The EMBO Journal vol.15 no.18 pp.5069–5076, 1996

Arginine aminoacylation identity is context-dependent and ensured by alternate recognition sets in the anticodon loop of accepting tRNA transcripts

Marie Sissler, Richard Giegé and Catherine Florentz[1]

UPR 9002 du CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 Strasbourg Cedex, France

[1]Corresponding author

Yeast arginyl-tRNA synthetase recognizes the non-modified wild-type transcripts derived from both yeast $tRNA^{Arg}$ and $tRNA^{Asp}$ with equal efficiency. It discriminates its cognate natural substrate, $tRNA^{Arg}$, from non-cognate $tRNA^{Asp}$ by a negative discrimination mechanism whereby a single methyl group acts as an anti-determinant. Considering these facts, recognition elements responsible for specific arginylation in yeast have been searched by studying the *in vitro* arginylation properties of a series of transcripts derived from yeast $tRNA^{Asp}$, considered as an arginine isoacceptor tRNA. In parallel, experiments on similar $tRNA^{Arg}$ transcripts were performed. Unexpectedly, in the $tRNA^{Arg}$ context, arginylation is basically linked to the presence of residue C35, whereas in the $tRNA^{Asp}$ context, it is deeply related to that of C36 and G37 but is insensitive to the nucleotide at position 35. Each of these nucleotides present in one host, is absent in the other host tRNA. Thus, arginine identity is dependent on two different specific recognition sets according to the tRNA framework investigated.

Keywords: arginine/aspartate/identity/recognition set/transfer RNA

Introduction

Specific aminoacylation of tRNAs by cognate aminoacyl-tRNA synthetases (aaRS) is linked to the presence of a set of specific positive signals on tRNAs, the recognition set which is perceived by amino acids of the cognate synthetase (e.g. Giegé *et al.*, 1993; McClain, 1993b; Saks *et al.*, 1994). This is complemented by negative signals hindering recognition by non-cognate synthetases (Muramatsu *et al.*, 1988; Pütz *et al.*, 1994). Searching positive signals by *in vitro* methods is based on synthesis of *in vitro* transcripts derived from the natural tRNA sequence and measurements of the kinetic parameters of their aminoacylation with the cognate synthetase. This approach has been widely used and has led to the establishment of a great number of recognition sets. These sets, also confirmed by the results of alternative *in vivo* approaches, are constituted by a limited number of nucleotides and are believed to be unique for a given tRNA–synthetase system in any particular cellular compartment (McClain, 1993a).

Nucleotides that are potential recognition elements are numerous and are spread all over the tRNA molecule. If one assumes that nucleotides conserved in all tRNAs do not participate in specificity, only ~16 among the 75–90 tRNA residues are not potential recognition elements. To limit the number of tRNA variants to be synthesized and tested individually for their aminoacylation properties in order to detect the recognition nucleotides, the choice of positions to be tested can be reduced by knowledge of several structural and functional features. In particular, as a paradigm, it is assumed that recognition elements have to be shared between isoacceptor species and thus, that they have to be present within the consensus sequence of these isoacceptors. Along these lines, any tRNA transcript efficiently aminoacylated by an aaRS can be considered as an isoacceptor species and the consensus sequence between this tRNA and the cognate tRNA should include the set of recognition elements. Potential recognition elements can also be searched by computer analysis of the tRNA sequences from a given organism, revealing unique features for individual identities (McClain and Nicholas, 1987; Nicholas and McClain, 1987). Although it is clear that these unique features may serve as signals for other functions of tRNAs, they may also contain the recognition elements involved in specific aminoacylation. Finally, since resolution of crystal structures of cognate tRNA–synthetase systems has shown, at least in two cases, that most recognition elements are in direct contact with the synthetase (Rould *et al.*, 1989, 1991; Cavarelli *et al.*, 1992, 1993), any nucleotide in contact with the enzyme, as revealed by footprinting experiments for non-crystallized complexes (reviewed in Giegé *et al.*, 1993), is a good candidate for a recognition element.

Here we describe the search of recognition elements for specific arginylation in yeast following the idea that specificity is linked to nucleotides common to tRNAs that are efficient substrates for the enzyme. We have already highlighted a relationship between arginine and aspartate systems in yeast. In particular it has been shown that the wild-type $tRNA^{Asp}$ transcript, deprived of modified bases, is not only efficiently aspartylated but also efficiently arginylated (Perret *et al.*, 1990a). Indeed, whereas wild-type natural $tRNA^{Asp}$ is 12 600-fold less efficiently arginylated than $tRNA^{Arg}$, the corresponding $tRNA^{Asp}$ transcript is only 30-fold less well arginylated. Recently, it has been shown that this cross-reaction can be hindered by a single methyl group at position 1 of guanosine 37 which protects the $tRNA^{Asp}$ transcript from being mischarged by ArgRS (Pütz *et al.*, 1994). These observations lead us to hypothesize that the set of recognition elements responsible for specific arginylation is present within the nucleotide sequence not only of $tRNA^{Arg}$ but also of $tRNA^{Asp}$. Since a large library of $tRNA^{Asp}$ variants was available (e.g. Pütz *et al.*, 1991, 1993; Perret *et al.*, 1992; Puglisi *et al.*, 1993), the search for arginine recognition elements was

Table I. Kinetic parameters for arginylation by yeast ArgRS of transcripts derived from yeast $tRNA^{Asp}$

Transcripts	k_{cat} (s^{-1})	K_m (nM)	k_{cat}/K_m (relative[a])	$\mathcal{L}^a$	$\mathcal{L}^b$
Argwt IIIb	0.68	550	1	1	
Aspwt	0.07	570	0.100	10	1
AspU73	0.19	82	1.95	0.5	0.05
*Asp***C32**	0.0082	77	0.0866	11.5	1.15
*Asp***U34**	0.083	64	1.05	0.95	0.095
*Asp***C35**	0.031	310	0.0820	12	1.2
*Asp***U36**	0.0006	1415	0.0003	2780	278
*Asp***A37**	0.0004	2560	0.0001	7230	723
*Asp***A38**	0.029	1275	0.0185	54	5.4

[a]Relative to wild-type $tRNA^{Arg}$ transcript.
[b]Relative to wild-type $tRNA^{Asp}$ transcript. Relative kinetic specificity constants are defined as $(k_{cat}/K_m)_{relative} = (k_{cat}/K_m)_{mutant}/(k_{cat}/K_m)_{wild\ type}$. $\mathcal{L}$-values are defined as inverses of $(k_{cat}/K_m)_{relative}$. $\mathcal{L}$-values for duplicates varied by <10%. For nomenclature of variants, see Figure 2.

started by studying the arginylation properties of the *in vitro* transcripts derived from the non-cognate $tRNA^{Asp}$. In parallel, $tRNA^{Arg}$ transcripts were studied for their arginylation properties.

