Molecular Phylogenetic Analysis of Tryptophanyl-tRNA Synthetase of *Actinobacillus actinomycetemcomitans*

Narayanan Rajendran[a,*], Rajendram V. Rajnarayanan[b], and Donald R. Demuth[c]

[a] Biology, MASC, Kentucky State University, Frankfort, KY-40601, USA.
Fax: (5 02) 5 97–68 26. E-mail: narayanan.rajendran@kysu.edu
[b] Chemistry, NSD, Tougaloo College, Tougaloo, MS-39174, USA
[c] Center for Oral Health and Systemic Disease – School of Dentistry, University of Louisville, Louisville, KY-40292, USA

* Author for correspondence and reprint requests

Z. Naturforsch. **63c**, 418–428 (2008); received November 8, 2007/January 7, 2008

Aminoacyl-tRNA synthetase family enzymes are of particular interest for creating universal phylogenetic trees and understanding the gene flow as these enzymes perform the basic and analogous biochemical function of protein synthesis in all extant organisms. Among them, tryptophanyl-tRNA synthetase (Trp-RS) plays a foremost role in phylogeny owing to the close relationship with tyrosine-tRNA synthetase. In this study, the sequence of the gene *Trp-RS* was amplified using degenerated adenylation domain primers in the periodontal bacterium *Actinobacillus actinomycetemcomitans*. The sequence of the cloned PCR amplicon confirmed the adenylation domain sequence with glutamic acid residue, which is absent in five other oral bacteria used in this study as well as in a number of other bacteria described in the database. The Trp-RS sequence analysis prevailed the identify elements such as Rossmann-fold sequence and $tRNA^{Trp}$ binding domains including acceptor stem and anticodon. A theoretical model of Trp-RS of *A. actinomycetemcomitans* was generated. Guided docking of the ligand tryptophanyl-5′-AMP revealed a highly identical active site in comparison with the bacterial template. The phylogenetic positioning of Trp-RS among a group of oral bacterial species revealed that *A. actinomycetemcomitans* is closely related to *Haemophilus influenzae*, *H. ducreyi* and *Pasteurella multocida*.

Key words: Aminoacyl-tRNA Synthetase (AARS), Tryptophanyl-tRNA Synthetase (Trp-RS), *Actinobacillus actinomycetemcomitans*, Phylogenetic Tree

Introduction

The aminoacyl-tRNA synthetase (AARS) family of enzymes is of particular interest for understanding vertical gene flow (parent to offspring) and horizontal gene transfer (between different species) for creating universal phylogenetic trees, as they perform the same fundamental biochemical function of ribosomal protein synthesis in all organisms (Donoghue and Luthey-Schulten, 2003). Previous studies revealed that the identities of such AARSs are pharmaceutically important as they are potentially valuable targets of antimicrobials (Kitabatake *et al.*, 2002) and could lead to the development of novel antibiotics that show no cross-resistance to other classical antibiotics (Kim and Choi, 2003).

The AARSs are essential to understand the bacterial evolution of translation and the transition from RNA to proteins (Donoghue and Luthey-Schulten, 2003) especially in oral community due to the existence of its diversity. Out of 500 species of microbial inhabitants that have been recorded in the human oral vicinity, about half of them have yet to be cultured (Harper-Owen *et al.*, 1999) and to be identified for their peptide products as well as for biochemical and phylogenetical significance. While there have been many studies on the 20 variable AARSs responsible for translation of the genetic message, only a few studies have targeted each specific group of AARSs in oral bacteria for phylogenetic studies (Farahi *et al.*, 2004; Donoghue and Luthey-Schulten, 2003; Woese *et al.*, 2000). Tryptophanyl-tRNA synthetase (Trp-RS) was not investigated individually for phylogenetic analysis among such groups of bacteria, especially in the periodontal bacterium *Actinobacillus actinomycetemcomitans*.

A. actinomycetemcomitans is a Gram-negative, non-motile, fermentative, oral coccobacillus yet to be recognized in Trp-RS phylogenetic studies. Over 200 genome projects have been completed (Duncan, 2003) along with *A. actinomycetemcomi-*

Table I. Peptide synthetase recognition primers used in this study.

Set	Primer	Degenerated primer sequence 5′–3′ (F = forward and R = reverse)	Reference
1	A3	F 5′ TAC ACS AGC GGS AGC ACS GG 3′	Marahiel
	A7	R 5′ AVG TCS CCS GTS CKG TAC ATS C 3′	*et al.,* 1997
2	A8	F 5′ CAG GTS AAG RTS MGS GGS TWC MG 3′	Marahiel
	E2	R 5′ GTC SAC SRM SAR GTG GTG 3′	*et al.*, 1997
3	PS-	F 5′ CCA ATC GGC AAA CCA ATC TCC 3′	Carnio
	specific	R 5′ GGT TTT AGT GCT TCT CCA CTA GC 3′	*et al.*, 2001
4	A2[a]	F 5′ GGA ATT CCT SAA GDC SGG CGG IGC CTA CGT SCC 3′	Healy
	A3[a]	R 5′ GGA ATT CCC TTS GGC IKS CCG GTS GIS CCG GAG G 3′	*et al.*, 2000
5	A7	F 5′ TAC CGI ACI GGI GAT CTI GTI CG 3′	Turgay and
	T	R 5′ ATI GAG TCI CCI CCI GGG IAA AAG AA 3′	Marahiel, 1994
6	A2	F 5′ GGA ATT CCT CAA GGC GGG CGG IGC CTA CGT GCC CIA CCC 3′	Rajendran,
	A3	R 5′ GGA ATT CCC TTG GGC IGG CCG GIC GIC CCG CAG GTG TAG A 3′	1999
7	PSF	F 5′ GGWCDACHGGHMANCCHAARGG 3′	Schulz
	PSF-2	R 5′ GGCAKCCATYTYGCCARGTCNCCKGT 3′	*et al.*, 2005

