Archaeal-type lysyl-tRNA synthetase in the Lyme disease spirochete Borrelia burgdorferi

MICHAEL IBBA*, JAMES L. BONO†, PATRICIA A. ROSA†, AND DIETER SOLL*‡§

Departments of *Molecular Biophysics and Biochemistry and †Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06520-8114; and ‡Laboratory of Microbial Structure and Function, Rocky Mountain Laboratories, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Hamilton, MT 59840

Contributed by Dieter Soll, October 22, 1997

ABSTRACT Lysyl-tRNAs are essential for protein biosynthesis by ribosomal mRNA translation in all organisms. They are synthesized by lysyl-tRNA synthetases (EC 6.1.1.6), a group of enzymes composed of two unrelated families. In bacteria and eukarya, all known lysyl-tRNA synthetases are subclass IIC-type aminoacyl-tRNA synthetases, whereas some archaea have been shown to contain an unrelated class I-type lysyl-tRNA synthetase. Examination of the preliminary genomic sequence of the bacterial pathogen Borrelia burgdorferi, the causative agent of Lyme disease, indicated the presence of an open reading frame with over 55% similarity at the amino acid level to archaeal class I-type lysyl-tRNA synthetases. In contrast, no coding region with significant similarity to any class II-type lysyl-tRNA synthetase could be detected. Heterologous expression of this open reading frame in Escherichia coli led to the production of a protein with canonical lysyl-tRNA synthetase activity in vitro. Analysis of B. burgdorferi mRNA showed that the lysyl-tRNA synthetase-encoding gene is highly expressed, confirming that B. burgdorferi contains a functional class I-type lysyl-tRNA synthetase. The detection of an archaeal-type lysyl-tRNA synthetase in B. burgdorferi and other pathogenic spirochetes, but not to date elsewhere in bacteria or eukarya, indicates that the gene that encodes this enzyme has a common origin with its orthologue from the archaeal kingdom. This difference between the lysyl-tRNA synthetases of spirochetes and their hosts may be readily exploitable for the development of anti-spirochete therapeutics.

The aminoacylation reaction plays a pivotal role in the essential cellular process of translation. It provides the only known biochemical means of interpreting the genetic code in terms of amino acids by esterifying tRNAs at their 3’ ends with an aminoacyl moiety which generally, but not always, corresponds to the appropriate anticodon. In the majority of cases, correct aminoacylation is accomplished in a single step by a particular aminoacyl-tRNA synthetase (AARS) that specifically selects the appropriate tRNA and amino acid from a large cellular pool of similar molecules. The accuracy of this process may be further enhanced by intrinsic proofreading and editing mechanisms. However, there also exist a number of examples where the AARS instead serves to activate a non-cognate amino acid, while the generation of the correct aminoacyl-tRNA pair is subsequently brought about by a second protein. The use of such pathways for the formation of Gln-tRNA\textsuperscript{Gln} (via Glu-tRNA\textsuperscript{Gln}) and selenocysteinyl-tRNA\textsuperscript{Sec} (via Ser-tRNA\textsuperscript{Sec}) is well documented in all the living kingdoms (4–6). It has also been found that in several archaea an additional aminoacyl-tRNA, Asn-tRNA\textsuperscript{Asn}, is formed by trans-

formation of a mischarged tRNA rather than by direct aminoacylation with asparaginyl-tRNA synthetase (7).

The use of two-step (indirect) aminoacylation pathways for the formation of Asn-tRNA\textsuperscript{Asn} and Gln-tRNA\textsuperscript{Gln} in some organisms circumvents the need for the enzymes that catalyze one-step formation of these molecules, the asparaginyl-tRNA synthetase (AsnRS) and glutaminyl-tRNA synthetase (GlnRS), respectively. Consequently, it is not surprising that genes encoding these enzymes have not been found in the completed genomic sequences of organisms that employ one or both of the indirect pathways (e.g., see refs. 8 and 9). However, in addition to lacking AsnRS and GlnRS, the genomic sequences of the euryarchaeons Methanococcus jannaschii and Methanobacterium thermoaerotrophicum do not contain genes encoding homologues of known cysteinyl-tRNA synthetase (CysRS) or lysyl-tRNA synthetase (LysRS; EC 6.1.1.6) (8, 9).

