4988–4998 *Nucleic Acids Research, 2013, Vol. 41, No. 9* *Published online 21 March 2013*
doi:10.1093/nar/gkt185

Leucine-specific domain modulates the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase

Wei Yan[1], Min Tan[1], Gilbert Eriani[2] and En-Duo Wang[1,*]

[1]State Key Laboratory of Molecular Biology, Center for RNA Research, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, the Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, PR China and [2]Architecture et Réactivité de l'ARN, Institut de Biologie Moléculaire et Cellulaire du CNRS, Université de Strasbourg, 15 rue René Descartes, 67084 Strasbourg, France

Received January 7, 2013; Revised February 26, 2013; Accepted February 27, 2013

ABSTRACT

The leucine-specific domain (LSD) is a compact well-ordered module that participates in positioning of the conserved KMSKS catalytic loop in most leucyl-tRNA synthetases (LeuRSs). However, the LeuRS from *Mycoplasma mobile* (*Mm*LeuRS) has a tetrapeptide GKDG instead of the LSD. Here, we show that the tetrapeptide GKDG can confer tRNA charging and post-transfer editing activity when transplanted into an inactive *Escherichia coli* LeuRS (*Ec*LeuRS) that has had its LSD deleted. Reciprocally, the LSD, together with the CP1-editing domain of *Ec*LeuRS, can cooperate when inserted into the scaffold of the minimal *Mm*LeuRS, and this generates an enzyme nearly as active as *Ec*LeuRS. Further, we show that LSD participates in $tRNA^{Leu}$ recognition and favours the binding of tRNAs harbouring a large loop in the variable arm. Additional analysis established that the Lys598 in the LSD is the critical residue for tRNA binding. Conversion of Lys598 to Ala simultaneously reduces the tRNA-binding strength and aminoacylation and editing capacities, indicating that these factors are subtly connected and controlled at the level of the LSD. The present work provides a novel framework of co-evolution between LeuRS and its cognate tRNA through LSD.

INTRODUCTION

Aminoacyl-tRNA synthetases (aaRS) are a large and diverse family of enzymes that catalyze the attachment of amino acids to their cognate tRNAs. Each aaRS specifically attaches its cognate amino acid to its corresponding tRNA isoacceptor. A two-step process is performed by the aaRS during aminoacylation: (i) activation of the amino acid by ATP hydrolysis to form an aminoacyl-adenylate intermediate; (ii) transfer of the aminoacyl moiety from the intermediate to the cognate tRNA isoacceptor to make the aminoacyl-tRNA (1). Based on sequence homology and the structures of the catalytic active sites, aaRSs are divided into two classes (2). Leucyl-tRNA synthetase (LeuRS) is a class I aaRS that has an active site folded to form a typical Rossmann dinucleotide-binding fold. According to evolutionary models, the primitive catalytic core of LeuRS was extended by the insertion and appendage of additional domains (also called modules) (3). Most LeuRSs carry a large insert called the connective polypeptide 1 (CP1) domain that is responsible for the amino acid-editing function. LeuRSs also exhibit tRNA-binding domains that recognize and bind $tRNA^{Leu}$ isoacceptors (4–7). A well-ordered module inserted into the catalytic domain, named the leucine-specific domain (LSD), is also found in most bacterial and some eukaryotic LeuRSs. LSD is connected to the KMSKS motif via a β-ribbon. The three-dimensional structure of the *Thermus thermophilus* LeuRS (*Tt*LeuRS) shows that the LSD contains five β-strands and two short α-helices (3,5). In comparison, the LSD of *Escherichia coli* LeuRS (*Ec*LeuRS) exhibits an additional extended β-hairpin (4). Crystal studies have also revealed that the LSD plays a critical role in positioning the conserved catalytic KMSKS loop during aminoacylation reactions (4).

Although the LSD is mainly found in prokaryotic LeuRSs, it is not highly conserved in sequence or length (3,4,8). The heterodimeric αβ-LeuRS from *Aquifex*

*To whom correspondence should be addressed. Tel: +86 21 5492 1241; Fax: +86 21 5492 1011; Email: edwang@sibcb.ac.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

aeolicus (*Aa*LeuRS) has one of the largest LSDs, and this also serves to split the enzyme into two subunits (3,9). The LSD can also be missing completely in some species, such as *Bacillus subtilis* or *Mycoplasma mobile*, in which LeuRS is remarkable for the complete absence of a CP1-editing domain (8,10). In addition, sequence alignment has shown that the LSD in the LeuRS from *M. mobile* (*Mm*LeuRS) is replaced by the tetrapeptide ^{398}GKDG^{401} (4,10).

A recent study revealed that the CP1 domain and LSD of *Ec*LeuRS both undergo large rotations when tRNA shifts from the synthetic site to the editing active site (Figure 1A) (4). The CP1 domain rotates by 12° to open up a passage for the translocation of the 3′ end of the tRNA, while the more dynamic LSD, together with the adjacent catalytically crucial KMSKS loop, is rotated by about 33° between the aminoacylation and editing conformations. Consistently, both the CP1 domain and LSD positions move by about 19° and 35° in the *Tt*LeuRS when comparing the aminoacylation and editing conformations (5). Another study indicated that the tRNA-triggered conformational rearrangement leads to inter-domain communication between the editing and synthetic domains of *Ec*LeuRS (11). All these data strongly suggest that both the CP1 domain and LSD are functionally connected and cooperate during the aminoacylation and editing reactions.

tRNA^{Leu}, together with tRNA^{Ser} and tRNA^{Tyr}, are class II tRNAs which are characterized by the presence of both a long variable stem and loop (12). Interactions between LeuRS and tRNA^{Leu} have been extensively investigated, and the conserved A73 nucleotide is considered to be the main element for identification. The amino acid-accepting end (CCA_{76}) of *Ec*tRNA^{Leu} is critical for both the aminoacylation and the editing processes (13). The tertiary interactions between the D- and T-loops that determine the tRNA folding are additional critical elements of the leucine identity (14). In addition, tRNA elements that are critical during the editing process have been detected in the anticodon arms of tRNA^{Leu} from *A. aeolicus* and *Saccharomyces cerevisiae* (15,16).

In this present study, the LSD of *Ec*LeuRS was substituted with the minimal tetrapeptide linker GKDG from the *Mm*LeuRS, and this created a chimeric mutant named *Ec*LeuRS-GKDG. In addition, the LSD and CP1 domain of *Ec*LeuRS were inserted into the minimal *Mm*LeuRS, and this produced another chimera termed *Mm*LeuRS-CP1/LSD. By comparing the catalytic performances of these chimeric enzymes, we found that the