[a] High-GC biased degenerated oligonucleotides with terminal *Eco*RI adaptors.

tans and other oral bacteria including *Prevotella intermedia*, *Fusobacterium nucleatum*, *Porphyromonas gingivalis*, *Streptococcus mutans* and *Treponema denticola* (www.oralgen.lanl.gov). Their Trp-RS identify elements have not yet been analyzed phylogenetically. Analysis of AARSs including completely sequenced bacterial, archaeal, and eukaryotic genomes shows a complex evolution involving a variety of horizontal gene transfers (Wolf *et al*., 1999). Among them Trp-RS, the class I AARS, catalyzes tryptophan activation in the absence of its cognate tRNA (Guo *et al*., 2007). Trp-RS also plays a foremost role in phylogeny owing to the close relationship with tyrosine-tRNA synthetase (Tyr-RS). In this study, an attempt was made to reveal the conserved motifs of Trp-RS in *A. actinomycetemcomitans* and analyze the Trp-RS phylogeny of *A. actinomycetemcomitans* with its related species.

Materials and Methods

Primer selection, PCR conditions, subcloning and DNA sequencing

Degenerated primers, derived from the sequences of non-ribosomal peptide synthetases (Table I), were synthesized at Integrated DNA Technologies, Coralville, IA, USA. All primers were employed against the genomic DNA of the periodontal bacteria *A. actinomycetemcomitans* 652*, Porphyromonas gingivalis,* and *Fusobacterium nucleatum* as well as the predominantly oral bacteria *Streptococcus mutans* (KPSK2) and *Streptococcus gordonii* (M5). The bacterial samples used in the present study were selected because of their major role in oral microbiota and availability from the laboratory stocks of Prof. Demuth. The work was carried out at the University of Louisville, KY, USA. The published PCR protocol (Rajendran, 1999) was modified using Invitrogen PCR-Supermix under the following conditions: 95 °C (5 min), 95 °C (1 min), 55 °C (2 min), and 72 °C (3 min) for 30 cycles in a Perkin-Elmer PCR machine. PCR amplicons were analyzed via 1% agarose gel electrophoresis using 100 bp DNA ladder as standard marker (catalog # N3231S) of New England BioLabs, Ipswich, MA, USA. The fragments were subcloned in pGEM-T Easy plasmid vector and transformed into DH5α competent (*E. coli*) cells according to the manufacturer's protocols of Promega, Madison, WI, USA. Based on IPTG/X-gal selection, the clones were cultured in LB medium with ampicillin. Plasmids were prepared using Promega mini-plasmid kit with a vacuum manifold. The plasmids of the confirmed clones were analyzed via 1% agarose gel electrophoresis using 100 bp DNA ladder and were sequenced using a T7 primer at the Center for Genetics and Molecular Medicine (CGeMM) DNA core facility of the University of Louisville, KY, USA.

Sequence analysis and construction of phylogenetic tree

After confirmation of the DNA sequence of the 1400 bp fragment of the *A. actinomycetemcomitans* clone (Pri2-*Aa*-1-T7), derived using primer set 2 (Table I), further computational analysis was

conducted. A homologous nucleotide sequence search was performed at NCBI's non-redundant database by using the BLASTX and protein BLASTP algorithms (http://www.ncbi.nlm.nih.gov/BLAST/). The newly determined *Trp-RS* nucleotide sequence was deposited in GenBank and an accession number was obtained using the BankIT: GenBank (www.ncbi.nlm.nih.gov/BankIT/) submission program. To confirm the signature sequence of *Trp-RS* in *A. actinomycetemcomitans*, the eMOTIFS search server at Stanford University (http://dna.stanford.edu/cgi-bin/emotif/nph-emotif-search) was used. The resulting sequences were aligned by using the CLUSTALW program (http://pir.georgetown.edu/pirwww/search/multaln.html). An initial sequence alignment was created with a list of 20 bacterial species that had homologous amino acid sequences. An alignment of Trp-RS sequences from 10 significant bacterial species on the basis of highest possible homology was then prepared. Identification of the published sequences of $tRNA^{Trp}$ binding domains in Trp-RS (Jia *et al*., 2002), as well as the basic Rossmann-fold sequence (Brown *et al*., 1997), was searched in the alignment. The ClustalX program was then used to generate the phylogenetic tree (Thompson *et al*., 1997). Theoretical homology models of Trp-RS were generated using Modeller7v7 (Marti-Renom *et al.*, 2000) and structural analyses were performed using a combination of SYBYL 7.0 (Tripos, St. Louis, MO, USA) and Insight II (Accelrys, Burlington, MA, USA). All the graphical images were generated using UCSF Chimera molecular visualization software (Pettersen *et al.*, 2004).

