The EMBO Journal Vol.17 No.1 pp.297–305, 1998

Genetic code in evolution: switching species-specific aminoacylation with a peptide transplant

Keisuke Wakasugi[1,2], Cheryl L.Quinn[3,4], Nianjun Tao[3,5] and Paul Schimmel[1,2,6]

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, [3]Cubist Pharmaceuticals, Inc., 24 Emily street, Cambridge, Massachusetts 02139, [2]The Scripps Research Institute, 10550 North Torry Pines Road, La Jolla, CA 92037, [4]Pharmacia and Upjohn, 301 Henrietta Street, Kalamazoo, MI 49007 and [5]Biogen, Inc., 12 Cambridge Center, Cambridge, MA 02142, USA

[6]Corresponding author

The genetic code is established in aminoacylation reactions whereby amino acids are joined to tRNAs bearing the anticodons of the genetic code. Paradoxically, while the code is universal there are many examples of species-specific aminoacylations, where a tRNA from one taxonomic domain cannot be acylated by a synthetase from another. Here we consider an example where a human, but not a bacterial, tRNA synthetase charges its cognate eukaryotic tRNA and where the bacterial, but not the human, enzyme charges the cognate bacterial tRNA. While the bacterial enzyme has less than 10% sequence identity with the human enzyme, transplantation of a 39 amino acid peptide from the human into the bacterial enzyme enabled the latter to charge its eukaryotic tRNA counterpart *in vitro* and *in vivo*. Conversely, substitution of the corresponding peptide of the bacterial enzyme for that of the human enabled the human enzyme to charge bacterial tRNA. This peptide element discriminates a base pair difference in the respective tRNA acceptor stems. Thus, functionally important co-adaptations of a synthetase to its tRNA act as small modular units that can be moved across taxonomic domains and thereby preserve the universality of the code.

Keywords: *E.coli*–human hybrid protein/modular unit/molecular evolution/RNA recognition/species-specific aminoacylation

Introduction

Aminoacyl-tRNA synthetases arose early in evolution and are believed to be ancient enzymes that were among the first proteins to emerge from the RNA world (Hountondji *et al*., 1986; Schimmel, 1987; Eriani *et al*., 1990; Cusack *et al*., 1991; Nagel *et al*., 1991; Moras, 1992; Carter, 1993; Schimmel and Ribas de Pouplana, 1995). The aminoacylation reactions establish the genetic code. In the aminoacylation reaction, each amino acid is matched with a tRNA that contains the trinucleotide (anticodon) that corresponds to that amino acid in the algorithm of the code. Although the code is the same in all species throughout evolution, an aminoacyl tRNA synthetase often displays species-specific recognition. That is, even though the anticodons within tRNAs are the same for all organisms, the synthetase from one species does not aminoacylate its cognate tRNA from another species.

For example, prokaryote and eukaryote cytoplasmic tyrosyl-tRNA synthetases (TyrRSs) can not cross-aminoacylate their respective $tRNA^{Tyr}$s (Chow and RajBhandary, 1993; Quinn *et al*., 1995). Similarly, glycyl-tRNA synthetases from *Escherichia coli* and human are unable to cross-aminoacylate human $tRNA^{Gly}$ and *E.coli* $tRNA^{Gly}$, respectively (Shiba *et al*., 1994a). And *E.coli* isoleucyl-tRNA synthetase fails to aminoacylate mammalian $tRNA^{Ile}$ with isoleucine (Shiba *et al*., 1994b). The reason for this behavior is that nucleotides other than, or in addition to, the anticodon trinucleotides are recognized by tRNA synthetases. These nucleotides are commonly found in the acceptor stems of the tRNAs, near the amino acid attachment sites (Giegé *et al*., 1993; Schimmel *et al*., 1993). It is thought that these changes in acceptor stem sequences are compensated by changes in the parts of the synthetases that interact with the acceptor stem during evolution. The sequences and structures of tRNA acceptor stems can be thought of as an operational 'RNA code' that is devoid of the anticodon trinucleotides of the genetic code (Schimmel *et al*., 1993).

The acceptor helices for tyrosine tRNAs of prokaryotes and eukaryotic organelles (mitochondria and chloroplasts) have the conventional G1:C72 base pair found in most tRNAs. However, the terminal base pair of $tRNA^{Tyr}$s of eukaryotic cytoplasm and archaebacteria is an uncommon C1:G72 (Steinberg *et al*., 1993) (Figure 1). The significance of the C1:G72 versus G1:C72 distinction between eukaryotic cytoplasmic and prokaryotic tyrosyl-tRNAs was established when a C1:G72 base pair was shown to be required for aminoacylation of an *E.coli* amber suppressor with tyrosine by yeast extracts (Lee and RajBhandary, 1991). In addition, an investigation with microhelix substrates, whose sequences are derived from the acceptor stem of eukaryote cytoplasmic $tRNA^{Tyr}$, elucidated that a simple change of C1:G72 to G:C abolished aminoacylation by a eukaryote TyrRS (Quinn *et al*., 1995).

TyrRS is a dimeric molecule which belongs to the class I synthetases, which are characterized by the Rossmann nucleotide binding fold of alternating β-strands and α-helices. Although the crystal structure of the *Bacillus stearothermophilus* TyrRS has been determined (Brick *et al*., 1989), only the active site domain has been resolved and no co-crystals with $tRNA^{Tyr}$ have been reported. Therefore, we cannot directly identify the locations of the residues in the TyrRS sequences which explain the G1:C72 versus C1:G72 selectivities of the *E.coli* and human enzymes. Among class I enzymes, the only co-crystal with a tRNA is that of *E.coli* glutaminyl-tRNA synthetase

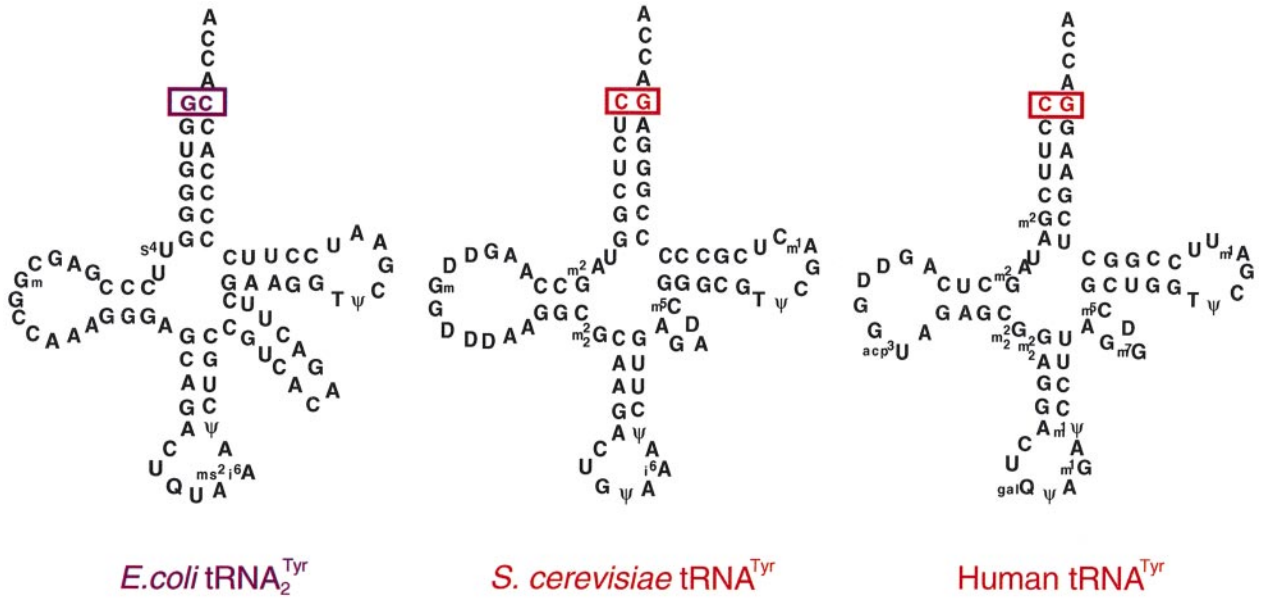

Fig. 1. Comparison of tyrosine tRNA structures. The *E.coli* (left), *S.cerevisiae* cytoplasmic (center), and human cytoplasmic (right) tyrosine tRNAs are shown in schematic cloverleaf form. Boxes enclose nucleotides shown by Lee and RajBhandary (1991) and Quinn *et al*. (1995), to be essential for species-specific aminoacylation by eukaryote cytoplasmic TyrRSs.