While no adequate explanation yet exists for the apparent absence of CysRS, it has recently been shown that several members of the archaea, including M. jannaschii, contain a LysRS encoded by a lys\textsuperscript{S} gene with no resemblance to known bacterial or eukaryal LysRS-encoding genes or any other sequences in the public database (10). This is in contrast to all other AARSs, which are highly conserved throughout the living kingdom. The high degree of conservation is exemplified by the invariant classification of AARSs into one of two classes defined by the presence of characteristic amino acid sequence motifs (11) and topologically distinct nucleotide binding folds (12, 13). This is not true of the recently identified archaeal LysRSs, which are class I-type AARSs and are thus easily distinguished from their known bacterial and eukaryal counterparts, which are class II enzymes.

Whereas it was initially assumed that this novel type of LysRS was confined to certain Archaea, continued genomic sequencing efforts have suggested that it may also be found in some bacteria. Specifically, open reading frames (ORFs) encoding proteins with significant similarity to the archaeal-type LysRS were found in two genera of spirochetes, Borrelia burgdorferi and Treponema pallidum, the causative agents in humans of Lyme disease (14) and syphilis (15), respectively. Spirochetes are a phylogenetically ancient bacterial group (16–18) that share a unique morphology. The apparent existence of archaeal-type genes in at least two spirochetal genera raises questions concerning their evolutionary origin and development. The complete analysis of the B. burgdorferi genome will reveal if, in addition to lys\textsuperscript{S}, other genes or unidentified ORFs resemble uniquely archaeal ORFs. Furthermore, the possibility that these pathogens contain an essential enzyme that bears no resemblance to the host protein that performs the same cellular function suggests that the

Abbreviations: AARS, aminoacyl-tRNA synthetase; LysRS, lysyl-tRNA synthetase. To whom reprint requests should be addressed at: Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208114, 266 Whitney Avenue, New Haven, CT 06520-8114. e-mail: soll@trna.chem.yale.edu.
archaeal-type LysRS may prove to be a novel therapeutic target. We therefore attempted to isolate and characterize the gene encoding the archaeal-type LysRS from \textit{B. burgdorferi}.

**MATERIALS AND METHODS**

\textbf{General.} Media for bacterial growth and molecular biology protocols were standard unless otherwise noted (19). \textit{Borrelia} were grown in liquid BS-K H medium (Sigma) supplemented with 6% rabbit serum at 23°C or 35°C as indicated. Aminoacylation assays were performed as previously described (10) except that \textit{Escherichia coli} tRNALys (1,400 pmol \textit{A}$_{260}$ Sigma) was used. The preparation and analysis of DNA from \textit{Borrelia} species were previously reported (20).

\textbf{Cloning of the \textit{B. burgdorferi} \textit{lysS} Gene.} The putative \textit{lysS} gene of \textit{B. burgdorferi} was amplified from genomic DNA (21) (kindly provided by D. R. Akins and J. D. Radolf) by PCR under standard conditions with annealing at 50°C. The primers used were BBKRS1 (TA\textit{AATAATGGACATATTGGT-GA\textit{AAAACACGCACTGTG}}; engineered NdeI site in bold, translation initiation site in italics), which is based upon the \textit{E. coli} codon of \textit{lysS}. The 5’ and 3’ termini of the cloned PCR product were sequenced to confirm that the \textit{lysS} gene would be transcribed in the desired reading frame.

