J Mol Evol (2002) 55:727–733
DOI: 10.1007/s00239-002-2368-3

JOURNAL OF MOLECULAR EVOLUTION

Intron Positions Delineate the Evolutionary Path of a Pervasively Appended Peptide in Five Human Aminoacyl-tRNA Synthetases

Kiyotaka Shiba

Department of Protein Engineering, Cancer Institute, Japanese Foundation for Cancer Research, 1-37-1, Kami-Ikebukuro, Toshima, Tokyo 170-8455, Japan

Received: 23 April 2002 / Accepted: 24 June 2002

Abstract. Recent progress in genome sequencing has revealed a correspondence between the evolution of multicellularity and the appending of new peptides onto age-old enzyme bodies. Indicative of the pervasive nature of these appended peptides, in some cases the same sequences have been appended to a number of different enzymes. By analyzing the positions of introns within one such roaming peptide, an approximately 50-amino acid motif appended to five human aminoacyl-tRNA synthetases, I have delineated its path in eukaryote evolution. The motif was first acquired as an N-terminal extension by histidyl- and glycyl-tRNA synthetases at a very early stage of eukaryote evolution. Later, but not less than 1200 million years ago, the motif spread from histidyl-tRNA synthetase to the C and N terminals of glutamyl- and prolyl-tRNA synthetase, respectively, and then spread further during the evolution of the Chordate lineage to the N terminal of tryptophanyl-tRNA synthetase. In similar fashion, the motif in glycyl-tRNA synthetase spread to the C terminal of methionyl-tRNA synthetase not later than 1000 million years ago.

Key words: Aminoacyl-tRNA synthetase — Appendix domain — Motif-N — Intron — Exon

Correspondence to: Kiyotaka Shiba; *email:* kshiba@jfcr.or.jp

Introduction

The family of aminoacyl-tRNA synthetases, enzymes catalyzing the attachment of amino acids to their cognate tRNAs, is believed to be among the oldest of enzyme families. Indeed, the observation that existing aminoacyl-tRNA synthetases from Bacteria, Eucarya, and Archaea all share sequence similarities (Brown and Doolittle 1995) suggests that the emergence of this catalyst was a prerequisite for the establishment of the modern genetic code system.

Human cells use 19 (or 20, when bifunctional glutamyl-prolyl-tRNA synthetase is counted as 2) cytoplasmic aminoacyl-tRNA synthetases to catalyze attachment of the 20 amino acids used to synthesize proteins. The primary structures of 18 of these enzymes share a high degree of sequence similarity with those of *E. coli*, indicating the enzymes from *H. sapiens* and *E. coli* to have evolved from common ancestors. The only exception, glycyl-tRNA synthetase, has a rather complicated evolutionary history—i.e., humans and *E. coli* use distinct forms of the enzyme that must have evolved from different origins (Shiba et al. 1994a). Nevertheless, the fact that orthologues of human glycyl-tRNA synthetase are found in the genomes of Archaea and some Bacteria (Shiba et al. 1997a) is indicative of its antiquity. Reflecting the structural conservation among these enzymes is the functional conservation revealed by cross-species complementation experiments. For example, we have shown that the gene encoding human lysyl-tRNA synthetases can support cell growth of a mutant

E. coli strain in which the chromosomal genes encoding its two lysyl-tRNA synthetases were deleted (Shiba et al. 1997b). Apparently, human lysyl-tRNA synthetase has all the necessary functionality of *E. coli* lysyl-tRNA synthetase, which is to be expected given its pivotal role in translation.

Besides aminoacylation, some aminoacyl-tRNA synthetases, especially those from eukaryotes, have acquired surprising new functions that are not observed in enzymes from other organisms. In many cases, these new functions were acquired by appending extra domains to the age-old core catalytic domain (Kisselev and Wolfson 1994; Mirande 1991). For example, human tyrosyl-tRNA synthetase has acquired an extra domain at its C terminus that is absent from prokaryotic and lower eukaryotic tyrosyl-tRNA synthetases (Kleeman et al. 1997). The new domain is approximately 170 amino acids in length and exerts cytokine-like effects in leukocytes and monocytes, including stimulation of chemotaxis and induction of myeloperoxidase, TNF-α, and tissue factor (Wakasugi and Schimmel 1999). Because orthologous ornamented tyrosyl-tRNA synthetases are found only in the mouse, bovine, and fruit fly, this domain would seem to have been acquired by the enzyme when the common ancestor of chordates and arthropods emerged.