Results

Trp-RS of A. actinomycetemcomitans

Degenerative nucleotide primers encoding conserved acylation and epimerization domains of non-ribosomal peptide synthetase were used in PCR to amplify corresponding regions on genomic DNA from *A. actinomycetemcomitans, Porphyromonas gingivalis, Fusobacterium nucleatum*, *Streptococcus mutans* (KPSK2) and *Streptococcus gordonii.* Among the studied organisms, the genomic DNA of the periodontal bacterium *A. actinomycetemcomitans* yielded a 1.4 kb in the size PCR fragment with primer set 2 (Fig. 1). The sequence of the fragment has been submitted to GenBank and obtained the accession number DQ143942. The BLAST searches reveal its identity as Trp-RS, and the signature sequence of Trp-RS was confirmed in *A. actinomycetemcomitans* by using the eMOTIFS search server, which yielded all four tryptophanyl-tRNA synthetase signature motifs: I (16-LTIGNYLGALRQWVK MQ-32), II (67-YLACGIDPAKSTIFIQSHV-85), III (144-VPVGEDQKQHLEITRDIAQR-163) and IV (195-KMSKSDEN-202) as depicted in Fig. 2.

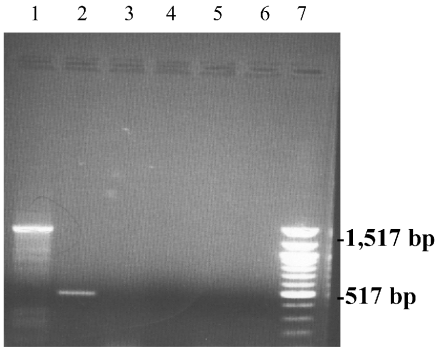

Fig. 1. Electrophoresis of PCR amplicons. Seven sets of degenerated primers (Table I) were employed against the genomic DNA of the periodontal bacterium *A. actinomycetemcomitans* 652 under the following conditions: 95 °C (5 min), 95 °C (1 min), 55 °C (2 min), and 72 °C (3 min) for 30 cycles in a Perkin-Elmer PCR machine. PCR amplicons were analyzed via 1% agarose gel electrophoresis. This figure indicates the amplified PCR fragment of interest at the size of 1.4 kb using primer set 2 (lane 1) and unconcerned fragment using primer set 3 (lane 2) and a 100 base pair DNA ladder (lane 7).

Seeking the adenylation domain

When the same degenerative primer set 2 was employed with other oral bacteria, none of them yielded a positive amplicon under the same PCR conditions and hence yielded no fragment for *Trp-RS* sequence. In order to seek further information and to confirm the observation, the nucleotide sequence obtained from *A. actinomycetemcomitans* was compared as query with other completed genome sequences of oral pathogens by blasting at www.oralgen.lanl.gov. The resulting *Trp-RS* sequences from some of the oral bacterial sequences obtained from the genome database showed no glutamic acid (Glu, E) in the GELTIGNYLG sequence at the residue 14–23 (Fig. 3B), which is a putative consensus adenylation domain sequence of one of the NRPS core motifs, GELxIxGxG(VL)ARGYL (Marahiel *et al.*, 1997).