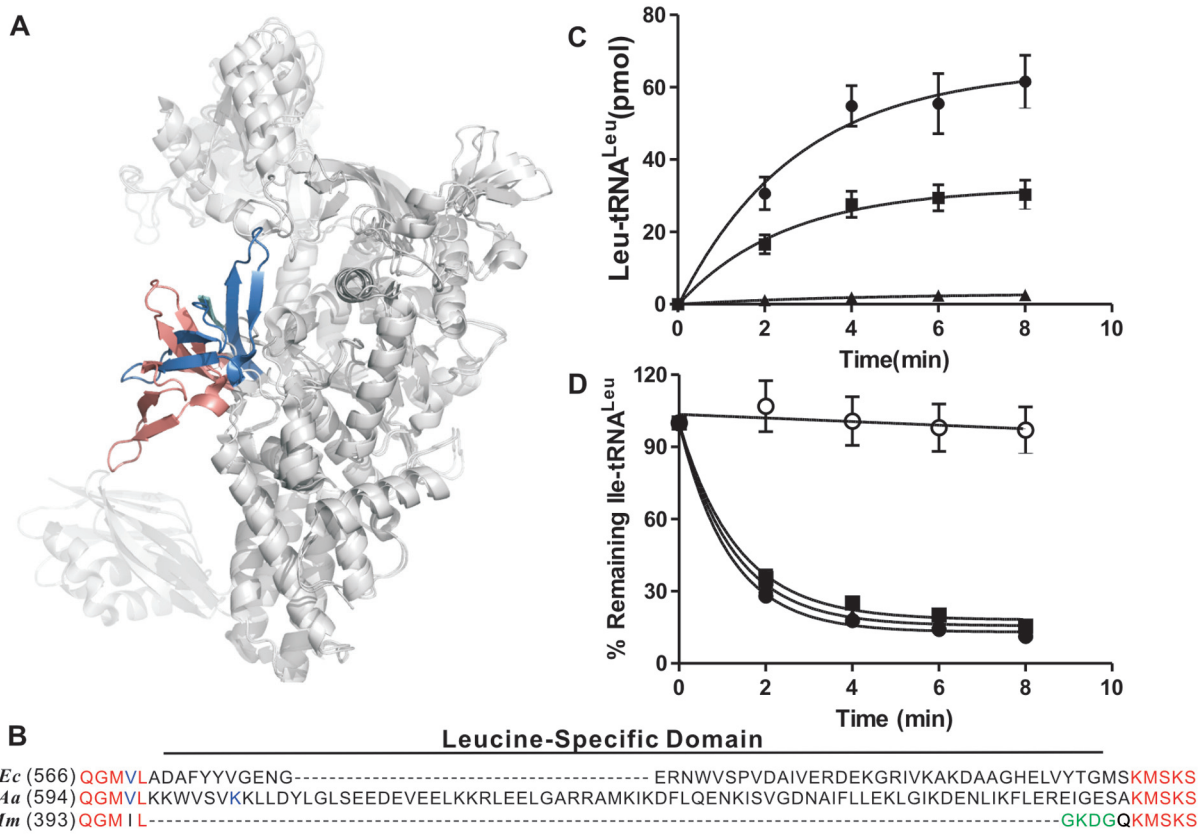

Figure 1. Impact of LSD mutations on aminoacylation and editing of *Ec*LeuRS. (**A**) Three-dimensional view of *Ec*LeuRS showing the LSD motion in the aminoacylation (blue) and editing state (red) (PDB entry 4AQ7 and 4ARC). (**B**) Sequence alignment based on structural elements of the LeuRS LSD; the tetrapeptide linker is highlighted in green. *Ec, Escherichia coli*; *Aa, Aquifex aeolicus; Mm, Mycoplasma mobile*. (**C**) Aminoacylation of 10 μM *Ec*tRNA^{Leu} by 5 nM of *Ec*LeuRS (black circle), *Ec*LeuRS-GKDG (black square) and *Ec*LeuRS-AAAA (black triangle). (**D**) Hydrolysis of 1 μM [^{3}H]-Ile-*Ec*tRNA^{Leu} by 5 nM of *Ec*LeuRS (black circle), *Ec*LeuRS-GKDG (black square), *Ec*LeuRS-AAAA (black triangle) and no enzyme (open circle) .

LSD is essential for neither aminoacylation nor editing functions of LeuRSs. However, LSD participates in tRNA binding, and it is able to discriminate between different $tRNA^{Leu}$ isoacceptors. Indeed, the LSD acts as a sensor that can measure the size of the V-arm loop and identify the nucleotide at position 20 of the $tRNA^{Leu}$. These results highlight the role of the LSD during tRNA recognition and suggest that interactions between the LSD and $tRNA^{Leu}$ might favour binding in both aminoacylation and editing catalytic steps. Altogether, these results emphasize the modular nature of the LSD as well as the important contribution played by the other synthetase modules in enhancing catalytic efficiency and tRNA specificity.

MATERIALS AND METHODS

Materials

L-leucine, L-norvaline (Nva), ATP, Tris-HCl buffer, $MgCl_2$ solution and dithiothreitol (DTT) were purchased from Sigma (USA). [^{3}H] L-leucine, [^{3}H] L-isoleucine and adenosine 5′-[α-^{32}P] triphosphate were obtained from PerkinElmer Life Sciences (USA). PEI Cellulose F plates for thin layer chromatography (TLC) were purchased from Merck (Germany). T4 DNA ligase and other restriction endonucleases were from MBI Fermentas (Lithuania). DEAE-Sepharose CL-6B and Superdex^{TM} 75 were purchased from GE Healthcare (USA). Ni^{2+}-NTA Superflow was purchased from Qiagen, Inc. (Germany). Plasmid pET30a was obtained from Novagen (USA), and *E. coli* strain BL21 (DE3) was from Invitrogen (USA). The expression vector pTrc99B and *E. coli* strain MT102 were gifts from Dr. J. Gangloff of the Institut de Biologie Moléculaire et Cellulaire du CNRS, Strasbourg, France.

Expression and purification of *Mm*LeuRS, *Ec*LeuRS and their mutants

The definition of the LSD in *Ec*LeuRS was based on the crystal structure of *Ec*LeuRS (PDB entry 4ARC) and sequence alignment. The LSD of *Ec*LeuRS spans from A_{571} to M_{617}. Each of the enzymes was expressed in *E. coli* BL21 (DE3) with a His_6-tag fused at the N-terminus. The enzymes were purified by affinity chromatography using Ni-NTA (Ni^{2+}nitrilotriacetate) Superflow resin, followed by gel-filtration chromatography with Superdex^{TM} 75. The final concentration was determined using a Bradford protein assay as described in the manufacturer's protocol (Bio-Rad, Hercules, CA, USA).

The genes encoding the various mutants were constructed using the KOD Plus Mutagenesis Kit (Toyobo Life Science) and confirmed by DNA sequencing (BioSun Bioscience). Insertion of the CP1 domain of *Ec*LeuRS into *Mm*LeuRS was performed as described previously (17). Insertion of the LSD of *Ec*LeuRS into *Mm*LeuRS-CP1 was performed in several steps. First, the ^{398}GKDG^{401} peptide was deleted from the *Mm*LeuRS-CP1, and then the 47 amino acid residues from the LSD of the *Ec*LeuRS (from A_{571} to M_{617}) were added progressively by five rounds of mutagenesis.

Preparation of RNA substrates

E. coli $tRNA^{Leu}_{GAG}$ (*Ec*$tRNA^{Leu}$) with an accepting activity of 1400 pM/A_{260} was prepared from overproducing strains constructed in our laboratory (18). *In vitro* transcription of *Mm*$tRNA^{Leu}$ and mutated derivatives was performed as described previously (17). The accepting activities of the *Mm*$tRNA^{Leu}_{UAA}$ and *Mm*$tRNA^{Leu}_{UAG}$ transcripts and the mutated derivatives (A6G, C20U, C67U, V-arm-4 nt, V-arm-5 nt, C20U + V-arm-5 nt, A6G + V-arm-5 nt) were all between 1200–1500 pM/A_{260}. [^{3}H]Ile-*Ec*$tRNA^{Leu}$, [^{3}H]Ile-*Mm*$tRNA^{Leu}$ and its mutants were obtained using the editing-deficient *Ec*LeuRS-Y330D mutant as described previously (19).

tRNA charging, misacylation and deacylation

Aminoacylation activities of *Mm*LeuRS, *Ec*LeuRS and their mutants were measured in a reaction mixture containing 100 mM Tris-HCl (pH 7.8), 30 mM KCl, 12 mM $MgCl_2$, 0.5 mM DTT, 4 mM ATP, 10 μM $tRNA^{Leu}$, 40 μM [^{3}H]Leu (11 Ci/mM) and the enzyme (5 nM *Ec*LeuRS or 20 nM *Mm*LeuRS and their mutants). Reactions were carried out at 30°C for *Mm*LeuRS and the mutants, while *Ec*LeuRS and derivatives were assayed at 37°C. For K_m determinations, tRNA concentrations ranged 0.5–30 μM. Misacylation assays were performed under similar conditions, except that 40 μM [^{3}H]Ile (30 Ci/mM; PerkinElmer) and 1 μM of enzyme were used. The deacylation reaction was measured by determining hydrolytic rates, and this was performed at 30°C in 100 mM Tris-HCl (pH 7.5), 30 mM KCl, 12 mM $MgCl_2$, 0.5 mM $MgCl_2$, 0.5 mM DTT and 1 μM [^{3}H]Ile-$tRNA^{Leu}$. Reactions were initiated with enzyme diluted to 20 nM. Because radioactive Nva is commercially unavailable, [^{3}H]Ile was used as a source to prepare mischarged $tRNA^{Leu}$.

AMP formation

The net effect of the editing reaction is the consumption of ATP. Therefore, editing can be measured through AMP formation in the presence of a non-cognate amino acid. AMP formation rates of *Mm*LeuRS, *Ec*LeuRS and their mutants were measured as described previously (19). The reaction mixture contained 100 mM Tris-HCl (pH 7.8), 30 mM KCl, 12 mM $MgCl_2$, 5 mM DTT, 5 U/ml pyrophosphatase (Roche), 3 mM ATP, 20 nM [α-^{32}P] ATP (3000 Ci/mM; PerkinElmer), 15 mM Nva and the presence or absence of 5 μM $tRNA^{Leu}$. The reaction was initiated by the addition of 1 μM *Mm*LeuRS-CP1/LSD (at 30°C), or 0.2 μM for *Ec*LeuRS and the mutant enzymes (at 37°C). At regular time intervals, aliquots of 1.5 μl were quenched in 6 μl of 200 mM sodium acetate (pH 5.0). Quenched aliquots (1.5 μl each) were spotted in duplicate on polyethyleneimine cellulose plates (PEI, Merck) that had been pre-washed with water. Separation of [^{32}P] aminoacyl-adenylate, [^{32}P]AMP and [^{32}P]ATP was performed by developing TLC plates in the presence of 0.1 M ammonium acetate and 5% acetic acid. Plates were visualized by phosphorimaging, and data were analyzed using Multi Gauge V3.0 software (Fujifilm). The grey densities of [^{32}P]AMP spots were compared

with those of known [^{32}P]ATP concentrations. Rate constants (k_{obs}) were obtained from graphs of [^{32}P]AMP formation plotted against time.