complexed with $tRNA^{Gln}$ and ATP (Rould *et al*., 1989). In glutaminyl-tRNA synthetase, connective polypeptide 1 (CP1), which is inserted between the two halves of the active site Rossmann nucleotide fold domain (Starzyk *et al*., 1987), serves to bind the acceptor stem of the complexed $tRNA^{Gln}$ and has been termed the acceptor helix-binding domain (Rould *et al*., 1989). By analogy, we imagined that contact with the 1:72 base pair of $tRNA^{Tyr}$ was made through residues in the CP1 insertion of TyrRS.

In this work, we chose *E.coli* and human cytoplasmic TyrRSs as representatives of prokaryotic and eukaryote cytoplasmic enzymes, respectively. Because the human enzyme contains an extra C-terminal domain compared with the *E.coli* enzyme (Kleeman *et al*., 1997), we used a 'truncated' version in which the human-specific domain is deleted. Although the identity of the amino acid sequences between the *E.coli* and 'truncated' human enzymes is less than 10%, we reasoned that only a small part of the respective proteins may be involved with recognition of a critical base pair difference (such as G1:C72 versus C1:G72). At first, we investigated whether human TyrRS complements a yeast null strain, in which the original cytoplasmic TyrRS gene is disrupted. Having established that the human but not the *E.coli* enzyme complements the null allele, we then tried to manipulate the species-specific recognition of TyrRSs. These manipulations established that a small segment within the CP1 domain was a critical determinant of species-specificity.

Results

Human but not E.coli TyrRS complements a TyrRS-deficient yeast strain

At first, to test the function of human or *E.coli* TyrRS in yeast, we prepared a TyrRS-deficient yeast strain. The whole gene encoding cytoplasmic TyrRS in yeast diploid strain Y93 was deleted and replaced by *TRP1*. The diploid strain was transformed with a maintenance plasmid ($pC^{3}679$) bearing the *Saccharomyces cerevisiae* cytoplasmic TyrRS gene and the selectable marker *URA3*. Transformants were sporulated and the resulting tetrads were dissected to generate the TyrRS-deficient yeast haploid strain QBY374.

The human and *E.coli* TyrRS genes were cloned into the high copy plasmid pQB169, in which expression is driven by the strong constitutive alcohol dehydrogenase promoter. After the second plasmid, containing the human or *E.coli* TyrRS gene and a different selectable marker, *LEU2*, was introduced, the maintenance plasmid was lost by growth on a 5-fluoro-orotic acid (5-FOA) plate (Boeke *et al*., 1984). Expression of the *E.coli* protein could not rescue the lethal phenotype caused by the knock-out mutation (see below). To show that the *E.coli* enzyme accumulated in yeast, we added a $(His)_6$ tag to the C-terminus of the protein. This tag did not disrupt the activity of the *E.coli* protein and, using a conjugate directed against the tag, we showed that the tagged enzyme was expressed with the expected apparent molecular weight of 47 kDa (data not shown). Thus, the failure of the *E.coli* enzyme to complement the yeast null strain was not due to degradation of the protein.

On the other hand, expression of human TyrRS resulted in complementation of the null strain (see below). Retention of the deletion/disruption of the null strain was shown by the $Leu^{+}Trp^{+}Ura^{-}$ phenotype, indicating that *S.cerevisiae* cytoplasmic TyrRS gene is disrupted with *TRP1* and that the maintenance plasmid was lost. As an additional test of plasmid dependence, plasmids isolated from the complementation plates were verified by restriction mapping to contain the genes encoding the human protein. Back-transformation of the yeast null strain with these plasmids confirmed our initial result (data not shown). These experiments demonstrate that the comple-

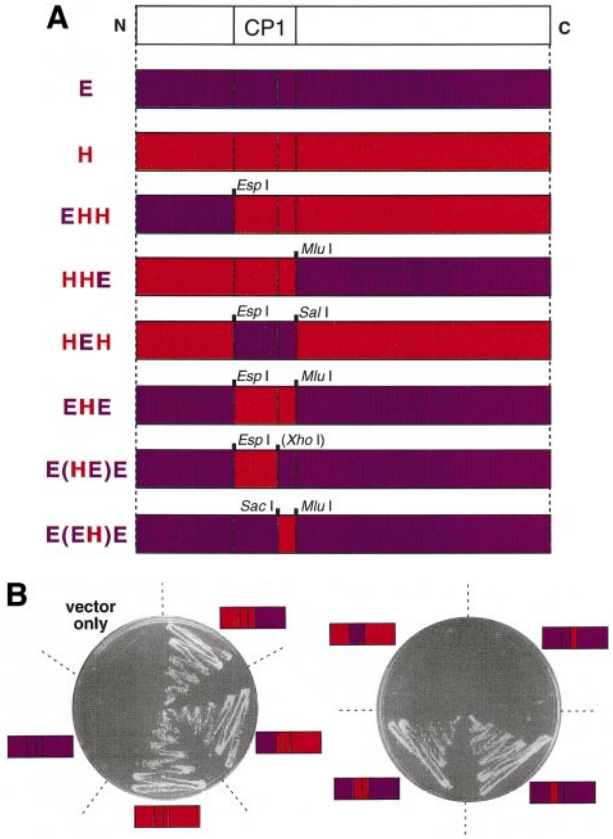

Fig. 2. Construction of six chimeric TyrRSs and their *in vivo* activities. (**A**) The natural or artificial restriction sites, which were used to prepare plasmids encoding chimeric enzymes, are shown. The introduced *Xho* I restriction site (in parentheses) was deleted by further site-directed mutagenesis to correct one amino acid change accompanied by introduction of the site. E designates *E.coli* and H designates human. (**B**) Growth at 30°C of yeast strains on 5-FOA (5-fluoroorotic acid) plates. Haploid yeast strains bearing the TyrRS gene-disrupted null allele, the pC[3]679 maintenance plasmid, and a tester plasmid encoding *E.coli*, human, or chimeric TyrRSs, were streaked on 5-FOA plates.

mentation depends on the presence of the gene for the human protein.

Complementation of chimeric TyrRSs in the TyrRS-deficient yeast strain

To address the importance of residues which make yeast- or human-specific tRNA contacts, we constructed chimeric EHH, HHE and HEH TyrRSs (Figure 2A). For example, the EHH protein consists of the N-terminal one-quarter of the *E.coli* (E) enzyme joined to the C-terminal three-quarters of the human (H) protein. Conversely, HHE has a C-terminal portion of the *E.coli* enzyme joined to a large N-terminal portion of the human enzyme.

