

Substrate recognition by class I lysyl-tRNA synthetases: A molecular basis for gene displacement

MICHAEL IBBA*[†], HEATHER C. LOSEY*, YUTAKA KAWARABAYASI[‡], HISASI KIKUCHI[‡], SHIPRA BUNJUN*,
AND DIETER SÖLL*^{§¶}

Departments of *Molecular Biophysics and Biochemistry and §Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520-8114; and
[‡]National Institute of Technology and Evaluation, 2-49-10 Nishihara, Shibuya, Tokyo 151-0066, Japan

Contributed by Dieter Söll, November 17, 1998

ABSTRACT Lysyl-tRNA synthetases (LysRSs) are unique amongst the aminoacyl-tRNA synthetases in being composed of unrelated class I and class II enzymes. To allow direct comparison between the two types of LysRS, substrate recognition by class I LysRSs was examined. Genes encoding both an archaeal and a bacterial class I enzyme were able to rescue an *Escherichia coli* strain deficient in LysRS, indicating their ability to functionally substitute for a class II LysRS *in vivo*. *In vitro* characterization showed lysine activation and recognition to be tRNA-dependent, an attribute of several class I, but not class II, aminoacyl-tRNA synthetases. Examination of tRNA recognition showed that class I LysRSs recognize the same elements in tRNA^{Lys} as their class II counterparts, namely the discriminator base (N73) and the anticodon. This sequence-specific recognition of the same nucleotides in tRNA^{Lys} by the two unrelated types of enzyme suggests that tRNA^{Lys} predates at least one of the LysRSs in the evolution of the translational apparatus. The only observed variation in recognition was that the G2-U71 wobble pair of spirochete tRNA^{Lys} acts as antideterminant for class II LysRS but does not alter class I enzyme recognition. This difference in tRNA recognition strongly favors the use of a class I-type enzyme to aminoacylate particular tRNA^{Lys} species and provides a molecular basis for the observed displacement of class II by class I LysRSs in certain bacteria.

The aminoacyl-tRNA synthetases have long been upheld as a paradigm of molecular evolution. This is because their products, aminoacyl-tRNAs, are essentially the same in all living organisms. The role of aminoacyl-tRNA synthetases is to interpret the genetic code in terms of amino acids, providing the essential link between RNA and protein without which translation would be impossible. This highly conserved function has been assumed to place constraints on the evolutionary variation of aminoacyl-tRNA synthetases beyond those enforced on most other protein families. A broader evolutionary interest in the aminoacyl-tRNA synthetases stems from their biological function being one of the core requirements for progression from the RNA world to the universal common ancestor in numerous schemes for the origin of life (reviewed in refs. 1 and 2). Thus, it could reasonably be assumed that in depth investigations of the aminoacyl-tRNA synthetases would provide insights into the origins of the genetic code and contemporary translation. Although studies of the aminoacyl-tRNA synthetases have provided a wealth of information in numerous areas, including specificity and catalysis in biological systems, protein-RNA molecular recognition, and diversification within protein families, they have not provided the expected bounty of evolutionary information such as has been garnered, for example, from ribosomal RNA. One of the main reasons underlying this has been the limited range of complete families of genes encoding aminoacyl-tRNA synthetase available,

namely those from *Escherichia coli*, yeast, and human. However, the recent boom in whole genome sequencing has vastly expanded the number of known synthetase families. This has revealed an unexpected diversity among the aminoacyl-tRNA synthetases, which now clearly can be seen to have evolved independently of each other (see, for example, ref. 3). The starkest examples of this are the replacement of asparaginyl- and glutaminyl-tRNA synthetases by tRNA-dependent transamidation pathways (4, 5) and the existence of unrelated lysyl-tRNA synthetases (LysRSs) with identical activities (6).

The aminoacyl-tRNA synthetases can be divided into two classes (I and II) of 10 members each based on the presence of mutually exclusive amino acid sequence motifs (7–9). This division reflects structurally distinct topologies within the active site; class I aminoacyl-tRNA synthetases contain a Rossmann fold, and class II synthetases contain a unique anti-parallel β fold. In addition, it has been observed that class I enzymes bind the acceptor helix of tRNA on the minor groove side whereas class II enzymes bind the major groove side (10). It generally is assumed that an aminoacyl-tRNA synthetase of particular substrate specificity will always belong to the same class regardless of its biological origin, reflecting the ancient evolution of this enzyme family (11, 12). The only known exceptions to this rule are the LysRSs, which are composed of two unrelated families, namely class I enzymes in certain archaea and bacteria and class II enzymes in all other organisms (13, 14). The LysRS family now provides an unprecedented opportunity to study the determinants of specificity during the evolution of molecular recognition because both class I and class II LysRSs are able to use the same amino acid and highly similar tRNA substrates. In addition, the unusual phylogenetic distribution of the class I LysRS-encoding *lysS* genes may reveal insights into the molecular basis of horizontal gene transfer (15) and displacement.

MATERIALS AND METHODS

General. *In vitro* transcription of tRNA genes and subsequent purification and refolding of transcripts were performed as detailed (16), and mature *E. coli* tRNA^{Lys} was purchased from Sigma. Plasmid preparation for transcription of tRNA genes was performed by standard techniques using precipitation with polyethylene glycol (17), and the tRNA gene sequences were confirmed by sequencing before transcription. Template DNA for transcription by T7 RNA polymerase was digested with *Bst*NI. Media for bacterial growth and molecular biology protocols were standard unless otherwise noted (17). *Methanococcus maripaludis*

Abbreviations: LysRS, lysyl-tRNA synthetase; Ap₄A, diadenosine tetraphosphate.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF108356).

A Commentary on this article begins on page 327.

