Nucleic Acids Research, 1999, Vol. 27, No. 3 721–729

SURVEY AND SUMMARY

tRNA recognition and evolution of determinants in seryl-tRNA synthesis

Boris Lenhard[1,2], Omar Orellana[3], Michael Ibba[4] and Ivana Weygand-Durasevic[1,2,*]

[1]Department of Chemistry, Faculty of Science, University of Zagreb, 10000 Zagreb, Croatia, [2]Department of Molecular Genetics, Rudjer Boskovic Institute, 10000 Zagreb, Croatia, [3]Programa de Biología Celular y Molecular, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Casilla 70086, Santiago 7, Chile and [4]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520-8114, USA

Received July 8, 1998; Revised and Accepted September 14, 1998

ABSTRACT

We have analyzed the evolution of recognition of tRNAs^{Ser} by seryl-tRNA synthetases, and compared it to other type 2 tRNAs, which contain a long extra arm. In Eubacteria and chloroplasts this type of tRNA is restricted to three families: tRNA^{Leu}, tRNA^{Ser} and tRNA^{Tyr}. tRNA^{Leu} and tRNA^{Ser} also carry a long extra arm in Archaea, Eukarya and all organelles with the exception of animal mitochondria. In contrast, the long extra arm of tRNA^{Tyr} is far less conserved: it was drastically shortened after the separation of Archaea and Eukarya from Eubacteria, and it is also truncated in animal mitochondria. The high degree of phylogenetic divergence in the length of tRNA variable arms, which are recognized by both class I and class II aminoacyl-tRNA synthetases, makes type 2 tRNA recognition an ideal system with which to study how tRNA discrimination may have evolved in tandem with the evolution of other components of the translation machinery.

A common characteristic of all cells and organelles is the presence of a highly accurate protein synthesis machinery. Aminoacyl-tRNA synthetases (aaRSs) and tRNAs are essential parts of this machinery as they define the amino acid identity of mRNA codons. These central components of translation show a high degree of sequence, structural and functional homology across kingdoms (1–4) and it is believed that synthetases and tRNAs have co-evolved during the evolution of extant organisms.

Twenty aaRSs, one enzyme specific for each amino acid found in proteins, constitute a minimum set for protein biosynthesis in most prokaryotic cells. Only exceptionally, the cell contains two forms of the enzyme with the same amino acid specificity (5), or lacks a particular synthetase, as will be discussed later. There are additional set(s) of aaRSs in eukaryotic cells, which function in the organelles. Since these enzymes catalyze the same overall reaction and utilize a common strategy for chemical activation of their amino acids, aminoacyl-AMP being formed at the expense of ATP, it has long been supposed that all the synthetases had a common ancestral root. However, this assumption initially seemed at odds with the large differences in polypeptide chain length and quaternary structure found for different synthetases. However, when the sequences of examples from all the aaRSs became available, two groups were identified based on sequence similarities and the universal conservation of two mutually exclusive sets of sequence motifs (6). X-ray crystallographic studies also supported the view that there are virtually no structural similarities between the two groups of enzymes (class I and class II) (7–10) which have thus evolved from independent roots. As a general rule enzymes specific for the same amino acid, regardless of source, could be readily aligned with one another, as they are more similar to each other than to enzymes specific for any other amino acid (11). These data suggest that the specialization of many aaRSs with respect to amino acid specificity occurred before the branching of the bacterial, archaeal and eukaryotic lineages (11). Within each class, enzymes specific for chemically similar amino acids tend to cluster together, indicating the importance of the adaptation of binding sites to particular amino acids during evolution. Although there are obvious similarities between members of each class, generally there are no sequence relationships between members of the two classes (6,12). To date there is only a single case known of class-switching by an aaRS during evolution: the presence of amino acid motifs characteristic of the Rossman dinucleotide-binding domain identifies lysyl-tRNA synthetases of certain archaea and bacteria as class I synthetases, in contrast to the other known examples of this enzyme, which are class II synthetases (13,14).

tRNA molecules may have descended from a single ancestor and evolved by gene duplication and subsequent mutations (15,16). They may also have evolved from more simple structures (17). The recently established recruitment model (18), which proposes that a tRNA gene can be recruited from one isoaccepting group to another by a point mutation that concurrently changes tRNA identity and mRNA coupling capacity, may account for the

*To whom correspondence should be addressed at: Department of Chemistry, Faculty of Science, University of Zagreb, Strossmayerov trg 14, 10000 Zagreb, Croatia. Tel: +385 1 456 1197; Fax: +385 1 456 1177; Email: weygand@rudjer.irb.hr

evolution of some tRNA genes. According to the length of the extra arms, tRNAs can be divided into two classes: those with a short extra arm of 4–5 nucleotides (type 1) and those of type 2 with a long extra arm of at least 11 nucleotides (20,21). [Numbering of nucleotides in tRNAs was according to Sprinzl *et al*. (19), since it was used in most references cited. The extra arm is placed between nucleotides 43 and 49.] Interestingly, the tRNAs with a long extra arm are phylogenetically well-conserved and are restricted to three families: $tRNAs^{Tyr}$, $tRNAs^{Leu}$ and $tRNAs^{Ser}$. No other tRNAs belong to type 2, thus suggesting that the long extra arm is involved in the recognition of these tRNAs by their cognate aaRSs and/or in discrimination against non-cognate synthetases. While $tRNA^{Leu}$ and $tRNA^{Ser}$ carry a long extra arm in all organisms and organelles except in animal mitochondria (22,23) the sequences of tyrosine-specific tRNAs have undergone tremendous evolutionary change. The long extra arms of $tRNAs^{Tyr}$ have been lost twice: early after separation of Bacteria from Archaea and eukaryotes, and later in parallel with comparable changes in $tRNAs^{Leu}$ and $tRNAs^{Ser}$ in animal mitochondria (20). Because of the phylogenetic variability of the length and orientation of the variable arm, the type 2 tRNA system is considered as one of the best targets to study how tRNA discrimination mechanisms have developed in association with the evolution of tRNA. The fact that type 2 tRNAs are recognized by either class I or class II synthetases makes such studies even more intriguing.

Each synthetase binds and aminoacylates only its cognate isoacceptor tRNAs, which are only a minor subset of the total cellular pool of tRNAs. In order to avoid misacylating tRNAs from any of the 19 non-cognate groups, within each tRNA sequence there exist elements that are unambiguously recognized only by a cognate synthetase. These recognition elements are most commonly located in the tRNA anticodon, the acceptor stem and the associated 'discriminator' base at position 73 (24–28). In addition, synthetases sometimes recognize nucleotides in the tRNA variable pocket (29) and certain aspects of tRNA structure (30). The analyses of the synthetase/tRNA complexes revealed that these macromolecules interact in a stereochemically complementary manner (31–33). Observed molecular complementarity raised the possibility that particular nucleotide sequences in tRNAs would mirror the class-defining amino-acid motifs in the enzymes. However, such relatedness has never been found (34). Recent studies have revealed that most major identity determinants are conserved during evolution, although minor identity elements are often changed (35–41). If the general locations of the specificity-determining nucleotides in one isoaccepting group are the same in bacterial and higher eukaryotic tRNAs and only a sequence variation occurred, this may imply a necessity for a similar co-variation in sequences/structure of the aaRSs. Thus, the interactions of the three-dimensional surfaces of protein–RNA pairs require mutual evolutionary adaptation of tRNAs and their cognate synthetases. The alteration of the structure of either synthetase or tRNA, introduced by such mutations, can result in misacylation. The same phenomenon can also occur as the consequence of the altered ratio between free aaRS and the synthetase complexed with its cognate tRNA (42,43). *In vivo* misacylation of tRNA is lethal to the cell because it introduces errors during ribosomal mRNA translation. However, there is a well-established case in nature for misacylation in the glutamine and asparagine systems. In Gram-positive Bacteria and Archaea as well as in the chloroplasts of higher plants there is no detectable GlnRS activity (44–46). $Gln\text{-}tRNA^{Gln}$ is formed by amidation of $Glu\text{-}tRNA^{Gln}$ which is synthesized by glutamyl-tRNA synthetase in a misacylation reaction of $tRNA^{Gln}$ with glutamate. Similarly, transamidation, and not direct acylation, provides $Asn\text{-}tRNA^{Asn}$ for protein biosynthesis in the halophilic archaebacterium *Haloferax volcanii* (46). Therefore, in these organisms glutamyl- and asparaginyl-tRNA synthetases are naturally occurring misacylating enzymes. Seryl-tRNA synthetase also charges with serine two families of tRNA isoacceptors: its cognate $tRNAs^{Ser}$ and the selenocysteine-inserting tRNA species ($tRNA^{Sec}$) (47). While $Ser\text{-}tRNA^{Ser}$ participates directly in mRNA translation, $Ser\text{-}tRNA^{Sec}$ first undergoes a tRNA-dependent serine modification to produce $Sec\text{-}tRNA^{Ser}$. This indicates that in some organisms aaRSs are not solely responsible for providing the pool of required aminoacyl-tRNA species.