Expression of chimera EHH and HHE TyrRSs rescued the lethal phenotype of the null strain on 5-FOA (Figure 2B). On the other hand, chimera HEH TyrRS did not complement the yeast null strain. Because the 'E' segment of the HEH enzyme spans CP1, these results suggested the importance of the CP1 region of the human enzyme for complementation by the EHH and HHE enzymes. Therefore, we next prepared chimera EHE, in which only the CP1 region of *E.coli* TyrRS is replaced by that of the human enzyme. This chimera also complemented the yeast null strain, suggesting that residues in CP1 of human TyrRS are responsible for complementation of the yeast null strain.

To delineate the region within CP1 of the human enzyme that is responsible for conferring complementation by the bacterial protein, we first examined a sequence alignment of the CP1 region of four TyrRSs (Figure 3). This alignment showed little conservation between the sequences of the CP1 regions of the bacterial and the eukaryotic enzymes, and thus offered no clear rationale for how to subdivide CP1 into small segments. Next, we considered the split gene structure of the closely related tryptophanyl-tRNA synthetase (Frolova *et al*., 1993). The genomic sequence of the human gene for TyrRS has not been determined. In the gene for human tryptophanyl-tRNA synthetase, an intron is inserted near the C-terminal end of CP1 at the position equivalent to that marked with an arrow in Figure 3 (Garret *et al*., 1991; Doublié *et al*., 1995; Landès *et al*., 1995; K.Wakasugi and P.Schimmel, unpublished data). This position divides CP1 of human TyrRS into a 39 amino acid N-terminal piece and a C-terminal segment of just 15 amino acids. In the alignment of the human and *E.coli* sequences of CP1, only four identities are present in the 39 amino acid N-terminal segment, while there are at least five identities in the 15 amino acid C-terminal peptide segment.

Chimeras were constructed that correspond to E(HE)E and E(EH)E, where the parentheses enclose the CP1 region which had been divided between the longer (39 amino acid) N-terminal and shorter (15 amino acid) C-terminal piece of CP1. We found that the E(HE)E, but not the E(EH)E, chimeric enzyme could complement the yeast null strain (Figure 2B). These observations show that the highly diverged 39 amino acid peptide segment of human CP1 is sufficient to confer the complementation phenotype on the *E.coli* enzyme.

In vitro aminoacylation assays with wild-type and chimeric synthetases

The complementation results imply that human and chimeric EHH, HHE, EHE, and E(HE)E TyrRSs are active on *S.cerevisiae* tRNA. To investigate these activities *in vitro*, we tested the ability of the proteins isolated from the rescued yeast deletion strain to charge yeast tRNA. As shown in Figure 4A, yeast tRNA is aminoacylated by human, but not by *E.coli*, TyrRS. On the other hand, *E.coli* $tRNA^{Tyr}$ is aminoacylated by the *E.coli*, but not the human, enzyme. These *in vitro* data demonstrate the species-specific aminoacylation by TyrRSs and are consistent with the complementation data shown in Figure 2B.

We next investigated the aminoacylation activities of the purified *E.coli* enzymes which were transplanted with the entire CP1 from the human enzyme, or with the 39 amino acid N-terminal section of human CP1. These two chimeric proteins [EHE and E(HE)E, respectively], were comprised mostly of sequences from *E.coli* TyrRS, complemented the null strain (Figure 2B). Consistent with this observation, and in contrast with the *E.coli* enzyme, the EHE and E(HE)E enzymes were active on yeast tRNA (Figure 4B). The k_{cat}/K_m values for the chimeric EHE and E(HE)E enzymes with yeast $tRNA^{Tyr}$ were reduced by ~5-fold compared with that for the human enzyme with

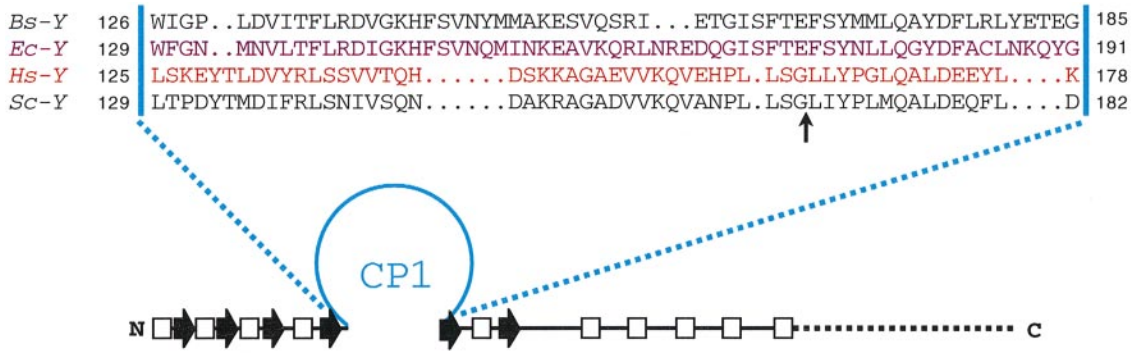

Fig. 3. Schematic diagram of the secondary structure of TyrRS and the partial sequence alignment of connective polypeptide 1. Elements of the secondary structure based on the three-dimensional structure of *B.stearothermophilus* TyrRS (Brick *et al*., 1989) are shown below. The locations of α-helices and β-strands are delineated with cylinders and solid arrows, respectively. The secondary structure for the C-terminal end of the protein was not resolved in the crystal structure and is drawn here with a broken line. The sequence alignment of *B.stearothermophilus* (Bs) (Winter *et al*., 1983), *E.coli* (Ec) (Barker *et al*., 1982), human (Hs) cytoplasmic (Ribas de Pouplana *et al*., 1996; Kleeman *et al*., 1997), and *S.cerevisiae* (Sc) cytoplasmic (Chow and RajBhandary, 1993) TyrRSs is also given in the CP1 region. Numbers on the left and right of the sequences correspond to those at the beginning and end of the CP1 sequences, respectively. Gaps in the sequences are indicated by dots. The location of an intron in the gene for the closely related human tryptophanyl-tRNA synthetase is marked with an arrow.