[†]Present address: Center for Biomolecular Recognition, The Panum Institute, Blegdamsvej 3C, DK2200N, Copenhagen, Denmark.

[¶]To whom reprint requests should be addressed at: Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, P.O. Box 208114, New Haven, CT 06520-8114. e-mail: soll@trna.chem.yale.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

genomic DNA was kindly provided by W. H. Whitman (University of Georgia), and *Borrelia burgdorferi* genomic DNA was provided by D. R. Akins and J. D. Radolf (University of Texas).

Strains and Plasmids. The strain PALΔSAUTR/pMAK705lysU and the plasmid pKS-lysS have been described (18). The pBR322-derived vector pCBS1-BstrpS, which contains the *Bacillus stearothermophilus* *trpS* gene under the control of the *E. coli* *trpS* promoter, was kindly provided by M. Kitabatake (Yale University). To allow the constitutive expression of the *B. burgdorferi* and *M. maripaludis* *lysS* genes in *E. coli*, they were subcloned behind the *trpS* promoter domain in pCBS1 to give pCBS1-BBlysS and pCBS1-MmlysS, respectively. The genes encoding tRNA^{Lys} from *B. burgdorferi* and *M. maripaludis* were amplified by PCR and were subcloned behind a synthetic T7 RNA polymerase promoter in the vector pUC119. These constructs then were transformed into the *E. coli* strain DH5α, and their sequences were confirmed. The gene encoding *E. coli* tRNA^{Lys} G2-U71 was generated by the same procedure by using a mutant oligonucleotide.

Cloning of the Gene Encoding *M. maripaludis* tRNA^{Lys}. *Methanococcus vannielii* (19) and *Methanococcus jannaschii* (20) contain identically ordered gene clusters encoding several tRNAs (including tRNA^{Lys}) and 5S ribosomal RNA. Based on the assumption that a similar cluster would exist in *M. maripaludis*, oligonucleotide primers complementary to conserved regions flanking the tRNA^{Lys} gene were used to amplify the corresponding region from genomic DNA. The primers used were MMtRNAK1 (GCGGACTGTAGATCCGCATGTCGCTG) and MMtRNAK2 (GAACCCGAGTCAC[AG]GGAGTGACAGTC). The resulting fragment was subcloned into the pBlue-script II KS (Stratagene)-derived T-vector (21) and was used to transform the *E. coli* strain DH5α. The cloned fragment was sequenced from two independent clones on both strands, all of which contained an identical sequence encoding tRNA^{Lys}.

Sequencing of the Gene Encoding *Aeropyrum pernix* LysRS. The complete genome of *A. pernix* (type strain: K1; JCM9820; ref. 22) was sequenced by using a modified whole genome shotgun strategy. Analysis of the nucleotide sequence showed the presence of an ORF encoding a class I- but not a class II-type LysRS. A detailed description of the sequencing protocol together with the complete analysis of the *A. pernix* genome can be found elsewhere (Y.K. and H.K., unpublished work).

Protein Purification. *B. burgdorferi* LysRS was purified as described (13). *M. maripaludis* LysRS was purified as before (14) except that, in the final purification step, the sample was applied to a Mono-Q column, which then was developed with a 0–500 mM NaCl gradient in standard buffer. *E. coli* LysRS was purified partially from the *E. coli* strain BL21 (DE3). A cell-free extract was prepared as described (13) and was applied to a Q-Sepharose Fast Flow column, which was developed in a 0–500 mM NaCl gradient. The fractions containing LysRS activity were pooled, were precipitated with ammonium sulfate, were resuspended in 20 mM Hepes, (pH 7.2), 1 mM MgCl₂, 10 mM NaCl, and 5 mM DTT, and finally were dialyzed against the same buffer containing 30% glycerol. The concentration of LysRS in this partially purified fraction then was determined by active site titration. T7 RNA polymerase was purified as described (16).

Kinetic Analyses. The pyrophosphate exchange reaction was performed at 37°C as described (23) in the presence of 100 mM Hepes (pH 7.2), 25 mM KCl, 10 mM MgCl₂, 4 mM DTT, 2 mM lysine, 100 nM LysRS, and 1.5 μM tRNA^{Lys}. Aminoacylation assays were performed as described (13, 14) with the concentration of the substrate under investigation varied over the range 0.2–5 × K_M.

Cyanogen Bromide Treatment of tRNA. tRNA (2,000 pmols) was resuspended in 90 μl STE (100 mM NaCl/10 mM Tris-Cl, pH 8.0/1 mM EDTA, pH 8.0) with 100 mM NaHCO₃ (pH 8.9), and the mixture was incubated with 225 μg CNBr for 10 minutes at room temperature with shaking. The pH was adjusted to 6.0 with HCl, and the reaction mixture was dialyzed against 0.5 mM

EDTA (pH 6.0) (24). The K_M for tRNA^{Lys} was determined by varying the concentration of CNBr-treated tRNA from 0.3–6 × K_M.

Synthesis of Diadenosine Tetraphosphate (Ap₄A). Samples were incubated in 100 mM Hepes (pH 7.2), 5 mM ATP, 10 mM MgCl₂, 20 mM KCl, 100 μM Lysine, 4 mM DTT, and 50 nM LysRS. After incubation at 37°C, the reaction was quenched with 10% perchloric acid and was neutralized with KOH. After removing the precipitate by centrifugation, ATP was depleted with 50 units of alkaline phosphatase (Boehringer–Mannheim) for 30 minutes at 37°C. Alkaline phosphatase then was inactivated by freeze/thaw. Luciferase (*Photinus pyralis*) was added to a final concentration of 1.2 μg/ml, and luciferin was added at a final concentration of 120 μM. Ten minutes after the addition of luciferin/luciferase, allowing for consumption of any remaining ATP, snake venom phosphodiesterase I was added to a final concentration of 8 μg/ml, and the light produced was quantified in a liquid scintillation counter at room temperature (25, 26). *E. coli* S100 was prepared in the same manner as described above and then was treated with RNase at a final concentration of 20 μg/ml for 5 min on ice.