tRNA RECOGNITION AND EVOLUTION OF DETERMINANTS IN SERYL-tRNA SYNTHESIS

Unlike the tyrosine system, the known serine-specific tRNAs have experienced no dramatic type switches in the course of evolution. Recognition studies performed with the components of serine systems from different species, including Bacteria (48–53), yeast (54–58) and human (20,59,60) revealed that some, but not all, of the determinants have been conserved during evolution. Subtle shortening of the long extra arm has accompanied the development of these organisms: in Bacteria this domain comprises a variable number of nucleotides, mostly between 16 and 20. In eukaryotes, the long extra arms of $tRNA^{Ser}$ isoacceptors are exactly 14 nucleotides long (20), while the extra arms of $tRNAs^{Ser}$ in Archaea vary in length from 13 (*H.volcanii)* to 18 (*Methanothermus fervidus*) nucleotides.

Considerable information has been accumulated concerning the specificity of interaction between bacterial $tRNAs^{Ser}$ and their cognate aaRSs, from various biochemical experiments (48–53,61,62) and intense crystallographic studies of several macromolecular complexes from *Escherichia coli* and *Thermus thermophilus* (33,63). Two kinds of identity determinants characterize bacterial $tRNA^{Ser}$: sequence elements at various locations in the macromolecule and nucleotides that determine the specific tertiary structure of particular regions within the whole tRNA (64).

The *E.coli* $tRNA^{Ser}$ isoacceptor set: recognition and identity

Escherichia coli SerRS recognizes five tRNA isoacceptors plus $tRNA^{Sec}$, none of whose anticodons are involved in recognition. Instead, the acceptor stem, D-arm and extra stem/loop nucleotides most strongly contribute to serine tRNA identity (50,65,66). Comparison of the sequences of the five *E.coli* serine tRNAs revealed that they share the absolutely conserved G1:C72 and G2:C71 base pairs, the conserved chemistry of purine–pyrimidine base pairs at positions 4:69, 5:68 and 7:66 and the AU and UA alternatives at position 3:70. The only bases that are absolutely conserved among the five $tRNA^{Ser}$ isoacceptors, but never present in other two tRNAs with a long extra arm, which are presumably the best candidates for being misacylated by serine, are D20, G20B, G47C and G73. The *in vivo* identity switch experiments of Normanly *et al*. (50) have shown the importance of the discriminator base and bases from the first three pairs of the acceptor stem for serine identity. This was further emphasized by

recent work on on $tRNA^{Ser}$ minihelices (53) which revealed that the relationship between acceptor stem recognition and aminoacylation specificity strongly depends on the content and accessibility of information presented by the base pairs of A-form RNA helices (67). Based on previous $tRNA^{Ser}$ footprinting (51) and X-ray crystallographic studies of the $tRNA^{Ser}$:SerRS complex (33) it was clear that SerRS primarily binds tRNA from its variable loop side. This orients SerRS toward the major rather than the minor groove of the acceptor stem helix. Regardless of the fact that the long variable arm of $tRNA^{Ser}$ shows variation in both length and sequence within isoacceptors, it makes the largest contribution to the specificity of serylation, as noted by different experimental approaches (52,64,68,69). The most interesting observation is that the mechanism of recognition of this tRNA domain varies in different species. In Bacteria, it is recognized stem-length specifically, but not sequence specifically, which is the consequence of the conserved stem pairing pattern of the variable arm within the $tRNA^{Ser}$ isoacceptors. Each member of the class 2 tRNA family has an unique orientation of the long variable arm. There are a few unpaired nucleotides between the possible stem of the variable arm and the base at position 48 (before the TψC stem starts): none in $tRNA^{Ser}$, one in $tRNA^{Leu}$ and two in $tRNA^{Tyr}$. While in the Leu system deletion of 1 bp caused only a small decrease in aminoacylation efficiency (70), the interaction with SerRS is impaired by shortening of the extra arm stem length in the cognate tRNA. There is direct interaction between the lower part of the stem and SerRS, which can be indirectly disturbed by a conformational change upon base pair deletion (64). Thus, regardless of the fact that the long variable arms of different $tRNA^{Ser}$ isoacceptors in Bacteria are not similar to each other, their orientation is preserved due to conservation of other structural features, as are some of the D-arm nucleotides. Each of the three type 2 tRNA species has a characteristic D-loop sequence following the invariant bases G18 and G19. Bases 20A and 20B are inserted into the D-loop in prokaryotic $tRNA^{Ser}$ and both play novel roles in tertiary interactions in the core of the tRNA. In particular, base 20B (which is absent in all eukaryotic cytoplasmic $tRNAs^{Ser}$, but present in archaeal serine isoacceptors) is stacked against the first base pair of the long variable arm and thus defines the direction of the latter. As shown by Himeno *et al.* (48), a base change at position 20B had a large effect on the specificity conversion from Tyr to Ser. Since neither the sequence in the D-loop, the stem-pairing pattern of the variable arm, the tertiary base pair 15:48, nor the nucleotide at position 59 in the TψC-loop are involved in base-specific recognition by the synthetase, it was concluded that SerRS selectively recognizes $tRNA^{Ser}$ on the basis of its characteristic tertiary structure (64). Thus, the tertiary structural variations among the type 2 tRNAs enable precise discrimination by their respective synthetases.

Recognition of $tRNA^{Ser}$ in eukaryotes reflects changes in the type 2 tRNA subset