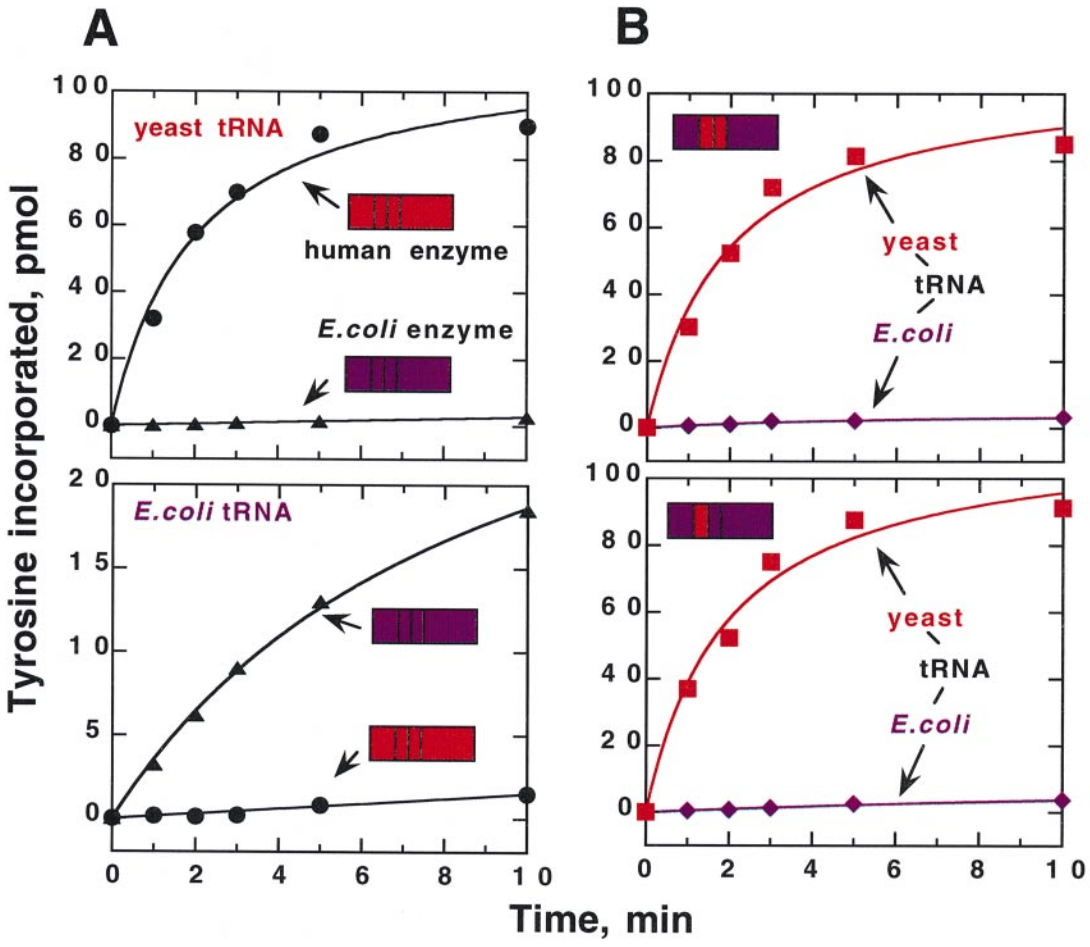

Fig. 4. Aminoacylation of *E.coli* or yeast tRNA by *E.coli*, human and chimeric TyrRSs. (**A**) Aminoacylation of yeast or *E.coli* tRNA with the human or *E.coli* enzyme. (**B**) Aminoacylation of yeast or *E.coli* tRNA with the EHE and E(HE)E chimeric enzymes.

yeast $tRNA^{Tyr}$. Interestingly, these two chimeric *E.coli* proteins were now specific for yeast tRNA. Thus, introduction of the N-terminal piece of CP1 conferred charging specificity for the eukaryote tRNA while simultaneously conferring loss of activity for its natural *E.coli* substrate.

To further characterize the transplanted *E.coli* proteins that are active on yeast tRNA, we compared their thermal stabilities with those of the wild-type proteins. In the three dimensional structure of *B.stearothermophilus* TyrRS (Brick *et al*., 1989), the C-terminal more conserved region of CP1 interacts with some of the residues in the Rossmann fold domain. We imagined that the substitution of these residues with those from the human protein could weaken the structure and that this could be most sensitively tested

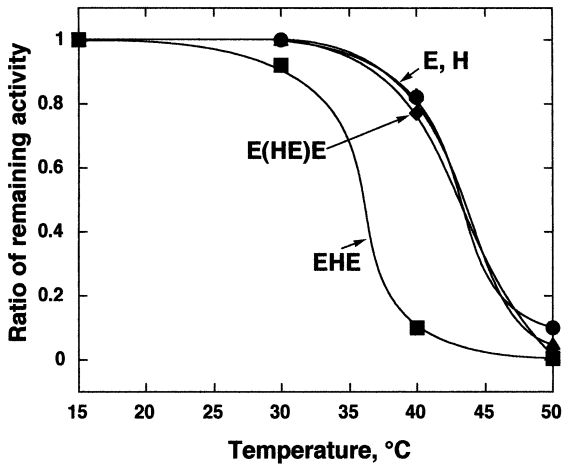

Fig. 5. Thermal stability of *E.coli*, human, EHE or E(HE)E TyrRSs. Portions of a cellular extract containing the indicated protein were heated at varying temperatures for 30 min. After cooling to ambient temperature (23°C) and clarification of the solution by centrifugation, the charging activity of the supernatant was measured. For each point on the curves, a fresh sample was used.

by investigating the effect of temperature on activity. The profiles of the activity remaining after exposure for 30 min at a series of temperatures showed a co-operative melting characteristic of an unfolding transition. The *E.coli*, human and E(HE)E proteins had similar profiles (Figure 5). In contrast the EHE enzyme lost activity at a significantly lower temperature. These observations are consistent with a weakening in the EHE enzyme of the interactions of the C-terminal region of the CP1 segment with the catalytic core in the nucleotide binding fold.

Aminoacylation of human microhelix^{Tyr}

To narrow down the site in tRNA where residues in CP1 would change the species-specificity, we concentrated on the acceptor stem, by analogy with the related class I glutaminyl-tRNA synthetase whose CP1 domain provides for contacts with the acceptor helix of tRNA^{Gln}. We were particularly interested in the 1:72 base pair, because this pair is known to be important for aminoacylations by eukaryote TyrRSs (Lee and RajBhandary, 1991; Quinn *et al*., 1995). For this purpose, we took advantage of chemically synthesized RNA microhelix oligonucleotide substrates. For at least ten tRNA synthetases, microhelices which reconstruct the acceptor stem of a tRNA are substrates for specific aminoacylations (Frugier *et al*., 1994; Hamann and Hou, 1995; Martinis and Schimmel, 1995). In the specific case of TyrRS, the eukaryotic *Pneumocystis carinii* enzyme has previously been shown to charge microhelices with tyrosine, with a specificity of charging that is sensitive to the nature of the 1:72 base pair (Quinn *et al*., 1995).

We synthesized RNA microhelices based on the sequence of human tRNA^{Tyr} and, in addition, a G1:C72 variant (Figure 6A). As shown in Figure 6B, human TyrRS charged human wild-type microhelix^{Tyr}. The single change from C1:G72 to G1:C72 of human microhelix^{Tyr} abolished aminoacylation by the human enzyme, suggesting that the C1:G72 base pair is essential for aminoacylation with human TyrRS. This result is consistent with previous studies showing the importance of the C1:G72 base pair

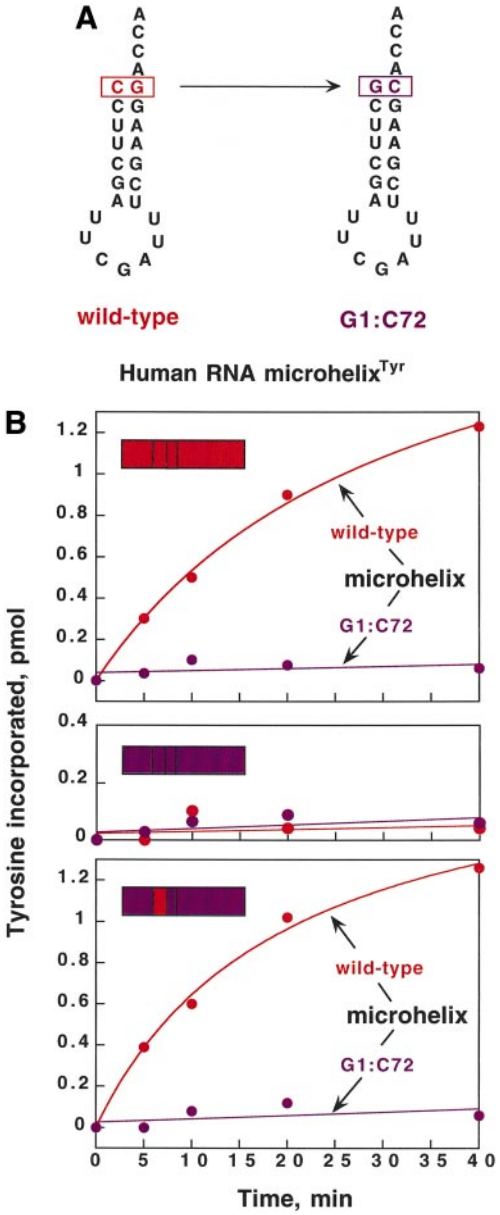

Fig. 6. Aminoacylation of RNA microhelices whose sequences are based on the acceptor stem of human tRNA^{Tyr}. (**A**) Microhelices that were used in the aminoacylation assays. The 1:72 base pair is boxed. (**B**) Aminoacylation of wild-type and G1:C72 microhelix^{Tyr} by *E.coli*, human or chimeric E(HE)E TyrRS. These assays were performed as described in the Materials and methods.

for other eukaryote TyrRSs (Lee and RajBhandary, 1991; Quinn *et al*., 1995). Reflecting the species-specificity seen with the full tRNAs (Figure 4), the human wild-type microhelix^{Tyr} was not aminoacylated with *E.coli* TyrRS.

