Published in "Journal of Biological Chemistry 281(50): 38217-38225, 2006"
which should be cited to refer to this work.

Trypanosoma seryl-tRNA synthetase is a metazoan-like enzyme with high affinity for $tRNA^{Sec}$

Renaud Geslain[#•], Eric Aeby[§•], Tanit Guitart[#], Thomas E. Jones[#], Manuel Castro de Moura[#], Fabien Charrière[§], André Schneider[§], and Lluís Ribas de Pouplana|[#]*
| Institució Catalana de Recerca i Estudis Avançats (ICREA) and
[#] Barcelona Institute for Biomedical Research, Barcelona Science Park
C/ Samitier 1-5, Barcelona 08015, Catalonia, Spain.
[§] Department of Biology/Cell & Developmental Biology, University of Fribourg,
Chemin du Musée 10, 1700 Fribourg, Switzerland
* corresponding author. Email: lluisribas@pcb.ub.es Tel.(+34)934034868, Fax.(+34)934034870
• These authors contributed equally to this work.

Trypanosomatids are important human pathogens that form a basal branch of eukaryotes. Their evolutionary history is still unclear, as are many aspects of their molecular biology. Here we characterize essential components required for the incorporation of serine and selenocysteine into the proteome of *Trypanosoma*. First, the biological function of a putative *Trypanosoma* seryl-tRNA synthetase is characterized *in vivo*. Secondly, the molecular recognition by *Trypanosoma* seryl-tRNA synthetase of its cognate tRNAs is dissected *in vitro*. The cellular distribution of $tRNA^{Sec}$ is studied, and the catalytic constants of its aminoacylation are determined. These are found to be markedly different from those reported in other organisms, indicating that this reaction is particularly efficient in trypanosomatids. Our functional data are analyzed in the context of a new phylogenetic analysis of eukaryotic seryl-tRNA synthetases that includes *Trypanosoma* and *Leishmania* sequences. Our results show that trypanosomatid seryl-tRNA synthetases are functionally and evolutionarily more closely related to their metazoan homologous enzymes than to other eukaryotic enzymes. This conclusion is supported by sequence synapomorphies that clearly connect metazoan and trypanosomatid seryl-tRNA synthetases.

The formation of aminoacyl-tRNA, catalyzed by aminoacyl-tRNA synthetases (ARS), is a crucial step in maintaining the fidelity of protein biosynthesis. In order to avoid misacylation of non-cognate tRNAs, each synthetase recognizes identity elements idiosyncratic to their cognate substrates. tRNA identity elements can be unique and universally distributed, as in the case of the G3:U70 base pair presented by all $tRNA^{Ala}$ (1,2). On the other hand, the evolutionary constraints imposed by the necessity of translation fidelity are not rigid, and the set of recognition elements for many tRNAs have changed during evolution (1,3). Thus, the comparison between sets of tRNA recognition elements in extant species could be used to estimate their evolutionary or biological relatedness.

Here we study the evolution of the serylation reaction of $tRNA^{Ser}$ and $tRNA^{Sec}$ by characterizing this activity in trypanosomatids, a group of protozoa normally considered to form a basal group within eukaryal (4). Intriguingly, the analysis of the evolutionary history of *Trypanosoma* SerRS, and the determination of the identity elements used by this enzyme to recognize $tRNA^{Ser}$, show that *Trypanosoma* SerRSs are closer relatives of metazoan SerRSs than plant, fungi, or other protozoan enzymes.

Seryl-tRNA synthetases (SerRSs) are class II ARS that contain the characteristic active site domain of this family of enzymes (5,6). In addition to this domain, the structures of most SerRSs include a coiled-coil domain at the amino end of their sequence and, in the case of eukaryotic enzymes, a C-terminal extension that plays a modest role on protein stability and amino acid recognition (7,8). The N-terminal coiled-coil domain is essential for the recognition of the elbow and, especially, the long variable loop of $tRNA^{Ser}$ (9,10).

Studies of the recognition modes between SerRS and $tRNA^{Ser}$ from different kingdoms of life have shown that some, but not all, $tRNA^{Ser}$ identity determinants have been conserved during evolution (for a review see (11)), and that different recognition modes exist. Thus, in bacteria and *Saccharomyces cerevisiae* the specific charging of $tRNA^{Ser}$ depends on the recognition of some base pairs in the acceptor

stem, but not of the discriminator base at position 73 (12-14). In these cases the sequence of the variable loop is also crucial. The crystallographic structures of *T. thermophilus* SerRS complexed with $tRNA^{Ser}$, confirm these observations (15,16).

A second type of recognition mode, seen in archaeal and human SerRS, is characterized by the crucial role of the discriminator base (17-19). In these cases the sequence of the variable loop is not important, but its size and orientation is fundamental for the interaction with the enzyme (20,21). The main common features between the bacterial and eukaryotic recognition strategies are the importance of the variable loop and the dispensable character of the anticodon stem-loop.

Finally, a third type of $tRNA^{Ser}$ recognition is seen in the SerRSs of methanogenic archaea, which use a idiosyncratic N-terminal domain to recognize the acceptor stem and the variable loop of their tRNA substrates (22). These unusual enzymes are also sensitive to the discriminator base position of their cognate tRNAs (23).

SerRS also aminoacylates $tRNA^{Sec}$ with serine, as the first step for the incorporation of selenocysteine into proteins. Components of the selenocysteine insertion machinery have been identified by computational methods in species of the three branches of the tree of life. Recently some of these components have also been identified in trypanosomatids, and the existence of a $tRNA^{Sec}$ in these organisms has been verified (24).

The serylation reaction of $tRNA^{Sec}$ has been characterized in detail in *E. coli* and *Homo sapiens* (25,26), but the system has not been characterized in protozoans or other basal eukaryotes. In *H. sapiens* and *E. coli*, $tRNA^{Sec}$ transcripts have been described as being, respectively, 10 and 100 fold less efficient substrates for SerRS than $tRNA^{Ser}$. It has been proposed that this difference is caused by the unique structure of $tRNA^{Sec}$ (26).

Here we report the first characterization of the serylation of $tRNA^{Sec}$ in trypanosomatids. We confirm the expression of $tRNA^{Sec}$ in *Trypanosoma brucei* and show that, unlike most trypanosomal tRNAs, it is exclusively localized in the cytosol. Furthermore we characterize its aminoacylation with serine by *Trypanosoma cruzi* SerRS. We show that the aminoacylation constants of $tRNA^{Sec}$ in *Trypanosoma* differ substantially from those reported in other organisms, suggesting that kinetic differences in the serylation activity of $tRNA^{Sec}$ may be species-specific, and that regulatory strategies may exist based on the efficiency of serine-$tRNA^{Sec}$ synthesis.

EXPERIMENTAL PROCEDURES

Materials - Oligonucleotides were synthesized by Sigma-Genosys. L-[^{3}H] serine and HisTrap nickel columns were from Amersham Biosciences. Restriction enzymes were from New England Biolabs, Pfu Ultra DNA polymerase was from Stratagene and vector pQE-70 from Quiagen. Novablue cells were from Novagen. *T. cruzi* genomic DNA was a gift from Dr. P. Bonay (Universidad Autónoma de Madrid, Spain).