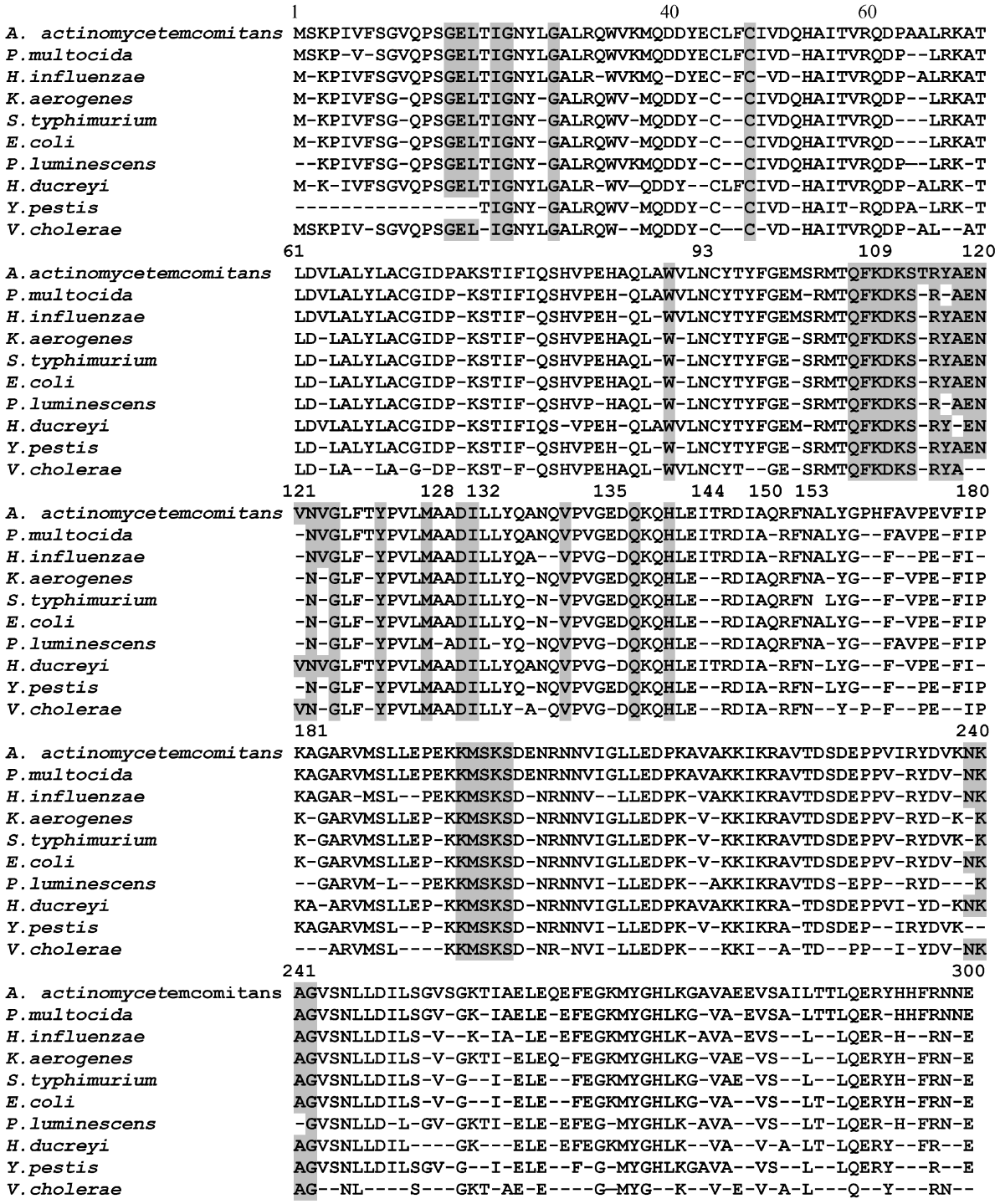

Fig. 2. Amino acid sequence alignment of Trp-RS. The results of the BLASTX sequences were aligned by using the CLUSTALW program. These sequence alignments were compared on the basis of highest percentage (60 or more) of homologous amino acid sequences from various bacterial species. Full names of the organisms are given in Table II. *A. actinomycetemcomitans* and a group of closely related bacterial species unusually have Gly-Glu-Leu (GEL) residues in their putative consensus adenylation domain sequence (GELTIGNYLG) in Trp-RS.

Table II. Homology sequence search results.

Highly matched bacterial sequences with the query Trp-RS of *A. actinomycetemcomitans* identified from the NCBI-BLAST algorithms search	Sequene identity (%)	GenBank accession number
Actinobacillus actinomycetemcomitans (query)	100	DQ143942[a]
Pasteurella multocida subsp. *multocida* str. PM70	88	AE006199
Haemophilus influenzae Rd KW20	81	U32746
Klebsiella aerogenes	81	AF308467
Salmonella typhimurium LT2	79	AE008860
Escherichia coli W3110	79	U38647
Photorhabdus luminescens subsp. *laumondii* TTO1	78	BX571859
Haemophilus ducreyi 35000HP	78	AE017151
Yersinia pestis biovar *Medievalis* str. 91001	76	AE017127
Vibrio cholerae O1 biovar *eltor* str. N16961	64	AE004329
Helicobacter pylori 26695	52	AE000630
Helicobacter hepaticus ATCC 51449	49	AE017145
Synechococcus elongatus PCC 6301	48	AP008231
Caulobacter crescentus CB15	46	AE005680
Prochlorococcus marinus str. MIT 9313	45	BX572096
Mesorhizobium loti MAFF303099	38	BA000012
Mycoplasma genitalium G-37	36	U39693
Xylella fastidiosa 9a5c	34	AE003894
Xylella fastidiosa Temecula1	34	AE012559
Pseudomonas syringae pv. *tomato* str. DC3000	33	AE016853

[a] Newly submitted GenBank accession number from this study.