RESULTS

The LSD is not essential for aminoacylation activity and post-transfer editing

*Mm*LeuRS is an exceptionally small LeuRS that lacks both the CP1 domain and the LSD, and sequence alignment shows that these two domains are replaced by a nonapeptide linker ^{227}KEEIDGKIT^{235} and a tetrapeptide linker ^{398}GKDG^{401}, respectively (Figure 1B). Previous studies have shown that the nonapeptide linker from *Mm*LeuRS can replace the CP1 domain of *Ec*LeuRS to permit aminoacylation (17). In this present study, we examined whether the tetrapeptide GKDG from *Mm*LeuRS could replace the LSD of the *Ec*LeuRS. The resulting mutant that lacked the LSD was called *Ec*LeuRS-GKDG. The catalytic efficiency (k_{cat}/K_m) of *Ec*LeuRS-GKDG for *Ec*tRNA^{Leu} aminoacylation was just more than half of that of the native *Ec*LeuRS (Figure 1C, Table 1), indicating that the GKDG sequence of *Mm*LeuRS could functionally replace the 47 amino acid residues of the LSD in *Ec*LeuRS. In parallel, we constructed a similar mutant to contain a tetra Ala peptide instead of the GKDG insertion but the resulting mutant (*Ec*LeuRS-AAAA) was inactive in the aminoacylation reaction (Figure 1C) despite intact folding as shown by CD-spectroscopy analysis (Supplementary Figure S1). Compared with results obtained in a previous study (8), the GKDG insertion led to much better recovery of activity in *Ec*LeuRS (55% aminoacylation activity of the wild-type enzyme, Table 1). However, the *Ec*LeuRS-AAAA mutant only displayed 0.55% of the aminoacylation activity of the wild-type enzyme (Table 1), and approximately 1/6 the activity of the previously reported ΔLSD-valRStt mutant (3.5%), which was obtained by using a seven-residue sequence (VLDEKGQ) from *T. thermophilus* ValRS instead of the LSD of *Ec*LeuRS (8). These results indicate that the ^{398}GKDG^{401} of *Mm*LeuRS is a kind of minimal functional domain. In addition, both *Ec*LeuRS-GKDG and *Ec*LeuRS-AAAA exhibited intact deacylation activity for mischarged Ile-tRNA^{Leu} (Figure 1D), further proving that the native LSD does not play a critical role during the deacylation of tRNA (17).

A mutagenesis study was carried out to further explore the role in *Mm*LeuRS of the residues of the GKDG peptide. Each of the residues in the tetrapeptide was mutated to Ala separately. All the mutants displayed altered tRNA-charging activity. Moreover, substitution of the flexible Gly398 and Gly401 to rigid Pro residues severely impaired aminoacylation activity to levels comparable with a full deletion of the tetrapeptide linker (Supplementary Table S1). These data suggest that the GKDG peptide of *Mm*LeuRS plays a critical role in providing flexibility to the catalytic site.

As the two catalytic activities of *Ec*LeuRS (aminoacylation and editing) do not require the presence of the 47-amino acid LSD, this raises questions concerning the conservation of this module in most prokaryotic LeuRS during evolution.

The LSD of *Ec*LeuRS favours aminoacylation but inhibits tRNA-independent pre-transfer editing when inserted into *Mm*LeuRS-CP1

In the next experiments, a series of insertion mutants was constructed to mimic a possible evolutionary process. Chimeric proteins were constructed based on the *Mm*LeuRS scaffold. First, the LSD of *Ec*LeuRS was inserted in place of the tetrapeptide GKDG in the *Mm*LeuRS (*Mm*LeuRS-LSD). The resulting *Mm*LeuRS-LSD mutant did not exhibit any detectable aminoacylation activity (data not shown). *Mm*LeuRS-CP1 was constructed by inserting the CP1 domain of *Ec*LeuRS into *Mm*LeuRS, and this chimeric enzyme had both aminoacylation and editing activities (17). When *Mm*LeuRS-CP1 was used as a scaffold to fuse the LSD of *Ec*LeuRS into its catalytic core, the resulting chimera (*Mm*LeuRS-CP1/LSD) had comparable aminoacylation activity to the native *Mm*LeuRS but demonstrated better catalytic efficiency due to greater affinity with tRNA as indicated by a decrease in K_m (Table 1). However, the LSD insertion severely decreased the tRNA-independent pre-transfer editing of *Mm*LeuRS and *Mm*LeuRS-CP1,

Table 1. Kinetic constants of various LeuRSs determined in the aminoacylation reaction

	*Ec*tRNA$^{Leu}_{GAG}$			*Mm*tRNA$^{Leu}_{UAG}$			D factor
Enzyme	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\,\mu M^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\,\mu M^{-1}$)	
*Ec*LeuRS[a]	2.2 ± 0.14	4.9 ± 0.30	2.2	10.6 ± 1.2	4.2 ± 0.40	0.40	5.5
*Ec*LeuRS-GKDG	1.2 ± 0.10	1.4 ± 0.11	1.2	1.6 ± 0.12	2.1 ± 0.13	1.3	0.92
*Ec*LeuRS-AAAA	1.6 ± 0.14	(1.9 ± 0.21) × 10^{-2}	0.012	0.60 ± 0.052	(1.4 ± 0.12) × 10^{-2}	0.023	0.52
*Mm*LeuRS[a]	7.5 ± 0.90	1.8 ± 0.21	0.24	7.6 ± 0.80	2.0 ± 0.30	0.26	0.92
*Mm*LeuRS-CP1[a]	4.3 ± 0.40	1.1 ± 0.15	0.25	4.6 ± 0.50	1.0 ± 0.17	0.22	1.1
*Mm*LeuRS-CP1/LSD	1.3 ± 0.12	1.4 ± 0.11	1.1	1.7 ± 0.13	0.51 ± 0.049	0.30	3.7

D is the discrimination factor of the different LeuRSs for the two bacterial tRNAs, and this value was calculated as follows: $D = k_{cat}/K_m$(*Ec*tRNA$^{Leu}_{GAG}$) / k_{cat}/K_m(*Mm*tRNA$^{Leu}_{UAG}$). Kinetic constants were determined using the tRNA charging assay described in the experimental section except the concentration was 100 nM for *Ec*LeuRS-AAAA and from 0.2 to 20 μM for tRNAs. All parameters represent the average of three trials with the standard deviations indicated.
[a]Data from Tan *et al.* (17).

and the observed rate constant for AMP formation in the presence of Nva (an analogue of Leu) dropped from 0.16 and 0.12 to 0.037 s^{-1} (Table 2). In the presence of *Ec*tRNA^{Leu} and Nva, the observed rate constant of *Mm*LeuRS-CP1/LSD for AMP formation was comparable with that of *Mm*LeuRS-CP1, and the rate was 3.6-fold (0.61 s^{-1}) greater than that of *Mm*LeuRS (0.17 s^{-1}). This shows that the tRNA-dependent editing pathway became the main editing pathway of *Mm*LeuRS-CP1/LSD, contributing to 94% of the total editing activity [(0.61 − 0.037)/0.61], whereas the corresponding value in *Mm*LeuRS-CP1 was just 78% [(0.55 − 0.12)/0.55] (Figure 2C). On the other hand, when the LSD of *Ec*LeuRS was replaced by the GKDG tetrapeptide of *Mm*LeuRS to form *Ec*LeuRS-GKDG, the observed rate constant for AMP formation in the presence of Nva of the mutant was 0.75 s^{-1} compared with 0.33 s^{-1} for the native *Ec*LeuRS (Table 2), indicating that the tRNA-independent pre-transfer editing of *Ec*LeuRS-GKDG contributed much more to total editing (28%; 0.75/2.69) than that of *Ec*LeuRS (9.6%; 0.33/3.42).