RNAi-mediated ablation of SerRS - RNAi-mediated ablation of the *T. brucei* SerRS was performed using stem loop constructs containing the puromycine resistance gene as described (27). As an insert we used a 498 bp fragment (nucleotides 1 to 498) of the *T. brucei* SerRS gene. Transfection of *T. brucei* (strain 29-13), selection with antibiotics, cloning and induction with tetracycline were done as described (28). To analyze the in vivo charging levels of the $tRNA^{Ser}$ and the $tRNA^{Sec}$ we isolated total RNA from uninduced and induced cells by using the acid guanidinium isothiocyanate procedure (29). The tRNAs remain aminoacylated during this procedure due to the low pH employed by the method. Subsequently, the RNA samples were analyzed on 50 cm long acid urea polyacrylamide gels as described (30), which can resolve aminoacylated from deacylated tRNAs. The gels were analyzed by Northern hybridization (31). The following [^{32}P] 5'-end labeled oligonucleotides were used as probes: 5'TGGCGTCACCAGCAGGATTC3' (for the $tRNA^{Ser}_{CGA}$) and 5'ACCAGCTGAGCTCATCGTGGC3' (for $tRNA^{Sec}$).

tRNA localization studies - To determine the intracellular localization of $tRNA^{Sec}$ we prepared mitochondria free of cytosolic RNAs by digitonin extraction and subsequent RNase A digestion from procyclic *T. brucei* as described (31). RNAs from total cells or isolated mitochondria were purified as indicated above (29). To test whether $tRNA^{Sec}$ concentrations were different in different life stages of *T. brucei*, total RNA of procyclic (strain 427) and bloodstream (strain AnTat1.1, a gift from Dr.

Seebeck, University of Bern, Switzerland) stages of *T. brucei* was purified. To test whether expression of $tRNA^{Sec}$ is induced in medium containing selenium, or under oxidative stress, total RNA was isolated from procyclic *T. brucei* grown in the absence (0) or the presence of 0.005 and 0.5 µg/ml of added Na_2SeO_3, or in the absence or presence of 10 and 25 µMol of added H_2O_2 (data not shown). For all RNA fractions the relative expression level of the $tRNA^{Sec}$ was analyzed by Northern hybridization using the $tRNA^{Sec}$-specific oligonucleotide indicated above. The $tRNA^{Ile}$ was detected using the oligonucleotide 5'TGCTCCCGGCGGGTTCGAA3'.

tRNA preparation - Constructions containing a T7 promoter followed by the gene encoding wild type or mutated tRNAs were assembled using six DNA oligonucleotides that were first annealed, and then ligated, between HindIII and BamHI restriction sites of plasmid pUC19. *In vitro* transcription using T7 RNA polymerase was performed according to standard protocols (32). Transcripts were separated on denaturing PAGE, full-length tRNAs were eluted from gel using an electroelution apparatus (Schleicher & Schüll) and refolded (2 min at 90 ºC followed by a gradual reduction of temperature in presence of 2.5 mM $MgCl_2$).

Enzyme cloning and purification - The 1.4-kbp intron-less gene coding for *T. cruzi* SerRS (Tc00.1047053511163.10) was amplified by PCR from genomic DNA (strain MC) using Pfu Ultra DNA polymerase and cloned in the vector pQE-70 for bacterial expression of a C-terminal His_6-tagged protein. The correct sequence of the gene was checked by sequencing it entirely. Novablue *E. coli* cells (Novagen) transformed with this construction were grown at 21 ºC up to A_{700nm}=0.3. Protein expression was then induced with 0.1 mM IPTG during 12 hours. Purification on nickel affinity columns was performed using standard procedures.

Aminoacylation assays - Aminoacylation was performed at 37 ºC in 50 mM Tris-HCl, pH 7.6, 15 mM $MgCl_2$, 4 mM DTT, 5 mM ATP, 10 mM NaCl, 100 µM L-[^{3}H] serine (300 Ci/mol) and varying concentrations of tRNA transcripts (1-80 µM). Reaction was initiated by addition of pure enzyme and samples of 20 µl were spotted onto Whatman 3MM discs at varying time intervals (usually 2 min). Radioactivity was measured by liquid scintillation. Enzyme concentration was experimentally determined for each tRNA in order to obtain linear velocities. Kinetic constants were obtained from Lineweaver-Burk plots using a minimum of two independent measurements and five tRNA concentrations.

Phylogenetic and sequence analyses - The sequences of seryl-tRNA synthetases used were pulled from GENBANK, either directly or by BLAST searches of available genomic sequences (33). The total set of sequences was aligned with CLUSTALX (34). All sequences were initially included in our analysis, but were later culled to focus our analysis on the eukaryotic clade that contains the *Trypanosoma* sequences. Coiled-coil predictions were done using the programs PAIRCOIL (35), COILS (36) and MULTICOIL (37). Phylogenetic distributions were calculated by parsimony, distance and maximum likelihood methods using PHYLIP 3.63 package (38). Maximum parsimony analysis (MP) was done using PROTPARS. The neighbour-joining method was applied using the programs NEIGHBOR and PROTDIST, with the Dayhoff 120 substitution matrix. Maximum likelihood (ML) phylogenies were calculated with the program PROML (38) using the JTT substitution matrix. The programs SEQBOOT and CONSENSE (38) were used to estimate the confidence limits of branching points from 1000 bootstrap replicates in the parsimony and distance calculations, while 100 bootstrap replicates were used for the maximum likelihood trees. Several combinations of sequences were used in our analyses to test the robustness of the trees obtained. All the groups of sequences used produced exactly the same tree topology reported here.

RESULTS

RNAi-mediated ablation of the trypanosomal SerRS.

To confirm that the putative *T. brucei* SerRS gene encodes the functional enzyme in charge of serylation of $tRNA^{Ser}$ and $tRNA^{Sec}$ we established a stable transgenic cell line, that allows tetracycline-inducible RNAi-mediated ablation of the protein. This cell line stops growing after induction of RNAi, confirming that the putative SerRS is essential for normal growth of the parasite (Fig. 1A). To determine the biochemical phenotype of the RNAi cell line we isolated total RNA from untreated cells and from cells grown in the presence of tetracycline. Subsequently, the RNAs were resolved on long acid urea polyacrylamide gels (30) which in

combination with Northern analysis allow to determine the ratio of charged to uncharged $tRNA^{Ser}$ and $tRNA^{Sec}$, respectively. The results in Fig. 1B and 1C show that ablation of the putative SerRS results in selective accumulation of both uncharged $tRNA^{Ser}$ and uncharged $tRNA^{Sec}$, indicating that the SerRS identified in our study is the enzyme responsible for *in vivo* serylation of both of these tRNAs.

Localization of the $tRNA^{Sec}$.

Unlike in other eukaryotes most tRNAs in trypanosomatids have a dual localization: Approx. 95% are found in the cytosol and function in cytosolic translation, however a small fraction (approx. 5%) are imported into the mitochondrion and function in organelar protein synthesis (31). Consequently, trypanosomal tRNAs are always encoded in nucleus and never on the mitochondrial DNA. *T. brucei* encodes a single *selC* gene, coding for the $tRNA^{Sec}$ (24). In order to confirm the expression of this tRNA and to determine its intracellular localization we carried out Northern blot analyses using total and mitochondrial RNA fractions from procyclic *T. brucei*. The result in Figure 2A shows that the $tRNA^{Sec}$ is only detected in the cytosol but not in the mitochondrion. Exclusive cytosolic localization is exceptional in trypanosomatids, the only cytosol-specific tRNA known to date being the initiator $tRNA^{Met}$ (31).

$tRNA^{Sec}$ expression is not influenced by the life cycle stage, selenium or H_2O_2.

In a next series of Northern blots, we tested whether the relative amount of $tRNA^{Sec}$ when compared to other tRNAs was dependent on the life cycle stage, or sensitive to the concentration of selenium or H_2O_2 in the culture media. Figure 2B shows that trypanosomal $tRNA^{Sec}$ is expressed to very similar levels in both the procyclic and bloodstream forms of *T. brucei*. This suggests that selenocysteine-containing proteins play a role throughout the life cycle of *T .brucei*. In contrast to reports in vertebrates (39), the expression of the $tRNA^{Sec}$ in procyclic *T. brucei* is not affected by the addition of selenium to the growth medium (Fig. 2C). A similar result was obtained when the parasites were grown in the presence of H_2O_2 (data not shown).

Characterization of the serylation reaction of $tRNA^{Ser}$ and $tRNA^{Sec}$ by T. cruzi SerRS .