The structure of $tRNA^{Ser}$ appears to vary extensively from species to species (21) which raises the question as to how the mode of tRNA discrimination has changed during evolution. There are significant differences between $tRNA^{Ser}$ from yeast and from *E.coli*, notably in the acceptor stem sequences and the determinants that affect the orientation of the long variable arm. This could explain the low cross-acylation of yeast $tRNA^{Ser}$ with *E.coli* SerRS. On the other hand, it has been demonstrated that yeast SerRS expressed in *E.coli* recognizes bacterial serine-specific tRNAs *in vivo* (55). Footprinting experiments (54) showed that, as in *E.coli*, the long variable arm of yeast $tRNA^{Ser}$ makes contact with the synthetase, though protection of the acceptor stem was not observed. This is in agreement with *in vitro* experiments performed with yeast $tRNA^{Ser}$ transcripts (57,58) which found the discriminator base to be unimportant for SerRS recognition; instead, its role in discrimination against misacylation by other synthetases, predominantly LeuRS, has been proposed. In contrast to the G2:C71 base pair which is conserved in *E.coli*, most *Saccharomyces cerevisiae* $tRNA^{Ser}$ isoacceptors contain a G:U wobble pair at position 2:71, which could be involved in different types of interactions with the cognate synthetase than observed in bacterial systems. In *Schizosaccharomyces pombe* $tRNA^{Ser}$ there is either a U:A or A:U base pair at the same position. The *S.pombe* $tRNA^{Ser}$ isoacceptor with A2:U71 is efficiently aminoacylated by *S.cerevisiae* SerRS, while the *E.coli* enzyme failed to recognize it *in vivo* (56). Footprinting experiments also revealed a strong protection of the upper part of the anticodon stem (54), which together with the absolute conservation of the A27:U43 base pair in *S.cerevisiae* $tRNA^{Ser}$ isoacceptors, suggests its involvement in the recognition process. It has not yet been experimentally proved that the interaction with the cognate synthetase requires the accessibility of particular chemical groups of the A27:U43 base pair, but we have previously shown that all the heterologous but cognate tRNAs that were efficiently recognized by *S.cerevisiae* SerRS, both *in vivo* and *in vitro*, possess this particular base pair. Another conserved nucleotide found in a pool of yeast and other eukaryotic serine specific tRNAs is U44. It is unpaired, located at the beginning of the extra arm and may influence the geometry of this domain and the interaction with cognate synthetases. On the other hand, different base pairing in the variable arm of prokaryotic and eukaryotic $tRNAs^{Ser}$, together with the absence of the nucleotide at position 20B in eukaryotes, may cause the formation of sufficiently different tertiary structures in eukaryotic serine-specific tRNAs, which cannot be recognized by bacterial SerRS enzymes. Recent work (48,57) suggests that in contrast to *E.coli*, where tertiary structural elements play a key role in discriminating from other class II tRNAs, such discrimination in yeast is more sequence dependent and less tertiary structure dependent. While in *E.coli*, every type 2 tRNA has a different number of unpaired nucleotides between the TψC-stem and the first base pair of the variable arm, together with a different tertiary 15:48 base pair at the base of this arm in $tRNA^{Ser}$ and $tRNA^{Leu}$, in *S.cerevisiae* $tRNA^{Leu}$ and $tRNA^{Ser}$ share the same number of unpaired nucleotides at the base of the long variable arm and the G15:C48 tertiary base pair. The simpler, i.e. less exclusive, recognition by yeast SerRS could be due to less constrained recognition by eukaryotic enzymes specific for type 2 tRNAs; they must reject only one kind of long-variable-arm tRNA ($tRNA^{Leu}$), while their bacterial counterparts have to reject two ($tRNA^{Leu}$ and $tRNA^{Tyr}$). Thus, evolutionary adaptation toward less stringent recognition of cognate tRNAs may be the consequence of tRNA type switching by $tRNAs^{Tyr}$. Comparison of higher eukaryotic cytoplasmic serine isoaccepting tRNAs, both from animals and plants, revealed the conservation of certain structural elements that might influence the orientation of the long variable arm (20). This domain in human $tRNA^{Ser}$ and $tRNA^{Sec}$ functions as the major identity element in an orientation-dependent but not sequence-specific manner (60). The orientation can also affect the interaction of the

bound synthetase with other identity elements. Some of the cytoplasmic leucine tRNAs in eukaryotes also contain a long extra arm in an orientation similar to that of $tRNA^{Ser}$, although their sequences are different from those of $tRNA^{Sec}$ and $tRNA^{Ser}$. It seems plausible that an interaction can occur between the extra arm of leucine tRNA and seryl-tRNA synthetase, since there is no requirement of sequence specificity in the extra arm for SerRS recognition. Therefore it seems that at least some elements of $tRNA^{Ser}$ tertiary structure, possibly different to those employed in eubacteria, are also important for recognition in eukaryotes. In agreement with this suggestion it has been observed that human amber suppressor $tRNA^{Ser}$ functions *in vivo* with the cognate enzyme from yeast, but not with that from *E.coli* (56). In contrast, recent experiments showed that *supF*, an *E.coli* amber supressor $tRNA^{Tyr}$ carrying the G73 mutation, fails to be misacylated with serine by yeast SerRS, but is a substrate for the *E.coli* enzyme, probably due to the greater similarity between tRNA determinants in the two bacterial tRNAs. *In vitro* experiments (59,71), revealed that exchange of the discriminator base A73 for G is alone sufficient to convert human $tRNA^{Leu}$ into a serine acceptor *in vitro*. Since G73 is the only requirement for the identity swap between leucine and serine accepting tRNAs, it is probably involved in a direct interaction with SerRS. In Bacteria (48,70) and lower eukaryotes (48,57) this base serves only as an antideterminant. Taken together, the body of work on $tRNA^{Ser}$ identity suggests that the mechanisms by which seryl-tRNA synthetases recognize the orientation of the variable arm as well as the acceptor stem have diverged during evolution. Such differences in the contributions of the various recognition elements to the specificity of aminoacylation and the mechanism of tRNA interaction with the cognate synthetases in various organisms have also been documented for several type 1 tRNAs (37–39). Perhaps the most notable gap in this area is the lack of any comparable data for archaeal $tRNA^{Ser}$ recognition.

Recognition of selenocysteine tRNAs

Selenocysteine-inserting tRNA species can be found in all three kingdoms: they were originally discovered in Bacteria, but during more recent surveys of the various branches of the Archaea and Eukarya, $tRNA^{Sec}$ genes have been detected in virtually all organisms examined (47). The discovery of selenocysteine tRNAs in the archaeal lineage (72) indicates that the principles of selenocysteine biosynthesis and co-translational insertion may have been established at a time before the divergence of the three lineages. The selenocysteine-inserting tRNAs are the longest tRNAs known to date, due to the length of their extra arms and their extended 8 bp acceptor stems. In addition, their primary structure deviates from the canonical tRNA consensus at several positions. Despite these differences, the three-dimensional model of the solution structure of $tRNA^{Sec}$ is similar to that of canonical tRNAs, with tertiary structure stabilization achieved by a set of novel tertiary interactions (73). These structural peculiarities provide the basis for the specialized biochemical functions of $tRNA^{Sec}$. Although the $tRNA^{Sec}$ species are charged with serine by seryl-tRNA synthetase (74) charging efficiencies were found to be 100-fold reduced compared with that of a canonical serine isoacceptor. The factors that may contribute to the low charging efficiency in *E.coli* are: the unusually long acceptor stem (8 bp), the deviation from the identity set of serine isoacceptors at position 11:24, and the special conformational features imposed by the unconventional mode of stacking the D–T loop region.

The primary structures of $tRNAs^{Sec}$ from Bacteria, Archaea and Eukarya bear little sequence similarity, but the crucial features of their structure necessary for function are highly conserved. Thus, the divergence of the selenocysteine systems within the two lineages is not greater than that observed for other components of the translational apparatus. Although the long variable arm of $tRNA^{Sec}$ reiterates the major recognition site for SerRS (60), the length of the acceptor stem is a major determinant for discrimination between $tRNA^{Ser}$ isoacceptors and $tRNA^{Sec}$ with respect to several other proteins which interact specifically with Sec-$tRNA^{Sec}$ such as selenocysteine synthase, and the selenocysteine-specific elongation factor SELB.