Taken together, these results show that LSD recruiting restricted internal tRNA-independent pre-transfer editing

Table 2. Observed rate constants for AMP synthesis in the presence of Nva

LeuRS	tRNA^{Leu}	Rate of AMP formation k_{obs}(s^{-1})
*Ec*LeuRS[a]	-	0.33 ± 0.040
	+*Ec*tRNA$^{Leu}_{GAG}$	3.42 ± 0.51
	+*Mm*tRNA$^{Leu}_{UAG}$	2.22 ± 0.29
*Ec*LeuRS-GKDG	-	0.75 ± 0.080
	+*Ec*tRNA$^{Leu}_{GAG}$	2.69 ± 0.40
	+*Mm*tRNA$^{Leu}_{UAG}$	2.89 ± 0.47
*Mm*LeuRS[a]	-	0.16 ± 0.025
	+*Ec*tRNA$^{Leu}_{GAG}$	0.16 ± 0.022
	+*Mm*tRNA$^{Leu}_{UAG}$	0.17 ± 0.030
*Mm*LeuRS-CP1[a]	-	0.12 ± 0.020
	+ *Ec*tRNA$^{Leu}_{GAG}$	0.55 ± 0.040
	+*Mm*tRNA$^{Leu}_{UAG}$	0.25 ± 0.032
*Mm*LeuRS-CP1/LSD	-	(3.7 ± 0.75) × 10^{-2}
	+*Ec*tRNA$^{Leu}_{GAG}$	0.61 ± 0.049
	+*Mm*tRNA$^{Leu}_{UAG}$	0.21 ± 0.030
*Mm*LeuRS-CP1/LSD-K598A	-	(9.3 ± 1.5) × 10^{-2}
	+*Ec*tRNA$^{Leu}_{GAG}$	0.14 ± 0.015
	+*Mm*tRNA$^{Leu}_{UAG}$	0.12 ± 0.013

All rates represent the average of three trials with the standard deviations indicated.
[a]Data from Tan *et al.* (17).

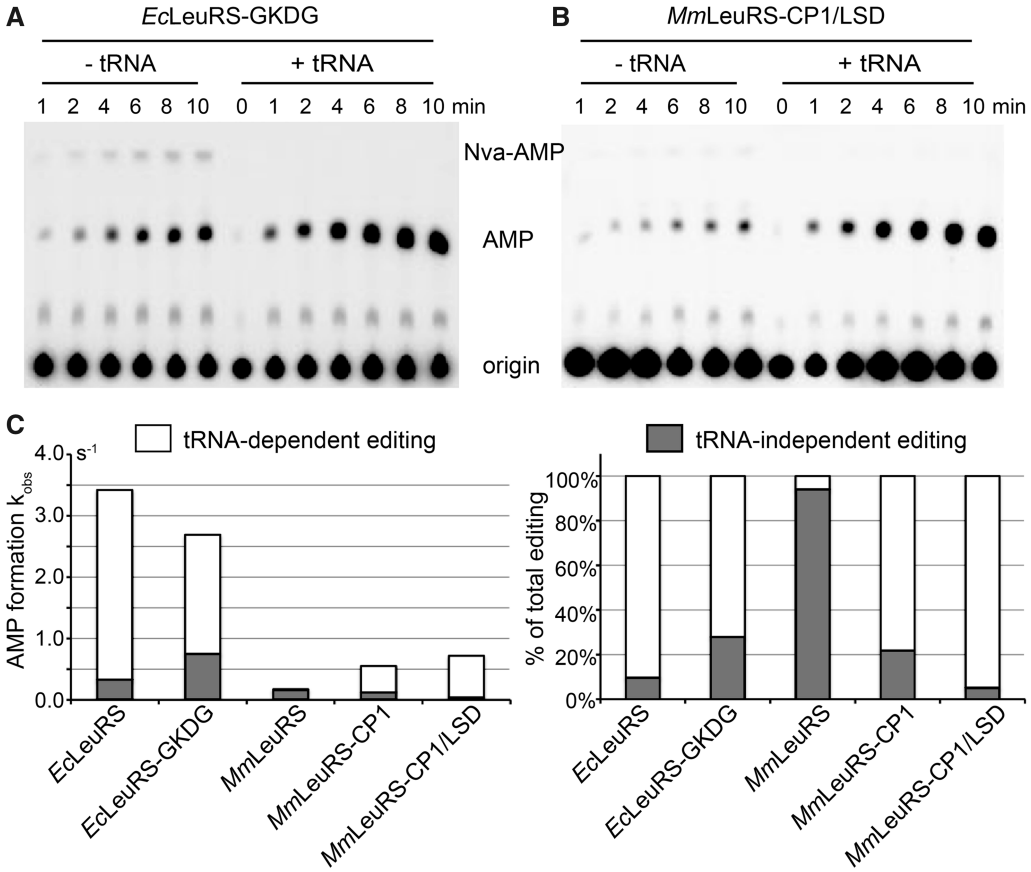

Figure 2. Effect of LSD mutations on tRNA-independent pre-transfer editing. (**A**) Total editing activity was measured using the AMP formation assay with 0.2 μM *Ec*LeuRS-GKDG in the absence or presence of 5 μM *Ec*tRNA^{Leu} and 15 mM Nva. (**B**) A similar assay was performed with 1 μM *Mm*LeuRS-CP1/LSD in the absence or presence of 5 μM *Ec*tRNA^{Leu} and 15 mM Nva. (**C**) Contributions of the different editing pathways for each protein: left, sum of the k_{obs} of different editing pathways; right, relative contributions of each pathway. Percentages were calculated from k_{obs} values of AMP formation reported in Table 1. tRNA-independent pre-transfer editing was measured in the absence of tRNA. tRNA-dependent editing was deduced by subtracting the tRNA-independent pre-transfer editing from total editing.

by the synthetic domain of LeuRS. As a consequence, the evolved LeuRS favoured tRNA-dependent pre-transfer editing, which was more effective in maintaining the catalytic fidelity. We propose that this is a possible reason why most prokaryotic LeuRSs have recruited and preserved LSD during their evolution.

The LSD is responsible for tRNA discrimination

In a previous study, it was found that *Mm*LeuRS-CP1 cross-leucylates *Ec*tRNA$^{Leu}_{GAG}$ with efficiency comparable with that of the *in vitro* transcript of *Mm*tRNA$^{Leu}_{UAG}$ (17). The present work showed that *Mm*LeuRS-CP1/LSD aminoacylates more efficiently *Ec*tRNA^{Leu} than *Mm*tRNA$^{Leu}_{UAG}$ with a discrimination factor (D factor) of 3.7 (according to k_{cat}/K_m) (Table 1). Therefore, *Mm*LeuRS-CP1/LSD had similar discriminatory properties as the native *Ec*LeuRS, which has a D factor of 5.5 (Table 1). These results suggest that the LSD may participate in tRNA binding and discrimination in some way. When the LSD of *Ec*LeuRS was replaced by the tetrapeptide GKDG, the mutant *Ec*LeuRS-GKDG leucylated *Ec*tRNA$^{Leu}_{GAG}$ and *Mm*tRNA$^{Leu}_{UAG}$ with similar catalytic efficiency (Table 1). The editing activity of *Ec*LeuRS-GKDG was also comparable in the presence of *Mm*tRNA$^{Leu}_{UAG}$ or *Ec*tRNA$^{Leu}_{GAG}$ (Table 2, Supplementary Figure S2).

Furthermore, when the CP1 domain and LSD of *Ec*LeuRS were inserted into *Mm*LeuRS, the mutant *Mm*LeuRS-CP1/LSD favoured *Ec*tRNA$^{Leu}_{GAG}$ not only in aminoacylation but also in editing. In the TLC-based AMP formation assay, *Mm*LeuRS-CP1/LSD had a rate constant for AMP formation in the presence of *Mm*tRNA$^{Leu}_{UAG}$ and Nva of 0.21 s^{-1}, while it was 0.61 s^{-1} and 0.56 s^{-1} for *Ec*tRNA$^{Leu}_{GAG}$ and *Mm*tRNA$^{Leu}_{UAA}$, respectively (Supplementary Figure S2, Tables 2 and 3). Consistently, *Ec*LeuRS also preferred *Ec*tRNA^{Leu} in the editing with AMP formation rate of 3.4 s^{-1}, with a corresponding value of 2.2 s^{-1}. However, LSD-deprived *Ec*LeuRS-GKDG showed no preference towards these two tRNAs during editing (k_{obs} 2.7 vs 2.9 s^{-1}) (Table 2). These results show that the LSD could confer tRNA discrimination properties to LeuRS, and this raises questions about how the LSD can distinguish tRNAs during aminoacylation and editing.