The cloning of SerRS gene was performed directly from genomic DNA. The 54 kDa protein was expressed in *E. coli*, and purified to homogeneity by affinity chromatography. Gel filtration experiments with the purified enzyme confirmed its ability to form dimers (data not shown), suggesting that *T. cruzi* SerRS is a classical class II aminoacyl-tRNA synthetase (5).

Transcripts of *Trypanosoma* $tRNA^{Ser}$ and $tRNA^{Sec}$ were used in aminoacylation assays and both were shown to be efficiently aminoacylated by SerRS (Fig. 3A). The kinetic constants for the serylation of both tRNAs by SerRS were then determined (Fig. 3B). The K_m for $tRNA^{Ser}$ of *T. cruzi* SerRS was calculated at 3,2 μM, a value comparable to those found in other systems but eight fold higher than that calculated for $tRNA^{Sec}$ with the same enzyme. On the other hand, the K_{cat} was essentially identical to that calculated for $tRNA^{Sec}$. As a result, the K_{cat}/K_m value for *Trypanosoma* $tRNA^{Ser}$ is seven fold lower than that of *T. cruzi* $tRNA^{Sec}$. Interestingly, these values are markedly different from those reported in other species. Thus, in the human system (25), the serylation of $tRNA^{Sec}$ is 10 fold less efficient than that for $tRNA^{Ser}$, and this difference grows to a 100 fold in favour of $tRNA^{Ser}$ in *E. coli* (26) (Fig. 3C).

Identity determinants in the acceptor stem of $tRNA^{Ser}$.

To understand how *T. cruzi* SerRS recognizes its cognate tRNAs, mutants of $tRNA^{Ser}_{CGA}$ were produced and their ability to be charged with serine was measured *in vitro* in presence of purified SerRS. Fourteen tRNA variants were generated, covering nearly 80 % of the primary structure. The design of the mutations was dictated both by the sequence conservation among *T. cruzi* $tRNA^{Ser}$ and $tRNA^{Sec}$ sequences (Fig. 4A), and by the distribution of identity elements in homologous $tRNA^{Ser}$ sets. A set of modifications targeted individual nucleotides or base pairs in the acceptor stem (variants 1 to 9) (Fig. 4B), whereas a second set of changes introduced large deletions or sequence swaps of discrete tRNA domains (variants 10 to 14) (Fig. 4C).

The change of the discriminator base G73 for pyrimidines had a dramatic effect on the charging of $tRNA^{Ser}$ (mutants 1 and 2) (Table 1), indicating that the enzyme strongly recognizes this position of the acceptor stem. This sensitivity to the discriminator base sequence is reminiscent of the recognition mechanisms

described for *H. sapiens* and archaeal SerRSs (17-19).

The first base pair (1:72) was found to be the most sensitive position in the acceptor stem of $tRNA^{Ser}$. Indeed, the simple inversion of the G1:C72 base pair (mutant 4) completely abolished the aminoacylation of this substrate by the enzyme. The introduction of a G1:U72 base pair in the $tRNA^{Ser}$ scaffold had no effect on the ability of this tRNA to be charged by SerRS (mutant 5), indicating that G1 may be contributing important contacts to the recognition by SerRS (Table 1). It should be noted that both the discriminator base G73 and the G1 base found to be essential for recognition of $tRNA^{Ser}$ by *Trypanosoma* SerRS are conserved in the sequences of $tRNA^{Sec}$ of the same species.

Additional mutations in other regions of the acceptor stem had no significant effect on the overall recognition of $tRNA^{Ser}$ by *T. cruzi* SerRS (mutant 6 to 9). Overall, our acceptor stem mutations affect mostly K_{cat} values, suggesting a defect in the reactive positioning of the acceptor extremity rather than a loss in binding energy.

Identity determinants beyond the acceptor stem of $tRNA^{Ser}$.

Next, we wanted to investigate the importance of the pseudoknot (D arm, T arm and variable loop) and anticodon domains of *T. cruzi* $tRNA^{Ser}$ for its recognition by SerRS. The three domains of the tRNA that form the pseudoknot were individually flipped (mutants 11 to 13) in order to vary their sequences without modifying the local secondary structure or the strength of their internal base pairs (Fig. 4C). None of the three corresponding mutants showed any catalytic defect, suggesting that the recognition mechanism for the elbow of the tRNA is based exclusively on interactions with the sugar-phosphate backbone (Table 1).

To further test this hypothesis the entire variable loop was excised, and replaced by the most commonly found class I variable loop (mutant 14) (Fig. 4C). This replacement had a dramatic effect on aminoacylation, causing a >8000 fold drop in K_{cat}/K_m. This mutation affected both catalytic and affinity constants. Our combined data indicates that this domain is recognized in a sequence-unspecific manner, provides binding energy, and acts as a guide for the reactive positioning of the tRNA (Table 1).

Finally, the entire anticodon loop was deleted and replaced by a stretch of four uridines (mutant 10) (Fig. 4C). This linker was used because it is the least likely to stabilize unanticipated structures in this region of the tRNA (13). This deletion mutant had only a modest effect on the velocity of the aminoacylation reaction. This is in agreement with other studies that have shown that the anticodon domain of eukaryotic $tRNA^{Ser}$ contributes poorly to the serine identity (for a review see (11)) (Table 1).

Phylogeny of Trypanosoma SerRS.

The known genomes of *Trypanosoma* code for only one SerRS gene. The corresponding protein is likely used both in the cytoplasm and the mitochondria, but no conventional targeting sequence can be identified in their sequences. The protein is 477 amino acids long, and it contains an N-terminal domain predicted to form a coiled coil domain (36), as seen in the available structural data for homologous enzymes (16, 40).

Multiple sequence alignments readily show the presence of a large synapomorphy that clusters trypanosomatid SerRSs with the rest of metazoan sequences (Fig. 5A). This synapomorphy is an insertion of twenty amino acids located at the center of the N-terminal coiled-coil motif of this group of SerRSs (Fig. 5B). Coiled-coil prediction of this group of sequences suggest that in metazoans and trypanosomatids this region may extend beyond the length seen in the structures solved so far (data not shown).

Our phylogenetic analysis of SerRS sequences is the first one of this enzyme to include kinetoplastid sequences. When sequences from all life domains were used, our results consistently agreed with previous analyses (41-44), supporting the conclusion that the overall evolution of this enzyme conforms to the canonical phylogenetic tree derived from ribosomal RNA sequences (Fig. 6A) (41,43). Although we were unable to obtain significant bootstrap support for the central branching points of the general tree, the trypanosomatid sequences strongly associated with metazoan sequences with strong bootstrap support, both in general trees and those limited to the major eukaryotic lineages (Fig. 6A and 6B). It should be stressed here that the region corresponding to the synapomorphy that also links *Trypanosoma* sequences with metazoan ones was not used in our phylogenetic analyses.

DISCUSSION

The constraints acting over tRNA recognition are strictly intra-specific and, for each species constitute a complex set of recognition and rejection elements that ensures faithful translation (45). Possibly this complex set of positive and negative identity elements contributes to the stability of the genetic code and limits its size, because the incorporation of new amino acids would require an increase in complexity of the recognition problem which may be impossible to assume without increasing the rate of aminoacylation errors (3).

In order to study the evolution of these sets of recognition elements we decided to characterize the recognition of $tRNA^{Ser}$ and $tRNA^{Sec}$ by *Trypanosoma* SerRS. In doing so we were expecting to extract information about the evolution of this recognition mechanism in the basal part of the eukaryotic phylogenetic tree. For convenience we chose *Trypanosoma brucei* as a model for the *in vivo* studies, while the *in vitro* studies were performed with *Trypanosoma cruzi* proteins and tRNAs. The overall identity between *T. cruzi* and *T. brucei* SerRSs is 80 %, and the sequences of their $tRNA^{Ser}$ are essentially identical.

We have shown that *Trypanosoma* SerRS recognizes its cognate tRNAs using a combination of structural signals in the variable loop and the sequence information of the discriminator base and the first base pair of the acceptor stem. The discriminator base G73 and the G1 base are essential for recognition by *Trypanosoma* SerRS. These bases are conserved among all serine tRNA isoacceptors in *Trypanosoma*, and strongly influence the velocity of the serylation reaction.