Complementation of an *E. coli* lysS lysU Double Mutant. The *E. coli* strain PALΔSAUTR/pMAK705lysU was transformed at 30°C with either pKS-lysS, pCBS1-EctrpS, pCBS1-BblysS, or pCBS1-MmlysS. The resulting transformants then were tested for growth on Luria–Bertani agar (100 μg/ml ampicillin) at 30°C, 40°C, and 42°C with and without the addition of L-lysine (5 mM).

RESULTS

Molecular Phylogeny of Class I LysRS and tRNA^{Lys}. The 12 known predicted amino acid sequences of class I LysRS proteins were aligned and used to construct a phylogenetic tree (Fig. 1). The class I LysRS is found in both Euryarchaeota (two species of *Methanococcus*, *Pyrococcus horikoshii*, *Pyrococcus furiosus*, *Meth-*

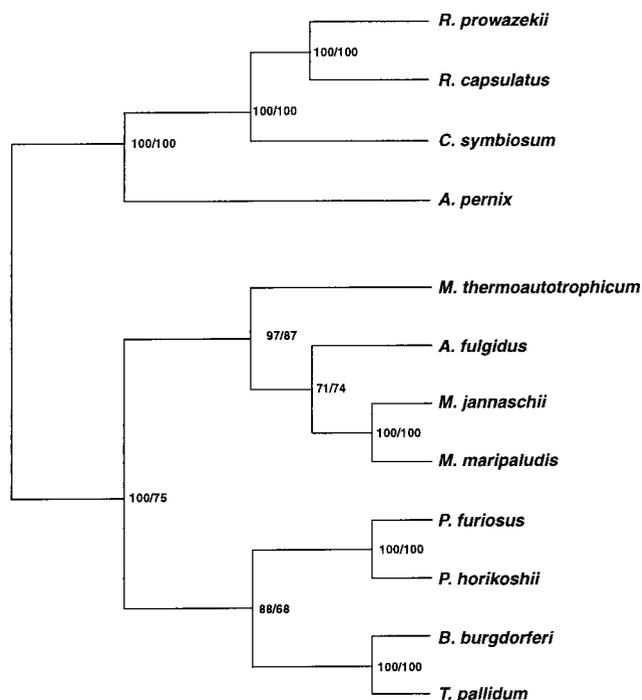


FIG. 1. Unrooted phylogeny of class I LysRSs. Phylogenies were constructed by using the maximum likelihood method (1,000 puzzling steps) implemented in the program PUZZLE 4.0 (48) and protein parsimony methods (200 bootstrap replicates) implemented in the Phylyp package version 3.5c (J. Felsenstein, University of Washington). Numbers represent the percentage occurrence of nodes from maximum likelihood and protein parsimony methods, respectively. Sequences were aligned by using the program CLUSTAL X (49).

anobacterium thermoautotrophicum, and *Archaeoglobus fulgidus*) and in Crenarchaeota (a species of *Cenarchaeum*, the ubiquitous marine archaeon, and *A. pernix*, an obligately aerobic hyperthermophile). The known bacterial class I *lysS* genes are found in two highly disparate taxa, the spirochetes [*B. burgdorferi* and *Treponema pallidum* (13)] and the α -proteobacteria (*Rickettsia prowazekii* and *Rhodobacter capsulatus*). The euryarchaeal versions of the enzyme form a distinctly coherent phylogenetic grouping in agreement with the 16s rRNA phylogeny (27, 28), from which the crenarchaeal examples are quite distant. In contrast, the bacterial LysRS proteins do not form a single grouping. This suggests that the bacterial class I LysRS proteins may have arisen through two distinct gene transfer events from the Archaea to Bacteria: LysRS in α -proteobacteria arising from the crenarchaeal LysRS and that in the spirochetes arising from the pyrococcal LysRS.

To compare the tRNA substrates of class I and class II LysRS enzymes, the tRNA^{Lys} gene of *M. maripaludis* was cloned and sequenced. Comparison with other tRNA^{Lys} species (Fig. 2) indicates a high degree of conservation, particularly at nucleotides 1–3, 34–36, and 70–73, all of which are known to be important during interaction with LysRS (see below). In contrast to the LysRS encoding genes, there appear to be no obvious divisions at the primary sequence level between the lysine tRNAs. The only exception is the G2:U71 wobble pair, which is found exclusively in spirochete lysine tRNA species, which are substrates of class I LysRS enzymes.

Biochemical Phylogeny of Class I LysRS. Sequence analysis indicates that a subgroup of the known LysRS enzymes are class

I aminoacyl-tRNA synthetases. Detailed comparison of these sequences with those of other class I aminoacyl-tRNA synthetases suggests that they are related, albeit distantly, to CysRS, ArgRS, GluRS, and GlnRS (M.I. and D.S., unpublished results; ref. 29). The last three enzymes are unique among the known aminoacyl-tRNA synthetases in that they require the presence of tRNA to perform the first step of the aminoacylation reaction, aminoacyl-adenylate synthesis (ref. 30 and Eq. 1).



where AA equals amino acid and E equals aminoacyl-tRNA synthetase.