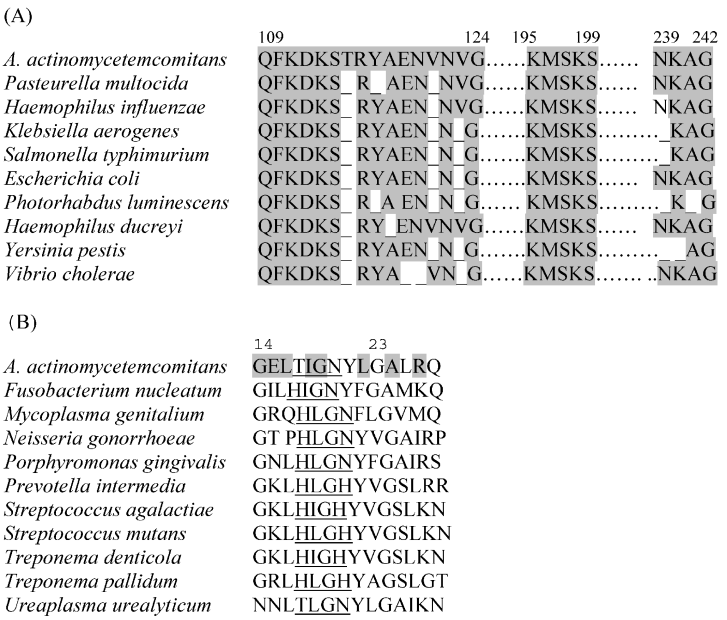

Fig. 3. Homologous special sequence comparisons of aligned amino acids in bacteria obtained from Oralgen and NCBI-BLAST. (A) The sequence alignment of this study matching with the published sequence of $tRNA^{Trp}$ binding domains in tryptophanyl-tRNA synthetase at residue 109–124 and 239–242, as well as the basic Rossmann-fold sequence at residue 195–199 in *A. actinomycetemcomitans* and other related bacterial species. (B) The residue glutamic acid (Glu or E), present in Trp-RS of *A. actinomycetemcomitans* and a group of closely related bacterial species, is apparently not present in these sequences of phylogenetically distant organisms. Instead of glutamic acid another residue, predominantly lysine (Lys or K), is found in those sequences.

Seeking other homologous residues

Out of many homologous prokaryotic and eukaryotic Trp-RS sequences obtained from BLAST search at NCBI, a group of 20 bacterial species was selected based on the closest identity starting with a 33% homology cut-off value (Table II). Most AARS from these species, including *A. actinomycetemcomitans*, are expressed from single copy genes where tryptophan (W) is conserved in the standard bacterial Trp-RS type-I (Buddha and Crane, 2005) at residue of 93 (Fig. 2). Other well conserved residues include: 40(C), 128(Y), 132(M), 135/136(D, I), 144(V), 150(Q), and 153(H). All of them are key residues in binding tryptophan (Buddha and Crane, 2005). However, two copies of same synthetases, like Trp-RS type-I and auxiliary Trp-RS II, are found in some bacteria such as *Streptomyces coelicolor* and *Deinococcus radiodurans* (Buddha and Crane, 2005). Our homology search for auxiliary Trp-RS II found no matches in *A. actinomycetemcomitans* as well as in other test bacteria. Similarly a special Trp-RS of *Streptococcus pyrogenes,* the only Trp-RS that is known to be more homologous to Trp-RS II (Buddha and Crane, 2005), was not found in *A. actinomycetemcomitans* or in other test bacteria in the present study. However, an interesting observation was made with *uvrA* gene expression in *A. actinomycetemcomitans.* As reported earlier, when the standard Trp-RS was inhibited, either by DNA damage or by inhibitors such as indolmycin, an inhibitor-resistant auxiliary Trp-RS II was induced, as in the case of *Deinococcus radiodurans* (Buddha and Crane, 2005). It was also observed that *uvrA* gene expression occurs when DNA damage happens due to ultraviolet light or other factors (Buddha and Crane, 2005). Although it was not demonstrated in *A. actinomycetemcomitans* before, a 250 bp fragment was amplified in the present study using primer set 2 in *A. actinomycetemcomitans* that corresponded to the *uvrA* gene as revealed by BLAST search at NCBI (data not shown).

Rossmann-fold and anticodon binding domain search

In the current study, the Rossmann-fold (RF) sequence, at residues 195–199 of *A. actinomycetemcomitans* and the other bacterial species aligned with query (Fig. 3), matches other published data on Trp-RS (Brown *et al.,* 1997). Besides this KMSKS (RF domain), which is a consensus motif (Brown *et al.*, 1997), a T(H)IGN domain that participates in ATP binding and typifies Class-Ic synthetases (Buddha and Crane, 2005) was identified in *A. actinomycetemcomitans.* The sequence alignment of the current study (Fig. 2) matched with the published sequence of $tRNA^{Trp}$ binding domains in Trp-RS at residues 109–124 and 239–242 (Fig. 3), indicating the presence of conserved regions of the acceptor stem and anticodon of variant $tRNA^{Trp}$.