We next investigated the recognition of eukaryote cytoplasmic microhelix^{Tyr} by the chimeric E(HE)E TyrRS. As shown in Figure 6B, the E(HE)E enzyme efficiently charged human wild-type microhelix^{Tyr} but showed no activity toward the mutated substrate (G1:C72). These results show that the 1:72 base pair is critical for microhelix aminoacylation by the chimeric enzyme as well as by the human enzyme. They also show that the 39 amino acid peptide segment in the CP1 region of human TyrRS is critical for recognition of the C1:G72 base pair.

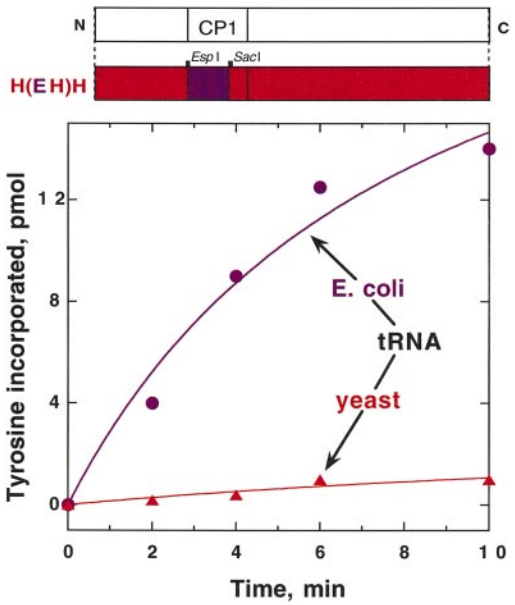

Fig. 7. Aminoacylation of *E.coli* or yeast tRNA with the H(EH)H chimeric TyrRS.

Reversing species-specific aminoacylation by a reverse peptide transplant

As shown in Figure 4A, *E.coli* and human TyrRSs show clear prokaryote versus eukaryote aminoacylation specificity. Given that a simple peptide transplant from the human to the *E.coli* protein enabled the *E.coli* enzyme to charge yeast tRNA, we wondered whether a reverse transplant would change the species-specific aminoacylation of the human enzyme. For this purpose, we constructed the chimeric H(EH)H TyrRS, which is the human enzyme transplanted with the 39 amino acid N-terminal section of *E.coli* CP1. This chimera is the reverse of the E(HE)E chimera which charges the eukaryote tRNA substrate.

We found that the H(EH)H chimera was stable and we were able to investigate its aminoacylation activity *in vitro*. The purified chimera H(EH)H enzyme charged *E.coli* $tRNA^{Tyr}$ but, in contrast to the human enzyme (HHH), had little activity on yeast tRNA (Figure 7). These data show that the peptide element from CP1 acts as a modular element that can be moved from one tyrosine enzyme to another to confer distinct charging phenotypes.

Discussion

This is the first report of a switch in species-specific aminoacylation achieved by manipulation of a tRNA synthetase. Previous work had shown that specific nucleotide differences between a prokaryote and eukaryote tRNA could account for species-specific acylation (Sampson *et al.*, 1989; Hipps *et al.*, 1995; Quinn *et al.*, 1995). In the present work, a simple peptide transplant from one synthetase into another was sufficient to switch species-specificity. We were surprised that a 39 amino acid peptide from the human enzyme, having only four sequence identities to the bacterial protein, could be accommodated into the structure of the *E.coli* enzyme without further mutagenesis. The success of these experiments suggests that the two peptide sequences from CP1, while highly dissimilar, are structurally homologous. Remarkably, the sequence differences are such as to allow the human peptide to be incorporated into the *E.coli* protein, and moreover, to allow the *E.coli* peptide to be incorporated into the human protein, without disruption of the surrounding structure. Consistent with these observations, the portion of CP1 that was effective in these experiments is a part that protrudes from the body of the enzyme and has few intramolecular contacts (Brick *et al.*, 1989).

In the related class I isoleucyl-, leucyl- and valyl-tRNA synthetases, the CP1 insertions are much larger (250–300 amino acids) than the CP1 insertions of TyrRSs (Burbaum and Schimmel, 1991; Schimmel *et al.*, 1992; Landès *et al.*, 1995). In addition to providing for acceptor helix contacts, in the three former enzymes the CP1 insertion harbors the activity for RNA-dependent amino acid recognition as manifested by the editing of mischarged and misactivated amino acids (Lin *et al.*, 1996; Hale *et al.*, 1997). However, TyrRS has no known editing activity (Fersht *et al.*, 1980). Probably the smaller size of the CP1 domain of TyrRS is related to its not having an editing activity. Thus, the 60 amino acids that make up CP1 of TyrRS may be designed solely for acceptor helix interactions. While no co-crystal of TyrRS with $tRNA^{Tyr}$ is available, our data with microhelix substrates (Figure 6) provide strong support for the role of CP1 in determining the specificity of acceptor helix recognition in this system.

The region of CP1 that was replaced by human sequences was one of those identified in a recent study to undergo changes in evolution that were correlated with changes in species-specific aminoacylation phenotype (Nair *et al.*, 1997). Structural modeling of the complex of the synthetase from *B.stearothermophilus* and $tRNA^{Tyr}$, based on the structure of the non-complexed enzyme and on mutagenesis data (Bedouelle and Winter, 1986), suggested that that C72 of the G1:C72 base pair and neighboring nucleotides are surrounded by some of the residues in this region of CP1 (amino acids 126–166 of the *B.stearothermophilus* enzyme) (Labouze and Bedouelle, 1989; Bedouelle, 1990; K.Wakasugi and P.Schimmel, unpublished data). These include the side chains of Glu152 and Gln155, for example. While these models of the complex are inherently limited, our data are consistent with the assignment of the region from 126–166 as being critical for the specificity of acceptor stem interactions.

The *in vivo* complementation results (Figure 2B) show that the *E.coli* enzyme containing a transplant of just 39 amino acids from the human enzyme can serve as the sole source of TyrRS in yeast. This experiment emphasizes the high specificity of aminoacylation conferred by the transplant, because misacylations are known to be associated with toxic phenotypes (Vidal-Cros and Bedouelle, 1992; Bedouelle *et al.*, 1993).