Our data show that arginylation is deeply dependent on nucleotides from the anticodon loop. Unexpectedly and interestingly, two recognition sets were found according to the tRNA context investigated. These sets are different at the level of the nucleotide nature and position within the anticodon loop. This is the first example where two alternative recognition sets are found for the same tRNA identity. Altogether, these data show that: (i) a tRNA recognition set is not necessarily unique, (ii) non-conserved nucleotides within isoacceptor tRNA species can be recognition elements and (iii) the mechanisms of expression of identity can be dependent on the nucleic acid context. These conclusions enlarge the present view on the concept of tRNA identity for specific aminoacylation. Hypotheses concerning possible mechanisms of expression of alternate recognition sets are proposed and evolutionary implications presented.

Results and discussion

Preliminary observations

Previous studies in our laboratory have revealed a strong relationship between the aspartic acid and arginine systems in yeast (Perret *et al.*, 1990a) since a transcript derived from $tRNA^{Asp}$ is a substrate for ArgRS. Its efficiency in arginylation is only 30-fold lower than that of native $tRNA^{Arg}$. This relationship has been investigated further by the search of arginine identity elements in both tRNA frameworks. Preliminary studies were performed in order to investigate a possible role of post-transcriptional modifications in arginine identity. A transcript derived from yeast $tRNA^{Arg}$IIIb (see below) was synthesized and used as reference. The charging level of this wild-type $tRNA^{Arg}$ transcript by yeast ArgRS is 100% and its kinetic parameters (K_m = 550 nM and k_{cat} = 0.68 s^{-1}) lead to a loss of catalytic efficiency of only 3-fold as compared with that of aminoacylation of native $tRNA^{Arg}$ (K_m = 400 nM and k_{cat} = 1.44 s^{-1}). Thus, post-transcriptional modifications of nucleotides are not required for efficient arginylation of yeast $tRNA^{Arg}$.

Comparison of the arginylation efficiency of an aspartate transcript and an arginine transcript shows that the $tRNA^{Asp}$ transcript is 10-fold less efficiently arginylated than the $tRNA^{Arg}$ transcript (Table I). This means that most nucleotides of the arginine recognition set are present in the $tRNA^{Asp}$ sequence and this encouraged us strongly to search for the arginine recognition elements via the *in vitro* analysis of the kinetic behaviour in arginylation of $tRNA^{Asp}$ variants. These variants were mainly expected to show decreased arginylation activities if arginine-specific recognition elements were removed, or alternatively, to show enhancements in activity if arginine recognition elements were inserted.

How is arginine identity expressed in yeast?

Yeast cells bear three cytosolic $tRNA^{Arg}$ isoacceptors (Weissenbach *et al.*, 1975; Keith and Dirheimer, 1980). The sequences of species II, IIIa and IIIb are displayed in Figure 1A; that of species I is not yet known. In addition to the 16 nucleotides conserved in all tRNA sequences, the three tRNAs share 32 nucleotides, mostly in the central part of the molecule but also in the D and anticodon arms as seen in the consensus sequence displayed in Figure 1B. Whereas the discriminator base is different in these molecules, being either G or C, the anticodon loop sequences point to four common nucleotides out of the seven forming the loop: C32, C35, A37 and A38. Residue U33 is common to all tRNAs and is thus *a priori* not a candidate for a recognition element involved in specificity. It is remarkable that the two isoacceptors present a structural difference at the level of the D-loop. Indeed, the number of nucleotides at the 3′ side of the two conserved residues G18 and G19, and defining the β domain (reviewed in Giegé *et al.*, 1993), is either three or two, suggesting that ArgRS can tolerate some structural variability at this level.

There is only one $tRNA^{Asp}$ species in yeast (Figure 1C) (Gangloff *et al.*, 1971). Eleven nucleotides are common to this tRNA and the consensus $tRNA^{Arg}$ sequence (Figure 1D). Unexpectedly, these common nucleotides are at locations inside the core of the tRNA molecule and are seldom found as identity positions (McClain, 1993a). It follows, according to the up-to-date accepted paradigm underlying the search of identity elements, that arginine identity would be of a special nature. In particular, one may prematurely conclude that arginine identity is not related to either anticodon loop residues or discriminator position. Indeed, the anticodon loop of $tRNA^{Asp}$ has no nucleotides in common with the consensus arginine anticodon loop, except conserved U33. However, since