Another example is human tryptophanyl-tRNA synthetase, which has acquired an extra domain of about 60 amino acids at its N terminus. This sequence, which is unrelated to the domain appended to tyrosyl-tRNA synthetase, has an inhibitory effect on the angiostatic activity that resides in the age-old core catalytic domain of the enzyme (Otani et al. 2002; Wakasugi et al. 2002). It is well known that two forms of tryptophanyl-tRNA synthetase are produced by alternative splicing in human cells; one has a complete N-terminal extra domain (full length), while the other lacks the first half of the extra sequence (Mini-TrpRS) (Tolstrup et al. 1995; Turpaev et al. 1996). Although both enzymes possess aminoacylation activity (Tolstrup et al. 1995), only Mini-TrpRS exhibits angiostatic activity (Otani et al. 2002; Wakasugi et al. 2002). Thus, the N-terminal appended domain somehow masks the potential angiostatic activity of the core domain of tryptophanyl-tRNA synthetase.

Of the 19 human cytoplasmic aminoacyl-tRNA synthetases, 5 have appended domains that share a common 50-amino acid motif, indicating that they all propagated from a single common ancestor. These include the N-terminal appendices of tryptophanyl-, histidyl-, and glycyl-tRNA synthetases; the C-terminal appendix of methionyl-tRNA synthetase; and the tandemly repeated (three times) sequences in the linker peptide that connects the glutamyl- and prolyl-tRNA synthetase domains in bifunctional glutamyl-prolyl-tRNA synthetase. This motif was first described as a conserved block in the N terminal of human histidyl-tRNA synthetase and in the repeated peptide of human glutamyl-prolyl-tRNA synthetase (Fett and Knippers 1991) and was later identified in the N-terminal appendices of human tryptophanyl- and *Bombyx mori* glycyl-tRNA synthetases (Nada et al. 1993) and at the C terminus of human methionyl-tRNA synthetase (Cahuzac et al. 2000). Analyses of circular dichroism (Raben et al. 1994) and NMR chemical shifts (Raben et al. 1994) showed the secondary structure of the motif to be rich in α-helix; moreover, the tertiary structure of the motif from hamster glutamyl-prolyl-tRNA synthetase was determined to be an antiparallel coiled-coil in solution (Cahuzac et al. 2000).

Although 5 of the 19 human aminoacyl-tRNA synthetases contain this motif, the number of enzymes in which it occurs varies in other organisms. For instance, there is no evidence that Bacteria or Archaea use this motif in any enzyme. In Eucarya, the motif is popular among vertebrates such as mammals, fish, and amphibians but is only occasionally found in fungus or plant genomes. Thus, the motif seems to have surged during the evolution to *H. sapiens*. In this paper, I summarize the distributions of the motif in five eukaryote genomes whose complete sequences have been reported and delineate the evolutional path of the motif using data on intron positions within the motif. I use the term "Motif-N" for the sequence, which was named to describe the appended motif in human glycyl-tRNA synthetase (Shiba et al. 1994a).

Methods

Database Analyses

Acquisition of Motif-N by various organisms was investigated in the NCBI "nr," "EST," and "Drosophila genome" databases using blastp and tbalstn (Altschul et al. 1990). Multiple sequence alignments were calculated using ClustalX (Thompson et al. 1997). Intron positions were identified from the genome sequences of *Schizosaccharomyces pombe* (Wood et al. 2002), *Caenorhabditis elegans* (The *C. elegans* Sequencing Consortium 1998), *Drosophilae melanogaster* (Adams et al. 2000), *Homo sapiens* (Lander et al. 2001; Venter et al. 2001), and *Arabidopsis thaliana* (The *Arabidopsis* Genome Initiative 2000). Genes for human glutamyl-prolyl-, histidyl-, glycyl-, methionyl-, and tryptophanyl-tRNA synthetase are located on chromosomes 1, 5, 7, 12, and 14, respectively.

Results

Distributions of Motif-N in Genome Space

Acquisition of Motif-N by histidyl-, glycyl-, bifunctional glutamyl-prolyl (or monofunctional glutamyl- and prolyl-)-, tryptophanyl-, and methionyl-aminoacyl-tRNA synthetases from various eukaryotes has been reported previously (Cahuzac et al.

2000; Cerini et al. 1991; Fett and Knippers 1991; Kaminska et al. 2001; Nada et al. 1993; Shiba et al. 1994a). In addition, blast searches using updated versions of the databases have added asparginyl-tRNA synthetase from *A. thaliana*, an organism whose genome contains four genes (SYNC0, SYNC1, SYNC2, and SYNC3) for asparginyl-tRNA synthetase (Peeters et al. 2000). Phylogenetic analysis showed that these four genes arose through repeated duplication of a gene transferred from an ancestral plastid genome (Peeters et al. 2000). SYNC0, which is closest to the ancestral gene, has a 60-amino acid extension at its N terminal and is believed to be targeted to both mitochondria and chloroplasts. The newly evolved SYNC1–SYNC3 lack this N-terminal extension; instead, they have acquired an approximately 90-amino acid insertion within their catalytic domains (Peeters et al. 2000). The sequence of the insert resembles that of other Motif-N's, leading to the conclusion that three of four *A. thaliana* asparginyl-tRNA synthetases contain Motif-N as an insertion domain (data not shown).