While for many tRNA synthetases the anticodon region is important for recognition and specificity (Schulman, 1991; Saks *et al.*, 1994), unfavorable acceptor stem interactions can dominate over favorable anticodon contacts so that aminoacylation is not observed. For example, methionyl-tRNA synthetase is known to have a strong dependence on the anticodon for aminoacylation, so that transfer of the CAU anticodon of $tRNA^{Met}$ into other tRNAs can confer aminoacylation with methionine (Schulman and Pelka, 1988). However, as shown by

Senger *et al*. (1992), the success of these experiments depends on having a host tRNA with an acceptor stem that is compatible with methionyl-tRNA synthetase. Thus, even when anticodon interactions are favorable, minimal substitutions in the acceptor stem are sufficient to block aminoacylation (Lee *et al*., 1992; Senger *et al*., 1992). Indeed, our results suggest that the anticodon interaction itself is not sufficient to confer charging of yeast $tRNA^{Tyr}$ with *E.coli* TyrRS. Instead, acceptor stem contacts with a small region of CP1 are also required and the mismatch of those contacts is sufficient to give rise to species-specific aminoacylation.

While the active site domain with the CP1 insertion has been elucidated for *B.stearothermophilus* TyrRS, the structure of the anticodon binding domain has not been determined. By analogy with other class I enzymes, this domain is at the C-terminal end of the polypeptide and detailed models of the docking of TyrRS with $tRNA^{Tyr}$ support the idea that this domain makes contact with the anticodon (Bedouelle, 1990). Therefore, in the E(HE)E chimeric protein, the anticodon binding domain is comprised entirely of sequences from *E.coli* TyrRS. Our data suggest that these *E.coli* TyrRS sequences accommodate the anticodon of either eukaryote or prokaryote $tRNA^{Tyr}$. Because the anticodons are conserved through evolution, this result might be anticipated, at least based on simple considerations. However, species-specific variations in base modifications occur and these modifications frequently occur in the anticodon and its surrounding nucleotides. In *E.coli* the major isoacceptor of $tRNA^{Tyr}$ has a QUA anticodon, where the wobble position Q is quenosine (a modified G). The Q-base is absent from yeast $tRNA^{Tyr}$ where the major isoacceptor has a GΨA anticodon (Ψ is pseudouridine). Thus, if the wobble position is a key contact for TyrRS, then atoms common to G and Q are likely to be the ones involved in contacts with the enzyme.

The N-terminal catalytic domain with the Rossmann nucleotide binding fold is believed to be the primordial, core tRNA synthetase (Schimmel *et al*., 1993; Schimmel, 1995). This domain is responsible for amino acid activation. Insertions such as the idiosyncratic (to the synthetase) CP1 polypeptides are thought to be later additions to the synthetase structure, and enabled the evolving enzymes to dock the 3′-end of minihelix-like RNA substrates near the activated amino acid. These insertions had to be accommodated in a way that would not perturb the active site. Our data show that this region of a class I tRNA synthetase is indeed amenable to an exchange of sequences without disruption of the active site. Thus, this region of the enzyme may be ideally suited for co-adaptations to changing sequences of tRNA acceptor stems in evolution.

Earlier work described a switch in anticodon recognition by a swap of a designed 10 amino acid peptide between two class I tRNA synthetases from *E.coli* (Auld and Schimmel, 1995, 1996). The peptide in this instance was located in the second major domain that is joined to the active site containing class-defining domain. This second domain is not conserved amongst all class I enzymes and, like the CP1 insertion into the catalytic domain, may have been added later in evolution to facilitate interactions with specific tRNAs. Thus, the present and the earlier example collectively demonstrate that two major points for synthetase-tRNA interactions—the acceptor stem and the anticodon triplet—are interactions controlled by small, idiosyncratic peptide elements which are not part of the core catalytic structure. This situation, in turn, provides an efficient way to diversify tRNA specificity around a core catalytic structure.

Materials and methods

Construction of maintenance plasmid pC^{3}679 for the S.cerevisiae null strain

Plasmid p13Gen containing *S.cerevisiae* cytoplasmic TyrRS gene was kindly provided by Professor U.RajBhandary (Chow and RajBhandary, 1993). Plasmid p13Gen is a genomic DNA clone of yeast cytoplasmic TyrRS gene isolated from a *S.cerevisiae* genomic DNA library. The ORF of the gene and its upstream 5′-end containing the promoter was obtained by the polymerase chain reaction (PCR), using a series of specially designed primers. The gene was cloned into the vector backbone of pQB173, a CEN plasmid with a selectable *URA3* marker, to yield plasmid pC^{3}679. This plasmid was used to maintain the null strain.

Construction of the null strain

The 5′- and 3′- non-coding flanking regions of *S.cerevisiae* cytoplasmic TyrRS gene were recovered by PCR and ligated, respectively, to the 5′- and 3′-ends of a *S.cerevisiae TRP1*-containing DNA fragment. This DNA fragment was isolated from plasmid YDp-W (Berben *et al*., 1991). The 5′-flanking sequences of the TyrRS gene extended from –430 to –6 (relative to +1 being the start of the coding sequence), while the 3′-non-coding sequences included the last 43 nucleotides of the ORF and extended from the first nucleotide (+1) after the TAA stop codon to +419. The entire construct was cloned into a pT7Blue T-Vector (Novagen, Madison, WI) to give plasmid pYY1. This plasmid has a selectable Trp marker from the *TRP1* insertion.

To delete the coding sequence of the TyrRS gene and replace it with an insertion of *TRP1*, *S.cerevisiae* diploid strain Y93 (also called FY83; *MAT***a**/α, *lys2-128/lys2-128, leu2*Δ*1/leu2*Δ*1, ura3-52/ura3-52, trp1*Δ*63/trp1*Δ*63*, obtained from Prof. F.Winston, Harvard Medical School), was transformed with linearized plasmid pYY1. Transformants were selected for a Trp^+ phenotype. PCR analysis established that one of the two copies of coding sequence of the TyrRS gene in the diploid had been deleted and replaced with *TRP1*. This heterologous diploid strain was designated QBY376. The diploid strain was first transformed with the maintenance plasmid (pC^{3}679), and then set for sporulation and tetrad dissection to generate a haploid strain QBY374 (*MAT*α, *lys2-128*Δ, *leu2*Δ*1, ura3-52, trp1*Δ*63, tyrrs*::*TRP1/pC*^{3}*679*).

Preparation of tester plasmids for human, E.coli and chimera TyrRSs

Expression vector pQB169 is a derivative of YEplac181 (Gietz and Sugino, 1988) bearing *LEU2* and contains the constitutive ADH promoter and the ILS1 transcriptional terminator. This vector was used to express in yeast the *E.coli*, human and chimeric TyrRSs. For human TyrRS, we prepared a 'truncated' enzyme, in which the coding sequence of the extra human-specific C-terminal domain (Kleeman *et al*., 1997) of native human TyrRS (528 amino acids) was removed. This domain is not homologous to that of any tRNA synthetase and is dispensable for aminoacylation (K.Wakasugi and P.Schimmel, unpublished data). The truncated enzyme has an extra six amino acids (GTELEF) that are attached to the C-terminus. It consists of 406 amino acids, which is almost the same length as the *E.coli* enzyme (424 amino acids). We confirmed that the extra 6 amino acids do not disrupt the activity of the human truncated enzyme *in vivo* or *in vitro*.

To express *E.coli* TyrRS with a C-terminal tag of six histidine residues in yeast, the gene encoding $(His)_6$ tag was introduced at the 3′-terminus of *tyrS* in the expression vector by PCR.

The tester plasmids encoding the seven chimera TyrRSs were constructed by using the natural or artificial restriction sites that are shown in Figures 2A and 7 together with mutagenic primers and PCR.