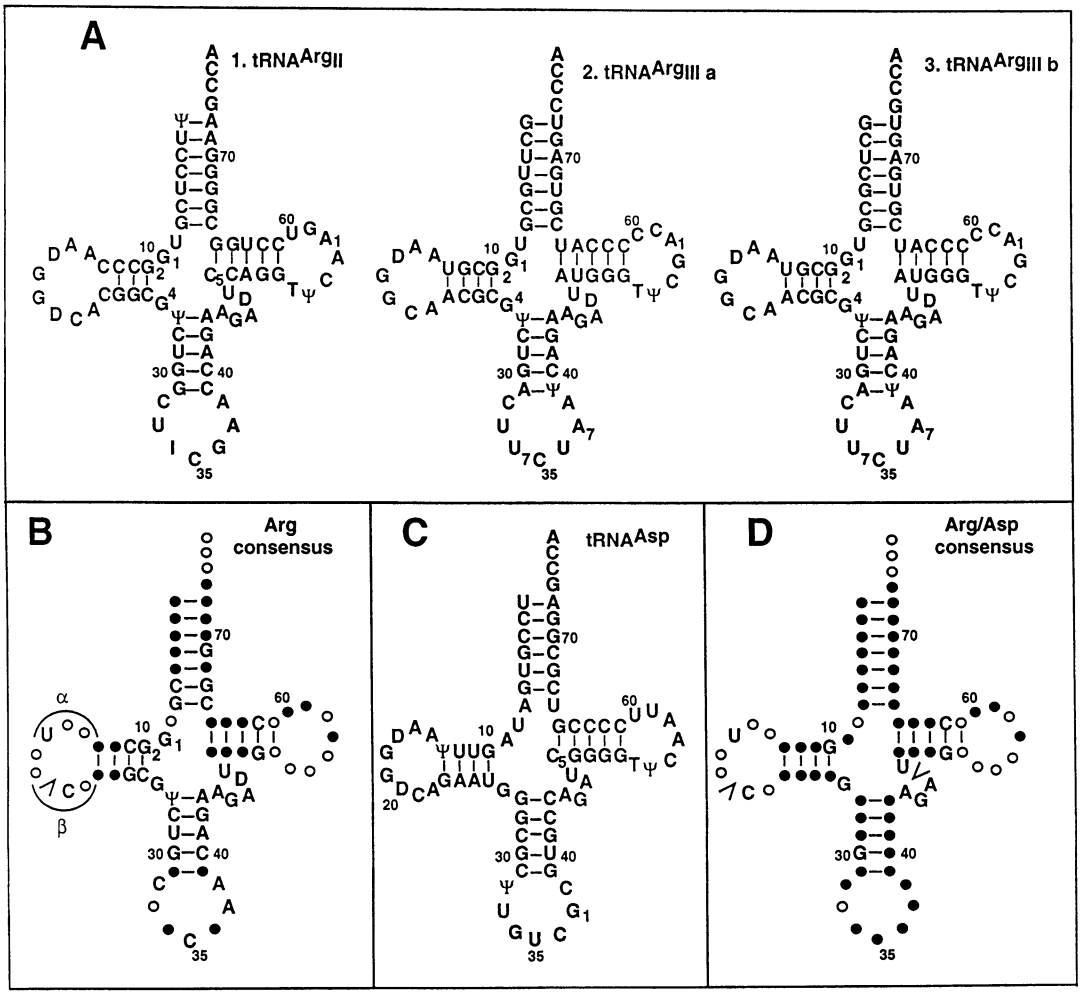

Fig. 1. Comparison of the primary structures of yeast $tRNA^{Arg}$ isoacceptors and yeast $tRNA^{Asp}$. (**A**) Sequences of $tRNA^{Arg}$ isoacceptors and (**B**) the corresponding consensus sequence. The α and β domains of the D-loops are localized in the consensus structure. (**C**) Sequence of the sole yeast $tRNA^{Asp}$ species. (**D**) Consensus sequence between yeast $tRNA^{Arg}$ and $tRNA^{Asp}$ species. Arrowheads indicate positions in the cloverleaf where insertions or deletions of residues occur. Nucleotide numbering and names of modified bases are according to Sprinzl *et al.* (1996). White dots correspond to nucleotides conserved in all tRNAs and black dots to non-conserved nucleotides.

the arginylation efficiency of the wild-type $tRNA^{Asp}$ transcript is sub-optimal and remains 10-fold lower than that of $tRNA^{Arg}$ (Perret *et al.*, 1990a), one may argue that some arginine-specifying nucleotides are missing in the anticodon loop of $tRNA^{Asp}$.

Arginylation properties of selected $tRNA^{Asp}$ variants

The discriminator base and anticodon loop residues are 'hot-spots' for identity elements in most systems studied so far (Shimizu *et al.*, 1992; Giegé *et al.*, 1993; McClain, 1993a). These domains have been investigated in the present study, even though they do not belong to the consensus set of arginine isoacceptor. Doing so was a way to verify whether the identity paradigm would hold or be broken. Besides these conceptual considerations, a number of experimental reasons prompted us also to test these 'hot-spot' positions. Indeed, the discriminator base is close to the catalytic site of aminoacylation and thus makes likely contacts with the synthetase, and the anticodon of $tRNA^{Arg}$ has been found in proximity to ArgRS by footprinting studies (Gangloff *et al.*, 1983). Moreover, the fact that the m1G37 anti-determinant within $tRNA^{Asp}$ hinders this tRNA from being arginylated strongly suggests that residue 37 in $tRNA^{Asp}$ is in the vicinity of the anticodon-binding domain of ArgRS. Finally the discriminator base and the central nucleotide of the anticodon have been found to be arginine identity elements in *Escherichia coli* $tRNA^{Arg}$ (McClain *et al.*, 1990; Tamura *et al.*, 1992) and it is likely that a functional relationship exists between prokaryotes and lower eukaryotes, since a prokaryotic ArgRS (from *Bacillus stearothermophilus*) has been found to charge easily yeast $tRNA^{Arg}$ (Grosjean *et al.*, 1976).

The discriminator base is either G or C in yeast $tRNA^{Arg}$ isoacceptors and G in $tRNA^{Asp}$. To investigate the contribution of this position in arginylation, a variant of $tRNA^{Asp}$ in which G73 is replaced by U73 was created (Figure 2A) and its arginylation properties analysed (Figure 2B, Table I). This variant, AspU73 (names of variants emphasize the basic host tRNA investigated and the nature and position of the nucleotide present in the variant), is chargeable to 100%, with an efficiency enhanced 2-fold as compared with wild-type $tRNA^{Arg}$ transcript, and thus with a 20-fold enhancement as compared with wild-type $tRNA^{Asp}$ transcript from which it is derived.

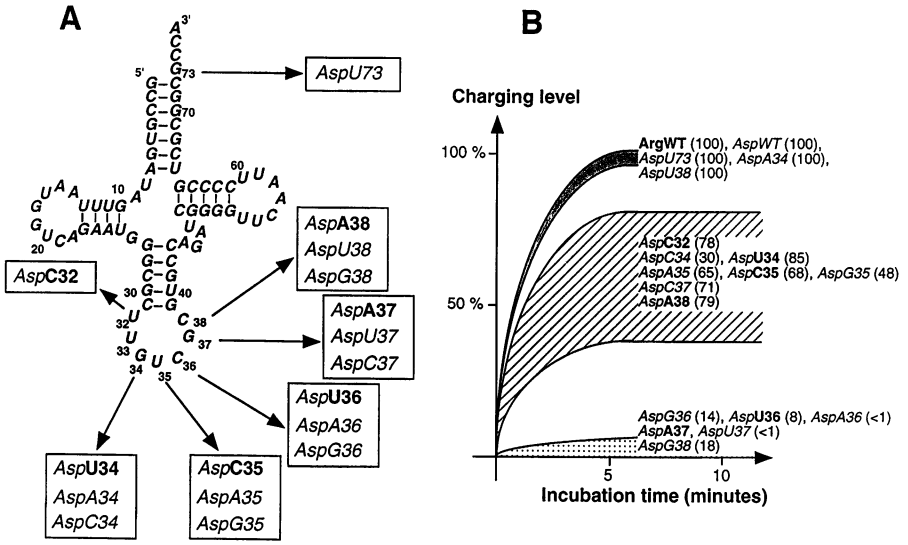

Fig. 2. Sequences and charging levels of $tRNA^{Asp}$ variants with yeast ArgRS. (**A**) Nomenclature of $tRNA^{Asp}$-derived variants. Nucleotides occurring naturally in the $tRNA^{Arg}$ anticodon loop are in bold characters. (**B**) Schematic summary of the arginylation plateaus observed for various transcripts. Values in brackets correspond to optimal charging levels, expressed as a percentage for 0.4 μM of transcript and 180 nM of ArgRS.