Table 3. Kinetic constants of *Mm*LeuRS-CP1/LSD for mutants of *Mm*tRNA$^{Leu}_{UAG}$ determined in the aminoacylation reaction

*Mm*tRNA$^{Leu}_{UAG}$	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} μM^{-1})	Relative catalytic efficiency
WT	1.7 ± 0.13	0.51 ± 0.049	0.30	1.0
A6G	1.2 ± 0.10	0.53 ± 0.041	0.44	1.5
C67U	1.9 ± 0.21	0.48 ± 0.032	0.25	0.83
C20U	1.4 ± 0.13	1.1 ± 0.10	0.75	2.5
V-arm-4nt	3.4 ± 0.32	1.1 ± 0.11	0.34	1.1
V-arm-5nt	1.1 ± 0.09	0.79 ± 0.080	0.70	2.3
A6G + V-arm-5nt	2.1 ± 0.20	1.7 ± 0.13	0.78	2.6
C20U + V-arm-5nt	0.88 ± 0.091	1.0 ± 0.10	1.1	3.7
*Mm*tRNA$^{Leu}_{UAA}$	2.8 ± 0.29	2.1 ± 0.24	0.78	2.6

Kinetic constants were determined using the tRNA charging assay described in the experimental section. The concentration of the tRNA^{Leu}s ranged from 0.5–30 μM. All parameters represent the average of three trials with the standard deviations indicated.

Identification of the critical nucleotides recognized by LSD

To identify the structural determinants of tRNA^{Leu} responsible for LeuRS ability to discriminate tRNA^{Leu}s from various species, we compared the tRNA^{Leu} sequences from *E. coli* and *M. mobile* and focused our attention on three differences between them: (i) the sixth base-pair in the acceptor stem of *Mm*tRNA$^{Leu}_{UAG}$ is a wobble base pair (A_{6}•••C_{67}), whereas it is a Watson Crick base pair G_6–C_{67} in *Ec*tRNA$^{Leu}_{GAG}$; (ii) the loop of the V-arm of *Mm*tRNA$^{Leu}_{UAG}$ contains three nucleotides; however, *Ec*tRNA$^{Leu}_{GAG}$ has a 4-nucleotide loop; (iii) nucleotide 20, located in the 'variable pocket' (20) of the D-loop, is always a U in *Ec*tRNA^{Leu}s but always a C in *Mm*tRNA$^{Leu}_{UAG}$ (Figure 3A). Therefore, a series of mutants of *Mm*tRNA$^{Leu}_{UAG}$ was constructed. Firstly, the A_{6}•••C_{67} pair was mutated to a Watson Crick base pair by introducing A6G or C67U mutations. Secondly, in the 'variable pocket', nucleotide C20 was changed to a U. Thirdly, the loop of the V-arm was enlarged from three nucleotides to four (V-arm-4 nt) or five (V-arm-5 nt), which are usual sizes for these loops in tRNA^{Leu}s. *Mm*LeuRS-CP1/LSD leucylated the C20U and V-arm-5 nt mutants at more than twice the catalytic efficiency of the wild-type *Mm*tRNA$^{Leu}_{UAG}$ (from 0.3 to 0.75 and 0.70 s^{-1} μM^{-1}, respectively). *Mm*LeuRS-CP1/LSD leucylated the double mutant (C20U + V-arm-5 nt), where the C20U mutation and V-arm-5 nt mutation were present, and catalytic efficiency (k_{cat}/K_m 1.1 s^{-1} μM^{-1}) (Table 3) almost reached the level of *Ec*tRNA$^{Leu}_{GAG}$ (1.1 s^{-1} μM^{-1} in Table 1). Similarly, *Mm*LeuRS-CP1/LSD charged another double mutant (A6G + V-arm-5 nt) and *Mm*tRNA$^{Leu}_{UAA}$ (another *Mm*tRNA^{Leu} isoacceptor) with the same catalytic efficiency (0.78 s^{-1} μM^{-1}) (Table 3). Interestingly, we found that *Mm*tRNA$^{Leu}_{UAA}$ naturally exhibits a large loop of 5 nucleotides in the V-arm according to the genomic tRNA database.

In the editing reaction, the seven *Mm*tRNA^{Leu} mutants showed various capacities to stimulate AMP formation. In the presence of Nva, *Mm*LeuRS-CP1/LSD had a rate constant for AMP formation of 0.21 s^{-1} for wild-type *Mm*tRNA$^{Leu}_{UAG}$; however, for *Mm*tRNA^{Leu}-C20U that was increased to 0.41 s^{-1}. In addition, the most efficient mutant leucylated by *Mm*LeuRS-CP1/LSD, *Mm*tRNA^{Leu}-(C20U + V-arm-5 nt), showed very similar effects on editing activity as *Mm*tRNA$^{Leu}_{UAA}$ in the presence of Nva (Figure 3B and C, Supplementary Figure S3, Table 4).

A key Lys residue of the LSD is responsible for tRNA discriminatory activity

It has been reported that *Ec*LeuRS contacts bases 10 and 27 of tRNA^{Leu} via the Arg595 and Arg600 residues located on the so-called β-hairpin of the LSD (4).

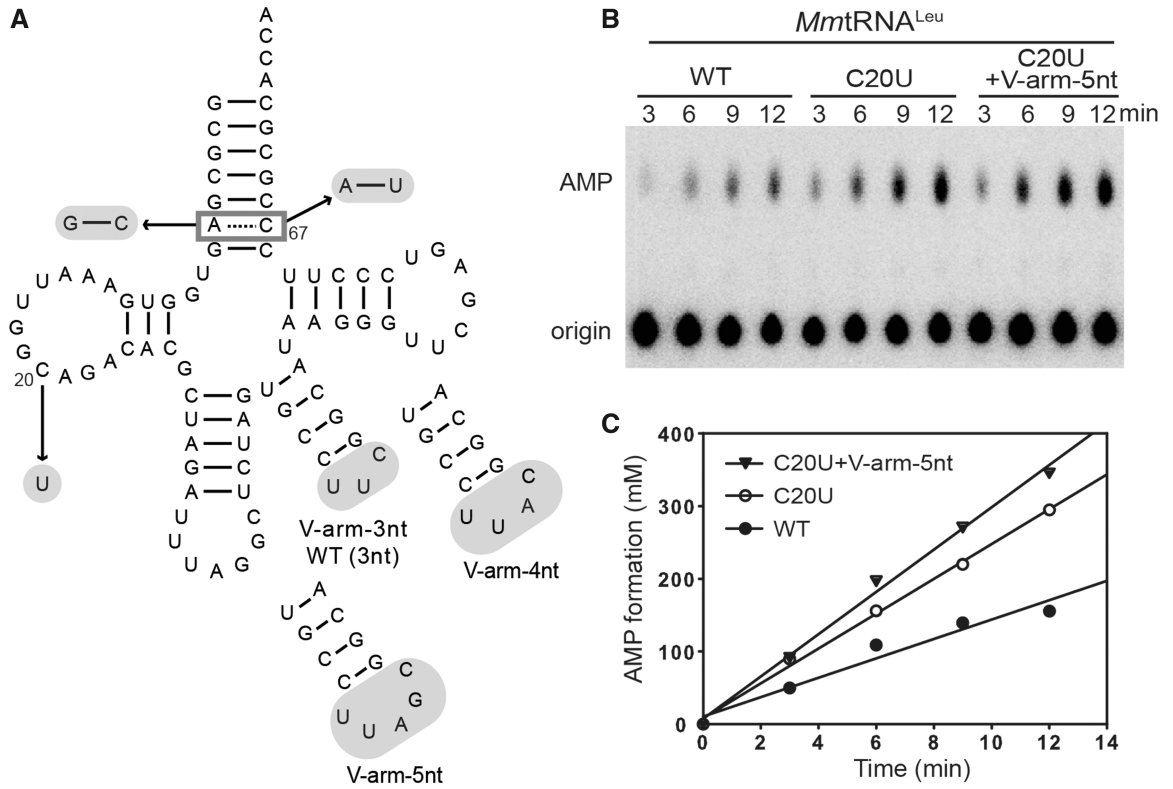

Figure 3. Mutations in *Mm*tRNA$^{Leu}_{UAG}$ that impact editing activity. (**A**) Cloverleaf structure of *Mm*tRNA$^{Leu}_{UAG}$ showing the mutations tested during the study. (**B**) AMP formation assay in the presence of 15 mM Nva catalyzed by 1 μM *Mm*LeuRS-CP1/LSD in the presence of 5 μM wild-type *Mm*tRNA$^{Leu}_{UAG}$, C20U and C20U + V-arm-5 nt. (**C**) Graphical representations of AMP formation as a function of time. k_{obs} values of AMP formation were calculated from the slopes, and these are shown in Table 4.