Phylogenetic positioning

Phylogenetic analysis of AARSs including completely sequenced bacterial, archaeal, and eukaryotic genomes shows a complex evolution involving a variety of horizontal gene transfers (Wolf *et al.*, 1999). In a recent sequence alignment study of three Trp-RS sequences from *Bacillus stearothermophilus, Deinococcus radiodurans,* and *Homo sapiens*, for which crystal structures are currently known, the highest sequence conservation between species was noticed within the signature KMSKS loop of Class-I synthetases (Buddha and Crane, 2005). The mechanism by which this KMSKS domain stages the adenylation reaction is also conserved (Buddha and Crane, 2005) in these species. In the present study 10 bacterial species were selected from the aligned 20 bacterial species (Table II) on the basis of highest possible homology (over 60% of Trp-RS sequence homology) with *A. actinomycetemcomitans.* The selection of highest possible homology minimizes putative errors in evaluating the reliability of a phylogenetic tree and ensures the closest possible phylogenetically linked species with *A. actinomycetemcomitans.* Interestingly, the signature KMSKS loop of Class-I synthetases is conserved in all 10 species (Fig. 3A) giving further evidence to substantiate the proposed phylogenetic tree (Fig. 4). The mechanism by which this KMSKS domain stages the adenylation reaction is thus likely to be conserved in *A. actinomycetemcomitans* and in other phylogenetically related bacterial species as shown in the phylogenetic tree (Fig. 4).

Discussion

Searching for peptide synthetases

Peptide synthetases in oral bacteria have biochemical and phylogenetic significance since these genes have not yet been fully investigated in most

species. In general every oral bacterial cell harbors 20 AARSs responsible for the synthesis of the set of 20 canonical aminoacyl-tRNA families. There are two ways of forming aminoacyl-tRNA (Woese *et al*., 2000). First is the direct acylation of tRNA by discriminating AARSs, which is an ATP-dependent reaction, and the second is indirect acylation of tRNA by non-discriminating AARSs, which is a tRNA-dependent amino acid modification (Guo *et al*., 2007). Woese *et al*. (2000) stated that, "the second indirect pathway depends on the acylation of tRNA with a precursor amino acid. This precursor amino acid is then converted while bound to tRNA to the correct amino acid by a second, non-synthetase enzyme. The current knowledge about these enzymes is still far from complete". Despite their significance, its process remains unknown like that of amidotransferase enzyme that converts Glu-tRNA to Gln-tRNA (Tumbula *et al.,* 2000). In an other case, building of pyrrolysine on tRNA is not required and might not occur as reported earlier (Polycarpo *et al.*, 2004; Blight *et al.*, 2004). However, an unknown enzyme responsible for the conversion of the charged amino acid has been reported earlier (Srinivasan and Krzycki, 2002), which indicates that there is an additional enzyme involved in this process. Many microbial cells also contain multienzyme complexes that make specific protein templates for a nucleic acid-independent biosynthesis of low-molecular weight peptides (Marahiel *et al.,* 1997). These non-ribosomally produced peptides include lipopeptides, depsipeptides, and peptidolactones (Schulz *et al.,* 2005; Rajendran *et al.,* 1999) and are assembled from an exceedingly diverse group of precursors including pseudo, non-proteogenic hydroxy, *N*-methylated, and D-amino acids (Marahiel *et al.,* 1997). The use of multiple sets of degenerated primers (Turgay and Marahiel, 1994; Marahiel *et al.,*1997; Rajendran, 1999; Healy *et al*., 2000; Carnio *et al.,* 2001), derived from the conserved domains of non-ribosomal peptide synthetase (NRPS), reveals promising peptides and their products in many soil bacteria. But no such extensive attempts were made in oral bacteria. Using such degenerated primers to attempt to reveal functional genes from oral bacteria is therefore more valuable in terms of the presence of diversified bacterial synthetases, their synthesis and their phylogenetic importance. Keeping this in mind, we applied these multiple sets of degenerated synthetase primers and found that one of the primers is capable to amplify the *Trp-RS* as confirmed by PCR and molecular analysis.

Glutamic acid residue in Trp-RS of A. actinomycetemcomitans

The present study indicates that *A. actinomycetemcomitans* and a group of closely related bacterial species unusually have Gly–Glu–Leu (GEL) residues in their Trp-RS sequence (Fig. 2). The residue glutamic acid (Glu or E) is not present in the sequences of phylogenetically distant organisms (Fig. 3B), which is apparently perceptible in the phylogenetic tree (Fig. 4). Instead of glutamic acid, another residue, lysine (Lys or K), was predominantly found in these sequences (Fig. 3B). The catalytic role of Glu or Lys has not yet been analyzed elsewhere. Similarly the role of the adenylation domain residue Glu, unusually it exists in *A. actinomycetemcomitans* and its closely related species but not in other distantly related oral species, has not yet been studied.

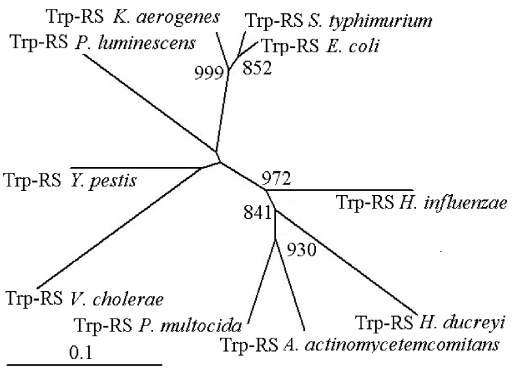

Fig. 4. The phylogenetic tree of tryptophanyl-tRNA synthetase. Aligned sequences of *A. actinomycetemcomitans* were compared with related species. The phylogenetic tree reveals that Trp-RS of *A. actinomycetemcomitans* is closely related to *Haemophilus influenzae, Haemophilus ducreyi* and *Pasteurella multocida.* Bootstrap values of more than 800 are indicated.