Media

Yeast cells were grown in YPD (rich medium + glucose) or defined synthetic medium supplemented with 2% glucose as a carbon source. The phenotype of yeast was always checked on YPD, SC (synthetic complete medium) minus lysine, SC minus leucine, SC minus uracil, SC minus tryptophan, SC minus histidine, SC minus adenine, and YPG (rich medium + glycerol) plates.

Fractionations of wild-type and chimeric TyrRSs

Cultures of yeast cells carrying plasmid pQB169 encoding wild-type, chimeric EHE or E(HE)E TyrRS from 5-FOA plates were harvested and resuspended in buffer A (25 mM Tris–HCl pH 7.5, 10 mM β-mercaptoethanol, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride). Cells were disrupted with a French press (SLM AMINCO Instruments, Rochester, NY) using a cell with a pressure of 20 000 atmospheres. The cell debris was removed by centrifugation at 18 000 *g* for 20 min. The soluble fraction was made 30% saturated with ammonium sulfate and centrifuged to remove insoluble material. The solution was adjusted to 70% saturated ammonium sulfate, and the 30–70% precipitate was recovered. The pellet was resuspended in buffer B (50 mM Tris–HCl pH 7.5, 10 mM β-mercaptoethanol, 5 mM $MgCl_2$, 1 mM phenylmethylsulfonyl fluoride), and was dialyzed against buffer B by using a dialysis membrane (Spectra/Por$^{\text{â}}$ 7 membrane, molecular weight cut off, 50 000). The solution was applied to a DEAE-Sephacel (Pharmacia, Piscataway, NJ) column, and the column was washed with buffer B and eluted with 350 mM KCl in buffer B. The fractions were analyzed by uv absorbance and enzymatic activity. The peak fractions were pooled, concentrated with Centricon-10 (Amicon, Beverly, MA) and dialyzed against buffer C (150 mM Tris–HCl pH 7.5, 150 mM KCl, 10 mM $MgCl_2$, 20 mM β-mercaptoethanol).

$(His)_6$-tagged *E.coli* TyrRS was purified on a Ni-NTA (nickel-nitrilotriacetic acid) affinity column (Qiagen, Santa Clarita, CA) from the supernatant of lysed cells. For detection of the $(His)_6$-tagged protein, Ni-NTA conjugate (Qiagen), which consists of Ni-NTA coupled to calf intestinal alkaline phosphatase, was used.

The chimeric H(EH)H TyrRS was expressed from a multicopy plasmid in yeast, using a construction to be described (K.Wakasugi and P.Schimmel, unpublished data), and was purified as described above.

Aminoacylation assays

Aminoacylation activity was assayed at ambient temperature (~23°C) in the following buffer: 150 mM Tris–HCl (pH 7.5), 150 mM KCl, 10 mM $MgCl_2$, 20 mM β-mercaptoethanol, 4 mM ATP, 20 μM tyrosine [2 μM [^{3}H]tyrosine (Amersham, Arlington Heights, IL)]. The reactions were initiated by adding enzyme samples (10–100 nM) to the buffer including 2 μM purified *E.coli* $tRNA^{Tyr}$, which was purchased from Sigma (St. Louis, MO) or 100 μM crude Brewer's yeast tRNA (Boehringer Mannheim, Indianapolis, IN). After incubation, the reaction was stopped by spotting the reaction mixture onto a Whatman 3MM paper filter and immediately placing the filter disc into cold 5% trichloroacetic acid. The filters were washed three times with cold 5% trichloroacetic acid, washed twice with ethanol and once with ether, and subsequently subjected to scintillation counting.

RNA microhelices based on the acceptor stem of human tRNA were chemically synthesized on a nucleic acids synthesizer (Gene Assembler Plus, Pharmacia, Piscataway, NJ) using phosphoramidites from ChemGene (Waltham, MA). The synthetic oligonucleotides were deprotected and purified as described (Musier-Forsyth *et al*., 1991). These substrates were used in assays at concentrations of 100–200 μM. Prior to assays, all RNA substrates were heated at 70°C for 2 min and reannealed at the ambient temperature for 30 min. The concentrations of the enzymes were 5–15 μM for assays with microhelices.

Acknowledgements

We thank Drs E.Fayelle Whelihan and Fariba Houman for helpful suggestions. This work was supported by Grant Number GM23562 from the National Institutes of Health. K.W. was supported by the JSPS Research Fellowships for Young Scientists and the JSPS Postdoctoral Fellowships for Research Abroad.

References

Auld,D.S. and Schimmel,P. (1995) Switching recognition of two tRNA synthetases with an amino acid swap in a designed peptide. *Science*, **267**, 1994–1996.

Auld,D.S. and Schimmel,P. (1996) Single sequence of helix-loop peptide confers functional anticodon recognition on two tRNA synthetases. *EMBO J.*, **15**, 1142–1148.

Barker,D.G., Burton,C.J. and Winter,G. (1982) The TyrRS from *Escherichia coli*. Complete nucleotide sequence of the structural gene. *FEBS lett.*, **150**, 419–423.

Bedouelle,H. (1990) Recognition of $tRNA^{Tyr}$ by TyrRS. *Biochimie*, **72**, 589–598.

Bedouelle,H. and Winter,G. (1986) A model of synthetase/transfer RNA interaction as deduced by protein engineering. *Nature*, **320**, 371–373.

Bedouelle,H., Guez-Ivanier,V. and Nageotte,R. (1993) Discrimination between transfer-RNAs by TyrRS. *Biochimie*, **75**, 1099–1108.

Berben,G., Dumont,J., Gillipuet,V., Bolle,P.A. and Hilger,F. (1991) The YDp plasmids: a uniform set of vectors bearing versatile gene disruption cassettes for *Saccharomyces cerevisiae*. *Yeast*, **7**, 475–477.

Boeke,J.D., LaCroute,F. and Fink,G.R. (1984) A positive selection for mutants lacking orotidine-5′-phosphate decarboxylase activity in yeast: 5-fluoro-orotic acid resistance. *Mol. Gen. Genet.*, **197**, 345–346.

Brick,P., Bhat,T.N., and Blow,D.M. (1989). Structure of TyrRS refined at 2.3 Å resolution. Interaction of the enzyme with the tyrosyl adenylate intermediate. *J. Mol. Biol.*, **208**, 83–98.

Burbaum,J.J. and Schimmel,P. (1991) Structural relationships and the classification of aminoacyl-tRNA synthetases. *J. Biol. Chem.*, **266**, 16965–16968.

Carter,C.W.,Jr. (1993) Cognition, mechanism, and evolutionary relationships in aminoacyl-tRNA synthetases. *Annu. Rev. Biochem.*, **62**, 715–748.

Chow,C.M. and RajBhandary,U.L. (1993) *Saccharomyces cerevisiae* cytoplasmic TyrRS gene. Isolation by complementation of a mutant *Escherichia coli* supressor tRNA defective in aminoacylation and sequence analysis. *J. Biol. Chem.*, **268**, 12855–12863.

Cusack,S., Härtlein,M. and Leberman,R. (1991) Sequence, structural and evolutionary relationships between class 2 aminoacyl-tRNA synthetases. *Nucleic Acids Res.*, **19**, 3489–3498.

Doublié,S., Bricogne,G., Gilmore,C. and Carter,C.W.,Jr. (1995) Tryptophanyl-tRNA synthetase crystal structure reveals an unexpected homology to TyrRS. *Structure*, **3**, 17–31.

Eriani,G., Delarue,M., Poch,O., Gangloff,J. and Moras,D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature*, **347**, 203–206.

Fersht,A.R., Shindler,J.S. and Tsui,W.C. (1980) Probing the limits of protein–amino acid side chain recognition with the aminoacyl-tRNA synthetases. Discrimination against phenylalanine by TyrRSs. *Biochemistry*, **19**, 5520–5524.

