl-tRNA synthetase reveals a distinct zinc-dependent mode of substrate recognition.** *EMBO J* 2006, **25**:2498–2509.
31. Liu RJ, Tan M, Du DH, Xu BS, Eriani G, et al: **Peripheral insertion modulates editing activity of the isolated CP1 domain of leucyl-tRNA synthetase.** *Biochem J* 2011, **440**:217–227.
32. Seiradake E, Mao W, Hernandez V, Baker SJ, Plattner JJ, et al: **Crystal structures of the human and fungal cytosolic Leucyl-tRNA synthetase editing domains: a structural basis for the rational design of antifungal benzoxaboroles.** *J Mol Biol* 2009, **390**:196–207.
33. Zhao MW, Zhu B, Hao R, Xu MG, Eriani G, et al: **Leucyl-tRNA synthetase from the ancestral bacterium *Aquifex aeolicus* contains relics of synthetase evolution.** *EMBO J* 2005, **24**:1430–1439.
34. Sokabe M, Okada A, Yao M, Nakashima T, Tanaka I: **Molecular basis of alanine discrimination in editing site.** *Proc Natl Acad Sci USA* 2005, **102**:11669–11674.
35. Fukunaga R, Yokoyama S: **Structure of the AlaX-M trans-editing enzyme from *Pyrococcus horikoshii*.** *Acta Crystallogr D Biol Crystallogr* 2007, **63**:390–400.
36. Ahel I, Korencic D, Ibba M, Söll D: **Trans-editing of mischarged tRNAs.** *Proc Natl Acad Sci USA* 2003, **100**:15422–15427.
37. Ruan B, Söll D: **The bacterial YbaK protein is a Cys-tRNA^{Pro} and Cys-tRNA^{Cys} deacylase.** *J Biol Chem* 2005, **280**:25887–25891.
38. An S, Musier-Forsyth K: **Trans-editing of Cys-tRNA^{Pro} by *Haemophilus influenzae* YbaK protein.** *J Biol Chem* 2004, **279**:42359–42362.
39. Zhang H, Huang K, Li Z, Banerjee L, Fisher KE, et al: **Crystal structure of YbaK protein from *Haemophilus influenzae* (HI1434) at 1.8 Å resolution: functional implications.** *Proteins* 2000, **40**:86–97.
40. Wong FC, Beuning PJ, Silvers C, Musier-Forsyth K: **An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing.** *J Biol Chem* 2003, **278**:52857–52864.
41. Lam H, Oh DC, Cava F, Takacs CN, Clardy J, et al: **D-aminoacids govern stationary phase cell wall remodeling in bacteria.** *Science* 2009, **325**:1552–1555.
42. Miyoshi Y, Hamase K, Tojo Y, Mita M, Konno R, Zaitu K: **Determination of D-serine and D-alanine in the tissues and physiological fluids of mice with various D-amino-acid oxidase activities using two-dimensional high-performance liquid chromatography with fluorescence detection.** *J Chromatogr B Analyt Technol Biomed Life Sci* 2009, **877**:2506–2512.
43. Bhatt TK, Yogavel M, Wydau S, Berwal R, Sharma A: **Ligand-bound structures provide atomic snapshots for the catalytic mechanism of D-amino acid deacylase.** *J Biol Chem* 2010, **285**:5917–5930.
44. Feng L, Tumbula-Hansen D, Toogood H, Söll D: **Expanding tRNA recognition of a tRNA synthetase by a single amino acid change.** *Proc Natl Acad Sci USA* 2003, **100**:5676–5681.
45. Charron C, Roy H, Blaise M, Giegé R, Kern D: **Non-discriminating and discriminating aspartyl-tRNA synthetases differ in the anticodon-binding domain.** *EMBO J* 2003, **22**:1632–1643.
46. Frechin M, Senger B, Brayé M, Kern D, Martin RP, Becker HD: **Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS.** *Genes Dev* 2009, **23**:1119–1130.
47. Liao CC, Lin CH, Chen SJ: **Trans-kingdom rescue of Gln-tRNA^{Gln} synthesis in yeast cytoplasm and mitochondria.** Wang CC: *Nucleic Acids Res*; 2012 [Epub ahead of print].
48. Roy H, Ibba M: **Bridging the gap between ribosomal and nonribosomal protein synthesis.** *Proc Natl Acad Sci USA* 2010, **107**:14517–14518.
49. Sissler M, Delorme C, Bond J, Ehrlich SD, Renault P, et al: **An aminoacyl-tRNA synthetase paralog with a catalytic role in histidine biosynthesis.** *Proc Natl Acad Sci USA* 1999, **96**:8985–8990.
50. Francklyn C: **tRNA synthetase paralogs: evolutionary links in the transition from tRNA-dependent amino acid biosynthesis to de novo biosynthesis.** *Proc Natl Acad Sci USA* 2003, **100**:9650–9652.
51. Roy H, Zou SB, Bullwinkle TJ, Wolfe BS, Gilreath MS, et al: **The tRNA synthetase paralog PoxA modifies elongation factor-P with (R)-β-lysine.** *Nat Chem Biol* 2011, **7**:667–669.
52. Gilreath MS, Roy H, Bullwinkle TJ, Katz A, Navarre WW, et al: **β-Lysine discrimination by lysyl-tRNA synthetase.** *FEBS Lett* 2011, **585**:3284–3288.
53. Ambrogelly A, O'Donoghue P, Söll D, Moses S: **A bacterial ortholog of class II lysyl-tRNA synthetase activates lysine.** *FEBS Lett* 2010, **584**:3055–3060.
54. Yanagisawa T, Sumida T, Ishii R, Takemoto C, Yokoyama S: **A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P.** *Nat Struct Mol Biol* 2010, **17**:1136–1143.