DIVERSITY AND DIVERGENCE OF SERYL-tRNA SYNTHETASES

All aaRSs contain two major functional domains (17,75): the class-defining catalytic core and a second region of highly variable sequence and structure. To a rough approximation, the two domains correspond to the two domains of the L-shaped tRNA molecules they recognize. The class-defining domain, and polypeptide insertions within it, make contact with the acceptor-TψC helix and the 3′ end of the tRNA, while the second domain interacts either with the anticodon-containing region (7,31,32, 76,77) or with the variable arm of the tRNA (33). Given the evolutionary changes in patterns of $tRNA^{Ser}$ recognition, it is of interest to look for corresponding evolutionary changes in seryl-tRNA synthetases. In eubacteria, SerRS uses a coiled-coil domain to bind the extra arm of $tRNA^{Ser}$ (33,61,62). Sequence alignments and modeling studies (78) indicate that this N-terminal antiparallel coiled-coil is apparently conserved throughout evolution as the major tRNA binding domain, in agreement with the observation that all $tRNA^{Ser}$ isoacceptors, except those in animal mitochondria, contain a long variable arm. Furthermore, shortening of the extra arm in eukaryotic $tRNA^{Ser}$ influences charging by the cognate synthetases (57,58,64). SerRS also interacts with the major groove of the A-form helix in the tRNA acceptor stem (67), in *E.coli* predominantly with the first 5 bp (53) and in the human system probably with the discriminator base G73 (20). SerRS binds $tRNA^{Ser}$ across the two subunits of the dimer. In this way the interaction of the variable arm with the N-terminal coiled-coil domain of one subunit serves to position the 3′ end of the same tRNA molecule into the active site on the other subunit. The loop inserted into motif 2 of SerRS is responsible for the specificity of the interaction in the major groove. This polypeptide is rather conserved in all seryl-tRNA synthetases. The *T.thermophilus* $tRNA^{Ser}$-SerRS co-crystal structure (79) shows that Ser 261 and Phe 262 make specific contacts with the first two base pairs of the $tRNA^{Ser}$ acceptor stem. The hydroxyl group of Ser 261 is involved in hydrogen bonding with the carbonyl oxygens of nucleotides G1 and G2. Phe 262 provides a hydrogen bond acceptor for interaction with the exocyclic amino group of C71, while the aromatic ring of the same residue participates in a hydrophobic interaction with its C5(H). The purine N7 and pyrimidine C5(H) functional groups on the outside of the major groove provide a basis for SerRS to distinguish between the RY and YR base pair combinations. A similar interaction is expected in the *E.coli* cognate complex between Y69 and the homologous Tyr 274. The correct positioning of the 3′ end of tRNA in the

active site of *T.thermophilus* SerRS causes a conformational switch in this protein domain (79). Recent results from our laboratory strongly suggest that yeast SerRS structurally and functionally resembles its prokaryotic counterparts in the active site. In addition to the involvement of conserved amino acids in the motif 2 loop of yeast SerRS in ATP binding, as seen in other class II synthetases, we have shown that mutations of several other residues belonging to this domain affect tRNA-dependent amino acid recognition (78). It is not known yet whether this phenomenon is a general characteristic of all seryl-tRNA synthetases, or a specific feature of the yeast enzyme. Residues Ser 261 and Phe 262 are positioned between two absolutely conserved glycines in the motif 2 loop of *T.thermophilus* SerRS. In all seryl-tRNA synthetases, except those of archaeal and organellar origin, the position which corresponds to 261 in *T.thermophilus* is occupied by serine, while the phenylalanine at the adjacent position is not conserved. All eukaryotic cytoplasmic SerRS enzymes of known primary structure contain a histidine at the comparable position, which is not expected to participate in hydrophobic interactions, although the purine–pyrimidine base pair is conserved at position 4:69 in many eukaryotic $tRNA^{Ser}$ isoacceptors, including *S.cerevisiae*, *S.pombe* and human. Furthermore, there is a G:U base pair at position 2:71 in yeast $tRNAs^{Ser}$, lacking the relevant exocyclic amino group for participation in hydrogen bonding, while there is a U:A pair in human $tRNAs^{Ser}$. The striking primary sequence similarity (72%) between the *S.cerevisiae* (P07284) and *S.pombe* (EMBLZ97210) seryl-tRNA synthetases, together with the well conserved putative $tRNA^{Ser}$ identity elements in the two yeasts, are in agreement with our experimental finding of efficient cross-charging between the two species (56). Consequently, based on the available $tRNA^{Ser}$ and SerRS sequences, eukaryotic seryl-tRNA synthetases probably employ a different mechanism of $tRNA^{Ser}$ acceptor stem recognition than their prokaryotic counterparts. Eukaryotic seryl-tRNA synthetases also differ in length from their bacterial counterparts since they contain basic C-terminal extensions between 18 and 48 amino acids in length. In yeast, this C-terminal extension affects both the stability of the enzyme and its interaction with various substrates (80). The existence of either N- or C-terminal extensions is characteristic of a great many eukaryotic aaRSs.

EVIDENCE FOR PARAPHYLETIC ORIGIN OF SerRS FROM THE PRESENT-DAY ARCHAEA

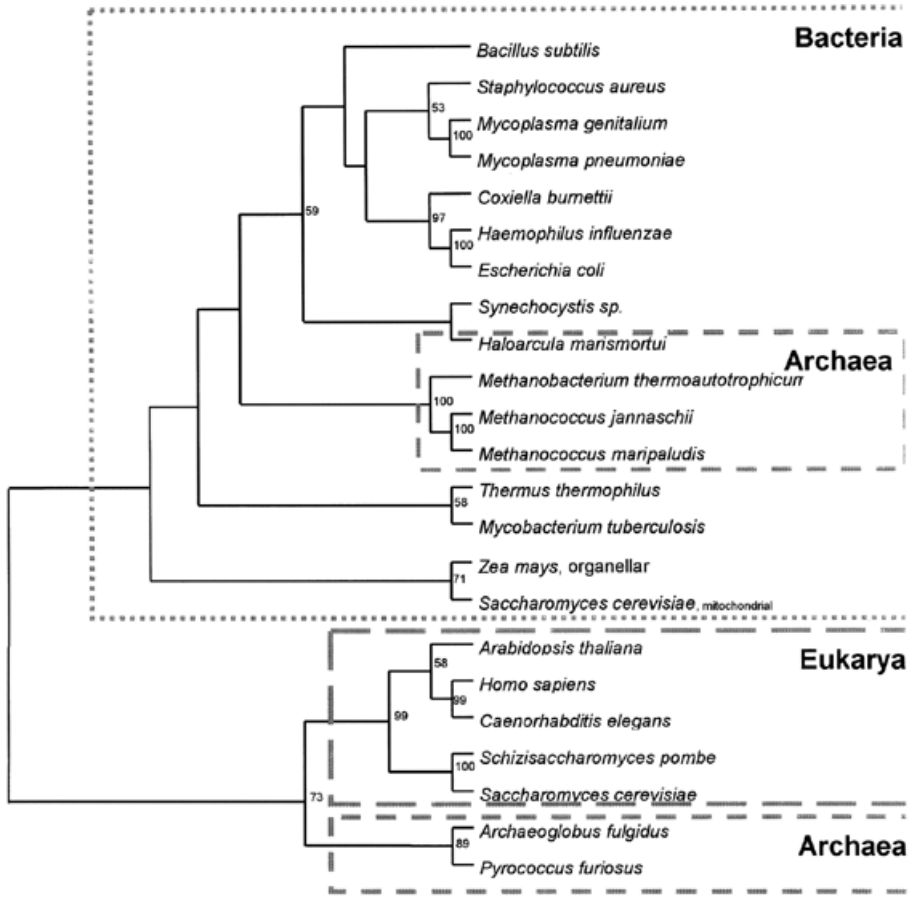
The first archaeal SerRS sequenced, isolated and characterized was that of the extreme halophile *Haloarcula marismortui* (81). As noted earlier (78) preliminary phylogenetic analyses showed this enzyme to be most similar to its counterparts from Gram-positive Bacteria, and not to eukaryotic SerRSs as expected from the universal tree of life (12). Taupin *et al*. (81) also showed that *H.marismortui* SerRS serylates $tRNA^{Ser}$ from *E.coli*, but not from yeast, which is in accord with both the structure of archaeal $tRNA^{Ser}$ and the observed phylogenetic position of the enzyme.