We therefore tested the ability of class I LysRS enzymes to catalyze aminoacyl-adenylate synthesis using the pyrophosphate exchange reaction (Fig. 3A). In the absence of tRNA, no significant turnover of ³²P from PP_i to ATP was observed whereas the addition of tRNA stimulated label exchange at a rate comparable to that observed for other aminoacyl-tRNA synthetases. These results provide experimental support for the classification of spirochete and archaeal LysRS proteins as class I aminoacyl-tRNA synthetases. To further investigate this classification, we tested the ability of the class I LysRS to support the synthesis of the alarmone Ap₄A, a compound produced both *in vivo* and *in vitro* by class II LysRS enzymes through subversion of the aminoacyl-adenylate biosynthesis pathway (26). After RNase treatment to inhibit Lys-tRNA^{Lys} biosynthesis, only an *E. coli* cell-free extract, but not *B. burgdorferi* or *M. maripaludis* LysRS, was able to synthesize Ap₄A (Fig. 3B). These clear differences in the process of aminoacyl-adenylate synthesis between the class I and class II LysRS enzymes provide phenotypic support for their

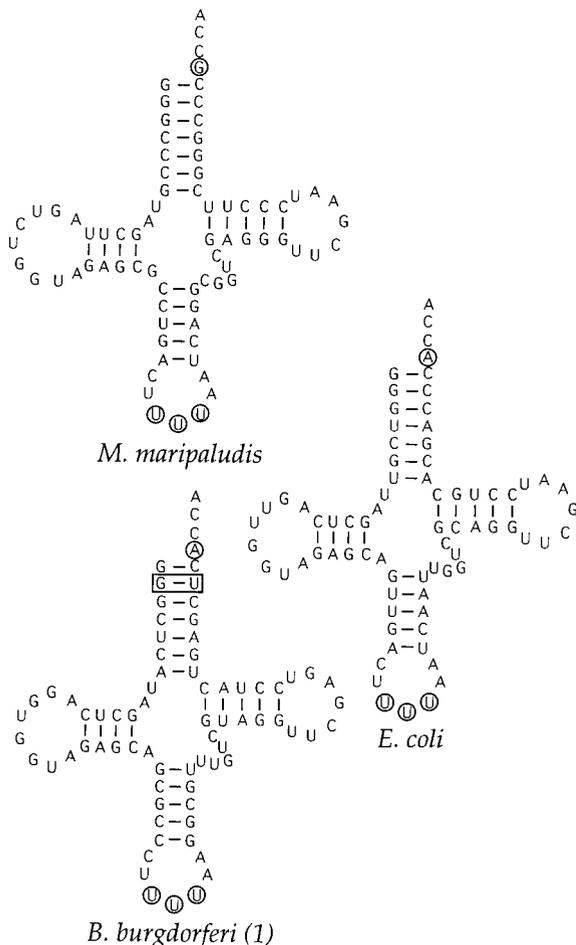


FIG. 2. Predicted secondary structure of unmodified tRNA^{Lys} from *B. burgdorferi*, *E. coli*, and *M. maripaludis*. The anticodon and discriminator base are circled; the G2:U71 wobble pair is boxed.

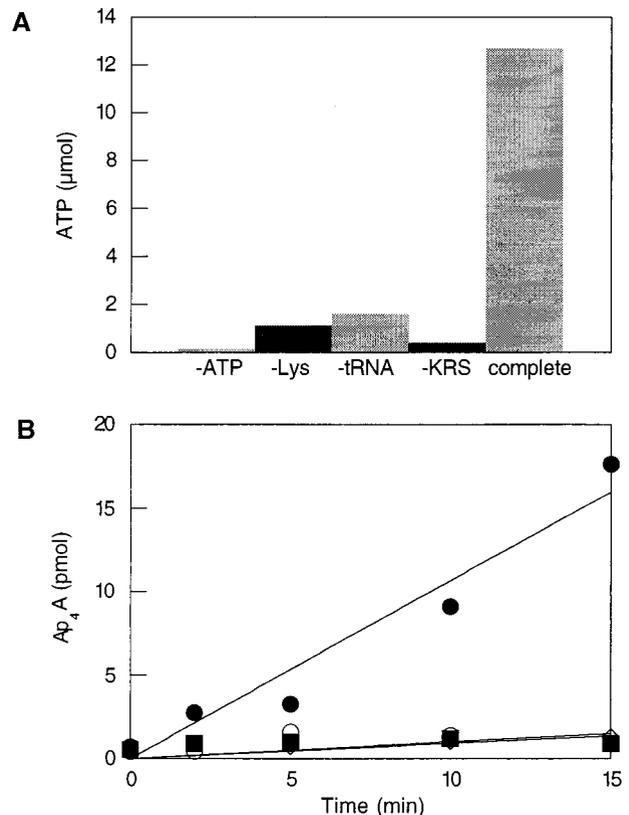


FIG. 3. (A) Pyrophosphate exchange by *M. maripaludis* LysRS. The reaction was incubated for 20 min at 37°C in the absence of various components as indicated. Values represent micromoles of ATP in a 20-μl aliquot. (B) *In vitro* Ap₄A biosynthesis. Reactions were performed in the presence of *E. coli* cell-free extract (●), *M. maripaludis* LysRS (○), or *B. burgdorferi* LysRS (■). Values represent picomoles of Ap₄A in a 50-μl aliquot.

classification within different divisions of the aminoacyl-tRNA synthetases.

Protein-tRNA Molecular Recognition by Class I LysRS. To investigate the ability of class I LysRS enzymes to recognize a tRNA^{Lys}, which is normally a substrate for a class II LysRS, we attempted to complement an *E. coli* strain in which the chromosomal copies of the *lysS* and *lysU* genes, which encode two isoforms of class II LysRS, have been disrupted [growth is maintained at permissive temperatures by a plasmid borne copy of *lysU* (ref. 18 and Table 1)]. Under normal growth conditions, only the *B. burgdorferi lysS* gene was able to rescue growth at a restrictive temperature, albeit less efficiently than *E. coli lysS*. Addition of exogenous lysine was necessary for phenotypic rescue by the *M. maripaludis lysS* gene, probably as a result of tRNA-dependent reduction in the lysine affinity of the enzyme (see below). These results indicate that *in vivo* class I LysRS enzymes can functionally substitute for their class II counterparts.