The analysis of the acquisition of Motif-N by the five eukaryotic lineages studied, along with the estimated divergence times, is summarized in Fig. 1. Motif-N's are found in methionyl- and tryptophanyl-tRNA synthetases from *H. sapiens* but not in those from either *D. melanogaster* or *C. elegans*. Blast searches using EST sequences from various organisms revealed that *Xenopus laevis* (frog) and *Danio rerio* (zebrafish) acquired the motif both at the C terminal of methionyl-tRNA synthetase and at the N terminal of tryptophanyl-tRNA synthetase (data not shown). This means that the acquisition of Motif-N by methionyl- and tryptophanyl-tRNA synthetase must have occurred after the divergence of Chordata and Arthropoda, about 1000 million years ago.

In human and fruit fly, the glutamyl- and prolyl-tRNA synthetases exist as a fused bifunctional glutamyl-prolyl-tRNA synthetase, and motif-N is present in the form of tandem repeats within the linker region that connects the two synthetase domains (three repeats in human and six repeats in fruit fly) (Cerini et al. 1991; Fett and Knippers 1991). Interestingly, Motif-N is appended at the C terminal (six repeats) of the monofunctional glutamyl-tRNA synthetase and at the N terminal (one repeat) of the prolyl-tRNA synthetase from *C. elegans* (Cahuzac et al. 2000). One simple interpretation of that observation is that homologous recombination between the C-terminal Motif-N of prolyl-tRNA synthetase and the N-terminal Motif-N of glutamyl-tRNA synthetase evolved a bifunctional glutamyl-prolyl-tRNA synthetase after divergence into Nematoda and the ancestor of Chordata and Arthropoda, about 1200 million years ago. However, additional sequence data from a variety of other organisms will need to be analyzed before one can draw a conclusion about the origin of the bifunctional glutamyl-prolyl-tRNA synthetase.

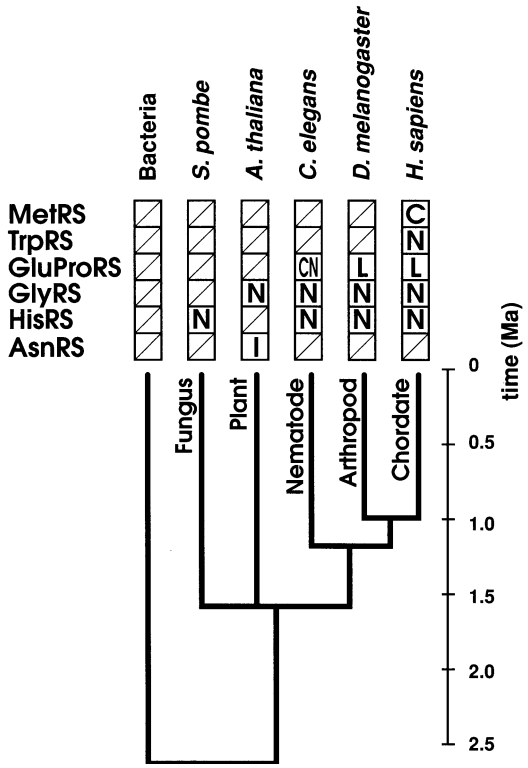

Fig. 1. Distributions of Motif-N in six aminoacyl-tRNA synthetases from five eukaryotes. The tree shows the relationships of eubacteria, fungi, plants, and animals. The scale on the right pane indicates the divergence time of these lineages expressed as thousands of millions of years ago (Ma) (Wang et al. 1999): MetRS, methionyl-; TrpRS, tryptophanyl-; GluProRS, bifunctional glutamyl-prolyl-; GlyRS, glycyl-; HisRS, histidyl-; and AsnRS, asparginyl-tRNA synthetase. "N" and "C" indicate that the motif is located at the N-terminal and C-terminal end of an enzyme, respectively. "CN" in *C. elegans* GluProRS indicates that the motif is appended at the C terminal of prolyl- and the N terminal of glutamyl-tRNA synthetase, respectively. "I" indicates that the motif has been acquired as an insertion peptide into the catalytic core. Bacteria and Archaea (not shown in the figure) do not use the motif in their proteome.