The appearance of five other archaeal SerRSs in sequence data banks, from *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Methanococcus maripaludis*, *Methanobacterium thermoautotrophicum* and, most recently, *Pyrococcus furiosus*, has now allowed a more stringent and reliable phylogenetic analysis of SerRS evolution. Inspection of preliminary multiple sequence alignments led to the conclusion that (i) parts of the alignment corresponding to the most variable parts of the sequence, the N-terminal coiled coil and the C-terminal extensions, are too ambiguous and should be excluded from the alignment; and (ii) due to the ambiguous positions of gaps in the alignment which are the result of intervening stretches of amino acids found in some sequences (especially those from the methanogens *M.jannaschii*, *M.maripaludis* and *M.thermoautotrophicum*) the positions with gaps were also excluded from the analysis. Under either maximum parsimony (Fig. 1A), neighbor-joining with correction for multiple substitutions invoked (Fig. 1B), and maximum likehood (Fig. 1C) methods the SerRSs from *A.fulgidus* and *P.furiosus* formed a clade with their eukaryotic counterparts, as expected. However, the SerRSs of *H.marismortui* and the three methanogens clustered with the bacterial enzymes 100% of the bootstrap replicates (82) (at least using the programs of the PHYLIP package, where this information is available), even though mere inspection of the sequence reveals that, while the SerRSs of methanogens have numerous 'loops' not found in their other counterparts, the *H.marismortui* SerRS has a 'normal' appearance. These sequences clustered most closely with SerRS of Gram-positive *B.subtilis* and the cyanobacterium *Synechocystis* sp. (from independent sources, Gram-positive bacteria and cyanobacteria are known to be sister taxa) although a reliable inference of their mutual positions in that part of the tree could not have been established by either method. These results indicate the paraphyletic origin of *serS* genes in today's Archaea. According to the present-day classification of Archaea, all six organisms (*A.fulgidus*, *P.furiosus*, *H.marismortui* and the three methanogens) belong to the group Euryarchaeota; the halophiles and the methanogens share a more recent common ancestor than the remaining two organisms. What seems most plausible is to postulate that this common ancestor of *H.marismortui* and the methanogens acquired its *serS* gene by horizontal (lateral) transfer from an eubacterial donor. It is possible that the original archaeal and the laterally transferred *serS* gene co-existed for a certain time in the course of evolution before the original gene was lost. The newly acquired enzyme was probably instantly functional, i.e. it serylated the host's $tRNA^{Ser}$ immediately after its acquisition, and after the transfer it evolved so as to optimize its function under the extreme conditions, the most obvious feature being an increased proportion of aspartate and glutamate residues in the sequence of the extreme halophile *H.marismortui* SerRS (for explanation see e.g. 81). The function of numerous 'appendages' acquired by the SerRS from methanogens is unknown.

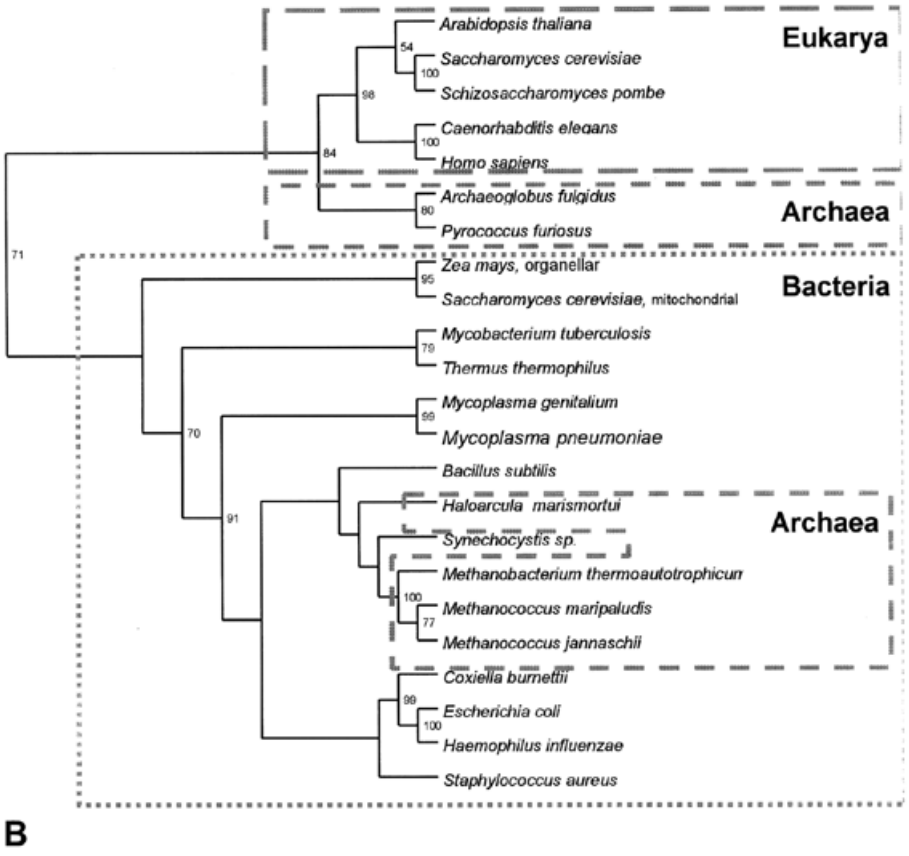
The proposed evolutionary event for the archaeal SerRS genes is not the sole example of such an event among genes coding for aaRSs. Other scenarios concerning the 'late' evolutionary acquisition of aaRSs have been proposed for GlnRS and GluRS, where horizontal genetic transfer of the GlnRS gene from eukaryotes to Bacteria has been postulated (12,83).

SERYLATION IN ORGANELLES

In agreement with the bacterial origin of organelles, components of the organellar translational apparatus are structurally most like their prokaryotic counterparts. All serine-specific tRNAs, except those in animal mitochondria, possess a long variable arm comprising usually more than 14 nucleotides, with no unpaired nucleotides at the base of the extra arm. Since these tRNAs normally contain both 20A and 20B nucleotides, these positions