Table 4. Observed rate constants for AMP synthesis of *Mm*LeuRS-CP1/LSD in the presence of Nva

*Mm*tRNA$^{Leu}_{UAG}$	Rate of AMP formation k_{obs} (s^{-1})	Relative activity
WT	0.21 ± 0.020	1.0
A6G	0.28 ± 0.026	1.3
C67U	0.31 ± 0.032	1.5
C20U	0.41 ± 0.051	2.0
V-arm-4 nt	0.12 ± 0.011	0.57
V-arm-5 nt	0.26 ± 0.029	1.2
A6G + V-arm-5 nt	0.34 ± 0.031	1.6
C20U + V-arm-5 nt	0.52 ± 0.063	2.5
*Mm*tRNA$^{Leu}_{UAA}$	0.56 ± 0.059	2.7

All rates represent the average of three trials with the standard deviations indicated.

To investigate whether these residues could be responsible of discrimination between *Ec*tRNA$^{Leu}_{GAG}$ and *Mm*tRNA$^{Leu}_{UAG}$, initially we mutated the Arg595 and Arg600 of *Ec*LeuRS to Ala residues. Both mutants, *Ec*LeuRS-R595A and *Ec*LeuRS-R600A, showed high catalytic efficiency preference for *Ec*tRNA$^{Leu}_{GAG}$ but neither reached the value of the wild-type *Ec*LeuRS (Table 5). However, another mutant on the β-hairpin, *Ec*LeuRS-K598A, exhibited a stronger effect on aminoacylation activity. *Ec*LeuRS-K598A displayed a considerably lower affinity for *Ec*tRNA$^{Leu}_{GAG}$ compared with wild-type *Ec*LeuRS (K_m increased about 4-fold), which resulted in a decrease of the catalytic efficiency by almost 3-fold from 2.2 to 0.84 $s^{-1}\,\mu M^{-1}$. On the other hand, *Ec*LeuRS-K598A bound more tightly with *Mm*tRNA$^{Leu}_{UAG}$, and this induced a significant increase in the catalytic efficiency for the leucylation of *Mm*tRNA$^{Leu}_{UAG}$ (1.1 $s^{-1}\,\mu M^{-1}$) compared with wild-type *Ec*LeuRS (0.4 $s^{-1}\,\mu M^{-1}$) (Table 5).

In the same way, when K598 in the LSD of chimeric *Mm*LeuRS-CP1/LSD was replaced with an Ala residue, the catalytic efficiency of the mutant for *Mm*tRNA$^{Leu}_{UAG}$ was greater than that for *Ec*tRNA$^{Leu}_{GAG}$; however, the catalytic efficiency of *Mm*LeuRS-CP1/LSD for *Mm*tRNA$^{Leu}_{UAG}$ was lower than that for *Ec*tRNA$^{Leu}_{GAG}$, indicating that mutant *Mm*LeuRS-CP1/LSD-K598A prefers to charge *Mm*tRNA$^{Leu}_{UAG}$, while *Mm*LeuRS-CP1/LSD prefers *Ec*tRNA$^{Leu}_{GAG}$. Thus, mutation at K598 changed the species preference of these enzymes for their tRNA^{Leu} substrates (Table 5). The results show that residue K598 in the β-hairpin of the LSD of *Ec*LeuRS contributes positively to the binding and aminoacylation of *Ec*tRNA$^{Leu}_{GAG}$ and acts as an antideterminant versus *Mm*tRNA$^{Leu}_{UAG}$. When K598 was mutated to an Ala residue, the antideterminant effect was suppressed and the specific recognition of *Ec*LeuRS LSD for *Ec*tRNA^{Leu} was extended to *Mm*tRNA^{Leu}. In parallel, there was a decrease of binding affinity of *Mm*LeuRS-CP1/LSD-K598A for

Table 5. Kinetic constants of *Ec*LeuRS and its mutants in the aminoacylation reaction

Enzyme	*Ec*tRNA$^{Leu}_{GAG}$			*Mm*tRNA$^{Leu}_{UAG}$			D factor
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\,\mu M^{-1}$)	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($s^{-1}\,\mu M^{-1}$)	
*Ec*LeuRS[a]	2.2 ± 0.14	4.9 ± 0.30	2.2	10.6 ± 1.2	4.2 ± 0.40	0.40	5.5
*Ec*LeuRS-R595A	3.4 ± 0.25	4.3 ± 0.31	1.3	4.2 ± 0.40	3.2 ± 0.33	0.75	1.7
*Ec*LeuRS-K598A	8.2 ± 0.81	6.9 ± 0.57	0.84	5.7 ± 0.41	6.3 ± 0.55	1.1	0.76
*Ec*LeuRS-R600A	3.9 ± 0.27	7.2 ± 0.65	1.9	8.6 ± 0.78	6.1 ± 0.59	0.70	2.7
*Mm*LeuRS-CP1/LSD	1.3 ± 0.12	1.4 ± 0.11	1.1	1.7 ± 0.13	0.51 ± 0.049	0.30	3.7
*Mm*LeuRS-CP1/LSD-K598A	6.6 ± 0.62	3.6 ± 0.37	0.54	5.7 ± 0.41	6.3 ± 0.55	1.4	0.39

D is the discrimination factor of the different LeuRSs for the two bacterial tRNAs, and this value was calculated as follows: $D = k_{cat}/K_m$(*Ec*tRNA$^{Leu}_{GAG}$) / k_{cat}/K_m(*Mm*tRNA$^{Leu}_{UAG}$). All parameters represent the average of three trials with the standard deviations indicated.
[a]Data from Tan *et al.*(17).