\textbf{Purification of \textit{B. burgdorferi} LysRS.} For overexpression of \textit{B. burgdorferi} lysS, pET15b-Bb\textit{lysS} was used to transform the strain BL21 (DE3) (23). Transformants were then grown in 5 ml of Luria broth (LB) containing 100 $\mu$g/ml ampicillin at 30°C for 15 h. This culture was used as inoculum for 500 ml of LB (100 $\mu$g/ml ampicillin), which was grown with moderate shaking (120 rpm) and the addition of further ampicillin (50 $\mu$g/ml) when the OD$_{600}$ reached 0.3 and 0.6. At the latter value, isopropyl $\beta$-D-thiogalactopyranoside (IPTG) was also added to a final concentration of 200 $\mu$M to induce expression of the His$_6$-tagged \textit{lysS} gene. After 40 min of induction, the culture was briefly chilled on ice (all subsequent manipulations were performed at 4°C) and then centrifuged at 10,000 $\times$ g for 10 min, and the cell pellet was washed in PBS and finally resuspended in buffer A [50 mM Tris-HCl, pH 8/500 mM NaCl/5 mM MgCl$_2$/5 mM 2-mercaptoethanol/0.1% Triton X-100/10% (vol/vol) glycerol] containing 1 mM phenylmethylsulfonyl fluoride and 10 $\mu$g/ml lysozyme. This suspension was sonicated and then centrifuged at 100,000 $\times$ g for 1 h, and the supernatant was retained and mixed with 2 ml of Ni-NTA-agarose (Qiagen). The resulting slurry was washed with buffer A containing 30 mM imidazole and the proteins of interest were then eluted at 750 mM imidazole. This fraction was exchanged into buffer B (50 mM Hepes, pH 7.2/25 mM KCl/10 mM MgCl$_2$/5 mM dithiothreitol/10% glycerol) by using a PD-10 column and applied to a Mono-S column (both Pharmacia). The column was developed with a 0–300 mM NaCl gradient in buffer B. Fractions containing solely His$_6$-LysRS (as judged by Coomassie blue staining after SDS-PAGE) were pooled, concentrated by ultrafiltration, and stored in buffer B containing 20% glycerol at $-70^\circ$C.

\textbf{Northern Blot Analysis.} Total borrelial RNA was extracted from logarithmic-phase cultures by using the Ultraspec RNA isolation system (Biotecx, Houston, TX) according to the manufacturer’s instructions. For the temperature-shift experiment, borreliae were grown at 23°C to mid-logarithmic phase, diluted 1:100 in fresh medium, and grown to mid-logarithmic phase at 35°C. RNA was denatured with glyoxal and dimethyl sulfoxide and electrophoresed in a 1% (wt/vol) gel. RNA was transferred to nylon membranes (Micron Separations), cross-linked with ultraviolet light, and air dried. Prehybridization and hybridization were at 55°C in 1% (wt/vol) BSA/7% (wt/vol) SDS/0.5 M sodium phosphate, pH 7.0/1 mM EDTA in rotating bottles in a hybridization oven (Belloco). Flagellin

\textbf{FIG. 1.} Structure of the archaeal-type \textit{lysS} ORF of \textit{B. burgdorferi}. A putative $\sigma^7$-dependent $-35/-10$ promoter region is presented (the consensus sequence is shown above the boxed regions (38)); a Shine-Delgano (SD) motif is indicated, as are the putative start (GTG) and stop (TAA) codons.

\textbf{FIG. 2.} Alignment of class I-type LysRS amino acid sequences. The sequences were aligned by using the CLUSTAL W program (26). The sequences shown are from \textit{Archaeoglobus fulgidus} (AF), \textit{B. burgdorferi} (BB), \textit{M. jannaschii} (MJ), \textit{M. maripaludis} (MM), \textit{M. thermoautotrophicum} (MT), and \textit{T. pallidum} (TP).
were searched by using the gapped BLAST algorithm (25). The extensive charging of purified lysS were probed either with burgdorferi were subjected to Northern analysis. RNA blots proteins was calculated by using the program BESTFIT from the degree of similarity at the amino acid level between particular proteins was calculated by using the program BESTFIT from the Wisconsin Package Version 8.1 (Genetics Computer Group). Multiple sequence alignments were generated with CLUSTAL W (26) and annotated remotely with BOXSHADE (http://ulrec3.unil.ch/software/BOX_form.html).