As described above, the histidyl-, glycyl-, glutamyl-, and prolyl-tRNA synthetases of an ancestral multicellular eukaryote seems to have acquired Motif-N before divergence into Nematoda, Arthropoda, and Chordata; it spread to the tryptophanyl- and methionyl-tRNA synthetases of chordates only after their separation from the arthropods. The pattern of spread of Motif-N within the genome of *A. thaliana* is particularly intriguing given that multicellularity may have evolved separately in plants and animals (Kaiser 2001). My analysis showed Motif-N to be appended to two aminoacyl-tRNA synthetases in Arabidopsis: one at the N terminal of glycyl-tRNA synthetase, as seen in animals, and the other within the catalytic domain of the three asparaginyl-tRNA synthetases described above. The latter must have been acquired

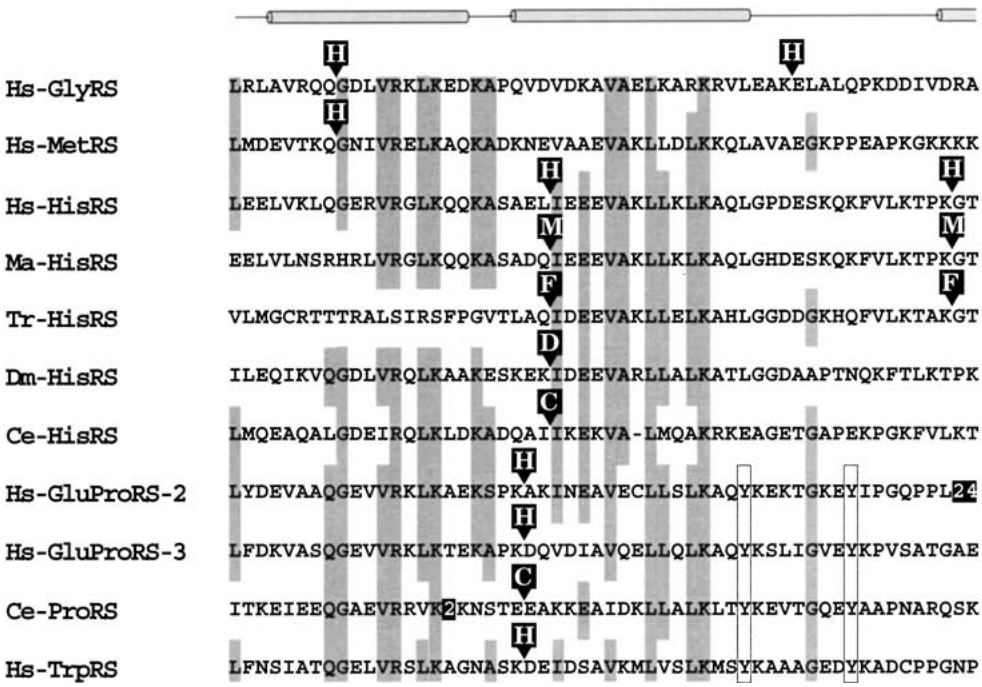

Fig. 2. Intron positions within Motif-N sequences. The amino acid sequences of 11 Motif-N's are shown. α-Helices within the motif, which were identified by NMR analysis of a peptide from hamster GluProRS, are shown at the *top* (Cahuzac et al. 2000). *Arrows* indicate intron insertion positions. When amino acid identities are conserved in seven sequences at a given position, they are *shaded*. *Boxes* show tyrosine residues involved in Ω-loop formation (Cahuzac et al. 2000). Hs, *Homo sapiens*; Ma, *Mesocricetus auratus* (golden hamster); Tr, *Takifugu rubripes* (Japanese pufferfish); Dm, *Drosophila melanogaster*; Ce, *Caenorhabditis elegans*. See the legend to Fig. 1 for abbreviations of enzyme names.

after separation of plant and animal, because *A. thaliana* asparaginyl-tRNA synthetases have an origin in the plastid genome, which is believed to have arisen through endosymbiosis of cyanobacteria after the separation of animal (Meyerowitz 2002). The genome of *A. thaliana* thus contains two Motif-Ns, one of which is plant lineage specific and another that may have its origin in the last common ancestor of plants and animals.

The above data suggest that glycyl-tRNA synthetase is the first enzyme to acquire Motif-N before the divergence of plant and animal. However, this conclusion would be inconsistent with the fact that whereas the single-celled eukaryote *S. pombe* expresses Motif-N at the N terminus of histidyl-tRNA synthetase (Cahuzac et al. 2000), glycyl-tRNA synthetase does not contain the motif (Wood et al. 2002). Moreover, Motif-N is absent from the genome of another single-celled eukaryote, *Saccharomyces cerevisiae* (Goffeau et al. 1996). Consequently, more data will be required before questions surrounding the antiquity of the Motif-N in glycyl- and histidyl-tRNA synthetases can be resolved.