Previous studies on $tRNA^{Ser}$ recognition in the three kingdoms of life show that acceptor stem recognition by SerRS fluctuates between the first four base pairs and the discriminator base of the acceptor stem (17, 46, 47). In the case of *Trypanosoma*, this recognition has shifted towards the CCA end of the molecule, and the discriminator base and the first base pair constitute the region recognized by SerRS.

We have established that the presence of a long variable loop is a pre-requisite to $tRNA^{Ser}$ recognition, but that this mechanism is not sequence specific. Moreover, the complete deletion of $tRNA^{Ser}$ anticodon domain does not affect its serylation by SerRS, confirming that the entire stem and anticodon loop is ignored by the enzyme during the recognition process. This feature is common to all serine systems studied so far with the only exception of *M. barkeri* SerRS, which displays a strong interaction with the G30:C40 base pair in the anticodon stem (46).

In summary, our data suggest that the tRNA specificity of *Trypanosoma* SerRS relies on a two-pronged mechanism based on interactions with the G73 discriminator base and the G1 base on one side, and the sequence-independent recognition of the variable loop on the other. Interestingly, this recognition mechanism functionally associates trypanosomatid SerRSs with their metazoan homologs, and separates them from other eukaryotic enzymes such as *S. cerevisiae* SerRS. This functional relationships are in agreement with the phylogenetic connections that we report here, and the presence of clear sequence synapomorphies that link these enzymes evolutionarily.

A possible explanation to this functional convergence would be a lateral gene transfer event between trypanosomatids and metazoans that could be favored by the parasitic nature of the former. A second possibility, more in agreement with polyphiletic views of eukaryote evolution (4), would be that the recognition solution displayed by metazoans and trypanosomatids is ancient, having persisted in these two groups of organisms since their radiation at the base of the eukaryotic evolutionary tree. Our phylogenetic studies do not allow us to discard either possibility.

Another interesting feature of trypanosomatid SerRSs is their apparently high affinity for $tRNA^{Sec}$. Indeed, the K_{cat}/K_m value for the aminoacylation of *Trypanosoma* $tRNA^{Ser}$ (which are comparable to values established for other species) is seven fold lower than that found for *T. cruzi* $tRNA^{Sec}$, a result that is in direct contrast with the values reported for the human system (25) (where the $tRNA^{Ser}$ is 10 fold more efficiently charged than $tRNA^{Sec}$) or in *Escherichia coli* (26) (100 fold difference in favour of $tRNA^{Ser}$) (Fig. 3B).

It should be noted here that all the studies where these values are reported (including this one) were performed with transcript tRNAs, thus excluding the potential effect of base modifications in either substrate. However, there seems to be a clear difference between the relative aminoacylation of $tRNA^{Sec}$ and $tRNA^{Ser}$ between trypanosomatids and other species. This may be explained by a high requirement for

selenoprotein synthesis in these species. Our attempts to increase the levels of $tRNA^{Sec}$ in *T. brucei* by the presence of oxidative stress or different growth conditions were unsuccessful, but these results are not necessarily contradictory with a high level of selenocysteine use by these organisms.

From our studies it is clear that $tRNA^{Sec}$ is not imported into the mitochondria of *T. brucei*. Some of us have recently shown that tRNA localization signals in this species are confined to two nucleotide pairs in the T-stem (48). In the cytosol-specific initiator $tRNA^{Met}$ these positions include the anti-determinants that, based on experiments with vertebrate initiator $tRNA^{Met}$ (49), are predicted to prevent interaction with translation elongation factor 1a (eEF-1a). The T-stem sequence of the cytosolic $tRNA^{Sec}$ is different to all other *T. brucei* tRNAs including the cytosolic initiator $tRNA^{Met}$. However, a feature that is shared between the two cytosol-specific tRNAs in *T. brucei* is that neither interacts with eEF-1a. Instead, the initiator $tRNA^{Met}$ interacts with initiation factor 2, and the $tRNA^{Sec}$ with the specialized elongation factor SelB/EFsec.

Thus, while the structural determinants that prevent interaction with EF-1a are different for both tRNAs, the mechanism for cytosolic localization might in both cases be achieved by exclusion of EF-1a binding. In other words, all tRNAs that interact with eEF-1a might be imported into mitochondria, and the ones which do not remain in the cytosol.

REFERENCES

1. Hou, Y.M. and P. Schimmel. (1989). *Biochemistry* 28, 6800-6804.
2. Gabriel, K., J. Schneider, and W.H. McClain. (1996). *Science* 271, 195-197.
3. Ribas de Pouplana, L. (2005). *Life* 57, 523-524.
4. Baldauf, S.L., A.J. Roger, I. Wenk-Siefert, and W.F. Doolittle. (2000). *Science.* 290, 972-977.
5. Eriani, G., M. Delarue, O. Poch, J. Gangloff, and D. Moras. (1990). *Nature* 347, 203-206.
6. Ribas de Pouplana, L. and P. Schimmel. (2001). *Cell* 104, 191-193.
7. Weygand-Durasevic, I., B. Lenhard, S. Filipic, and D. Söll. (1996). *J. Biol. Chem.* 271, 2455-2461.
8. Lenhard, B., S. Filipic, I. Landeka, I. Skrtic, D. Söll, and I. Weygand-Durasevic. (1997). *J. Biol. Chem.* 272, 1136-1141.
9. Borel, F., C. Vincent, R. Leberman, and M. Härtlein. (1994). *Nucleic Acids Res.* 22, 2963-2969.
10. Lenhard, B., M. Praetorius-Ibba, S. Filipic, D. Soll, and I. Weygand-Durasevic. (1998). *FEBS Lett.* 439, 235-240.
11. Lenhard, B., O. Orellana, M. Ibba, and I. Weygand-Durasevic. (1999). *Nucleic Acids Res.* 27, 721-729.
12. Himeno, H., S. Yoshida, A. Soma, and K. Nishikawa. (1997). *J. Mol. Biol* .268, 704-711.
13. Sampson, J.R. and M.E. Saks. (1993). *Nucleic Acids Research* 21, 4467-4475.
14. Normanly, J., T. Ollick, and J. Abelson. (1992). *Proc. Natl. Acad. Sci. USA* 89, 5680-5684.
15. Cusack, S., A. Yaremchuk, and M. Tukalo. (1996). *EMBO J.* 15, 2834-2842.
16. Biou, V., M. Yoremchuk, M. Tukalo, and S. Cusack. (1994). *Science* 263, 1404-1410.
17. Metzger, A.U., M. Heckl, D. Willbold, K. Breitschopf, U.L. RajBhandary, P. Rosch, and H.J. Gross. (1997). *Nucleic. Acids. Res.* 25, 4551-4556.
18. Breitschopf, K. and H.J. Gross. (1996). *Nucleic Acids Res* 24, 405-410.
19. Achsel, T. and H.J. Gross. (1993). *EMBO J.* 12, 3333-3338.
20. Wu, X.Q. and H.J. Gross. (1993). *Nucleic Acids Research* 21, 5589-5594.
21. Heckl, M., K. Busch, and H.J. Gross. (1998). *FEBS Lett.* 427, 315-319.
22. Bilokapic, S., T. Maier, D. Ahel, I. Gruic-Sovulj, D. Soll, I. Weygand-Durasevic, and N. Ban. (2006). *Embo J* 4, 4.
23. Bilokapic, S., D. Korencic, D. Soll, and I. Weygand-Durasevic. (2004). *Eur. J. Biochem.* 271, 694-702.
24. Cassago, A., E.M. Rodrigues, E.L. Prieto, K.W. Gaston, J.D. Alfonzo, M.P. Iribar, M.J. Berry, A.K. Cruz, and O.H. Thiemann. (2006). *Mol. Biochem. Parasitol.* 23, 23.
25. Amberg, R., T. Mizutani, X.Q. Wu, and H.J. Gross. (1996). *J. Mol. Biol.* 263, 8-19.
26. Baron, C. and A. Böck. (1991). *Journal of Biological Chemistry* 266, 20375-20379.