RESULTS

Genomic Analysis. Searching of the genomic sequence of B. burgdorferi revealed a region sharing a high degree of similarity with the lysS gene of Methanococcus maripaludis (10) but did not identify an ORF encoding a protein homologous to any known class II-type LysRS. Further examination showed that this region contains a putative canonical bacterial ORF encoding a gene similar in size to M. maripaludis lysS (Fig. 1). Translation of this ORF gives rise to an amino acid sequence with extensive similarity to euryarchaeal class I-type LysRSs (Fig. 2, Table 1). A comparable ORF was found in the genome of another spirochete, T. pallidum (Fig. 2, Table 1), although it should be noted that in this case a sequence was also detected that encodes a 345-amino acid protein similar (32% identity and 52% similarity at the amino acid level) to the carboxy-terminal 70% of the class II-type LysRS of Thermus thermophilus (27).

Cloning and Heterologous Expression of B. burgdorferi lysS. To determine whether B. burgdorferi lysS encodes a functional LysRS enzyme the gene was directly cloned from B. burgdorferi genomic DNA and heterologously expressed in E. coli with a 5’ His6 epitope tag. Induction of expression led to the production of a His6-tagged protein that could readily be detected in whole cell extracts (but not in control extracts) by using a Ni-NTA-alkaline phosphatase conjugate (data not shown). Subsequent purification gave rise to an electrophoretically pure protein of the expected molecular weight (Fig. 3). The ability of this protein to aminoacylate tRNA with lysine was then investigated in vitro (Fig. 4). B. burgdorferi LysRS showed extensive charging of purified E. coli tRNA[lys] at catalytic enzyme concentrations; the inhibition of this activity by the addition of unlabeled lysine confirmed that the lysS gene encodes a protein with canonical LysRS activity.

Expression of B. burgdorferi lysS. To investigate whether the lysS gene is transcribed, total RNA preparations from B. burgdorferi were subjected to Northern analysis. RNA blots were probed either with lysS or, as a control, the flagellin gene, which has previously been shown to be constitutively expressed in B. burgdorferi (24). A transcript of the expected length was detected in RNA prepared from B. burgdorferi grown at both 23°C and 35°C, with the full-length transcript being considerably more stable at the higher temperature (Fig. 5). Comparison to the abundant flagellin transcript indicates that the lysS gene is expressed at high levels.

Distribution of lysS in Borrelia Species. Although the above data clearly indicate the presence of an archaeal-type lysS gene in a single isolate of B. burgdorferi sensu strictu, they provide no indication of how ubiquitous the gene is in borreliae. Southern analysis of genomic DNA from multiple isolates of three genospecies of Lyme disease spirochetes as well as the relapsing fever spirochete Borrelia hermsii, using B. burgdorferi lysS as a probe, showed the lysS gene to be present on the chromosome in all borreliae tested (Fig. 6 and data not shown). The lysS probe did not hybridize to any of the circular or linear plasmids. These data strongly suggest that lysS is common to all members of this genus and the gene is present on the chromosome.

DISCUSSION

B. burgdorferi Contains an Archaeal-Type LysRS. Preliminary analysis of the sequence of the linear chromosome of B.
burgdorferi indicated the presence of an ORF encoding a protein with over 55% similarity (34% identity) at the amino acid level to the class I-type LysRSs previously identified in archaea (Table 1). In contrast, no coding region with significant similarity to any class II-type LysRSs could be detected. The borrelial genome is unusual, as it is composed of a linear chromosome and numerous linear and circular plasmids and is therefore sometimes regarded as a segmented genome (28, 29). Although it has been shown that some essential genes, such as those required for purine biosynthesis (30), are found on plasmids rather than the linear chromosome, this does not appear to be the case for the AARSs. Regions of the linear chromosome could be found encoding putative proteins with 50–75% similarity at the amino acid level to known AARSs, allowing the identification of orthologues for 18 of the 20 members of this family (data not shown). The nonorthologous replacement of LysRS reported here accounts for one of the two missing orthologues, the synthesis of Gln-tRNAGln by a transamidation pathway (4), rather than by GlnRS, for the other. In the latter case, an operon encoding an orthologue of the Glu-tRNA^Glu amidotransferase was found on the linear chromosome, indicating that all of the enzymatic activities necessary for aminoacyl-tRNA synthesis are encoded in this part of the genome. Thus, although it cannot be excluded, it seems unlikely that an ORF encoding a class II-type LysRS exists on one of the plasmids rather than on the chromosome. If this is indeed the case, then Lys-tRNALys for ribosomal mRNA translation in B. burgdorferi must be exclusively synthesized by the product of the class I-type lysS gene.