55. Nakatsu T, Kato H, Oda J: **Crystal structure of asparagine synthetase reveals a close evolutionary relationship to class II aminoacyl-tRNA synthetase.** *Nat Struct Biol* 1998, **5**:15–19.
56. Richards NG, Schuster SM: **An alternative mechanism for the nitrogen transfer reaction in asparagine synthetase.** *FEBS Lett* 1992, **313**:98–102.
57. Boehlein SK, Richards NG, Schuster SM: **Glutamine-dependent nitrogen transfer in *Escherichia coli* asparagine synthetase B. Searching for the catalytic triad.** *J Biol Chem* 1994, **269**:7450–7457.
58. Charron C, Roy H, Blaise M, Giegé R, Kern D: **Crystallization and preliminary X-ray diffraction data of an archaeal asparagine synthetase related to asparaginyl-tRNA synthetase.** *Acta Crystallogr D Biol Crystallogr* 2004, **60**:767–769.
59. Blaise M, Fréchin M, Oliéric V, Charron C, Sauter C, Lorber B, Roy H, Kern D: **Crystal structure of the archaeal asparagine synthetase: interrelation with aspartyl-tRNA and asparaginyl-tRNA synthetases.** *J Mol Biol* 2011, **412**:437–452.
60. Nakamura M, Yamada M, Hirota Y, Sugimoto K, Oka A, et al: **Nucleotide sequence of the *asnA* gene coding for asparagine synthetase of *E. coli* K-12.** *Nucleic Acids Res* 1981, **9**:4669–4676.
61. Aurrecochea C, Brestelli J, Brunk BP, Fischer S, Gajria B, et al: **EuPathDB: a portal to eukaryotic pathogen databases.** *Nucleic Acids Res* 2010, **38**:D415–D419.
62. Cusack S, Härtlein M, Leberman R: **Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases.** *Nucleic Acids Res* 1991, **19**:3489–3498.
63. Sauter C, Lorber B, Cavarelli J, Moras D, Giegé R: **The free yeast aspartyl-tRNA synthetase differs from the tRNA(Asp)-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon-binding domain.** *J Mol Biol* 2000, **299**:1313–1324.
64. Laskowski RA: **PDBsum new things.** *Nucleic Acids Res* 2009, **37**:D355–D359.
65. Aslett M, Aurrecochea C, Berriman M, Brestelli J, Brunk BP, et al: **TriTrypDB: a functional genomic resource for the Trypanosomatidae.** *Nucleic Acids Res* 2010, **38**:D457–D462.
66. Krogh A, Brown M, Mian IS, Sjölander K, Haussler D: **Hidden Markov models in computational biology. Applications to protein modeling.** *J Mol Biol* 1994, **235**:1501–1531.
67. Consortium UP: **Ongoing and future developments at the Universal Protein Resource.** *Nucleic Acids Res* 2011, **39**:D214–D219.
68. Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, et al: **Infrastructure for the life sciences: design and implementation of the UniProt website.** *BMC Bioinformatics* 2009, **10**:136–154.
69. Katoh K, Misawa K, Kuma K, Miyata T: **MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform.** *Nucleic Acids Res* 2002, **30**:3059–3066.
70. Altschul SF, Wootton JC, Gertz EM, Agarwala R, Morgulis A, et al: **Protein database searches using compositionally adjusted substitution matrices.** *FEBS J* 2005, **272**:5101–5109.
71. Horton P, Nakai K: **Better prediction of protein cellular localization sites with the k nearest neighbors classifier.** *Proc Int Conf Intell Syst Mol Biol* 1997, **5**:147–152.
72. Buchan DW, Ward SM, Lolley AE, Nugent TC, Bryson K, et al: **Protein annotation and modelling servers at University College London.** *Nucl Acids Res* 2010, **38**:W563–W568.
73. Finn RD, Mistry J, Tate J, Coghill P, Heger A, et al: **The Pfam protein families database.** *Nucleic Acids Res* 2010, **38**:D211–D222.

74. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, et al: **CDD: a Conserved Domain Database for the functional annotation of proteins.** *Nucleic Acids Res* 2011, **39**:D225–D229.
75. Gough J, Chothia C: **SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments.** *Nucleic Acids Res* 2002, **30**:268–272.
76. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673–4680.
77. Jones DT, Taylor WR, Thornton JM: **The rapid generation of mutation data matrices from protein sequences.** *Comput Appl Biosci* 1992, **8**:275–282.
78. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al: **MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods.** *Mol Biol Evol* 2011, **28**:2731–2739.
79. Eswar N, Eramian D, Webb B, Shen MY, Sali A: **Protein structure modeling with MODELLER.** *Methods Mol Biol* 2008, **426**:145–159.
80. DeLano WL: *The PyMOL Molecular Graphics System.* San Carlos, CA, USA: DeLano Scientific LLC.

doi:10.1186/1471-2164-13-621

Cite this article as: Gowri et al.: Unusual domain architecture of aminoacyl tRNA synthetases and their paralogs from *Leishmania major*. *BMC Genomics* 2012 **13**:621.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