27. Bochud-Allemann, N. and A. Schneider. (2002). *J. Biol. Chem.* 277, 32849-32854. Epub 32002 Jul 32842.
28. McCulloch, R., E. Vassella, P. Burton, M. Boshart, and J.D. Barry. (2004). *Methods Mol. Biol.* 262, 53-86.
29. Chomczynski, P. and N. Sacchi. (1987). *Anal Biochem.* 162, 156-159.
30. Varshney, U., C.P. Lee, and U.L. RajBhandary. (1991). *J. Biol. Chem.* 266, 24712-24718.
31. Tan, T.H., R. Pach, A. Crausaz, A. Ivens, and A. Schneider. (2002). *Mol. Cell. Biol.* 22, 3707-3717.
32. Sampson, J.R. and O.C. Uhlenbeck. (1988). *Proc. Natl. Acad. Sci. USA* 85, 1033-1037.
33. Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, and D.L. Wheeler. (2005). *Nucleic Acids Res.* 33, D34-38.
34. Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. (1997). *Nucleic Acids Res.* 25, 4876-4882.
35. Berger, B., D.B. Wilson, E. Wolf, T. Tonchev, M. Milla, and P.S. Kim. (1995). *Proc Natl Acad Sci U S A.* 92, 8259-8263.
36. Lupas, A., M. Van Dyke, and J. Stock. (1991). *Science.* 252, 1162-1164.
37. Wolf, E., P.S. Kim, and B. Berger. (1997). *Protein Sci.* 6, 1179-1189.
38. Felsenstein, J. (1988). *Annu Rev Genet.* 22, 521-565.
39. Choi, I.S., A.M. Diamond, P.F. Crain, J.D. Kolker, J.A. McCloskey, and D.L. Hatfield. (1994). *Biochemistry.* 33, 601-605.
40. Chimnaronk, S., M. Gravers Jeppesen, T. Suzuki, J. Nyborg, and K. Watanabe. (2005). *Embo J* 15, 15.
41. Taupin, C.M. and R. Leberman. (1999). *J Mol Evol.* 48, 408-420.
42. Kim, H.S., U.C. Vothknecht, R. Hedderich, I. Celic, and D. Soll. (1998). *J Bacteriol* 180, 6446-6449.
43. Wolf, Y.I., L. Aravind, N.V. Grishin, and E.V. Koonin. (1999). *Genome Res.* 9, 689-710.
44. Woese, C.R., G.J. Olsen, M. Ibba, and D. Söll. (2000). *Microbiol Mol Biol Rev* 64, 202-236.
45. Giegé, R., M. Sissler, and C. Florentz. (1998). *Nucleic Acids Res.* 26, 5017-5035.
46. Korencic, D., C. Polycarpo, I. Weygand-Durasevic, and D. Soll. (2004). *J. Biol. Chem.* 279, 48780-48786.
47. Saks, M.E. and J.R. Sampson. (1996). *EMBO J.* 15, 2843-2849.
48. Crausaz Esseiva, A., L. Marechal-Drouard, A. Cosset, and A. Schneider. (2004). *Mol Biol Cell.* 15, 2750-2757. Epub 2004 Apr 2752.
49. Drabkin, H.J., M. Estrella, and U.L. Rajbhandary. (1998). *Mol Cell Biol.* 18, 1459-1466.

FOOTNOTES

This work was supported by grants BIO2003-02611 from the Spanish Ministry of Science and Education, and 3100-067906 of the Swiss National Foundation (A. S.), a Marie Curie International Reintegration Fellowship (L.R.P.), and a Fellowship of the Novartis Foundation (F. C.). We are grateful to Dr. Bonay (Hospital de la Princesa) for the gift of *T. cruzi* genomic DNA and to Dr. Seebeck (University of Bern) for bloodstream stage *T. brucei* cells. We thank Drs. Guigó and Chappel (Pompeu Fabra University), Dr. Gladyshev (University of Nebraska), Dr. Eriani (CNRS, Strasbourg), and Dr. Roy (Ohio State University) for useful discussions.

FIGURE LEGENDS

Fig. 1. SerRS is essential for growth of procyclic *T. brucei* and is responsible for the serylation of both $tRNA^{Ser}$ and $tRNA^{Sec}$. (A) Growth curve in the presence and absence of tetracycline (+, -Tet) of a representative clonal *T. brucei* RNAi cell line ablated for the trypanosomal SerRS homologue. (B) Northern blot analysis of total RNA isolated under acidic conditions from the SerRS RNAi cell line. Hours of induction by tetracycline are indicated at the top. The blots were probed for the *T. brucei* $tRNA^{Ser}$ and $tRNA^{Ile}$. The latter serves as controls that is not affected by the RNAi. The RNA fractions were resolved on long acid urea gels that allow to separate aminoacylated (aa) from deacylated (dea) tRNAs. The relative amounts of deacylated $tRNA^{Ser}$ and $tRNA^{Ile}$ are indicated at the bottom. For each lane the sum of aminoacylated and deacylated tRNA was set to 100% percent. (C)

Same as (B) but analysis was done for the $tRNA^{Sec}$. The x refers to an unidentified band probably corresponding to phosphoserine.

Fig. 2. Expression of $tRNA^{Sec}$ in *T. brucei*. (A) $tRNA^{Sec}$ is not imported into the mitochondrion. Left panel: 0.1 x 10^8 cell equivalents of total (Tot.) and 2 x 10^8 cell equivalents of mitochondrial (Mito.) RNA, isolated by digitonin extraction from procyclic *T. brucei*, were separated on a 10% polyacrylamide/8 M urea gel and stained with ethidium bromide. Right panel: the gel was processed for Northern hybridization and probed for $tRNA^{Sec}$ and imported $tRNA^{Ile}$. (B) $tRNA^{Sec}$ is expressed in both life cycle stages. Total RNA, 5 and 7 µg each, of procyclic (Proc.) and bloodstream stage *T. brucei* was analyzed by Northern hybridization as in (A). The relative expression levels (set to be 1 for procyclic RNA) of the $tRNA^{Sec}$ when compared to the expression of tRNAIle are indicated at the bottom. (C) Expression of $tRNA^{Sec}$ is not induced in medium containing selenium. Five µg of total RNA isolated from procyclic *T. brucei* grown in the absence (0) or the presence of 0.005 and 0.5 µg/ml of added Na_2SeO_3 were analyzed by Northern hybridization as in (A) and (B). The relative expression levels are indicated.

Fig. 3. Comparison of charging efficiency of $tRNA^{Sec}$ and $tRNA^{Ser}$ from various origins. (A) Aminoacylation plot of trypanosomal $tRNA^{Ser}$ (■) and $tRNA^{Sec}$ (♦) in the presence of purified seryl-tRNA synthetase and [^{3}H] serine. Three independent measures have shown only minimal variations of the charging activity, ranging from 1 to 15% (data not shown on the graph for visual clarity). (B) Comparison of the kinetic values for the *in vitro* serylation of trypanosomal, human and *E. coli* transcribed $tRNA^{Ser}$ and $tRNA^{Sec}$ (25, 26, this work). N.D., not determined (C) Histogram showing the efficiency of aminoacylation of $tRNA^{Sec}$ with respect to $tRNA^{Ser}$ for *T. cruz*i, *H. sapiens*, and *E. coli*. The level of aminoacylation efficiency of $tRNA^{Ser}$ is indicated by a dashed line. The gray columns represents the level of aminoacylation efficiency of $tRNA^{Sec}$. *T. cruzi* $tRNA^{Sec}$ is 7 fold more efficiently aminoacylated than *T. cruzi* $tRNA^{Ser}$. * The data concerning *E. coli* $tRNA^{Ser}$ and $tRNA^{Sec}$ were reported from experiments using total protein extract as the source of SerRS (26).