Sixteen other $tRNA^{Asp}$ variants, mutated at the level of the anticodon loop nucleotides (Figure 2A) have also been tested for their charging levels with ArgRS (Figure 2B). Their aminoacylation behaviour distributes them within three families. A first family concerns molecules charged to completion or almost, a second family of transcripts contains molecules charged between 30 and 80%, and finally, a third family contains transcripts with very poor charging levels. Variants concerning a same position of the loop fall generally in either of these families, except variants of position 36 which all present very poor aminoacylation levels and variants of position 35 which all present medium aminoacylation levels. Aminoacylation of two variants of position 37 are not detectable (AspA37 and AspU37) whereas that of the corresponding third variant (AspC37) is charged to 71%.

A more precise kinetic analysis of the aminoacylation properties has been performed for six of the variants, namely those corresponding to $tRNA^{Asp}$ molecules into which one of the original $tRNA^{Arg}$III anticodon nucleotides has been transplanted (Figure 2A and Table I). Within these transplants, variants AspC32, AspC35 and AspA38 show no significant changes in arginylation efficiency as compared with wild-type $tRNA^{Asp}$ transcript ($\mathcal{L}$ = 1.15–5.4). Variant AspU34 is 10-fold more efficiently arginylated than the $tRNA^{Asp}$ transcript, and unexpectedly, the arginylation efficiency of variants AspU36 and AspA37 are diminished; losses in arginylation efficiency are of 280- and 720-fold respectively when compared with the wild-type $tRNA^{Asp}$ transcript and of 2780- and 7230-fold when compared with the wild-type $tRNA^{Arg}$ transcript. These effects are linked mainly to large decreases (1100- to 1700-fold) in k_{cat} and increments (2.5- to 4.5-fold) in K_m values.

According to these data it appears clearly that nucleotides C36 and G37 are arginine-specific recognition nucleotides in $tRNA^{Asp}$, since their replacement leads to severe losses in arginylation. This is in contradiction to the expected effects of the presence of nucleotides occurring in natural $tRNA^{Arg}$ species, since C36 and G37 residues are not consensus residues in $tRNA^{Arg}$. The contradiction becomes more striking if one considers that replacement of positions 36 and 37 in $tRNA^{Asp}$ by U36, present in two $tRNA^{Arg}$ isoacceptors, and especially of A37 present in the three $tRNA^{Arg}$ isoacceptors, has a strong negative effect. These replacements were expected to be beneficial, or at least neutral, for arginylation of the host $tRNA^{Asp}$. Another unexpected effect concerns the stimulation of arginylation of the aspartate transcript by introduction of residue U34, a nucleotide present in $tRNA^{Arg}$III but not in the other arginine isoacceptors. Finally, our data show that transplantation of conserved arginine nucleotides C32, C35 or A38 has no effect on arginylation of $tRNA^{Asp}$, whereas it could be expected that these transplantations might overcome the gap in arginylation efficiency between the aspartate and wild-type arginine transcripts.

Altogether, these data are indicative of novel rules underlying tRNA identity.

Arginylation efficiencies of relevant $tRNA^{Arg}$ variants

To understand the unexpected results observed for anticodon loop nucleotides within the $tRNA^{Asp}$ context, analogous studies were performed on a complete set of anticodon loop variants within the $tRNA^{Arg}$ framework. The role of the discriminator base was also tested in the cognate framework (Figure 3A). The arginylation levels of the $tRNA^{Arg}$-derived transcripts are schematized in Figure 3B and the kinetic parameters summarized in Table II. Here, variants of position 35 are the worst substrates. Levels of charging of variants ArgG35, ArgA35 and ArgU35 are 8%, 5% and 12% and losses in catalytic efficiency vary between 2670- and 4920-fold. Variants of position 36 lead to arginylation levels of 20–46%. Whereas the efficiency of aminoacylation of AspG36 is quite as good as that of wild-type $tRNA^{Arg}$ transcript, a loss of 45-fold was observed for ArgC36 and a loss of 123-fold for ArgA36. Variants of position 38 show the same kind of effect as those of position 36. Indeed, introduction of C38 or U38 does not affect arginylation ($\mathcal{L}$ = 1 and $\mathcal{L}$ =

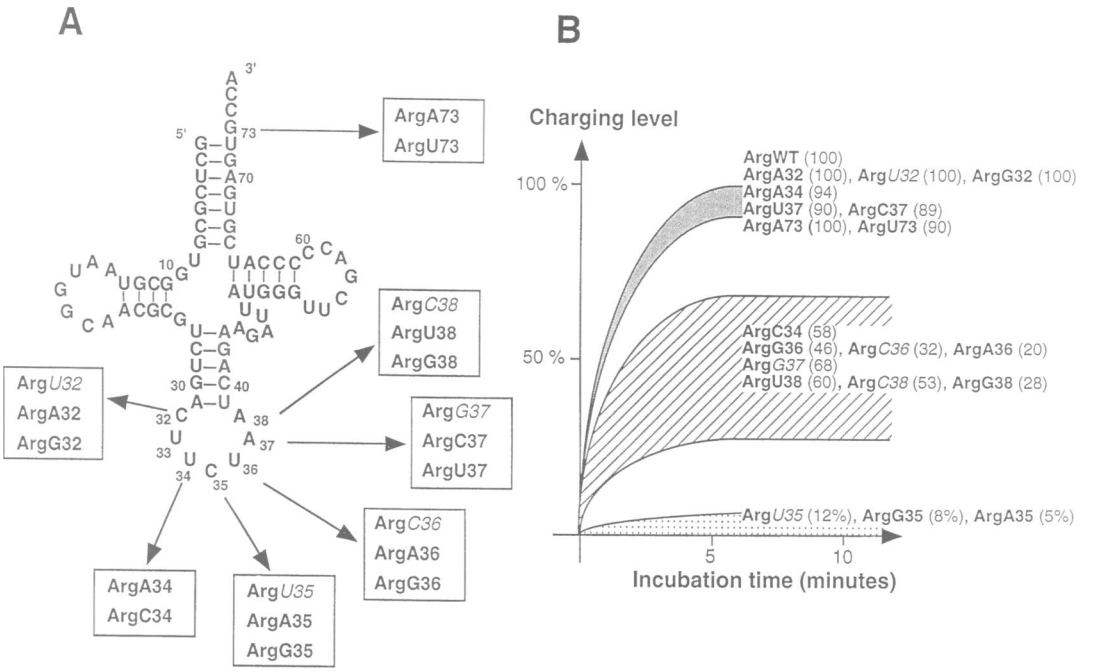

Fig. 3. Sequences and charging levels of $tRNA^{Arg}$ variants with yeast ArgRS. (**A**) Nomenclature of $tRNA^{Arg}$-derived variants. Nucleotides occurring naturally in the $tRNA^{Asp}$ anticodon loop are in italics. (**B**) Schematic summary of the arginylation plateaus observed for various transcripts. Values in brackets correspond to optimal charging levels, expressed as a percentage, for 0.4 μM of transcript and 180 nM of ArgRS.