A



B

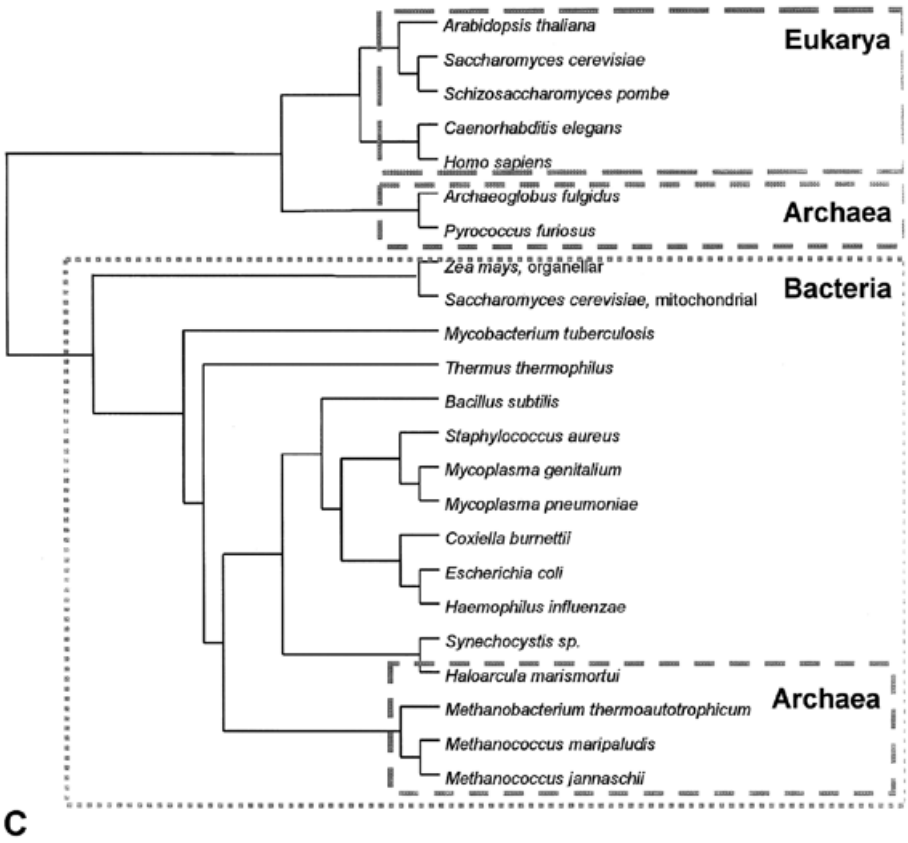


Figure 1. Results of phylogenetic inferences based on the sequences of seryl-tRNA synthetases from different organisms. The alignment included 23 different sequences over their globular domains performed with Clustal_X (94). The gaps were excluded from the analysis in order to be able to include the SerRS sequences of methanogens, which contain numerous extra loops in comparison with other sequences. The organisms, with total length (in amino acids) of SerRS given in parentheses, are as follows: *Arabidopsis thaliana* (451), *A.fulgidus* (446), *Bacillus subtilis* (425), *Caenorhabditis elegans* (487), *Coxiella burnetii* (423), *E.coli* (430), *Haemophilus influenzae* (429), *H.marismortui* (460), *Homo sapiens* (514), *M.thermoautotrophicum* (513), *M.jannaschii* (521), *M.maripaludis* (514), *Mycobacterium tuberculosis* (419), *Mycoplasma genitalium* (417), *Mycoplasma pneumoniae* (420), *P.furiosus* (455), *S.cerevisiae* (462), *S.cerevisiae*, putative mitochondrial (446) *S.pombe* (450), *Staphylococcus aureus* (428), *Synechocystis* sp. (430), *T.thermophilus* (421), *Zea mays*, organellar (489). (**A**) Most parsimonious unrooted tree constructed from a bootstrap analysis of the SerRS alignment (100 data sets). The programs used were SEQBOOT, PROTPARS and CONSENSE of the PHYLIP package (95). Numbers at the branches correspond to percentage bootstrap frequencies for each branch (82). Only values >50 are shown. (**B**) Unrooted neighbor-joining tree constructed from a bootstrap analysis of the SerRS alignment using Clustal_X with correction for multiple substitution (1000 replicates; percentages >50% are shown in the figure). (**C**) Maximum likehood tree constructed using PROTML (96); the tree was obtained by star decomposition followed by evaluation of 16 user trees constructed by permuting sequence positions at the most uncertain nodes.

may influence the orientation of the variable arm as in eubacteria. The results of phylogenetic analysis of $tRNA^{Ser}$ sequences (B.Lenhard, unpublished results) suggest an especially interesting branching pattern for the organellar $tRNAs^{Ser}$: while mitochondrial $tRNA^{Ser}$ of fungi and animals cluster together, plant organellar $tRNAs^{Ser}$ form a separate clade that includes both mitochondrial and chloroplast sequences, as if there was selection in favor of common identitiy elements. It will be of great interest to find out the interrelation of the corresponding seryl-tRNA synthetases, or even to find out whether two separate organellar enzymes exist.

The primary structures of only two organellar seryl-tRNA synthetases are known thus far: one is from yeast mitochondria (P38705) and the other belongs to maize organelles, and is probably also mitochondrial (84). Phylogenetic analyses revealed that these organellar synthetases cluster together, and are phylogenetically closer to the bacterial SerRS enzymes. The structural resemblance between organellar seryl-tRNA synthetases and their bacterial counterparts is also evident from the lack of any C-terminal extension, which characterizes all cytoplasmic eukaryotic seryl-tRNA synthetases. We have recently shown the complementation of an *E.coli serS* mutant strain with the gene encoding maize organellar SerRS. Furthermore, its mature protein product overexpressed in *E.coli* efficiently aminoacylated bacterial $tRNA^{Ser}$ *in vitro*, while yeast tRNA was a poor substrate (84).

Serylation in yeast mitochondria is especially intriguing, since the organelle contains three $tRNA^{Ser}$ isoacceptors, which differ considerably in primary structure (85,86). $tRNA_1^{Ser}$ has only 39 bp in common with $tRNA_2^{Ser}$. $tRNA_2^{Ser}$ and $tRNA_3^{Ser}$ are encoded by the same gene, but the mature tRNAs differ in their modification pattern. Phylogenetic analyses of the relation of $tRNA_1^{Ser}$ to the other two isoacceptors tends to yield ambiguous results (B.Lenhard, unpublished observation), although both genes are also known to exist in some other fungi (based upon inspection of tRNA sequences). Both $tRNA_2^{Ser}$ and $tRNA_3^{Ser}$

should recognize all four UCN codons. As demonstrated earlier (87), only these two isoacceptors (UCN) were aminoacylated by *E.coli* SerRS, while $tRNA_1^{Ser}$ (AGY) was not. Both types of mitochondrial $tRNA^{Ser}$ isoacceptors contain a long variable arm, whose base pairing pattern and orientation probably differ due to a very different primary structure in the arm *per se* and in the D-loop of the tRNAs. It would be interesting to learn what the mechanism is of recognition of such diverse tRNA substrates with mitochondrial SerRS, which is, according to sequence alignment and modeling studies, rather similar to its bacterial counterparts. Such a unilateral aminoacylation specificity also exists between bovine mitochondria and eubacteria. It has been shown that e.g. bovine mitochondrial SerRS charges cognate *E.coli* tRNA species and misacylates non-cognate *E.coli* species, whereas their bacterial counterparts do not efficiently charge cognate mitochondrial tRNAs (88). The latter is probably due to a very specific $tRNA^{Ser}$ structure in these organelles: unlike all other serine-specific tRNAs, only those from animal mitochondria do not possess a long variable arm, and in some cases (those for the AGY codons) the D-stem/loop is missing (22). The T-loop of this truncated $tRNA^{Ser}$ is the main recognition site for the mitochondrial SerRS (89). Serylation in animal mitochondria has recently gained much attention, since the necessity for evolutionary adaptation may accordingly produce seryl-tRNA synthetases which differ substantially from all other SerRS enzymes, especially in the N-terminal domain.

IDENTITY SWITCHES AND TYPE-CONSTRAINED MISACYLATION AMONG TYPE 2 tRNAs

Since the long variable arm is an important recognition element for three aminoacyl-tRNA synthetases in *E.coli* (LeuRS, SerRS and TyrRS), it is apparent that these enzymes have to recognize particular sequences and/or orientations of this domain, as well as the additional determinants in their cognate tRNAs in order to avoid misacylation. The complete specificity change of $tRNA^{Tyr}$ to $tRNA^{Ser}$ *in vitro* was facilitated by insertion of three nucleotides into the variable arm (stem) of $tRNA^{Tyr}$ plus two nucleotide changes at positions 9 and 73. Both G73 in $tRNA^{Ser}$ and A73 in $tRNA^{Leu}$ and $tRNA^{Tyr}$ are phylogenetically conserved, implying an important role for this base during aminoacylation of all three tRNAs (70). In this context it is not surprising that a G73 mutant in the *E.coli* $tRNA^{Tyr}$ amber suppressor, required no other structural changes in order to be misacylated with serine *in vivo* after increasing the level of bacterial SerRS (M.Nalaskowska and I.Weygand-Durasevic, unpublished results). Interestingly, similar misacylation was not observed upon overexpression of yeast SerRS in *E.coli*. While bacterial suppressor $tRNA^{Tyr}$ inserts only tyrosine at amber codons in *E.coli*, this tRNA inserts leucine in *S.cerevisiae* (insertion of serine is probably prevented by an A73 antideterminant). Thus, the *E.coli* tyrosine tRNA is functionally a leucine tRNA in yeast cytoplasm, indicating that in evolution $tRNA^{Tyr}$ is more closely related to a tRNA of different acceptor specificity, but of the same class type, than to one with the same amino acid specificity, but of a different class type. An obvious explanation would be that, after $tRNA^{Tyr}$ switched from type 2 to type 1, eukaryotic (and possibly archaeal) LeuRS lost its ability to discriminate against $tRNA^{Tyr}$ of type 2. However, since the determinants for identity between the *E.coli* and yeast mitochondrial tyrosine tRNAs are conserved (i.e. yeast mitochondrial $tRNAs^{Tyr}$ are bacteria-like, comprising a long extra arm), *E.coli* tyrosyl-tRNA synthetase can substitute for yeast mitochondrial enzyme function *in vivo* (90). The ability of overexpressed GlnRS to misacylate variants of both *E.coli* $tRNA^{Tyr}$ and $tRNA^{Ser}$ amber suppressors (91) suggests that the long variable arm of type 2 tRNAs does not necessarily prevent misrecognition by synthetases specific for type 1 tRNAs, in the presence of other overlapping determinants (amber anticodon and/or altered acceptor stem).