Fig. 4. Secondary structure of $tRNA^{Ser}_{CGA}$ and its mutants. (A) The cloverleaf structure of $tRNA^{Ser}_{CGA}$ is presented. The circles denote nucleotides conserved among the four *T. cruzi* $tRNA^{Ser}$ isoacceptors. The squares denote nucleotides conserved both in $tRNA^{Ser}$ and $tRNA^{Sec}$. (B) Detail of the mutations performed on the acceptor stem of $tRNA^{Ser}_{CGA}$. Each mutant is labeled in parenthesis. (C) Detail of the mutation performed beyond the acceptor stem. Curved arrows and dashed lines are used to denote the flipping, and the deletion of sequences, respectively.

Fig. 5. Sequence alignment of the region of SerRS sequences that contains the synapomorphy that links metazoan and trypanosomatid enzymes. (A) The alignment was performed using CLUSTAL X (34). *T. cruzi* and *T. brucei* SerRS sequences are in bold. Helix 1 and helix 2 indicate the long helices that form the coiled-coil arm found in *T. thermophilus* SerRS. The insertion that is present in metazoan and trypanosomatid sequences is boxed and labeled accordingly. (B) Crystallographic structure of one monomer of *T. thermophilus* SerRS (16). The coiled-coil arm important for the stabilization of the $tRNA^{Ser}$ pseudoknot is highlighted in green. The apparent insertion point of the metazoan and trypanosomatid synapomorphy is marked by an arrow and labeled accordingly.

Fig. 6. Phylogenetic analysis of trypanosomal SerRS.
(A) Maximum likelihood tree of archaeal, bacterial, and eukaryotic seryl-tRNA synthetase sequences. The overall architecture of this tree was obtained with all methods used, but bootstrap support was weak (not shown). (B) Consensus tree of eukaryotic seryl-tRNA synthetase sequences. The association between trypanosomatid and metazoan SerRSs is highlighted. Numbers correspond to percentages of bootstrap support for each node from the parsimony, distance, and ML analyses.

Table 1. Serylation parameters of tRNASec, $tRNA^{Ser}$ and its variants in presence of SerRS. The rate of aminoacylation of the wild type $tRNA^{Ser}$ is 9.56 s^{-1}. For better readability the *kcat* of each mutant is expressed as a percentage of the *kcat* value of the wild type $tRNA^{Ser}$ (i.e. a rate of 9.56 s^{-1} corresponds to a *kcat* of 100%). The overall efficiency of serylation has been determined as the ratio between the *Km* and the *kcat*. To enhance the readability the efficiency of the wild type $tRNA^{Ser}$ have been set at 1.

Values higher than 1 denote tRNA variants that are less efficiently aminoacylated and *vice versa*. Numbers in parentheses refers to the mutations detailed on figure 4B and 4C.
[a]ND, not determined

Figure 1

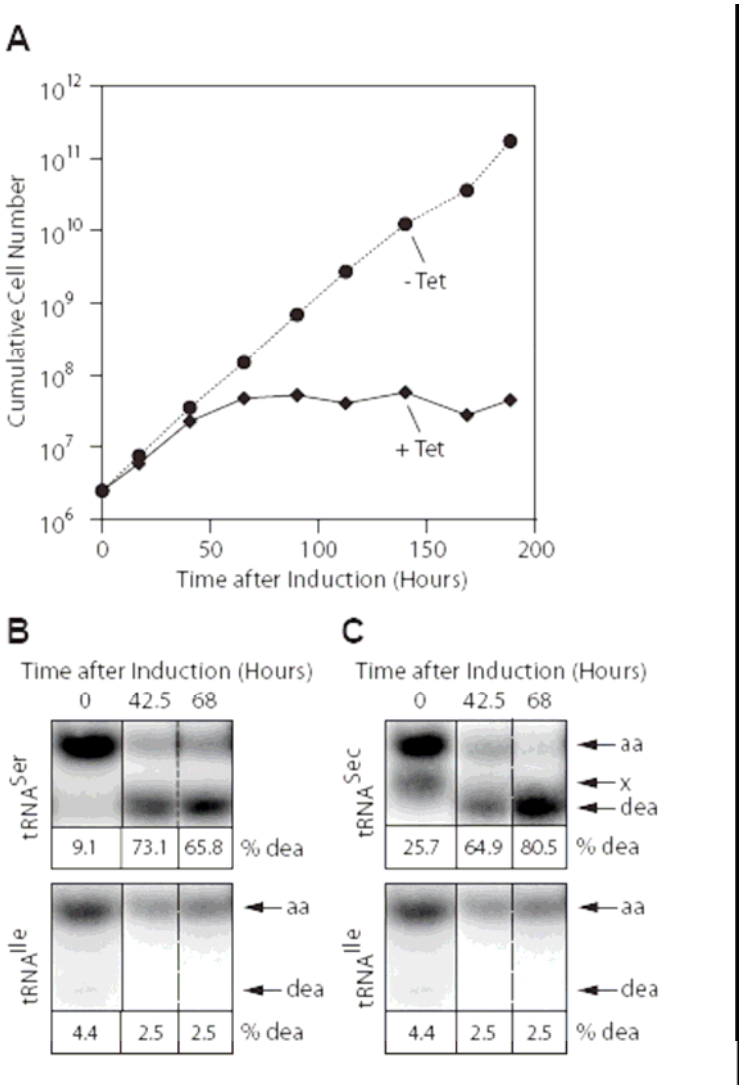


Figure 2

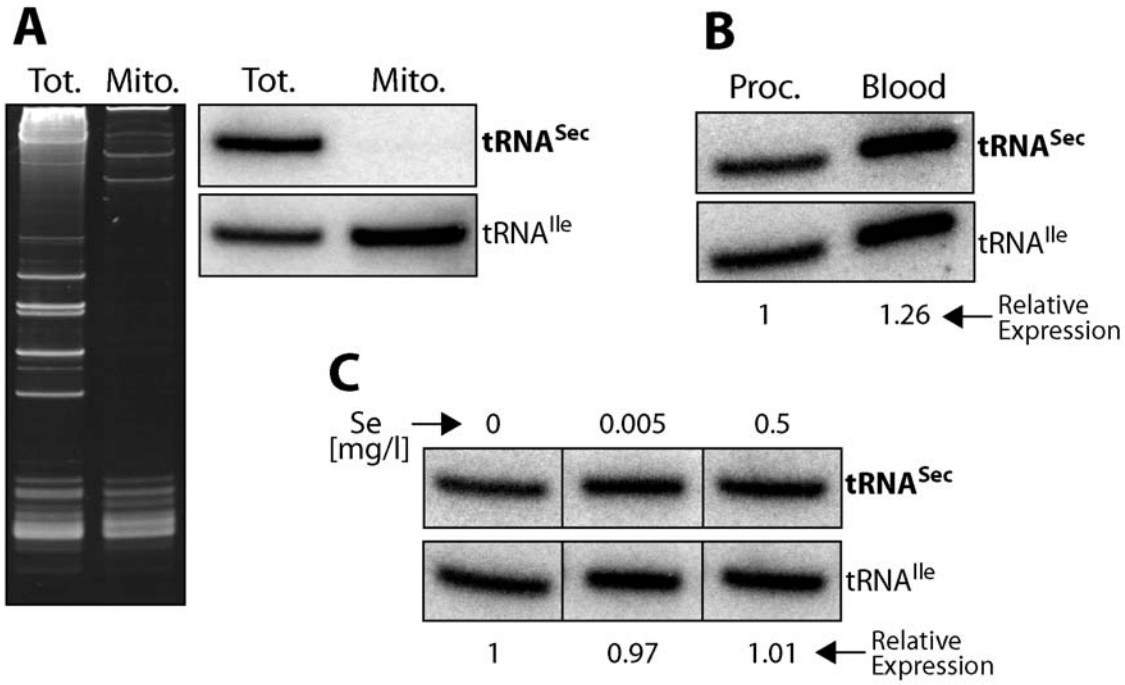
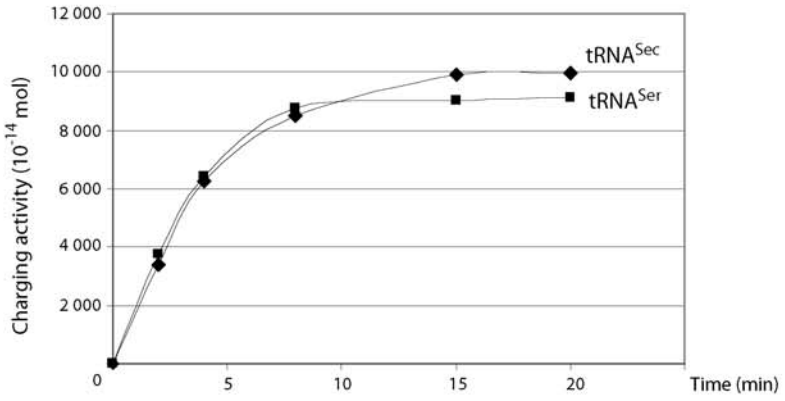