3.5 respectively) whereas the third mutant, where G38 has been inserted, leads to a loss of 270-fold. Mutations at positions 32, 34 and 37 do not affect arginylation, since losses in arginylation efficiencies are between 0.2- and 2.5-fold (in fact, up to a 5-fold increase for some variants). Mutations of the discriminator base (G or C in the wild-type $tRNA^{Arg}$ sequences) into A or U, have no effect on arginylation ($\mathcal{L}$ = 1.8 and $\mathcal{L}$ = 0.95).

Arginine identity in yeast is defined by two different recognition sets in the anticodon loops of host tRNAs

Comparative analysis of the kinetic properties of two series of tRNA variants, one in the $tRNA^{Asp}$ framework, the other in the $tRNA^{Arg}$ framework, clearly highlights that: (i) arginylation of $tRNA^{Arg}$ and $tRNA^{Asp}$ transcripts rely on the presence of different combinations of nucleotides, (ii) contrary to what is predicted by simple sequence comparisons of arginylable $tRNA^{Arg}$ and $tRNA^{Asp}$ species, the anticodon loop nucleotides are in fact involved in identity. In the $tRNA^{Arg}$ context, C35 is the major recognition element of the anticodon loop. In addition U36 or G36 present in $tRNA^{Arg}$ isoacceptors also contributes, but to a lesser extent, to this identity. Moreover, a G residue is not permitted at position 38. Alternatively, analysis of $tRNA^{Asp}$ variants transplanted with single $tRNA^{Arg}$ anticodon loop nucleotides, shows that residues C36 and G37 are fundamental for arginylation. These major anticodon loop arginine recognition elements are highlighted in Figure 4. This is the first case described in which identity of a tRNA is correlated with two distinct combinations of nucleotides. These combinations are distinct at the level both of the nature of the nucleotides and of their position.

The contribution of the nucleotides previously discussed in arginylation emerged from the negative effects of their removal. Our study suggests also a contribution of two additional nucleotides, since their introduction into the $tRNA^{Asp}$ framework stimulates arginylation to the level of wild-type $tRNA^{Arg}$. These nucleotides are U34 and U73. They are non-conserved or even absent in $tRNA^{Arg}$ isoacceptors. Thus, the positive effects brought by these non-conserved nucleotides may be related to the presence of other nucleotides either proximal or at a long distance, or to characteristics of the structural framework of $tRNA^{Asp}$.

Table II. Kinetic parameters for arginylation by yeast ArgRS of transcripts derived from yeast $tRNA^{Arg}$

Transcripts	k_{cat} (s^{-1})	K_m (nM)	k_{cat}/K_m	$\mathcal{L}$ relative[a]
Argwt IIIb	0.68	550	1	1
ArgA73	0.29	430	0.55	1.8
ArgU73	0.42	323	1.05	0.95
ArgA32	0.22	132	1.3	0.75
ArgG32	0.5	280	1.4	0.7
Arg*U32*	0.93	139	5.4	0.2
ArgA34	0.46	70	5.3	0.2
ArgC34	1.8	695	2.1	0.5
ArgG35	0.0008	3485	0.0002	4920
ArgA35	0.0002	690	0.0004	3330
Arg*U35*	0.0003	735	0.0005	2670
ArgG36	0.52	370	1.15	0.8
ArgA36	0.0065	650	0.0081	123
Arg*C36*	0.014	520	0.022	45
ArgU37	1.20	1205	0.81	1.2
ArgC37	1.34	410	2.6	0.4
Arg*G37*	0.30	605	0.40	2.5
ArgU38	0.18	500	0.29	3.5
ArgG38	0.00067	147	0.0037	270
Arg*C38*	0.37	300	1	1

[a]Relative to wild-type $tRNA^{Arg}$ transcript. Relative kinetic specificity constants are defined as $(k_{cat}/K_m)_{relative} = (k_{cat}/K_m)_{mutant}/(k_{cat}/K_m)_{wild\text{-}type}$. $\mathcal{L}$-values are defined as inverses of $(k_{cat}/K_m)_{relative}$. $\mathcal{L}$-values for duplicates varied by <10%. For nomenclature of variants, see Figure 2.

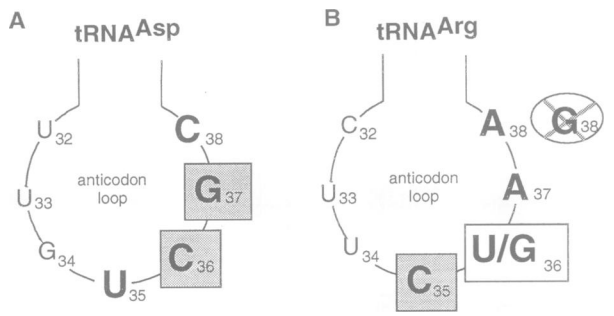

Fig. 4. Nucleotides of $tRNA^{Asp}$ (**A**) and $tRNA^{Arg}$ (**B**) anticodon loops, important for efficient arginylation. All nucleotides have been investigated except residues U33. Very important elements are in grey boxes, less important ones are in a white box. A guanosine at position 38 is detrimental in $tRNA^{Arg}$. Data concerning $tRNA^{Asp}$ emerge from studies of a single mutant for each position, i.e. nucleotides imported from $tRNA^{Arg}$III, whereas data concerning $tRNA^{Arg}$ correspond to the outcome of a systematic mutational analysis at each position.

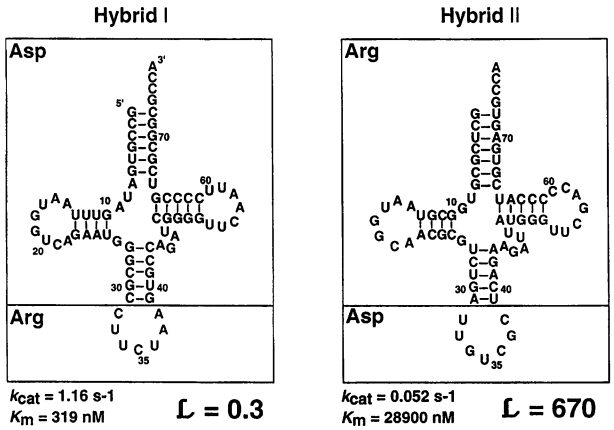

Fig. 5. Hybrid transcripts tested for arginylation. Hybrid I is constituted of yeast $tRNA^{Asp}$ sequence except for the anticodon loop which corresponds to that of yeast $tRNA^{Arg}$III. Hybrid II is constituted of yeast $tRNA^{Arg}$IIIb sequence except for the whole anticodon loop which corresponds to that of yeast $tRNA^{Asp}$.