The molecular basis of tRNA recognition by class I LysRS was investigated by using both the archaeal enzyme from *M. maripaludis* and the bacterial enzyme from *B. burgdorferi*. It has been shown that *M. vanielli* tRNA^{Lys} contains the modified nucleotide methyl-aminomethyl-2-selenouridine (mnm⁵Se²U) at position 34 (31). Because this organism is the closest known relative of *M. maripaludis*, with which it shares an identical gene sequence for tRNA^{Lys}, we assumed that *M. maripaludis* tRNA^{Lys} was modified similarly. In addition, it has been shown that *E. coli* tRNA^{Lys} contains an analogous modification of U34, methyl-aminomethyl-2-thiouridine (mnm⁵S²U). The reactivity of the thio- and seleno-moieties toward cyanogen bromide is well documented (Eq. 2; refs. 24 and 31) and was used as the basis for probing anticodon recognition by *M. maripaludis* LysRS (Table 2).



where X equals S or Se. Modification of both *M. maripaludis* and *E. coli* tRNA led to small reductions in the *k*_{cat} for Lys-tRNA^{Lys} synthesis whereas modification of the *E. coli* tRNA caused a significant fall in the *K*_M for tRNA. This improvement in the apparent affinity for tRNA^{Lys} on modification is in contrast to the loss of recognition by *E. coli* LysRS of the same substrate (data not shown), indicating that, although both the class I and class II LysRS enzymes recognize the anticodon, they do so differently. Changes also were observed in the *K*_M for lysine, which depended strongly on both the source of the tRNA and modification of the anticodon. Anticodon recognition has been implicated in tRNA-dependent amino acid activation in *E. coli* GlnRS (32) and correlates with the requirement for tRNA during lysyl-adenylate synthesis reported above.

The sequence-specific recognition of tRNA^{Lys} was investigated *in vitro* by using unmodified RNA transcripts of wild-type and mutant tRNA genes. These studies were confined to the *B. burgdorferi* LysRS because the transcript of *M. maripaludis* tRNA^{Lys} could not be aminoacylated by its cognate synthetase under normal conditions (data not shown). To allow direct comparison of tRNA recognition by class I and class II LysRS

Table 1. Complementation of *E. coli lysS lysU* double mutants by genes encoding class I LysRS proteins

Plasmid	Growth conditions*				
	30°C	40°C	40°C + Lys	42°C	42°C + Lys
pCBS1- <i>Bs trpS</i>	++	-	-	-	-
pKS- <i>Ec lysS</i>	++	+	+	±	±
pCBS1- <i>Mm lysS</i>	++	-	±	-	-
pCBS1- <i>Bb lysS</i>	++	+	+	-	-

++, growth after 24 h; +, growth after 48 h; ±, growth after 72 h; -, no growth.

*Growth was on Luria-Bertani agar (100 μg/ml ampicillin) with the addition of 5 mM L-lysine where indicated.

Table 2. Recognition of tRNA by *M. maripaludis* LysRS

tRNA	<i>K</i> _M tRNA, μM	<i>K</i> _M Lys, μM	<i>k</i> _{cat} , S ⁻¹
<i>M. maripaludis</i> *	0.11 ± 0.03	2.2 ± 0.2	0.2 ± 0.05
<i>M. maripaludis</i> -CNBr†	0.08 ± 0.01	25.0 ± 8.0	0.05 ± 0.01
<i>E. coli</i> ‡	3 ± 0.8	350 ± 47	0.004 ± 0.0004
<i>E. coli</i> -CNBr†	0.25 ± 0.05	230 ± 48	0.002 ± 0.0001

**M. maripaludis* total tRNA (40 pmol tRNA^{Lys}/A₂₆₀).

†tRNA was treated with CNBr as described.

‡*E. coli* tRNA^{Lys}.

enzymes, the tRNA^{Lys} mutant set previously studied in the *E. coli* system was investigated (ref. 33; Table 3). The class I LysRS shows a nearly identical recognition pattern to its class II counterpart. The anticodon is recognized primarily at U35 and U36 and, to a slightly lesser extent, at U34. This correlates with the observation that *B. burgdorferi* contains two tRNA^{Lys} isoacceptors with the anticodons UUU and CUU and thus would not be expected to specifically recognize U34. The identity of the discriminator base (N73) also was shown to be important during interaction with tRNA^{Lys}, although mutation of this position was less detrimental to recognition by class I than by class II LysRS. Of interest, this increased ability to tolerate mutation of the discriminator base also has been described for the human class II-type LysRS (34). The ability of *B. burgdorferi* LysRS to aminoacylate a transcript of one of its cognate tRNA genes (tRNA^{Lys}) also was investigated (Table 4). In contrast to the *M. maripaludis* and *E. coli* tRNA^{Lys} transcripts, *B. burgdorferi* tRNA^{Lys} was an excellent substrate for aminoacylation, indicating that modified nucleotides do not play a significant role in the recognition of this tRNA. Furthermore, the 160-fold difference in the catalytic efficiency (*k*_{cat}/*K*_M) of aminoacylation between the *E. coli* and *B. burgdorferi* tRNA transcripts indicates that nucleotides, in addition to the anticodon and discriminator base, play a role in recognition by the bacterial class I LysRS.