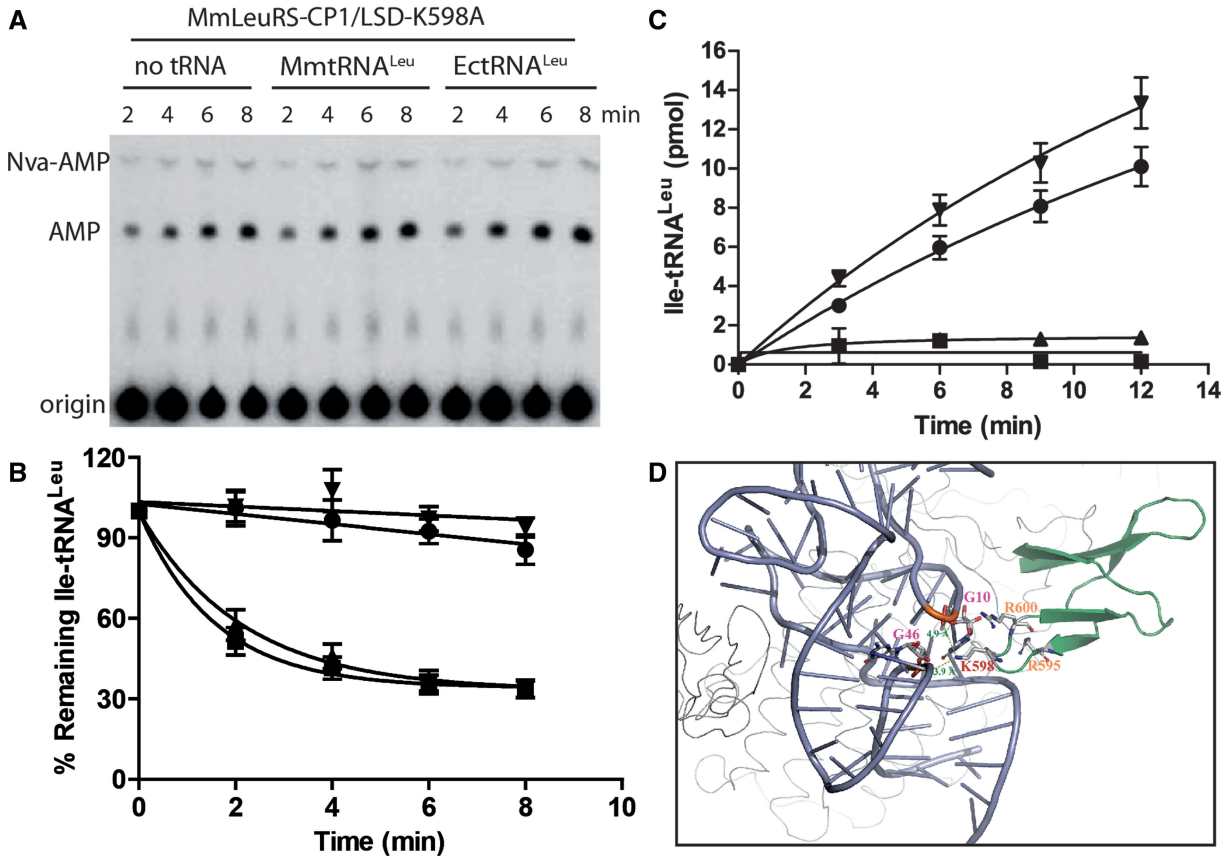

Figure 4. Editing and mischarging properties of *Mm*LeuRS-CP1/LSD-K598A. (**A**) Total editing activity was measured by the AMP formation assay with 1 μM of *Mm*LeuRS-CP1/LSD-K598A and 15 mM Nva in the absence or presence of 5 μM *Ec*tRNA$^{Leu}_{GAG}$ or *Mm*tRNA$^{Leu}_{UAG}$. (**B**) Deacylation of [^{3}H]-Ile-*Ec*tRNA^{Leu} (1 μM) by 20 nM of *Mm*LeuRS (black circle), *Mm*LeuRS-CP1/LSD-K598A (inverted black triangle), *Mm*LeuRS-CP1/LSD (black square) and *Mm*LeuRS-CP1 (black triangle). (**C**) Mischarging of *Ec*tRNA$^{Leu}_{GAG}$ (20 μM) with Ile catalyzed by 1 μM of *Mm*LeuRS (black circle), *Mm*LeuRS-CP1/LSD-K598A (inverted black triangle), *Mm*LeuRS-CP1/LSD (black square) and *Mm*LeuRS-CP1(black triangle). (**D**) Crystal structure of tRNA^{Leu} (light blue in the cartoon model) in complex with *Ec*LeuRS (grey) during the editing conformation (PDB ID code 4ARC, Ref.4). Residues R595, K598 and R600 of LSD (green) are numbered and shown in stick representation with labelling. Both G10 and G46 of tRNA^{Leu} were also highlighted with the stick model with their distances to K598 labelled.

*Ec*tRNA^{Leu}, which reduced its catalytic efficiency to a lower level (k_{cat}/K_m 0.54 $s^{-1}\,\mu M^{-1}$) than for *Mm*tRNA$^{Leu}_{UAG}$ (k_{cat}/K_m 1.4 $s^{-1}\,\mu M^{-1}$). These data show that K598 is a key residue that controls the cross-recognition of tRNA^{Leu}s from different species.

Consistently, in the chimeric *Mm*LeuRS-CP1/LSD, the K598A mutation controlled post-transfer editing, as there was a drop in the AMP synthesis rate (Figure 4A, Table 2) and an absence of deacylation activity towards Ile-*Ec*tRNA$^{Leu}_{GAG}$ (Figure 4B). The loss of deacylation properties was further confirmed by a loss of aminoacylation specificity as illustrated by the Ile mischarging of *Ec*tRNA$^{Leu}_{GAG}$ (Figure 4C). Both *Mm*LeuRS-CP1/LSD-K598A and *Mm*LeuRS were able to mischarge Ile

in contrast to *Mm*LeuRS-CP1/LSD and *Mm*LeuRS-CP1 that could not catalyze this substrate. Taken together, these results suggested that the crucial K598 residue of LSD mediated the interaction with tRNA and involved in tRNA recognition.

In three of the four crystallographic structures that describe the aminoacylation and proofreading states of LeuRS (4), the ε-amino group of K598 is located in the vicinity of the phosphate group of G10. K598 approaches the tRNA bound in the editing conformation at distances from 4.6 to 4.9 Å according to the different tertiary structures (in 4ASI, 4ARC and 4ARI). In addition, in the editing complex bound with leucine (4ARC), the ε-amino group of K598 forms a potential interaction with the phosphate group of G46 at a distance of 3.9 Å (Figure 4D). However, these putative interactions with the phosphate backbone of tRNA can hardly explain the new discriminating properties of the K598A mutant for *Mm*tRNA^{Leu} and *Ec*tRNA^{Leu}. Nevertheless, we cannot exclude the possibility that they play a role during the transition of the 3′ end of tRNA between the aminoacylation and the editing states, and thus favour the aminoacylation of one isoacceptor.

DISCUSSION

With genome sizes <1 Mb, bacteria from the genus *Mycoplasma* have been described as the 'smallest free-living organisms', and thus are considered to be the best representatives for the concept of a minimal cell. The *M. mobile* genome encodes only 635 proteins (21), and includes 27 tRNA genes, one of the lowest abundances reported for any organism. Strong evidence suggests that mycoplasmas evolved by a process of reductive evolution that was made possible by adopting a parasitic lifestyle. During this process, the mycoplasmas lost considerable portions of their ancestral chromosomes but retained the genes essential for life. Genome compaction in mycoplasmas is often reflected by the presence of reduced intergenic spacers and by the shortness of most putative proteins relative to their orthologues (22). aaRSs genes did not escape this size reduction, and several of these enzymes have lost key residues in their editing domains, and in the extreme case of LeuRS, the CP1-editing domain has been deleted completely (10,17). Therefore, mycoplasmas are following a kind of reverse evolution that consists of selecting minimalist proteins that mimic the primitive proteins. Primitive aaRSs have followed an opposite evolutionary pathway by progressively adding domains to improve efficiency and fidelity and to conserve the genetic code and proteome in its present form.

LeuRSs from various species are very complex enzymes that are amongst the largest aaRSs. These enzymes have an unusually high number of modules appended to the catalytic core that participate in a concerted way in tRNA binding, aminoacylation and proofreading. Recent X-ray analysis of tRNA^{Leu}–LeuRS complexes in the aminoacylation or editing conformation has provided the structural basis and dynamics of the aminoacylation and proofreading functional cycle (4). LeuRS produces error-free Leu-tRNA^{Leu} by coordinating the translocation of the CCA-end of mischarged tRNAs from its synthetic site to the separate proofreading site where the editing occurs. Such translocation involves correlated rotations of four LeuRS domains that are linked to the catalytic core. These motions drive the CCA sequence of the tRNA from the aminoacylation site to the editing site. During this process, the CP1-editing domain stabilizes the tRNA during aminoacylation, while a large rotation of the LSD positions the conserved KMSKS loop of the LeuRS to bind the CCA end of the tRNA, thereby promoting catalysis (4).

The absence of both CP1 and LSD in *Mm*LeuRS offers the opportunity to investigate the mechanism of insertion of these additional modules and explore the plasticity of the catalytic core to acquire new functions. Previously, it was shown that insertion of the CP1 domain into the minimal *Mm*LeuRS did not change synthetic efficiency (17). CP1 insertion does improve affinity for the tRNA but it decreases k_{cat}, suggesting that the tighter binding of the substrate is deleterious for its subsequent reactivity or release. The fusion of the domains of *Ec*LeuRS with *Mm*LeuRS also provided the post-transfer editing function to the chimeric enzyme *Mm*LeuRS-CP1, and this enzyme demonstrated greater activity for *E. coli* tRNA^{Leu}. Although the post-transfer editing activity of *Mm*LeuRS-CP1 remained modest compared with that of *Ec*LeuRS, this observation supports the theory that the aaRS evolved by fusion with additional modules (23).

Here, we showed that insertion of the LSD of *Ec*LeuRS into the pre-existing chimeric protein *Mm*LeuRS-CP1 further improved tRNA binding, leading to a protein with greater catalytic efficiency. In contrast, the editing activity of the double insertion mutant was increased only rather poorly, and a decrease of the pre-existing pre-transfer editing activity of *Mm*LeuRS was observed. Therefore, fusion with the second insertion domain improved not only tRNA binding and the synthetic activity of the enzyme but it also conferred greater importance to post-transfer editing relative to pre-transfer editing. This change might be explained by adenylate molecules reacting faster with tRNA to synthesize aminoacyl-tRNAs, thereby reducing their opportunity to be edited by the pre-transfer editing process in the synthetic site.

These data provide evidence that the CP1 domain and LSD cooperate for greater synthetic and proofreading properties when inserted in the *Mm*LeuRS framework, and these observations suggest how these enzymes could have evolved from primitive aaRSs. In this manner, the editing domain, or another domain, could have been distributed amongst different aaRSs before their fine adjustment to the new substrate through the accumulation of mutations. In this present work, we further simulate evolution and show that single mutation events could significantly improve enzyme activity. For instance, mutations could take place in *trans* in the genes of the corresponding tRNAs. We showed that a mutation at position 20 of *Mm*tRNA^{Leu} (C20U) doubled the relative activity of *Mm*LeuRS-CP1/LSD in the aminoacylation and proofreading compared with the wild-type *Mm*tRNA^{Leu} (Tables 3 and 4). Residue 20 is located in the 'variable

pocket', and it is known to be a recognition element in different aminoacylation systems (20,24–26). In LeuRS, the only putative interaction of nucleotide 20 occurs with Lys813 that is located in the C-terminal domain, and this can occur only during the editing state (PDB entry 4ARC). Therefore, modifying a specific interaction of the editing state with a distinct module of the enzyme may have improved both synthetic and editing activities. As these activities contribute to a unique functional cycle, any mutation impacting one step may have repercussions on other activities. A similar improvement of catalytic properties was also observed with a double mutant that contained mutations in the acceptor arm and variable arm (A6G + V-arm-5 nt). Here also restoration of activity may occur through the C-terminal domain of LeuRS, which interacts with several nucleotides of the V-arm. Enlarging the loop might have reorganized tRNA binding and pivoting during the catalytic cycle. In addition, *Mm*tRNA$^{Leu}_{UAA}$ with the natural 5-nt loop exhibited much greater aminoacylation and editing activities and was endowed with the most codon usage in *M. mobile*. The second mutation (A6G; located at a 50-Å distance in the acceptor arm) might have amplified the first effect (4).