Towards an understanding of the dual specificity of ArgRS

We show here that arginine identity is defined by two different sets of nucleotides within the anticodon loops of substrate tRNAs. This discovery may only be in apparent contradiction with the present understanding of tRNA identity. Indeed, although the sequences of the nucleotides involved are different, identity signals may be restricted to common chemical groups. However, a simple search for such chemical groups in homologous nucleotides was unsuccessful. A new possibility would be that the anticodon loops of $tRNA^{Arg}$ and $tRNA^{Asp}$ transcripts share common nucleotides or chemical groups which are at non-homologous positions, but which are nevertheless recognized by the synthetase. According to this view, the global plasticity of $tRNA^{Asp}$ transcript would be such that ArgRS 'reads' nucleotides 36 and 37 from $tRNA^{Asp}$ as it 'reads' nucleotides 35 and 36, from cognate $tRNA^{Arg}$ transcript. This hypothesis implies thus a 'slipped' recognition set in which a cytosine residue, either at position 35 or 36, would be the major arginine recognition element. Alternatively, one may also imagine that two different combinations of chemical groups within the anticodon loops of $tRNA^{Asp}$ and $tRNA^{Arg}$ transcripts form 'recognition pockets' which would both be perceived equivalently by ArgRS to trigger function. It is presently premature to discriminate between the two hypotheses.

Whatever the mechanism of the anticodon loop recognition will be, the question of the contribution of the rest of the tRNA molecules to the optimal expression of identity has to be raised. In other words, is arginine identity in both tRNAs restricted solely to the anticodon loops or are other parts of the tRNAs involved in complementary processes? In the $tRNA^{Arg}$ context, residue C35 cannot be the only information perceived by ArgRS, since this nucleotide is also present in the anticodons of $tRNA^{Cys}$, $tRNA^{Trp}$, $tRNA^{Ser}$ and $tRNA^{Gly}$ (Sprinzl *et al.*, 1996). It is probable that additional nucleotides contribute to identity, as is the case in the *E.coli* arginine system, where residues A20 and the discriminator base are involved in the specificity process (McClain *et al.*, 1990; Tamura *et al.*, 1992). The unexpected effects of the anticodon loop nucleotides in the $tRNA^{Asp}$ context may be linked to cooperative, anti-cooperative or additive effects of elements present in the rest of the molecule, as already found for expression of aspartate identity in yeast (Pütz *et al.*, 1993). These features may be either additional recognition elements, and/or structural characteristics conferring a particular plasticity to the $tRNA^{Asp}$ transcript.

With these considerations in mind, and to search for a functional role of tRNA features outside of the anticodon loop, we have designed and tested the arginylation activity of hybrid tRNA molecules (Figure 5). Hybrid I contains the $tRNA^{Asp}$ body (all nucleotides of $tRNA^{Asp}$ except the anticodon loop) and the $tRNA^{Arg}$III anticodon loop. The arginylation ability of this molecule is efficient. It is charged to 83%, the K_m is unchanged and the k_{cat} is ~1.7-fold increased, leading to a loss of arginylation efficiency of 0.3-fold which is considered as non-significantly different from wild-type $tRNA^{Arg}$ activity (Table I). This shows either that the arginine anticodon loop contains all required specific information or that the body of $tRNA^{Asp}$ contains the same complementary recognition elements as those present in the body of $tRNA^{Arg}$. Hybrid II contains all nucleotides of $tRNA^{Arg}$III except the anticodon loop which is from $tRNA^{Asp}$. This hybrid molecule is charged up to 70% but here the arginylation efficiency is significantly impaired ($\mathcal{L}$ = 670-fold). This loss is the result of a decreased k_{cat} (13-fold) and an increased K_m (52-fold) and indicates that the positive information brought by the aspartate anticodon loop is prohibited by the $tRNA^{Arg}$ framework. The functional analysis of this hybrid clearly demonstrates a contribution of the body of the tRNA in optimal expression of arginine identity. Noteworthy and contrarily to what could be anticipated, hybrid I, structurally resembling $tRNA^{Asp}$ and just bearing the seven-nucleotide arginine anticodon loop, functionally resembles a wild-type $tRNA^{Arg}$ molecule. Thus, expression of the anticodon recognition sets is dependent on the host tRNA framework. Optimal arginylation requires information within the tRNA body to enable communication between the anticodon domain of the tRNA and the CCA-end.

In conclusion, the present data show that there are alternative non-conventional ways to assess a given aminoacylation identity in different tRNAs. Our discovery contrasts with the current concept of tRNA identity, according

to which identity is conferred by a unique set of nucleotides, i.e. a distinct and explicit set, and with its consequence which stipulates that simple transplantation of recognition sets from one tRNA framework into another is directly correlated with the acquisition of the new identity by the host tRNA. We show that, at least in some instances, expression of specificity in tRNA aminoacylation is more complex and emerges from subtle combinations of events based on sequence, context and more generally on tRNA structure and plasticity. It is worth recalling the work of Hou and Schimmel (1992) who found that AlaRS can recognize a substrate in which the absence of the major identity element can be compensated by a second site mutation at long distance. We also show that non-consensus nucleotides between isoacceptor tRNAs (any tRNA efficiently charged by a given synthetase) may be very important for specificity. Studies on the aminoacylation properties of minihelices have already highlighted this point (Martinis and Schimmel, 1993).

Evolutionary considerations

Though unexpected, the fact that alternative recognition sets for a particular synthetase occur in a particular biological context (i.e. within yeast) indicates that Nature can utilize different strategies to solve the same problem. This is in line with conclusions of *in vitro* genetic studies that led to the discovery of RNA aptamers with alternative structures able to recognize proteins (e.g. Tuerk and Gold, 1990; Bartel and Szostak, 1994; Eckland *et al.*, 1995). Thus, the novel functional features described here, related to arginine identity in yeast, raise the problem of their evolutionary origin.