Rejection of tRNA^{Lys} by Class II LysRS. All tRNA^{Lys} isoacceptors encoded in the genomes of *B. burgdorferi* and *T. pallidum* contain a G2·U71 wobble pair, in contrast to all other known lysine tRNAs, which contain a canonical Watson-Crick pair at this position. To probe the function of this base pair, position 71 of *E. coli* tRNA^{Lys} was mutated, resulting in a change from G2·C71 → G2·U71. The *in vitro* recognition of a transcript of this mutant gene then was investigated by using *B. burgdorferi* and *E. coli* LysRS (Table 4). Although the presence of the wobble pair has little effect on recognition by *B. burgdorferi* class I LysRS, it acts as a strong antideterminant for the *E. coli* class II LysRS. The G2·U71 variant shows a 73-fold reduction in the catalytic efficiency for aminoacylation compared with wild-type, which cor-

Table 3. Recognition of tRNA^{Lys} by *B. burgdorferi* LysRS

tRNA ^{Lys}	<i>B. burgdorferi</i> LysRS <i>V</i> _{max} / <i>K</i> _M , relative*	<i>E. coli</i> LysRS <i>V</i> _{max} / <i>K</i> _M , relative†
Wild-type	1	1
Anticodon		
UCU	<0.04	<0.01
UAC	<0.04	<0.01
UUA	<0.04	<0.01
UUG	0.08	0.22
GUU	0.34	0.059
Discriminator		
G73	0.28	0.11
C73	0.28	0.10
U73	0.29	0.12

*For the wild-type tRNA^{Lys}, *k*_{cat} = 0.22 ± 0.03 min⁻¹ and *K*_M = 1.5 ± 0.6 μM.

†Values taken from ref. 33 (for wild-type, *k*_{cat} = 3.32 min⁻¹ and *K*_M = 1.9 μM).

Table 4. Recognition of *in vitro* transcribed tRNA^{Lys} variants by *B. burgdorferi* and *E. coli* LysRS

LysRS	tRNA	K_M tRNA, μM	k_{cat} , min^{-1}	k_{cat}/K_M , R^*
<i>B. burgdorferi</i>	<i>B. burgdorferi</i>	1.98 ± 0.32	47.4 ± 4.2	1
	<i>E. coli</i> wild type	1.5 ± 0.6	0.22 ± 0.003	160
	<i>E. coli</i> G2·U71	3.9 ± 0.6	0.21 ± 0.02	432
<i>E. coli</i>	<i>B. burgdorferi</i>	4.3 ± 0.4	0.21 ± 0.01	490
	<i>E. coli</i> wild type [†]	1.9	3.32	14
	<i>E. coli</i> G2·U71	6.3 ± 0.8	0.15 ± 0.008	1022

*Relative value expressed as fold decrease in k_{cat}/K_M compared to aminoacylation of *B. burgdorferi* tRNA^{Lys} transcript by *B. burgdorferi* LysRS.

[†]From ref. 33.

responds to a difference in Gibbs' free energy for tRNA binding ($\Delta\Delta G_{\text{tRNA}}$) of $2.7 \text{ kcal}\cdot\text{mol}^{-1}$.

DISCUSSION

Molecular Phylogeny Shows an Archaeal Origin of Bacterial Class I LysRS. The phylogenetic analysis of class I LysRS amino acid sequences shows a relationship between the archaeal examples, which reiterates phylogenies deduced from small subunit rRNA sequences. There is a clear separation of the crenarchaeal and euryarchaeal kingdoms, with the *Pyrococci* forming a coherent group within the latter (35). The bacterial class I LysRS proteins do not group together, instead grouping with the pyrococcal sequences in the case of the spirochetes and with *Cenarchaeum symbiosum* in the case of the α -proteobacteria. This separation of the bacterial sequences, which are not deeply rooted in the class I LysRS phylogeny, indicates that they may have arisen relatively recently after two separate horizontal gene transfer events from archaea. The absence of class I LysRS sequences from the genomic sequences of more deeply rooted bacteria such as *Aquifex aeolicus* (36) and *Thermotoga maritima* indicates that the class I LysRS was absent during the early evolution of bacteria, supporting an archaeal origin for the contemporary bacterial examples of this protein. Furthermore, preliminary investigation of additional partial genomic sequences indicates the presence in other bacteria (e.g., *Streptomyces coelicolor*) of class I LysRS proteins, suggesting that they may be distributed even more widely than originally anticipated (13).

Given that the molecular phylogeny shows an archaeal origin for bacterial class I LysRS proteins, this now raises the question of the origin of the enzyme itself. Two possibilities exist (13): either the class I and class II LysRS proteins coexisted in the universal ancestor (37) or the class I LysRS arose early in the archaeal lineage after duplication and diversification of another class I aminoacyl-tRNA synthetase. Comparison of LysRS with other class I aminoacyl-tRNA synthetases shows that these enzymes form a separate group that is not closely related to any other single group of synthetases (M.I. and D.S., unpublished results; ref. 29). If the class I LysRS had arisen in the archaeal lineage, then it would be expected to be related to an archaeal cluster of synthetases of particular amino acid specificity, in much the same way that all glutaminyl-tRNA synthetases are most closely related to eukaryotic, but not bacterial, glutamyl-tRNA synthetases (38). The lack of a close homology between LysRS and other class I aminoacyl-tRNA synthetases clearly indicates that the class I-type LysRS was present in the universal ancestor of extant life whereas phylogenetic data suggests that it was only retained during the evolution of the archaeal lineage.