Intron Positions of Motif-N's

The distributions of Motif-N within the five eukaryote genomes studied indicate that the motif was first acquired by glycyl- and/or histidyl-tRNA synthetases at an early stage in the evolution of eukaryotes and then spread to glutamyl-, prolyl-, tryptophanyl-, and methionyl-RNA synthetases as multicellularity continued to evolve. The spread to tryptophanyl- and methionyl-tRNA synthetases occurred only in the Chordate lineage and, compared to the spread to glutamyl- and prolyl-tRNA synthetases, is a relatively recent event. The evolution of Motif-N within newer synthetases must have been completed *via* genetic events, including recombination and duplication of preexisting chromosomal fragments that code for Motif-N in other aminoacyl-tRNA synthetases. To determine the immediate parent for each Motif-N in glutamyl-prolyl-, tryptophanyl-, and methionyl-RNA synthetases, I compared the positions of the introns within Motif-N's from various organisms; intron position is known to be a reliable marker of evolutionary paths (Iwabe et al. 1990; Kersanach et al. 1994).

A total of 11 examples in which one or more introns were located within Motif-N's was found in six enzymes from five organisms (Fig. 2). These include histidyl-tRNA synthetases from *Takifugu rubripes* (Japanese pufferfish) (Brenner and Corrochano 1996) and *Mesocricetus auratus* (golden hamster) (Tsui and Siminovitch 1987), in addition to enzymes from *H. sapiens*, *D. melanogaster*, and *C. elegans*. The

results showed the intron positions to be well conserved within subclasses of Motif-N. For instance, the Motif-N in histidyl-tRNA synthetase from *H. sapiens* has two introns, one in the middle of the motif and one in the C-proximal region (Fig. 2). The middle position is precisely conserved in the same enzyme from *M. auratus*, *T. rubripes*, *D. melanogaster*, and *C. elegans*. Likewise, the intron position in prolyl (or glutamyl-prolyl-)-tRNA synthetase is conserved between *H. sapiens* and *C. elegans*. Thus, the intron positions appear to be fairly stable within the relevant time frame (i.e., 1200 million years), despite the fact that the positions are believed to "slide" over long periods of evolutionary history (Stoltzfus et al. 1997). Assuming that the introns did not slide, the conservation of intron positions within Motif-N's from different enzymes is indicative of the close evolutionary relationship between the appended motifs. As shown in Fig. 2, the intron locations within Motif-N from human glycyl-tRNA synthetase and methionyl-tRNA synthetase are strictly conserved, as are the locations within human tryptophanyl-tRNA synthetase and human glutamyl-prolyl-tRNA synthetase. This means that the immediate parent of the Motif-N in methionyl-tRNA synthetase is also the parent of the Motif-N in glycyl-tRNA synthetase, and the parent of the Motif-N of tryptophanyl-tRNA synthetase is also the parent of the Motif-N in glutamyl-prolyl-tRNA synthetase.

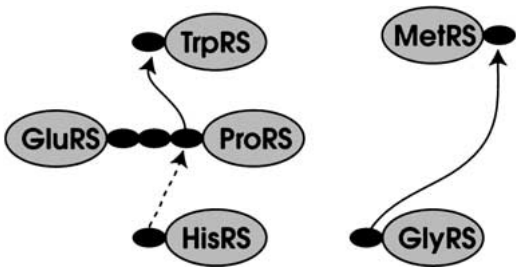

Fig. 3. Inferred history of five Motif-N's in human aminoacyl-tRNA synthetases. *Black* ovals represent Motif-N's. See the legend to Fig. 1 for abbreviations of enzyme names.

Discussion

Inference of the Evolutionary History of Motif-N in H. sapiens *Based on Intron Location*

The evolutionary history of Motif-N inferred from the present analysis is diagramed in Fig. 3. At an early stage of eukaryote evolution (before the divergence of animal and plant), histidyl-tRNA synthetase and glycyl-tRNA synthetase first acquired the motif at their N termini. Why and how this first motif arose is unclear. Blast searches using current versions of the databases identified Motif-N's in aminoacyl-tRNA synthetases but not in any other enzymes. It is possible that the sequence emerged in histidyl- and glycyl-tRNA synthetase as a targeting signal for the transport of a protein translated in cytosol into mitochondria. In fact, the evolution of such a targeting signal would have been required for the transfer of mitochondrial genes to the nuclear chromosome, which is known to have occurred (Gray 1992). Moreover, Motif-N (Cahuzac et al. 2000) is now known to share an amphipathic α-helical structure with a mitochondrial targeting signal (Roise and Schatz 1988). This means that a single gene for histidyl-tRNA synthetase in the genome of *S. pombe* (Wood et al. 2002) must somehow be localized in both the cytosol and the mitochondria. *S. cerevisiae* accomplishes the dual targeting of histidyl-tRNA synthetase from a single *HTS1* gene by using two in-frame translation start sites (Natsoulis et al. 1986), and a similar mechanism might be employed for dual targeting of histidyl-tRNA synthetase in *S. pombe*—i.e., the longer enzyme, which includes Motif-N, would be transported into mitochondria, while the shorter version is localized in the cytosol. Likewise, one of the two glycyl-tRNA synthetases from *A. thaliana* that contain Motif-N's is localized in both the mitochondria and the cytosol (Duchêne et al. 2001).