Conserved residue search and analysis

During translation of the genetic code an error-free molecular recognition of 21 amino acids and their attachments to tRNAs that bear the appropriate anticodon triplets is critical (Woese *et al*., 2000). There are two classes of AARSs, based on the structure and reactivity, involved in the catalytic reaction. Class-I AARS are comprised of a catalytic Rossmann-fold domain with KMSKS

residues and a helical tRNA anticodon binding domain (Buddha and Crane, 2005). The RF domain makes ATP interaction possible in bacterial species. The existence of the signature motifs, which are involved in ATP binding, such as the consensus motif KMSKS and a T(H)IGN domain, reveals that the Trp-RS sequences of *A. actinomycetemcomitans* and other bacterial sequence in the present study belong to Class-Ic synthetases. A previous study showed that Trp-RS from various species shared a high homology of residues, especially within two short regions, "QFKDKS_RYAENVNVG" and "NKAG". These two short regions were predicted to bind to acceptor stem and anticodon of variant $tRNA^{Trp}$ (Jia *et al*., 2002). The presence of such a domain in *A. actinomycetemcomitans* indicates the conserved nature of the Rossmann-fold.

Compared to ribosomal protein synthesis, non-ribosomal protein synthesis has less specificity. The substrate activation for the ribosomal origin is tRNA synthetase, whereas adenylation is the substrate activation domain for the non-ribosomal origin. This is likely due to the evolutionary force to maintain a given substrate selection, and "it is not very strict" as predicted earlier (Doekel and Marahiel, 2001). The adenylation domain is the specific "main gatekeeper" of NRPS that catalyzes the activation of cognate carboxy acids as adenylates by hydrolysis of ATP (Doekel and Marahiel, 2001). In Fig. 3, residues 14–23 (GELTIGNYLG) indicate the presence of a component of the putative consensus adenylation domain sequence which is of one of the non-ribosomal synthetase core motifs GELxIxGxG(VL)ARGYL (Marahiel *et al.,*1997). The primer set 2 which used in the present study (Table I) represents one of the adenylation domains of the NRPS. This adenylation domain also happens to contain the residues of one of the signature motifs that participates in ATP binding; T(H)IGN (Buddha and Crane, 2005). This was observed in *A. actinomycetemcomitans* by applying that primer for the first time. Adenylation domains in NRPS, that are about 550 amino acids in size, indicate that it is a member of the superfamily of adenylate-forming enzymes like acyl-CoA-ligase. As reported earlier (Doekel and Marahiel, 2001), this domain in NRPS is a part of a larger protein frame, but could exhibit a comparable catalytic activity also when expressed heterologously as a separate unit (Marahiel *et al.,*1997; Doekel and Marahiel, 2001). Searching for other conserved motifs in the current amino acid sequence of *A. actinomycetemcomitans* by using the NRPS-PKS BLAST server (http://www.nii.res.in/nrps-pks.html) yielded no matching sequence as expected, since the query has only 340 residues. However, it could be possible that the adenylation sequence observed in *A. actinomycetemcomitans* is the Trp-RS signature sequence and has no relation to NRPS. As explained elsewhere (Doekel and Marahiel, 2001) nature could divulge various alternatives to produce variations through tRNA, NRPS, polyketide synthases (PKS), mixed peptide synthetases (NRPS-PKS), and/or engineered peptide synthetases. Therefore, it is likely that the sequence of the adenylation domain used in the present study, which reveals the Trp-RS identity, could have some phylogenetic relation with early NRPS origin, and may reveal the Trp-RS identity in other closely related species. This is under investigation.

Phylogenetic analysis

Trp-RS is the smallest protein in the family of bacterial tRNA synthetases (Jia *et al.,* 2002) and less complex to compare. Phylogenetic analysis of Trp-RS from various oral bacterial sequences is not available elsewhere. It was the main rationale to select Trp-RS for phylogenetic analysis in the present study. For the first time, the amino acid sequence was analyzed with the goal of phylogenetically positioning the Trp-RS of *A. actinomycetemcomitans* among other closely related bacterial species by using phylogenetic tree reconstruction methods. The compatible data set, without bias from a stochastic effect, indicates that Trp-RS is highly conservative in *A. actinomycetemcomitans* and is closely related to *Haemophilus influenzae, Haemophilus ducreyi* and *Pasteurella multocida* (Fig. 4). In a previous genome sequence study, it was predicted that the proteins with highest homologies to other organismshave orthologs in *H. influenzae*, approx. 43%, and in *P. multocida*, 41% (Duncan, 2003). In the current study, the codon usage indicates that *A. actinomycetemcomitans* is closely related to *H. influenzae*, *H. ducreyi* and *P. multocida* (Table II). The previous study on 16S ribosomal RNA analysis between *A. actinomycetemcomitans* and *H. influenzae* also indicated their close relationship (Duncan, 2003). The BLASTX search of the present study on Trp-RS reveals that *A. actinomycetemcomitans* has 88% of gene sequence homology with *P. multocida* subsp. *multo-*

cida str. PM70, more than with *H. influenzae* and *H. ducreyi*, and they share close phylogenetic characteristics as depicted in the phylogenetic tree (Fig. 4).