Proteobacteria Can Accommodate a Class I LysRS. The horizontal gene transfer scenario described above is strongly supported by the observation that class I LysRS encoding-genes can substitute functionally for their class II counterparts in *E. coli*. The archaeal gene from *M. maripaludis* was considerably less efficient at rescuing the growth phenotype of the *E. coli* mutant than the bacterial gene from *B. burgdorferi*. Because both the *B. burgdorferi* and *M. maripaludis* *lysS* genes can be expressed in *E. coli* to comparable levels (data not shown), this results from better recognition of *E. coli* tRNA^{Lys} by the spirochete enzyme, as

confirmed by *in vitro* analyses. This difference in recognition is accentuated by the tRNA-dependent recognition of lysine by class I LysRS, which renders the *E. coli* mutant transformed with the archaeal enzyme auxotrophic for lysine. Nevertheless, the observation that, under appropriate conditions *in vivo*, an archaeal class I-type LysRS can functionally replace a bacterial class II-type LysRS provides experimental support for the horizontal gene transfer hypothesis outlined above.

Molecular Basis for Gene Displacement in Bacteria. The completion of a number of microbial genomic sequences recently has shown that horizontal gene transfer is widespread and thus may have contributed significantly to the evolution of many organisms (e.g., ref. 39). Horizontal gene transfer generally is regarded as an event that initially neither advantages nor handicaps the recipient cell but simply increases the size of the host's gene pool. The presence of two LysRS proteins in the ancestor of spirochetes after horizontal gene transfer would mimic both the universal ancestor and contemporary bacteria such as *E. coli*, which contain two isoforms of the class II enzyme (40). Whether the maintenance of two LysRS enzymes offers a selective advantage is unclear but may relate to an improvement in cadaverine resistance under certain conditions (41). Examination of the molecular determinants of tRNA^{Lys} recognition by class I and class II LysRS enzymes suggests that loss of the class II enzyme-encoding gene could have resulted from a point mutation in the recipient's tRNA^{Lys} gene. If this mutation gave rise to a non-canonical G2·U71 wobble pair in the acceptor helix of the folded tRNA (which to date has been found in lysine tRNAs exclusively from spirochetes), then recognition by the class II LysRS would be impaired severely. In contrast, such a mutation would have no significant effect on recognition by the class I LysRS enzyme, which then would be preferred by the cell to act as the provider of lysyl-tRNA, the substrate for translation of lysine codons. Under these conditions, the class II LysRS would be functionally displaced and there would be no selective advantage to retaining an accurate copy of the gene that encoded it, resulting in its degeneracy and eventual loss. Thus, differences in tRNA recognition between class I and class II LysRS enzymes provide a molecular basis for gene displacement. The recruitment of a heterologous aminoacyl-tRNA synthetase by a mutated tRNA is analogous to the "tRNA gene recruitment" hypothesis in which tRNAs with anticodon mutations are proposed to become substrates for homologous aminoacyl-tRNA synthetases of different amino acid specificity (42).

Substrate Conservation Leads to Functional Convergence of Divergent Enzymes. *In vitro* analysis of molecular recognition clearly indicates that both class I and class II LysRS proteins recognize the discriminator base (N73) and anticodon of tRNA^{Lys}. However, the marked differences in their abilities to recognize a tRNA with a chemically modified anticodon suggest that they bind tRNA differently. This was confirmed by the strikingly different effects on recognition of introducing a G2·U71 wobble pair into the acceptor helix of *E. coli* tRNA^{Lys}. Although this mutation had no significant effect on aminoacylation by the class I enzyme, it led to a 73-fold drop in the efficiency (k_{cat}/K_M) of the class II LysRS. An explanation for these effects can be found by comparison to the alanine system.

All known alanine tRNA species contain a G3:U70 wobble pair whose function is to distort the acceptor helix such that functional groups in the minor groove can be specifically recognized by alanyl-tRNA synthetase, a class II enzyme (43, 44). This correlates with observations from structural (8, 10) and biochemical (45) data that class I and class II synthetases approach the RNA acceptor helix from different sides. This is apparently also the case for the two types of LysRS, in which the ability of the class II LysRS to recognize tRNA^{Lys} is impaired by the exposure of functional groups in the minor groove of the G2:U71 mutant. Because the class I LysRS approaches the opposite side of the acceptor stem, its aminoacylation efficiency is not impaired by this mutation. Thus, although the class I and class II LysRS enzymes recognize the same sites in their RNA substrates, they do so differently (similarly, they differ in their exact mechanisms of lysine activation). Only small changes are observed in lysine tRNA sequences from organisms in different kingdoms; thus, there are no obvious gross differences that distinguish the substrates of class I and class II LysRS enzymes. (M.I. and D.S., unpublished results; ref. 29).

The functional convergence of two unrelated LysRS enzymes strongly suggests that the evolution of specificity in the aminoacyl-tRNA synthetases was driven by conserved, preexisting substrates. The possibility that tRNA^{Lys} existed before one class of LysRS is consistent with earlier suggestions that tRNAs are derived from more primitive structures not involved in translation (46) and that tRNAs in general predate aminoacyl-tRNA synthetases (47). In view of the experimental data presented here, it now seems reasonable to assume that transfer RNAs did indeed appear before at least some of the aminoacyl-tRNA synthetases during the evolution of the contemporary translational apparatus.

We thank Sylvain Blanquet, Pierre Plateau, and Koji Tamura for kindly providing strains and plasmids and Siv Andersson for access to the *R. prowazekii* *lysS* sequence before publication. We are greatly indebted to Alan Curnow, Alan Weiner, and Carl Woese for critical reading of the manuscript. Work in the author's laboratory was supported by a grant from the National Institute of General and Medical Sciences (to D.S.).