The ancestral Motif-N in histidyl-tRNA synthetase appears to have next spread to the N terminal of prolyl-tRNA synthetase before the emergence of nematodes about 1200 million years ago). The proposed sibling relationship between the Motif-N in histidyl-tRNA synthetase and that in prolyl-tRNA synthetase is legitimated by the proximity of the intron positions of these motifs (Fig. 2). That their positions are shifted by two amino acids (six nucleotides) could be explained by "intron sliding" (Stoltzfus et al. 1997). The evolutional proximity of the motif in histidyl- and glutamyl-prolyl-tRNA synthetases has been discussed elsewhere (Brenner and Corrochano 1996). Glutamyl-tRNA synthetase must have acquired the motif at its C terminal during this period, but because the motif in *C. elegans* glutamyl-tRNA synthetase has no intron, its immediate parent is unclear from the present data. Still, one simple story for the origin of the bifunctional glutamyl-prolyl-tRNA synthetase found in chordates and arthropods would be that it arose through the fusion of the two enzymes at their homologous regions, the C-terminal and N-terminal Motif-N's.

The complete retention of intron positions in the Motif-N's in tryptophanyl- and glutamyl-prolyl-tRNA synthetases indicates the former to have originated from the latter (or prolyl-tRNA synthetase), and not from histidyl- or glycyl-tRNA synthetase (Fig. 2). In addition, the conservation of two tyrosine

residues involved in Ω-loop formation in the Motif-N in glutamyl-prolyl-tRNA synthetase (Cahuzac et al. 2000) and tryptophanyl-tRNA synthetase is consistent with their being closely related (Fig. 2). In a similar fashion, the Motif-N in the C terminal of methionyl-tRNA synthetase was propagated from glycyl-tRNA synthetase. The surge of Motif-N's in aminoacyl-tRNA synthetases has thus occurred in parallel with multicellularity within the animal lineage. In the plant lineage, in contrast, only glycyl- and asparaginyl-tRNA synthetases acquired Motif-N's.

It is intriguing that more than one biological function has been attributed to Motif-N. As mentioned above, Motif-N in human tryptophanyl-tRNA synthetase acts to hinder the angiostatic activity embedded in the core domain (Otani et al. 2002; Wakasugi et al. 2002), though it is unnecessary for the enzyme's aminoacylation activity (Tolstrup et al. 1995). Similarly, deletion of Motif-N does not abolish the *in vitro* aminoacylation activity of glutamyl-prolyl-tRNA synthetase (Cerini et al. 1991; Stehlin et al. 1998). On the other hand, the aminoacylation activity of human histidyl-tRNA synthetase is dependent on the presence of Motif-N (Raben et al. 1994), and the motif is also required for *B. mori* glycyl-tRNA synthetase to bind tRNAs (Wu et al. 1995). The involvement of Motif-N in aminoacylation activity is most likely related to its ability to bind RNA—e.g., isolated Motif-N from human (Rho et al. 1998) or hamster (Cahuzac et al. 2000) glutamyl-prolyl-tRNA synthetase binds RNA *in vitro*, and Motif-N in human methionyl-tRNA synthetase is involved in a tRNA-sequestering activity (Kaminska et al. 2001).

The other potential function of Motif-N is mediating protein–protein interactions. A yeast two-hybrid screening experiment in which the C-terminal repeated appendix domain of isoleucyl-tRNA synthetase—a sequence unrelated to Motif-N (Shiba et al. 1994b)—was used as bait showed that Motif-N mediates the protein–protein interaction in human glutamyl-prolyl-tRNA synthetase (Rho et al. 1996). In metazoans, moreover, eight different aminoacyl-tRNA synthetases and three non-aminoacyl-tRNA synthetase components form high molecular weight complexes, and it has been suggested that it is the appendix domains, including Motif-N, that mediate the formation of such complexes (Dang 1986; Kisselev and Wolfson 1994; Mirande 1991). Similarly, isoleucyl-, glutamyl-prolyl-, and methionyl-tRNA synthetases interact to form a multi-aminoacyl-tRNA synthetase complex, perhaps via Motif-N (Quevillon et al. 1999; Rho et al. 1996, 1999).