Frolova,L.Y., Grigorieva,A.Y., Sudomoina,M.A. and Kisselev,L.L. (1993) The human gene encoding tryptophanyl-tRNA synthetase: interferon-response elements and exon-intron organization. *Gene*, **128**, 237–245.

Frugier,M., Florentz,C. and Giegé,R. (1994) Efficient aminoacylation of resected RNA helices by class II aspartyl-tRNA synthetase dependent on a single nucleotide. *EMBO J.*, **13**, 2219–2226.

Garret,M. *et al*. (1991) A mammalian tryptophanyl-tRNA synthetase shows little homology to prokaryotic synthetases but near identity with mammalian peptide chain release factor. *Biochemistry*, **30**, 7809–7817.

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

Gietz,R.D. and Sugino,A. (1988) New yeast-*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six-base pair restriction sites. *Gene*, **74**, 527–534.

Hale,S.P., Auld,D.S., Schmidt,E. and Schimmel,P. (1997) Discrete determinants in transfer RNA for editing and aminoacylation. *Science*, **276**, 1250–1252.

Hamann,C.S. and Hou,Y.-M. (1995) Enzymatic aminoacylation of tRNA acceptor stem helices with cysteine is dependent on a single nucleotide. *Biochemistry*, **34**, 6527–6532.

Hipps,D., Shiba,K., Henderson,B. and Schimmel,P. (1995) Operational RNA code for amino acids: species-specific aminoacylation of minihelices switched by a single nucleotide. *Proc. Natl Acad. Sci. USA*, **92**, 5550–5552.

Hountondji,C., Dessen,P. and Blanquet,S. (1986) Sequence similarities among the family of aminoacyl-tRNA synthetase. *Biochimie*, **68**, 1071–1078.

Kleeman,T.A., Wei,D., Simpson,K.L. and First,E.A. (1997) Human TyrRS shares amino acid sequence homology with a putative cytokine. *J. Biol. Chem.*, **272**, 14420–14425.

Labouze,E. and Bedouelle,H. (1989) Structural and kinetic bases for the recognition of $tRNA^{Tyr}$ by TyrRS. *J. Mol. Biol.*, **205**, 729–735.

Landès,C., Perona,J.J., Brunie,S., Rould,M.A., Zelwer,C., Steitz,T.A. and Risler,J.L. (1995) A structure-based multiple sequence alignment of all class I aminoacyl-tRNA synthetases. *Biochimie*, **77**, 194–203.

Lee,C.P. and RajBhandary,U.L. (1991) Mutants of *Escherichia coli* initiator tRNA that suppress amber codons in *Saccharomyces cerevisiae* that are aminoacylated with tyrosine by yeast extracts. *Proc. Natl Acad. Sci. USA*, **88**, 11378–11382.

Lee,C.P., Dyson,M.R., Mandal,N., Varshney,U., Bahramian,B. and RajBhandary,U.L. (1992) Striking effects of coupling mutations in the acceptor stem on recognition of tRNAs by *Escherichia coli* Met-tRNA synthetase and Met-tRNA transformylase. *Proc. Natl Acad. Sci. USA*, **89**, 9262–9266.

Lin,L., Hale,S.P. and Schimmel,P. (1996) Aminoacylation error correction. *Nature*, **384**, 33–34.

Martinis,S.A. and Schimmel,P. (1995) Small RNA oligonucleotide substrates for specific aminoacylations. In Söll,D. and RajBhandary,U. (eds), *tRNA: Structure, Biosynthesis and Function*. American Society for Microbiology, Washington, DC, pp. 349–370.

Moras,D. (1992) Structural and functional relationships between aminoacyl-tRNA synthetases. *Trends Biochem. Sci.*, **17**, 159–164.

Musier-Forsyth,K., Scaringe,S., Usman,N. and Schimmel,P. (1991) Enzymatic aminoacylation of single-stranded RNA with an RNA cofactor. *Proc. Natl Acad. Sci. USA*, **88**, 209–213.

Nagel,G.M. and Doolittle,R.F. (1991) Evolution and relatedness in two aminoacyl-tRNA synthetase families. *Proc. Natl Acad. Sci. USA*, **88**, 8121–8125.

Nair,S., Ribas de Pouplana,L., Houman,F., Avruch,A., Shen,X. and Schimmel,P. (1997) Species-specific tRNA recognition in relation to tRNA synthetase contact residues. *J. Mol. Biol.*, **269**, 1–9.

Quinn,C.L., Tao,N. and Schimmel,P. (1995) Species-specific microhelix aminoacylation by a eukaryotic pathogen tRNA synthetase dependent on a single base pair. *Biochemistry*, **34**, 12489–12495.

Ribas de Pouplana,L., Frugier,M., Quinn,C.L. and Schimmel,P. (1996) Evidence that two present-day components needed for the genetic code appeared after nucleated cells separated from eubacteria. *Proc. Natl Acad. Sci. USA*, **93**, 166–170.

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.

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

Sampson,J.R., DiRenzo,A.B., Behlen,L.S. and Uhlenbeck,O.C. (1989) Nucleotides in yeast $tRNA^{Phe}$ required for the specific recognition by its cognate synthetase. *Science*, **243**, 1363–1367.

Schimmel,P. (1987) Aminoacyl tRNA synthetases: general scheme of structure-function relationships in the polypeptides and recognition of transfer tRNAs. *Annu. Rev. Biochem.*, **56**, 125–158.

Schimmel,P. (1995) An operational RNA code for amino acids and variations in critical nucleotide sequences in evolution. *J. Mol. Evol.*, **40**, 531–536.

Schimmel,P. and Ribas de Pouplana,L. (1995) Transfer RNA: from minihelix to genetic code. *Cell*, **81**, 983–986.

Schimmel,P., Shepard,A. and Shiba,K. (1992) Intron locations and functional deletions in relation to the design and evolution of a subgroup of class I tRNA synthetases. *Protein Science*, **1**, 1387–1391.

Schimmel,P., Giegé,R., Moras,D. and Yokoyama,S. (1993) An operational RNA code for amino acids and possible relationship to genetic code. *Proc. Natl Acad. Sci. USA*, **90**, 8763–8768.

Schulman,L.H. (1991) Recognition of tRNAs by aminoacyl-tRNA synthetases. *Prog. Nucleic Acid Res. Mol. Biol.*, **41**, 23–87.

Schulman,L.H. and Pelka,H. (1988) Anticodon switching changes the identity of methionine and valine transfer RNAs. *Science*, **242**, 765–768.

Senger,B., Despons,L., Walter,P. and Fasiolo,F. (1992) The anticodon triplet is not sufficient to confer methionine acceptance to a transfer RNA. *Proc. Natl Acad. Sci. USA*, **89**, 10768–10771.

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

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

Starzyk,R.M., Webster,T.A. and Schimmel,P. (1987) Evidence for dispensable sequences inserted into a nucleotide fold. *Science*, **237**, 1614–1618.

Steinberg,S., Misch,A. and Sprinzl,M. (1993) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **21**, 3011–3015.

Vidal-Cros,A. and Bedouelle,H. (1992) Role of residue Glu152 in the discrimination between transfer RNAs by TyrRS from *Bacillus stearothermophilus*. *J. Mol. Biol.*, **223**, 801–810.

Winter,G., Koch,G.L.E., Hartley,B.S. and Barker,D. (1983) The amino acid sequence of the TyrRS from *Bacillus stearothermophilus*. *Eur. J. Biochem.*, **132**, 383–387.

Received October 6, 1997; accepted October 23, 1997