Here we are faced with structural and functional correlations between aspartate and arginine aminoacylation systems. Aspartic acid is considered as a primordial amino acid that appeared in chemical evolution much before arginine (e.g. Crick, 1968; Miller and Orgel, 1974). Thus, the primitive protein synthesis machinery likely involved an aspartic acid system whereas arginine-accepting systems appeared later. To achieve early evolutionary tinkering, primordial systems likely had neither optimal nor strict specificities. Thus it can be hypothesized that primordial $tRNA^{Asp}$ molecules were by chance such as to become arginine acceptors as soon as the arginine aminoacylation system appeared, in addition of being aspartate acceptors. Initially, weak discrimination between positively- and negatively-charged amino acids would have been a driving force for molecular evolution processes, in allowing the emergence of novel proteins from libraries of variant proteins. Improvement of aspartate specificity and separation from arginine specificity of tRNAs, most likely occurred through mutation effects at both protein and nucleic acid levels, combined with epigenetic phenomena such as base modification. In this view it is interesting to recall that unmodified tRNA transcripts bear more conformational flexibility than the post-transcriptionally modified natural tRNAs (e.g. Hall *et al.*, 1989; Perret *et al.*, 1990a; Derrick and Horowitz, 1993) and that modifications were shown to be anti-determinants hindering incorrect aminoacylation (Muramatsu *et al.*, 1988; Pütz *et al.*, 1994). Thus, it can be proposed that the post-transcriptional modification process is a way retained by evolution to restrict the specificity potential of tRNA molecules. To verify this hypothesis it becomes important to check whether introduction of modified bases within the structure of different transcripts studied in this work will modulate their functionality, for instance by freezing a conformational plasticity. Alternatively, using combinatorial methods it will become interesting to search for additional sequence possibilities not selected by natural evolution but able to sustain arginine specificity. Experiments are underway to explore these various possibilities.

Materials and methods

Materials

Oligonucleotides were synthesized on an Applied Biosystems 381 DNA synthesizer using the phosphoramidite method and purified by HPLC on Nucleosyl 125-5-C18 column (Bischoff Chromatography, Zymark-France, Paris). L-[^{3}H]Arginine (57 Ci/mmol) was from Amersham France (Les Ulis). Yeast ArgRS was purified according to Gangloff *et al.* (1976). T7 RNA polymerase was purified according to Wyatt *et al.* (1991). Restriction enzymes *Bst*NI, *Hin*dIII and *Bam*HI, and T4 polynucleotide kinase were from New England Biolabs (Beverly, MA, USA). T4 DNA ligase was from Boehringer Mannheim (Meylan, France).

Cloning and in vitro transcription

Most tRNAs used in this work have been obtained by *in vitro* transcription of synthetic genes. Each of these genes corresponds to the T7 RNA polymerase promoter region directly upstream of the tRNA sequence. The $tRNA^{Asp}$ genes were constructed as described in Perret *et al.* (1990b), the wild-type $tRNA^{Arg}$ gene was constructed accordingly from 10 oligonucleotides and cloned into plasmid pTFMa (Perret *et al.*, 1990b) linearized at *Bam*HI and *Hin*dIII sites. Plasmids containing genes for $tRNA^{Arg}$ variants were obtained by site-directed mutagenesis of pTFSMArgWT according to Sayers and Eckstein (1989). TG1 cells were transformed. A *Bst*NI site coincidental with the 3′-end of the tRNA sequences, allowed synthesis of tRNAs ending with the expected CCA sequence. Experimental procedures were as described in Perret *et al.* (1990b). *In vitro* transcription and purification of transcripts to single nucleotide resolution were performed as described in Frugier *et al.* (1993). Concentration of stock solutions of transcripts have been determined by OD measurements at 260 nm.

Aminoacylation reactions

Aminoacylation reactions of transcripts derived from $tRNA^{Asp}$ and $tRNA^{Arg}$ have been performed as described by Frugier *et al.* (1993) in a medium containing 50 mM HEPES–NaOH pH 7.5, 30 mM KCl, 15 mM $MgCl_2$, 10 mM ATP, 2.5 mM glutathione, 50 μM ^{3}H-labelled arginine and adequate amounts of transcript (0.2 μM–4 μM) and yeast ArgRS (0.3–900 nM). Before aminoacylation, transcripts were renatured by heating at 65°C for 90 s and slow cooling to room temperature. Assays were performed in the conventional way (Perret *et al.*, 1990b) with incubation at 30°C. The kinetic constants were derived from Lineweaver–Burk plots. Since the concentration of amino acid was subsaturating, only apparent kinetic parameters are given. They represent an average of at least two independent experiments.

Acknowledgements

We thank P.Walter for fruitful discussions and A.Hoeft for the synthesis of oligonucleotides. This work was partly supported by grants from the Centre National de la Recherche Scientifique (CNRS), the Ministère de l'Enseignement Supérieur et de la Recherche (MESR) and Université Louis Pasteur (Strasbourg). M.S. was supported by a MESR grant.

References

Bartel,D.P. and Szostak,J.W. (1994) Study of RNA–protein recognition by *in vitro* selection. In Nagai,K. and Mattaj,I.W. (eds), *RNA–Protein Interactions*. Oxford University Press, Oxford, UK, pp. 248–268.

Cavarelli,J., Rees,B., Ruff,M., Poterszman,A., Thierry,J.-C. and Moras,D. (1992) Class II aminoacyl-tRNA synthetases seen through the crystal structure of the complex between yeast $tRNA^{Asp}$ aspartyl-tRNA synthetase. In *Structural Tools for the Analysis of Protein-Nucleic Acid Complexes*. Birkhäuser Verlag, Basel, pp. 287–298.

Cavarelli,J., Rees,B., Ruff,M., Thierry,J.-C. and Moras,D. (1993) Yeast $tRNA^{Asp}$ recognition by its cognate class II aminoacyl-tRNA synthetase. *Nature*, **362**, 181–184.

Crick,F.H.C. (1968) The origin of the genetic code. *J. Mol. Biol.*, **38**, 367–379.

Derrick,W.B. and Horowitz,J. (1993) Probing structural differences between native and *in vitro* transcribed *Escherichia coli* valine transfer RNA: evidence for stable base modification-dependent conformers. *Nucleic Acids Res.*, **21**, 4948–4953.

Eckland,E.H., Szostak,J.W. and Bartel,D.P. (1995) Structurally complex and highly active RNA ligases derived from random RNA sequences. *Science*, **269**, 364–370.

Frugier,M., Florentz,C., Schimmel,P. and Giegé,R. (1993) Triple aminoacylation specificity of a chimerized transfer RNA. *Biochemistry*, **32**, 14053–14061.

Gangloff,J., Keith,G., Ebel,J.-P. and Dirheimer,G. (1971) Structure of aspartate tRNA from brewer's yeast. *Nature New Biol.*, **230**, 125–127.

Gangloff,J., Schutz,A. and Dirheimer,G. (1976) Arginyl-tRNA synthetase from baker's yeast. Purification and some properties. *Eur. J. Biochem.*, **65**, 177–182.