It thus appears that a constellation of functions, including hindrance of angiostatic activity, binding to tRNA, protein–protein interactions, and signaling of mitochondrial targeting, may have evolved in Motif-N. And it is likely that it was its capacity to evolve that led to the surge of Motif-N during animal evolution. However, allocation of these functions to the respective Motif-N's may not explain the differentiation of evolutionary paths into the two main branches shown in Fig. 3. For instance, while the Motif-N's in histidyl-, glutamyl-prolyl-, and tryptophanyl-tRNA synthetases are closely related, their functions are diverse, varying from hindrance of angiostatic activity to mediation of tRNA binding and protein–protein interaction.

Acknowledgments. This work was supported by an HFSP Research Grant.

References

Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, George RA, et al. (2000) The genome sequence of *Drosophila melanogaster*. Science 287:2185–2195

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:4035–410

The *Arabidopsis* Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Brenner S, Corrochano LM (1996) Translocation events in the evolution of aminoacyl-tRNA synthetases. Proc Natl Acad Sci USA 93:8485–8489

Brown JR, Doolittle WF (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc Natl Acad Sci USA 92:2441–2445

Cahuzac B, Berthonneau E, Birlirakis N, Guittet E, Mirande M (2000) A recurrent RNA-binding domain is appended to eukaryotic aminoacyl-tRNA synthetases. EMBO J 19:445–452

The *C. elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans:* A platform for investigating biology. Science 282:2012–2018

Cerini C, Kerjan P, Astier M, Gratecos D, Mirande M, Sémériva M (1991) A component of the multisynthetase complex is a multifunctional aminoacyl-tRNA synthetase. EMBO J 10:4267–4277

Dang CV (1986) Multienzyme complex of aminoacyl-tRNA synthetases: An essence of being eukaryotic. Biochem J 239:249–255

Duchêne AM, Peeters N, Dietrich A, Cosset A, Small ID, Wintz H (2001) Overlapping destinations for two dual targeted glycyl-tRNA synthetases in *Arabidopsis thaliana* and *Phaseolus vulgaris*. J Biol Chem 276:15275–15283

Fett R, Knippers R (1991) The primary structure of human glutaminyl-tRNA synthetase: A highly conserved core, amino acid repeat regions, and homologies with translation elongation factors. J Biol Chem 266:1448–1455

Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG (1996) Life with 6000 genes. Science 274:546

Gray MW (1992) The endosymbiont hypothesis revisited. Int Rev Cytol 141:233–357

Iwabe N, Kuma K, Kishino H, Hasegawa M, Miyata T (1990) Compartmentalized isozyme genes and the origin of introns. J Mol Evol 31:205–210

Kaiser D (2001) Building a multicellular organism. Annu Rev Genet 35:103–123

Kaminska M, Shalak V, Mirande M (2001) The appended C-domain of human methionyl-tRNA synthetase has a tRNA-sequestering function. Biochemistry 40:14309–14316

Kersanach R, Brinkmann H, Liaud MF, Zhang DX, Martin W, Cerff R (1994) Five identical intron positions in ancient duplicated genes of eubacterial origin. Nature 367:387–389

Kisselev LL, Wolfson AD (1994) Aminoacyl-tRNA synthetases from higher eukaryotes. Prog Nucl Acid Res Mol Biol 48:83–142

Kleeman TA, Wei DB, Simpson KL, First EA (1997) Human tyrosyl tRNA synthetase shares amino acid sequence homology with a putative cytokine. J Biol Chem 272:14420–14425

Lander ES, Linton LM, Birren B, et al. (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921

Meyerowitz EM (2002) Plants compared to animals: The broadest comparative study of development. Science 295:1482–1485

Mirande M (1991) Aminoacyl-tRNA synthetase family from prokaryotes and eukaryotes: Structural domains and their implications. Prog Nucl Acid Res Mol Biol 40:95–142

Nada S, Chang PK, Dignam JD (1993) Primary structure of the gene for glycyl-transfer RNA synthetase from *Bombyx mori*. J Biol Chem 268:7660–7667

Natsoulis G, Hilger F, Fink GR (1986) The *HTS1* gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of *S. cerevisiae*. Cell 46:235–243

Otani A, Slike BM, Dorrell MI, Hood J, Kinder K, Ewalt KL, Cheresh D, Schimmel P, Friedlander M (2002) A fragment of human TrpRS as a potent antagonist of ocular angiogenesis. Proc Natl Acad Sci USA 99:178–183

Peeters NM, Chapron A, Giritch A, Grandjean O, Lancelin D, Lhomme T, Vivrel A, Small I (2000) Duplication and quadruplication of *Arabidopsis thaliana* cysteinyl- and asparaginyl-tRNA synthetase genes of organellar origin. J Mol Evol 50:413–423