Cloning and heterologous expression of the B. burgdorferi class I-type lysS gene allowed the production and purification of the encoded protein. In vitro aminoacylation assays confirmed that this protein could catalyze the specific charging of tRNALys with lysine (Fig. 4), and is by definition a LysRS (31). Analysis of total RNA showed that although there is more full-length lysS transcript in B. burgdorferi grown at 35°C than at 23°C, the gene is highly expressed at both temperatures, as would be expected of an essential gene (Fig. 5). This temperature shift of in vitro grown organisms in part mimics the in vivo situation in the tick midgut after a bloodmeal, which correlates with the peak in spirochetal cell division and a concomitant elevation in the level of translation (32). In addition to more rapid growth, tick feeding (and temperature shift in vitro)
correlates temporally with a switch in outer surface protein gene expression (33–35), presumably facilitating transmission from the tick vector and adaptation to the mammalian host. Taken together these data confirm that the class I-type LysRS is functional in Lys-tRNALys synthesis in *B. burgdorferi*, effectively compensating for the lack of a conventional bacterial class II LysRS.

**Evolutionary Origin of Lys-tRNA<sup>lys</sup> Synthesis.** The occurrence of class I- and class II-type LysRS-encoding genes in living organisms is shown in Fig. 7. While a number of archaeal orders contain class I-type orthologues, Spirochaetaceae is the only bacterial “family” known to date that contains these genes, indicating two possible scenarios for the evolution of Lys-tRNALys synthesis. The occurrence of distinct groups of orthologues capable of carrying out the same core process in gene expression suggests that both of these activities could have been present in the universal ancestor (36), with different lineages retaining or losing either the class I- or the class II-type *lysS* gene during their evolution. The possibility that an ancestral organism may have contained both LysRS activities, but only a single group of lysine-specific tRNAs, is supported by the observation that the species specificity of tRNA<sup>lys</sup> recognition does not correlate with the utilization of a particular LysRS orthologue; for example *E. coli* tRNA<sup>lys</sup> is an excellent substrate for *B. burgdorferi* LysRS (Fig. 4). An alternative possibility is that the class I-type LysRS was not present in the universal ancestor, but instead arose later in evolution through duplication and subsequent divergence of another class I AARS gene such that it acquired the substrate specificity characteristic of LysRS. A comparable scheme has previously been proposed for the development of glutamyl-tRNA synthetase from glutamyl-tRNA synthetase (37). If such an event occurred early in the archaean lineage it could readily explain the broad use of the class I-type LysRS by contemporary archaea, and would in turn be indicative of an archaean origin for the class I gene in spirochetes, possibly as a result of gene transfer. A comparable pattern of codon usage, particularly the abundance of normally rare Arg codons, in *B. burgdorferi* and *M. maripaludis* could facilitate such a transfer (data not shown).

**Practical Implications of Nonorthologous Gene Replacement.** *B. burgdorferi* contains a functional class I-type LysRS, in contrast to its mammalian hosts, which contain class II-type enzymes. This fundamental difference between pathogen and host in an essential enzyme suggests that LysRS may be exploitable as a target for the development of Lyme disease antibiotics. A broader application of such antibiotics as anti-spirochetal agents is suggested by the presence of ORFs encoding a full-length class I, but only a significantly truncated and thus possibly inactive class II, LysRS in the genome of the spirochete *T. pallidum*, the causative agent of syphilis.

We thank the Institute for Genomic Research for availability of sequence data before publication. We are greatly indebted to D. R. Akins and J. D. Radolf for the initial gift of *B. burgdorferi* genomic DNA. We also thank A. W. Curnow and C. R. Woese for their insightful comments on the manuscript, M. Kitabatake for help with cloning, and P. P. Ibba for his cooperation throughout this work. This study was supported by a grant from the National Institute of General Medical Sciences (to D.S.).