- Gesteland, R. F. & Atkins, J. F. (1993) *The RNA World* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Schimmel, P. & Ribas de Pouplana, L. (1995) *Cell* **81**, 983–986.
- Shiba, K., Motegi, H. & Schimmel, P. (1997) *Trends Biochem. Sci.* **22**, 453–457.
- Ibba, M., Curnow, A. W. & Söll, D. (1997) *Trends Biochem. Sci.* **22**, 39–42.
- Curnow, A. W., Tumbula, D., Pelaschier, J., Min, B. & Söll, D. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 12838–12843.
- Doolittle, R. F. (1998) *Nature (London)* **392**, 339–342.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990) *Nature (London)* **347**, 203–206.
- Cusack, S. (1997) *Curr. Opin. Struct. Biol.* **7**, 881–889.
- Nagel, G. M. & Doolittle, R. F. (1995) *J. Mol. Evol.* **40**, 487–498.
- Arnez, J. G. & Moras, D. (1997) *Trends Biochem. Sci.* **22**, 211–216.
- Brown, J. R. & Doolittle, W. F. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 2441–2445.
- Brown, J. R. (1998) in *Thermophiles: The Keys to Molecular Evolution and the Origin of Life?*, eds. Wiegand, J. & Adams, M. H. W. (Taylor & Francis, London), pp. 217–230.
- Ibba, M., Bono, J. L., Rosa, P. A. & Söll, D. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 14383–14388.
- Ibba, M., Morgan, S., Curnow, A. W., Pridmore, D. R., Vothknecht, U. C., Gardner, W., Lin, W., Woese, C. R. & Söll, D. (1997) *Science* **278**, 1119–1122.
- Syvanen, M. (1994) *Annu. Rev. Genet.* **28**, 237–261.
- Curnow, A. W., Garcia, G. A., Kung, F. L. & Koch, K. (1993) *Biochemistry* **32**, 5239–5246.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1990) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
- Chen, J., Brevet, A., Lapadat-Tapolosky, M., Blanquet, S. & Plateau, P. (1994) *J. Bacteriol.* **176**, 2699–2705.
- Wich, G., Jarsch, M. & Böck, A. (1984) *Mol. Gen. Genet.* **196**, 146–151.
- Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., *et al.* (1996) *Science* **273**, 1058–1072.
- Kwak, J. H. & Kim, M. Y. (1995) *Anal. Biochem.* **228**, 178–180.
- Sako, Y., Nomura, N., Uchida, A., Ishida, Y., Morii, H., Koga, Y., Hoaki, T. & Maruyama, T. (1996) *Int. J. Syst. Bacteriol.* **46**, 1070–1077.
- Wells, T. N., Knill-Jones, J. W., Gray, T. E. & Fersht, A. R. (1991) *Biochemistry* **30**, 5151–5156.
- Agris, P. F., Söll, D. & Seno, T. (1973) *Biochemistry* **12**, 4331–4337.
- Ogilvie, A. (1981) *Anal. Biochem.* **115**, 302–307.
- Brevet, A., Chen, J., Leveque, F., Plateau, P. & Blanquet, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8275–8279.
- Olsen, G. J., Woese, C. R. & Overbeek, R. (1994) *J. Bacteriol.* **176**, 1–6.
- Barns, S. M., Delwiche, C. F., Palmer, J. D. & Pace, N. R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 9188–9193.
- Ribas de Pouplana, L., Turner, R. J., Steer, B. A. & Schimmel, P. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 11295–11300.
- Schimmel, P. R. & Söll, D. (1979) *Annu. Rev. Biochem.* **48**, 601–648.
- Ching, W. M., Alzner-DeWeerd, B. & Stadtman, T. C. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 347–350.
- Ibba, M., Hong, K.-W., Sherman, J. M., Sever, S. & Söll, D. *Proc. Natl. Acad. Sci. USA* **93**, 6953–6958.
- Tamura, K., Himeno, H., Hasegawa, T. & Shimizu, M. (1992) *Nucleic Acids Res.* **20**, 2335–2339.
- Shiba, K., Stello, T., Motegi, H., Noda, T., Musier-Forsyth, K. & Schimmel, P. (1997) *J. Biol. Chem.* **272**, 22809–22816.
- Pace, N. R. (1997) *Science* **276**, 734–740.
- Deckert, G., Warren, P. V., Gaasterland, T., Young, W. G., Lenox, A. L., Graham, D. E., Overbeek, R., Snead, M. A., Keller, M., Aujay, M., *et al.* (1998) *Nature (London)* **392**, 353–358.
- Woese, C. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6854–6859.
- Siatecka, M., Rozek, M., Barciszewski, J. & Mirande, M. (1998) *Eur. J. Biochem.* **256**, 80–87.
- Doolittle, W. F. & Logsdon, J. M. (1998) *Curr. Biol.* **8**, R209–R211.
- Leveque, F., Plateau, P., Dessen, P. & Blanquet, S. (1990) *Nucleic Acids Res.* **25**, 305–312.
- Brevet, A., Chen, J., Leveque, F., Blanquet, S. & Plateau, P. (1995) *J. Biol. Chem.* **270**, 14439–14444.
- Saks, M. E., Sampson, J. R. & Abelson, J. (1998) *Science* **279**, 1665–1670.
- Gabriel, K., Schneider, J. & McClain, W. H. (1996) *Science* **271**, 195–197.
- Beuning, P. J., Yang, F., Schimmel, P. & Musier-Forsyth, K. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 10150–10154.
- Sissler, M., Eriani, G., Martin, F., Giege, R. & Florentz, C. (1997) *Nucleic Acids Res.* **25**, 4899–4906.
- Weiner, A. M. & Maizels, N. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 7383–7387.
- Woese, C. (1967) *The Genetic Code* (Harper & Row, New York).
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **25**, 4876–4882.
- Strimmer, K. & von Haeseler, A. (1996) *Mol. Biol. Evol.* **13**, 964–969.