Quevillon S, Robinson JC, Berthonneau E, Siatecka M, Mirande M (1999) Macromolecular assemblage of aminoacyl-tRNA synthetases: Identification of protein–protein interactions and characterization of a core protein. J Mol Biol 285:183–195

Raben N, Nichols R, Dohlman J, Mcphie P, Sridhar V, Hyde C, Leff R, Plotz P (1994) A motif in human histidyl-tRNA synthetase which is shared among several aminoacyl-tRNA synthetases is a coiled-coil that is essential for enzymatic activity and contains the major autoantigenic epitope. J Biol Chem 269:24277–24283

Rho SB, Lee KH, Kim JW, Shiba K, Jo YJ, Kim S (1996) Interaction between human tRNA synthetases involves repeated sequence elements. Proc Natl Acad Sci USA 93:10128–10133

Rho SB, Lee JS, Jeong EJ, Kim KS, Kim YG, Kim S (1998) A multifunctional repeated motif is present in human bifunctional tRNA synthetase. J Biol Chem 273:11267–11273

Rho SB, Kim MJ, Lee JS, Seol W, Motegi H, Kim S, Shiba K (1999) Genetic dissection of protein–protein interactions in multi-tRNA synthetase complex. Proc Natl Acad Sci USA 96:4488–4493

Roise D, Schatz G (1988) Mitochondrial presequences. J Biol Chem 263:4509–4511

Shiba K, Schimmel P, Motegi H, Noda T (1994a) Human glycyl-tRNA synthetase: Wide divergence of primary structure from bacterial counterpart and species-specific aminoacylation. J Biol Chem 269:30049–30055

Shiba K, Suzuki N, Shigesada K, Namba Y, Schimmel P, Noda T (1994b) Human cytoplasmic isoleucyl-tRNA synthetase: Selective divergence of the anticodon-binding domain and acquisition of a new structural unit. Proc Natl Acad Sci USA 91:7435–7439

Shiba K, Motegi H, Schimmel P (1997a) Maintaining genetic code through adaptations of tRNA synthetases to taxonomic domains. TIBS 22:453–457

Shiba K, Stello T, Motegi H, Noda T, Musier-Forsyth K, Schimmel P (1997b) Human lysyl-tRNA synthetase accepts nucleotide 73 variants and rescues *Escherichia coli* double-defective mutant. J Biol Chem 272:22809–22816

Stehlin C, Burke B, Liu H, Shiba K, Musier-Forsyth K (1998) Species-specific differences in the operational RNA code for aminoacylation of $tRNA^{Pro}$. Biochemistry 37:8605–8613

Stoltzfus A, Logsdon Jr., JM, Palmer JD, Doolittle WF (1997) Intron "sliding" and the diversity of intron positions. Proc Natl Acad Sci USA 94:10739–10744

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl Acids Res 25:4876–4882

Tolstrup AB, Bejder A, Fleckner J, Justesen J (1995) Transcriptional regulation of the interferon-g-inducible tryptophanyl-tRNA synthetase includes alternative splicing. J Biol Chem 270:397–403

Tsui FW, Siminovitch L (1987) Isolation, structure and expression of mammalian genes for histidyl-tRNA synthetase. Nucl Acids Res 15:3349–3367

Turpaev KT, Zakhariev VM, Sokolova IV, Narovlyansky AN, Amchenkova AM, Justesen J, Frolova LY (1996) Alternative processing of the tryptophanyl-tRNA synthetase mRNA from interferon-treated human cells. Eur J Biochem 240:732–737

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA et al. (2001) The sequence of the human genome. Science 291:1304–1051

Wakasugi K, Schimmel P (1999) Two distinct cytokines released from a human aminoacyl-tRNA synthetase. Science 284:147–151

Wakasugi K, Slike BM, Hood J, Otani A, Ewalt KL, Friedlander M, Cheresh DA, Schimmel P (2002) A human aminoacyl-tRNA synthetase as a regulator of angiogenesis. Proc Natl Acad Sci USA 99:173–177

Wang DY, Kumar S, Hedges SB (1999) Divergence time estimates for the early history of animal phyla and the origin of plants, animals and fungi. Proc R Soc Lond B Biol Sci 266:163–171

Wood V, Gwilliam R, Rajandream MA, et al. (2002) The genome sequence of *Schizosaccharomyces pombe*. Nature 415:871–880

Wu H, Nada S, Dignam JD (1995) Analysis of truncated forms of Bombyx mori glycyl-tRNA synthetase: Function of an N-terminal structure in RNA binding. Biochemistry 34:16327–16336