Gangloff,J., Jaozara,R. and Dirheimer,G. (1983) Study of the interaction of yeast arginyl-tRNA synthetase with yeast $tRNA^{Arg}2$ and $tRNA^{Arg}3$ by partial digestion with cobra venom nuclease. *Eur. J. Biochem.*, **132**, 629–637.

Giegé,R., Puglisi,J.D. and Florentz,C. (1993) tRNA structure and aminoacylation efficiency. *Progr. Nucleic Acid Res. Mol. Biol.*, **45**, 129–206.

Grosjean,H., Charlier,J., Darte,C., Dirheimer,G., Giegé,R., De Henau,S., Keith,G., Parfait,R. and Takada,C. (1976) Purification, characterization and mechanism of action of several aminoacyl-tRNA synthetases from *Bacillus stearothermophillus*. *Experientia*, **32**, 347–362.

Hall,K.B., Sampson,J.R., Uhlenbeck,O.C. and Redfield,A.G. (1989) Structure of an unmodified tRNA molecule. *Biochemistry*, **28**, 5794–5801.

Hou,Y.M. and Schimmel,P. (1992) Functional compensation of a recognition-defective transfer RNA by a distal base pair substitution. *Biochemistry*, **31**, 10310–10314.

Keith,G. and Dirheimer,G. (1980) Reinvestigation of the primary structure of brewer's yeast $tRNA^{Arg}3$. *Biochem. Biophys. Res. Commun.*, **92**, 116–119.

Martinis,S.A. and Schimmel,P. (1993) Microhelix aminoacylation by a class I tRNA synthetase. Non-conserved base pairs required for specificity. *J. Biol. Chem.*, **268**, 6069–6072.

McClain,W.H. (1993a) Rules that govern tRNA identity in protein synthesis. *J. Mol. Biol.*, **234**, 257–280.

McClain,W.H. (1993b) Transfer RNA identity. *FASEB J.*, **7**, 72–78.

McClain,W.H. and Nicholas,H.B.J. (1987) Differences between transfer RNA molecules. *J. Mol. Biol.*, **194**, 635–642.

McClain,W.H., Foss,K., Jenkins,R.A. and Schneider,J. (1990) Nucleotides that determine *Escherichia coli* $tRNA^{Arg}$ and $tRNA^{Lys}$ acceptor identities revealed by analyses of mutant opal and amber suppressor tRNAs. *Proc. Natl Acad. Sci. USA*, **87**, 9260–9264.

Miller,S.L. and Orgel,L.E. (1974) The Origins of Life on the Earth. Prentice-Hall, London.

Muramatsu,T., Nishikawa,K., Nemoto,F., Kuchino,Y., Nishimura,S., Miyazawa,T. and Yokoyama,S. (1988) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature*, **336**, 179–181.

Nicholas,H.B. and McClain,W.H. (1987) An algorithm for discriminating sequences and its application to yeast transfer RNA. *Comput. Applic. Biosci.*, **3**, 177–181.

Perret,V., Garcia,A., Grosjean,H., Ebel,J.-P., Florentz,C. and Giegé,R. (1990a) Relaxation of transfer RNA specificity by removal of modified nucleotides. *Nature*, **344**, 787–789.

Perret,V., Garcia,A., Puglisi,J.D., Grosjean,H., Ebel,J.-P., Florentz,C. and Giegé,R. (1990b) Conformation in solution of yeast $tRNA^{Asp}$ transcripts deprived of modified nucleotides. *Biochimie*, **72**, 735–744.

Perret,V., Florentz,C., Puglisi,J.D. and Giegé,R. (1992) Effect of conformational features on the aminoacylation of tRNAs and consequences on the permutation of tRNA specificities. *J. Mol. Biol.*, **226**, 323–333.

Puglisi,J.D., Pütz,J., Florentz,C. and Giegé,R. (1993) Influence of tRNA tertiary structure and stability on aminoacylation by yeast aspartyl-tRNA synthetase. *Nucleic Acids Res.*, **21**, 41–49.

Pütz,J., Puglisi,J.D., Florentz,C. and Giegé,R. (1991) Identity elements for specific aminoacylation of yeast $tRNA^{Asp}$ by cognate aspartyl-tRNA synthetase. *Science*, **252**, 1696–1699.

Pütz,J., Puglisi,J.D., Florentz,C. and Giegé,R. (1993) Additive, cooperative and anti-cooperative effects between identity nucleotides of a tRNA. *EMBO J.*, **12**, 2949–2957.

Pütz,J., Florentz,C., Benseler,F. and Giegé,R. (1994) A single methyl group prevents the mischarging of a tRNA. *Nature Struct. Biol.*, **1**, 580–582.

Rould,M.A., Perona,J.J., Söll,D. and Steitz,T.A. (1989) Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science*, **246**, 1135–1142.

Rould,M.A., Perona,J.J. and Steitz,T.A. (1991) Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. *Nature*, **352**, 213–218.

Saks,M.E., Sampson,J.R. and Abelson,J.N. (1994) The transfer RNA identity problem: a search for rules. *Science*, **263**, 191–197.

Sayers,J.R. and Eckstein,F. (1989) Site-directed mutagenesis, based on the phosphorothioate approach. In Creighton,T.E. (ed.), *Protein Function: A Practical Approach.* IRL Press at Oxford University Press, Oxford, UK, pp. 279–295.

Shimizu,M., Asahara,H., Tamura,K., Hasegawa,T. and Himeno,H. (1992) The role of anticodon bases and the discriminator nucleotide in the recognition of some *E. coli* tRNA by their aminoacyl-tRNA synthetases. *J. Mol. Evol.*, **35**, 436–443.

Sprinzl,M., Steegborn,C., Hübel,F. and Steinberg,S. (1996) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **24**, 68–72.

Tamura,K., Himeno,H., Asahara,H., Hasegawa,T. and Shimizu,M. (1992) *In vitro* study of *E. coli* $tRNA^{Arg}$ and $tRNA^{Lys}$ identity. *Nucleic Acids Res.*, **20**, 2335–2339.

Tuerk,C. and Gold,L. (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science*, **249**, 505–510.

Weissenbach,J., Martin,R. and Dirheimer,G. (1975) The primary structure of $tRNA^{Arg}$II from brewers' yeast. 2. Partial digestion with ribonuclease T1 and derivation of the complete sequence. *Eur. J. Biochem.*, **56**, 527–532.

Wyatt,J.R., Chastain,M. and Puglisi,J.D. (1991) Synthesis and purification of large amounts of RNA oligonucleotides. *BioTechniques*, **11**, 764–769.

Received on April 11, 1996; revised on June 3, 1996