Figure 3

A



B

Species	Transcript	Kcat (s^{-1})	Km (µM)	Ref.
Trypanosoma cruzi	$tRNA^{Ser}_{CGA}$	9.56	3.2	(this work)
	$tRNA^{Sec}_{UCA}$	8.69	0.4	
Homo sapiens	$tRNA^{Ser}_{UGA}$	0.11	1.2	(25)
	$tRNA^{Sec}_{UCA}$	0.01	11.2	
Escherichia coli	$tRNA^{Ser}_{UGA}$	n.d.	0.2	(26)
	$tRNA^{Sec}_{UCA}$	n.d.	1.3	

C

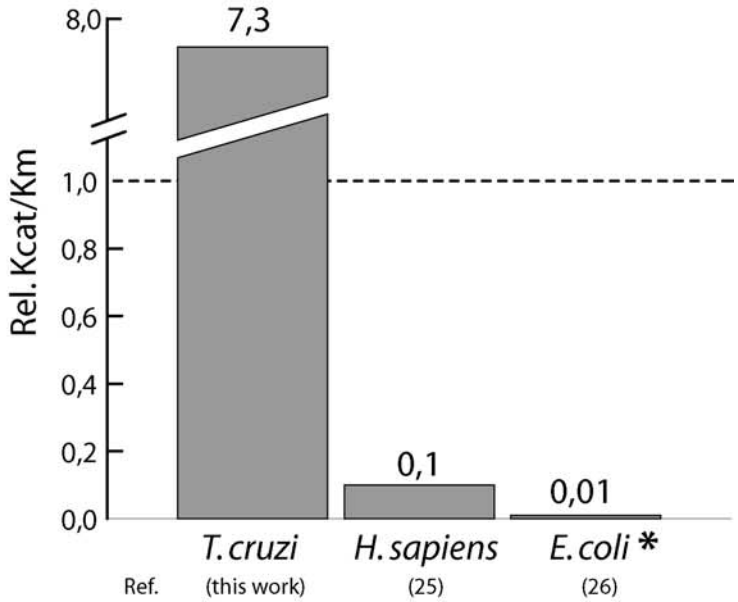


Figure 4

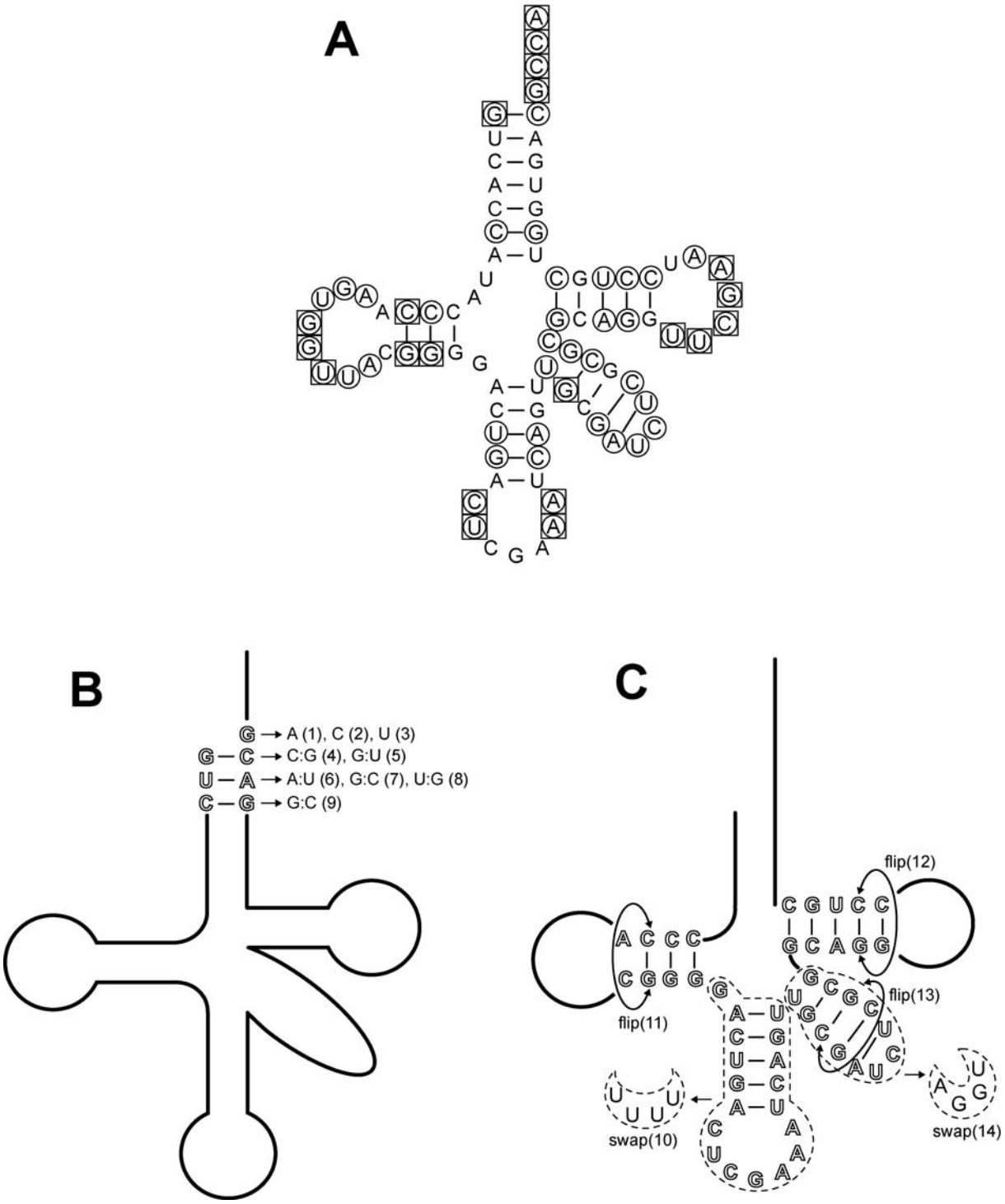


Figure 5

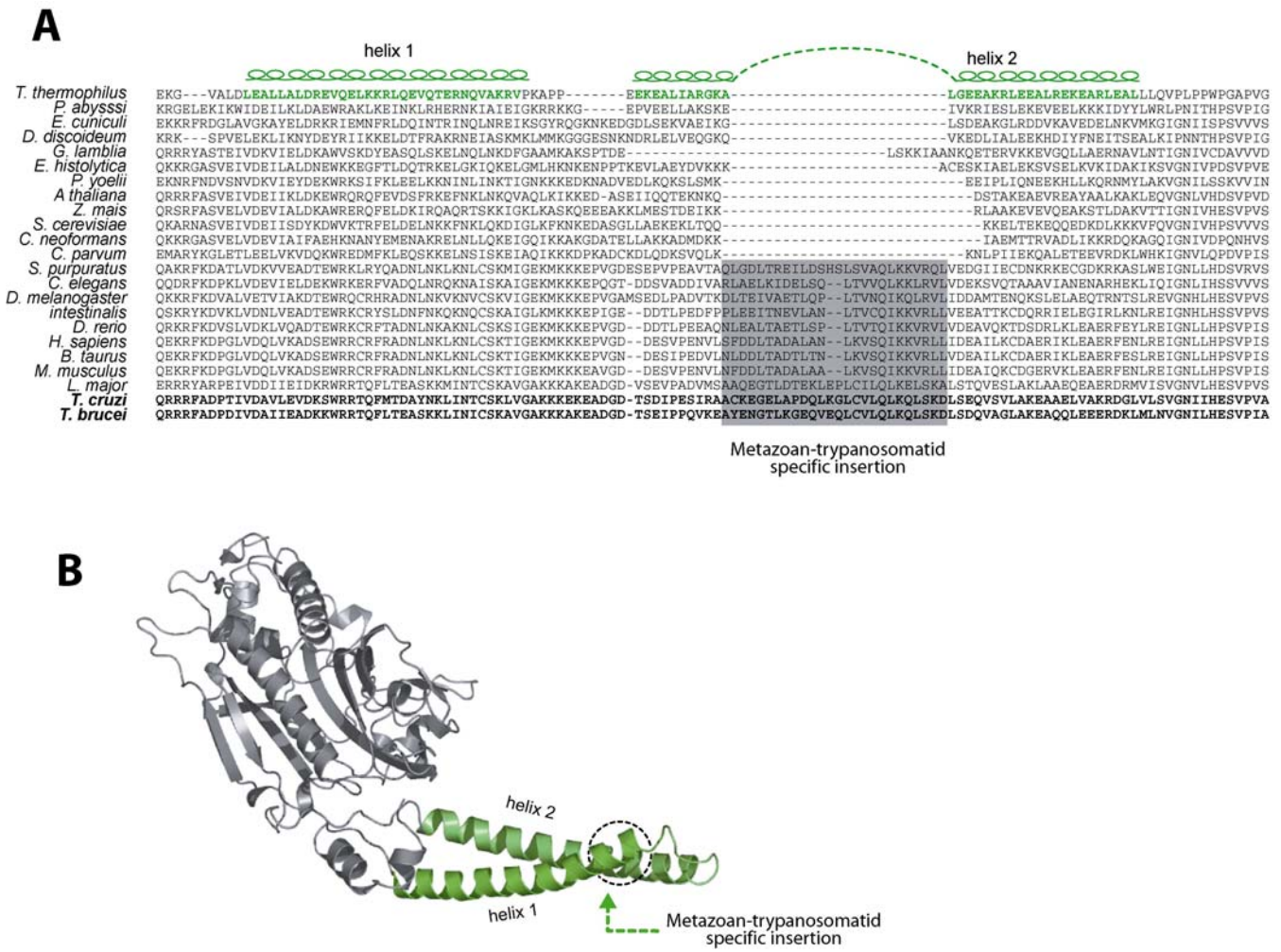


Figure 6

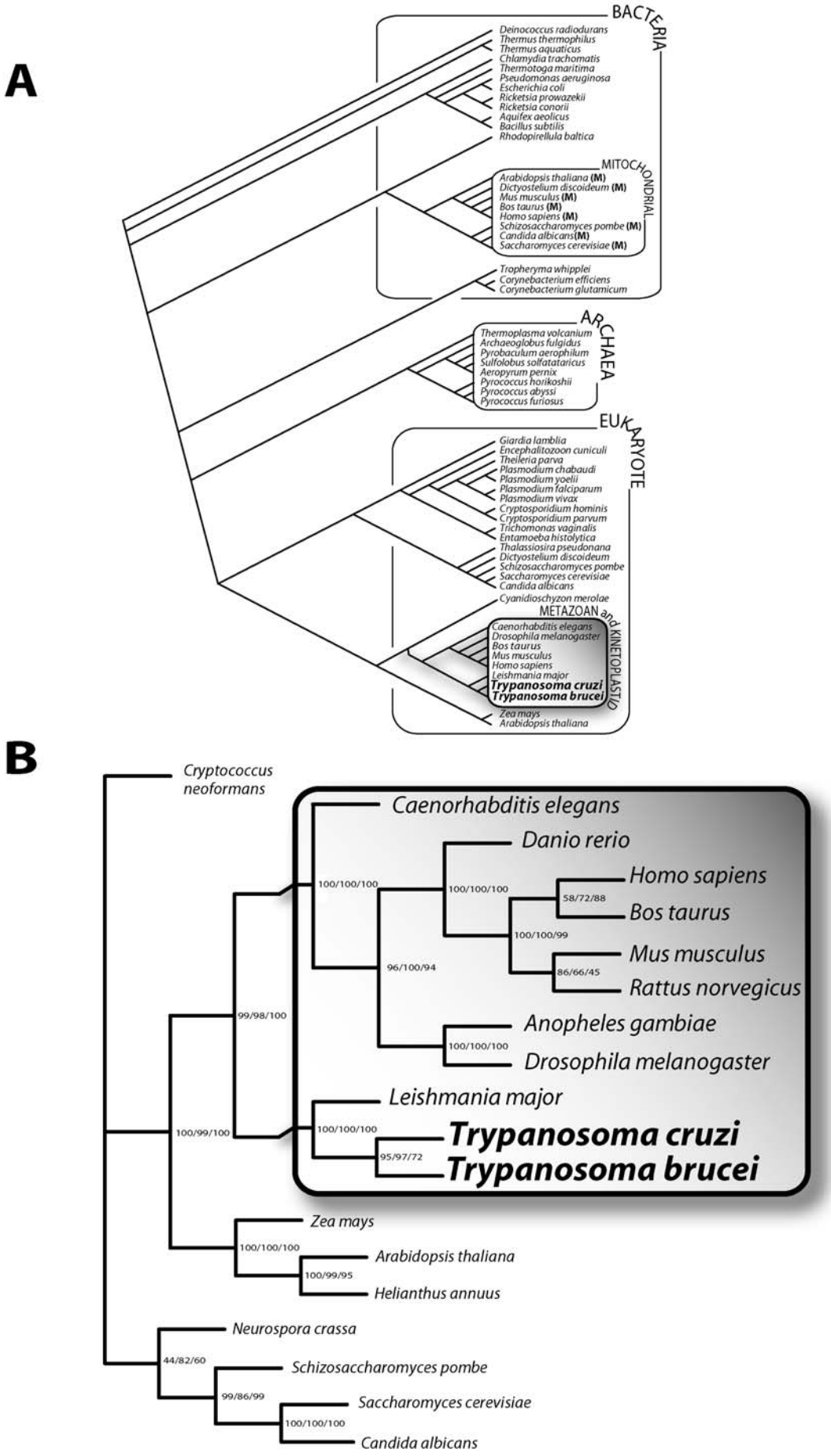


Table 1

Transcripts	kcat (%)	Km (μM)	rel Km/kcat
Wild-type $tRNA^{Ser}_{CGA}$	100	3.2	1
Wild-type $tRNA^{Sec}_{UCA}$	91	0.4	0.1
G73→A (1)	16	4.7	9.1
G73→C (2)	1	13	416.6
G73→U (3)	0.01	[a]nd	[a]nd
G1:C72→C:G (4)	<0.001	[a]nd	[a]nd
G1:C72→G:U (5)	89	2.0	0.7
U2:A71→A:U (6)	82	0.9	0.3
U2:A71→G:C (7)	165	4	0.7
U2:A71→U:G (8)	114	2.6	0.7
C3:G70→G:C (9)	56	5.6	3.1
Δ anticodon domain (10)	15	2.5	5.3
D arm Flip (11)	152	3.5	0.7
T arm Flip (12)	101	3.9	1.2
Variable loop Flip (13)	85	6.4	2.4
Δ variable loop (14)	0.25	71	8 928