CONCLUSIONS

The tRNAs with long extra arms are likely to have originated with the closure of the genetic code. One explanation for their role was that they allowed for specific recognition by the corresponding aaRSs without resorting to the anticodon (i.e. Ser and Leu have six anticodons each). However, the ancient existence of a long extra arm in bacterial $tRNA^{Tyr}$ requires another explanation. One imaginable scenario might reside in the fact that TyrRS and TrpRS are structural isomers which separated much later than other aaRSs (92,93), and that tRNA might have played a more active role in establishing the amino acid specificity of these enzymes. The later evolutionary divergence of the two systems would then allow the loss of the extra arm in $tRNA^{Tyr}$ of Eukarya.

The archaeal systems are especially interesting, with tRNAs in general being more bacteria-like than their corresponding aaRS. The case of SerRS in methanogens and halophiles mentioned above is a noteworthy exception, suggestive of early events in evolution. Furthermore, it is likely that co-evolution of aaRSs and their cognate tRNAs might account for some of their idiosyncratic features that cannot be otherwise explained.

ACKNOWLEDGEMENTS

This work was supported by grants from National Institutes of Health, NIH/FIRCA, International Centre for Genetic Engineering and Biotechnology (Trieste, Italy), Fondecyt (Chile) and Ministry of Science and Technology (Croatia).

REFERENCES

1 Cusack,S. (1993) *Biochimie*, **75**, 1077–1081.
2 Carter,C.W. (1993) *Annu. Rev. Biochem.*, **62**, 715–746.
3 Nagel,G.M. and Doolittle,R.F. (1995) *J. Mol. Evol.*, **40**, 487–498.
4 Arnez,G.J. and Moras,D. (1997) *Trends Biochem. Sci.*, **22**, 211–216.
5 Leveque,F., Plateau,P., Dessen,P. and Blanquet,S. (1990) *Nucleic Acids Res.*, **18**, 305–312.
6 Eriani,G., Delarue,M., Poch,O., Gangloff,J. and Moras,D. (1990) *Nature*, **347**, 203–206.
7 Rould,M.A., Perona,J.J., Söll,D. and Steitz,T.A. (1989) *Science*, **246**, 1135–1142.
8 Cavarelli,J., Rees,B., Ruff,M., Thierry,J.-C. and Moras,D. (1993) *Nature*, **362**, 181–184.
9 Cavarelli,J., Eriani,G., Rees,B., Ruff,M., Boeglin,M., Mitschler,A., Martin,F., Gangloff,J., Thierry,J.-C. and Moras,D. (1994) *EMBO J.*, **13**, 327–337.
10 Cusack,S. (1995) *Nat. Struct. Biol.*, **2**, 824–831.
11 Nagel,G.M. and Doolittle,L.F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 8121–8125.
12 Brown,J.R. and Doolittle,W.F. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 2441–2445.
13 Ibba,M., Bono,J.L., Rosa,P.A. and Söll,D. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 14383–14388.
14 Ibba,M., Morgan,S., Curnow,A.W., Pridmore,D.R., Vothknecht,U.C., Gardner,W., Lin,W., Woese,C.R. and Söll,D. (1997) *Science*, **278**, 1119–1122.
15 Eigen,M., Lindemann,B.F., Tietze,M., Winkler-Oswatitsch,R., Dress,A. and von Haeseler,A. (1989) *Science*, **244**, 673–679.
16 Jukes,T.H. (1995) *J. Mol. Evol.*, **40**, 537–540.
17 Schimmel,P., Giege,R., Moras,D. and Yokoyama,S. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 8763–8768.