Additionally, we showed that the synthetic performance of the chimeric enzyme could be improved in *cis* by a single mutation in the inserted LSD. We have found that Lys598 is an antideterminant for *Mm*tRNA^{Leu}, but negative effects could be cancelled by Ala mutation. Therefore, this mutant shows that there are at least two alternative ways to improve the aminoacylation–proofreading functional cycle: one way consists of adapting the enzyme by mutating critical amino acids residues, while the second way consists of adjusting the tRNA structure in keeping with the newly inserted modules and the resulting conformation changes that occur during the catalytic processes.

Altogether, our results support the theory that fusion of additional modules to the ancient catalytic core of aaRSs during evolution introduced new catalytic functions to improve fidelity and catalytic performance (27). Moreover, this present study shows that the minimalist *Mm*LeuRS is an ideal platform for further studies to understand the evolution of the aaRSs family through the acquisition of complementary modules.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Table 1 and Supplementary Figures 1–3.

FUNDING

National Key Basic Research Foundation of China [2012CB911000]; The Natural Science Foundation of China [30930022 and 31130064]; Committee of Science and Technology in Shanghai [12JC1409700]; Visiting Professorship for Senior International Scientists from the Chinese Academy of Sciences [2011T2S10]. Funding for open access charge: The Natural Science Foundation of China [30930022 and 31130064].

Conflict of interest statement. None declared.

REFERENCES

1. Ling,J., Reynolds,N. and Ibba,M. (2009) Aminoacyl-tRNA synthesis and translational quality control. *Annu. Rev. Microbiol.*, **63**, 61–78.
2. 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.
3. Cusack,S., Yaremchuk,A. and Tukalo,M. (2000) The 2 A crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO J.*, **19**, 2351–2361.
4. Palencia,A., Crepin,T., Vu,M.T., Lincecum,T.L. Jr, Martinis,S.A. and Cusack,S. (2012) Structural dynamics of the aminoacylation and proofreading functional cycle of bacterial leucyl-tRNA synthetase. *Nat. Struct. Mol. Biol.*, **19**, 677–684.
5. Tukalo,M., Yaremchuk,A., Fukunaga,R., Yokoyama,S. and Cusack,S. (2005) The crystal structure of leucyl-tRNA synthetase complexed with tRNA^{Leu} in the post-transfer-editing conformation. *Nat. Struct. Mol. Biol.*, **12**, 923–930.
6. Fukunaga,R. and Yokoyama,S. (2005) Crystal structure of leucyl-tRNA synthetase from the archaeon *Pyrococcus horikoshii* reveals a novel editing domain orientation. *J. Mol. Biol.*, **346**, 57–71.
7. Fukunaga,R. and Yokoyama,S. (2005) Aminoacylation complex structures of leucyl-tRNA synthetase and tRNA^{Leu} reveal two modes of discriminator-base recognition. *Nat. Struct. Mol. Biol.*, **12**, 915–922.
8. Vu,M.T. and Martinis,S.A. (2007) A unique insert of leucyl-tRNA synthetase is required for aminoacylation and not amino acid editing. *Biochemistry*, **46**, 5170–5176.
9. Ma,J.J., Zhao,M.W. and Wang,E.D. (2006) Split leucine-specific domain of leucyl-tRNA synthetase from the hyperthermophilic bacterium Aquifex aeolicus. *Biochemistry*, **45**, 14809–14816.
10. Li,L., Boniecki,M.T., Jaffe,J.D., Imai,B.S., Yau,P.M., Luthey-Schulten,Z.A. and Martinis,S.A. (2011) Naturally occurring aminoacyl-tRNA synthetases editing-domain mutations that cause mistranslation in Mycoplasma parasites. *Proc. Natl Acad. Sci. USA*, **108**, 9378–9383.
11. Tan,M., Zhu,B., Liu,R.J., Chen,X., Zhou,X.L. and Wang,E.D. (2013) Inter-domain communication modulates the tRNA-dependent pre-transfer editing of leucyl-tRNA synthetase. *Biochem. J.*, **449**, 123–131.
12. Steinberg,S., Misch,A. and Sprinzl,M. (1993) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **21**, 3011–3015.
13. Zhou,X.L., Du,D.H., Tan,M., Lei,H.Y., Ruan,L.L., Eriani,G. and Wang,E.D. (2011) Role of tRNA amino acid-accepting end in aminoacylation and its quality control. *Nucleic Acids Res.*, **39**, 8857–8868.
14. Du,X. and Wang,E.D. (2003) Tertiary structure base pairs between D- and TpsiC-loops of Escherichia coli tRNA(Leu) play important roles in both aminoacylation and editing. *Nucleic Acids Res.*, **31**, 2865–2872.
15. Yao,P., Zhou,X.L., He,R., Xue,M.Q., Zheng,Y.G., Wang,Y.F. and Wang,E.D. (2008) Unique residues crucial for optimal editing in yeast cytoplasmic Leucyl-tRNA synthetase are revealed by using a novel knockout yeast strain. *J. Biol. Chem.*, **283**, 22591–22600.
16. Huang,Q., Yao,P., Eriani,G. and Wang,E.D. (2012) *In vivo* identification of essential nucleotides in tRNALeu to its functions by using a constructed yeast tRNA^{Leu} knockout strain. *Nucleic Acids Res.*, **40**, 10463–10477.
17. Tan,M., Yan,W., Liu,R.J., Wang,M., Chen,X., Zhou,X.L. and Wang,E.D. (2012) A naturally occurring nonapeptide functionally compensates for the CP1 domain of leucyl-tRNA synthetase to modulate aminoacylation activity. *Biochem. J.*, **443**, 477–484.
18. Li,Y., Wang,E.D. and Wang,Y.L. (1998) Overproduction and purification of *Escherichia coli* tRNA(Leu). *Sci. China C Life Sci.*, **41**, 225–231.

19. Tan,M., Zhu,B., Zhou,X.L., He,R., Chen,X., Eriani,G. and Wang,E.D. (2010) tRNA-dependent pre-transfer editing by prokaryotic leucyl-tRNA synthetase. *J. Biol. Chem.*, **285**, 3235–3244.
20. McClain,W.H. and Foss,K. (1988) Changing the acceptor identity of a transfer RNA by altering nucleotides in a ''variable pocket''. *Science*, **241**, 1804–1807.
21. Jaffe,J.D., Stange-Thomann,N., Smith,C., DeCaprio,D., Fisher,S., Butler,J., Calvo,S., Elkins,T., FitzGerald,M.G., Hafez,N. *et al.* (2004) The complete genome and proteome of *Mycoplasma mobile*. *Genome Res.*, **14**, 1447–1461.
22. Katinka,M.D., Duprat,S., Cornillot,E., Metenier,G., Thomarat,F., Prensier,G., Barbe,V., Peyretaillade,E., Brottier,P., Wincker,P. *et al.* (2001) Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature*, **414**, 450–453.
23. Schimmel,P. and Ribas de Pouplana,L. (1995) Transfer RNA: from minihelix to genetic code. *Cell*, **81**, 983–986.
24. McClain,W.H., Foss,K., Jenkins,R.A. and Schneider,J. (1991) Four sites in the acceptor helix and one site in the variable pocket of tRNA(Ala) determine the molecule's acceptor identity. *Proc. Natl Acad. Sci. USA*, **88**, 9272–9276.
25. 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**, 1363–1366.
26. Liu,W., Huang,Y., Eriani,G., Gangloff,J., Wang,E. and Wang,Y. (1999) A single base substitution in the variable pocket of yeast tRNA(Arg) eliminates species-specific aminoacylation. *Biochim. Biophys. Acta*, **1473**, 356–362.
27. SternJohn,J., Hati,S., Siliciano,P.G. and Musier-Forsyth,K. (2007) Restoring species-specific posttransfer editing activity to a synthetase with a defunct editing domain. *Proc. Natl Acad. Sci. USA*, **104**, 2127–2132.