18 Saks,M.E., Sampson,J.R. and Abelson,J. (1998) *Science*, **279**, 1665–1670.
19 Sprinzl,M., Dank,N., Nock,S. and Schön,A. (1991) *Nucleic Acids Res.*, **19** (suppl), 2127–2171.
20 Achsel,T. and Gross,H.J. (1993) *EMBO J.*, **12**, 3333–3338.
21 Sprinzl,M., Horn,C., Brown,M., Ioudovitch,A. and Steinberg,S. (1998) *Nucleic Acids Res.*, **26**, 148–153.
22 Steinberg,S., Gautheret,D. and Cedergren,R.J. (1994) *J. Mol. Biol.*, **236**, 982–989.
23 Härtlein,M. and Cusack,S. (1995) *J. Mol. Evol.*, **40**, 519–530.
24 Pallanck,L. and Schulman,L.H. (1992) In Hatfield,D., Lee,B.J. and Pirtle,R. (eds), *Transfer RNA in Protein Synthesis*. CRC, Boca Raton, FL, pp. 273–318.
25 Giege,R., Puglisi,J.D. and Florentz,C. (1993) *Prog. Nucleic Acid Res. Mol. Biol.*, **45**, 129–206.
26 McClain,W.H. (1993) *J. Mol. Biol.*, **234**, 257–280.
27 McClain,W.H. (1993) *FASEB J.*, **7**, 72–78.
28 Saks,M.E., Sampson,J.R. and Abelson,J.N. (1994) *Science*, **263**, 191–197.
29 McClain,W.H. and Foss,K. (1988) *Science*, **241**, 1804–1807.
30 Saks,M.E. and Sampson,J.R. (1995) *J. Mol. Evol.*, **40**, 509–518.
31 Rould,M.A., Perona,J.J. and Steitz,T.A. (1991) *Nature*, **352**, 213–218.
32 Ruff,M., Krishnaswamy,S., Boeglin,M., Poterszman,A., Mitschler,A., Podjarny,A., Rees,B., Thierry,J.-C. and Moras,D. (1991) *Science*, **252**, 1682–1689.
33 Biou,V., Yaremchuk,A., Tukalo,M. and Cusack,S. (1994) *Science*, **263**, 1404–1410.
34 Nicholas,H.B.,Jr and McClain,W.H. (1995) *J. Mol. Evol.*, **40**, 482–486.
35 Hou,Y.-M. and Schimmel,P.R. (1989) *Biochemistry*, **28**, 4942–4947.
36 Trezeguet,V., Edwards,H. and Schimmel,P.R. (1991) *Mol. Cell. Biol.*, **11**, 2744–2751.
37 Nameki,N., Tamura,K., Himeno,H., Asahara,H., Hasegawa,T. and Shimizu,M. (1992) *Biochem. Biophys. Res. Commun.*, **189**, 856–862.
38 Nameki,N. (1995) *Nucleic Acids Res.*, **23**, 2831–2836.
39 Nameki,N., Asahara,M., Shimizu,M., Okada,N. and Himeno,H. (1995) *Nucleic Acids Res.*, **23**, 389–394.
40 Nazarenko,I.A., Peterson,E.T., Zakharova,O.Z., Lavrik,O.I. and Uhlenbeck,O.C. (1992) *Nucleic Acids Res.*, **20**, 475–478.
41 Moor,N.A., Ankilova,V.N. and Lavrik,O.I. (1995) *Eur. J. Biochem.*, **234**, 897–902.
42 Swanson,R., Hoben,P., Sumner-Smith,M., Uemura,H., Watson,L. and Söll,D. (1988) *Science*, **242**, 1548–1551.
43 Rogers,M.J. and Söll,D. (1990) *Prog. Nucleic Acid Res. Mol. Biol.*, **39**, 185–208.
44 Schön,A., Kannangara,C.G., Gough,S.P. and Söll,D. (1988) *Nature*, **331**, 187–190.
45 Verkamp,E., Kumar,A.M., Lloyd,A., Martins,O., Stange-Thomann,N. and Söll,D. (1995) In Söll,D. and RajBhandary,U. (eds), *tRNA: Structure, Biosynthesis and Function*. American Society for Microbiology, Washington, DC, pp. 545–550.
46 Curnow,A.W., Ibba,M. and Söll,D. (1996) *Nature*, **382**, 589–590.
47 Baron,B. and Bock,B. (1995) In Söll,D. and RajBhandary,U. (eds), *tRNA: Structure, Biosynthesis and Function*. American Society for Microbiology, Washington, DC, pp. 529–544.
48 Himeno,H., Hasegawa,T., Ueda,T., Watanabe,K. and Shimizu,M. (1990) *Nucleic Acids Res.*, **18**, 6815–6819.
49 Normanly,J., Masson,J.-M., Kleina,L.G., Abelson,J.N. and Miller,J.H. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 6548–6552.
50 Normanly,J., Ollick,T. and Abelson,J.N. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 5680–5684.
51 Schatz,D., Leberman,R. and Eckstein,F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 6132–6136.
52 Sampson,J.R. and Saks,M.E. (1993) *Nucleic Acids Res.*, **21**, 4467–4475.
53 Saks,M.E. and Sampson,J.R. (1996) *EMBO J.*, **15**, 2843–2849.
54 Dock-Bregeon,A.-C., Garcia,A., Giege,R. and Moras,D. (1990) *Eur. J. Biochem.*, **188**, 283–290.
55 Weygand-Durasevic,I., Ban,N., Jahn,D. and Söll,D. (1993) *Eur. J. Biochem.*, **214**, 869–877.
56 Weygand-Durasevic,I., Nalaskowska,M. and Söll,D. (1994) *J. Bacteriol.*, **176**, 232–239.
57 Soma,A., Kumagai,R., Nishikawa,K. and Himeno,H. (1996) *J. Mol. Biol.*, **263**, 707–714.
58 Himeno,H., Yoshida,S., Soma,A. and Nishikawa,K. (1997) *J. Mol. Biol.*, **268**, 704–711.
59 Breitschopf,K. and Gross,H.J. (1994) *EMBO J.*, **13**, 3166–3169.
60 Wu,X.-Q. and Gross,H.J. (1993) *Nucleic Acids Res.*, **193**, 5589–5594.
61 Vincent,C., Borel,F., Willson,J.C., Leberman,R. and Härtlein,M. (1995) *Nucleic Acids Res.*, **23**, 1113–1118.
62 Borel,F., Vincent,C., Leberman,R. and Härtlein,M. (1994) *Nucleic Acids Res.*, **22**, 2963–2969.
63 Price,S., Cusack,S., Borel,F., Berthet-Colominas,C. and Leberman,R. (1993) *FEBS Lett.*, **324**, 167–170.
64 Asahara,H., Himeno,H., Tamuru,K., Nameki,N., Hasegawa,T. and Shimizu,M. (1994) *J. Mol. Biol.*, **236**, 738–748.
65 Normanly,J. and Abelson,J.N. (1989) *Annu. Rev. Biochem.*, **58**, 1029–1049.
66 Normanly,J., Kleina,L.G., Masson,J.-M., Abelson,J.N. and Miller,J.H. (1990) *J. Mol. Biol.*, **231**, 719–726.
67 Weeks,K.M. and Crothers,D.M. (1993) *Science*, **261**, 1574–1577.
68 Himeno,H., Hasegawa,T., Asahara,H., Tamura,K. and Shimizu,M. (1991) *Nucleic Acids Res.*, **19**, 6379–6382.
69 Asahara,H., Himeno,H. and Shimizu,M. (1991) *Chem. Lett.*, 363–366.
70 Asahara,H., Himeno,H., Tamura,K., Hasegawa,T., Watanabe,K. and Shimizu,M. (1993) *J. Mol. Biol.*, **231**, 219–229.
71 Breitschopf,K., Achsel,T., Busch,K. and Gross,H.J. (1995) *Nucleic Acids Res.*, **23**, 3633–3637.
72 Bult,C.J., White,O., Olsen,G.J., Zhou,L., Fleischmann,R.D., Sutton,G.G., Blake,J.A., FitzGerald,L.M., Clayton,R.A., Gocayne,J.D., Kerlavage,A.R., Dougherty,B.A., Tomb,J.F., Adams,M.D., Reich,C.I., Overbeek,R., Kirkness,E.F., Weinstock,K.G., Merrick,J.M., Glodek,A., Scott,J.L., Geoghagen,N.S.M. and Venter,J.C. (1996) *Science*, **273**, 1058–1073.
73 Baron,C., Westhof,E., Boeck,A. and Giege,R. (1993) *J. Mol. Biol.*, **231**, 274–292.
74 Leinfelder,W., Zehelein,E., Mandrand-Berthelot,M.A. and Boeck,A. (1988) *Nature*, **331**, 723–725.
75 Buechter,D.D. and Schimmel,P. (1993) *CRC Crit. Rev. Biochem.*, **28**, 309–322.
76 Perona,J.J., Rould,M.A., Steitz,T.A., Risler,J.-L., Zelwer,C. and Brunie,S. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2903–2907.
77 Brunie,S., Zelwer,C. and Risler,J.-L. (1990) *J. Mol. Biol.*, **216**, 411–424.
78 Lenhard,B., Filipic,S., Landeka,I., Skrtic,I., Söll,D. and Weygand-Durasevic,I. (1997) *J. Biol. Chem.*, **272**, 1136–1141.
79 Cusack,S., Yaremchuk,A. and Tukalo,M. (1996) *EMBO J.*, **15**, 2834–2842.
80 Weygand-Durasevic,I., Lenhard,B., Filipic,S. and Söll,D. (1996) *J. Biol. Chem.*, **271**, 2455–2461.
81 Taupin,C.M.-J., Härtlein,M. and Leberman,R. (1977) *Eur. J. Biochem.*, **243**, 141–150.
82 Li,W.-H. (1997) *Molecular Evolution*. Sinauer Associates, Inc., Sunderland, MA.
83 Lamour,V., Quevillon,S., Diriong,S., N′Guyen,V.C., Lipinski,M. and Mirande,M. (1994) *Proc. Natl Acad. Sci. USA*, **91**, 8670–8674.
84 Rokov,J., Söll,D. and Weygand-Durasevic,I. (1998) *Plant Mol. Biol.*, **38**, 497–502.
85 Martin,N.C. and Rabinowitz,M. (1978) *Biochemistry*, **17**, 1628–1634.
86 Miller,J.H. and Albertini,A.M. (1983) *J. Mol. Biol.*, **164**, 59–71.
87 Baldacci,G., Falcone,C., Frontali,L., Macino,G. and Palleschi,C. (1976) *Genetics and Biogenesis of Chloroplasts and Mitochondria*. North-Holland Publishing Co., Amsterdam.
88 Kumazawa,Y., Yokogawa,T., Hasegawa,E., Miura,K. and Watanabe,K. (1989) *J. Biol. Chem.*, **264**, 13005–13011.
89 Ueda,T., Yotsumoto,Y., Ikeda,K. and Watanabe,K. (1992) *Nucleic Acids Res.*, **20**, 2217–2222.
90 Edwards,H. and Schimmel,P.R. (1987) *Cell*, **51**, 643–649.
91 Sherman,J.M., Rogers,M.J. and Söll,D. (1995) In Söll,D. and RajBhandary,U. (eds), *tRNA: Structure, Biosynthesis and Function*. American Society for Microbiology, Washington, DC.
92 Doublie,S., Bricogne,G., Gilmore,C. and Carter,C.W.J. (1995) *Structure*, **3**, 17–31.
93 Brown,J.R., Robb,F.T., Weiss,R. and Doolittle,W.F. (1997) *J. Mol. Evol.*, **45**, 9–16.
94 Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) *Nucleic Acids Res.*, **25**, 4876–4882.
95 Felsenstein,J. (1989) *Cladistics*, **5**, 164–166.
96 Adachi,J. and Hasegawa,M. (1994) *MOLPHY: Programs for Molecular Phylogenetics* v2.2. Available at ftp://sunmh.ism.ac.jp/pub/molphy/