In another approach, synthetases of other kinds closely related to Trp-RS were searched among the sequences in order to find the close relationship between the closest AARSs. As reported earlier, a comparative study indicated that the phylogeny of *tRNA* genes matched with that of *Tyr-RS* and *Trp-RS* and hence Tyr-RS and Trp-RS may have close relationship (Ribas *et al.,* 1996). However, it was later overlooked and put forward that Tyr-RS and Trp-RS have an early divergence in the phylogeny (Brown *et al*., 1997). Our current study observes no congruent amino acid alignment between Tyr-RS and Trp-RS sequences. This finding hence supports the latter observation that a possible divergence could exist between these two synthetases, and Trp-RS could putatively maintain a highly conserved nature thereafter in *A. actinomycetemcomitans* and its closely related species such as *Haemophilus influenzae, Haemophilus ducreyi* and *Pasteurella multocida.*

Besides the high sequence similarities among these groups of bacteria and the divergence from the nearest synthetases, the presence of glutamic acid (Glu or E) in Trp-RS of *A. actinomycetemcomitans* and in a group of closely related bacterial species (Fig. 2) reveals its phylogenetic value, since it is apparently not present in the sequences of distant organisms. The unrooted phylogenetic tree, constructed in the present study (Fig. 4), clearly depicts this observation for the first time, and it could be the putative reason for those closely related species (Fig. 2) to have phylogenetic relationships rather than distant species (Fig. 3B). Instead of glutamic acid another residue, lysine (Lys or K), is predominantly found in the sequences of those distant species. This study was helpful in terms of identifying Trp-RS in an oral bacterium and guided to the further investigation on the unusual existence of an adenylation domain component with glutamic acid in *A. actinomycetemcomitans* and its closely related species. Further investigation in this direction will reveal more insight into their relationships and help to construct a universal Trp-RS phylogenetic tree.

Structural analysis of Trp-RS

In order to analyze the structural homology and conservation of the Trp-binding site, we constructed theoretical homology models of all Trp-RS sequences of Fig. 4 using Modeller7v7. A high-resolution bacterial Trp-RS X-ray crystal structure (PDB ID: 1I6K, *B. stearothermophilus*) with 57% sequence identity to Trp-RS *A. actinomycetemcomitans* was used as the structural template. The final models were energy-minimized using the Discover module of Insight II prior to further structural analyses. Superposition of the Trp-RS *models @ Cα* with the template had a RMS deviation in the range 0.15–0.35 Å showing the high homology of this structural domain. Guided docking (SYBYL 7.0) of the ligand tryptophanyl-5′-AMP into the binding site of Trp-RS resulted in an identical binding pose in comparison with the template (Fig. 5). A closer look at the binding site of Trp-RS of *A. actinomycetemcomitans* and that of its template revealed a highly identical active site (a ~90% identity) in comparison with the bacterial

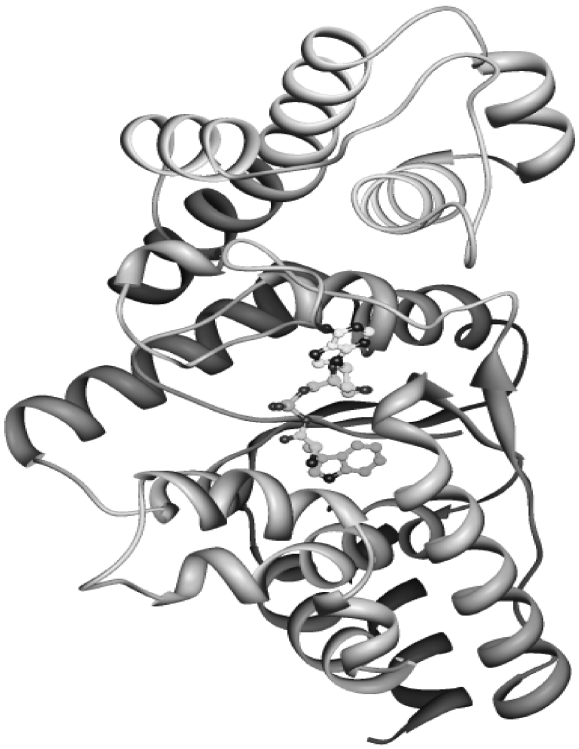

Fig. 5. Structure of Trp-RS from *A. actinomycetemcomitans.* A theoretical model of Trp-RS was generated using Modeller7v7 with a high-resolution bacterial Trp-RS template (PDB ID: 1I6K, 57% sequence identity). The structure of Trp-RS of *A. actinomycetemcomitans* is shown as a ribbon model. The ball-and-stick model with carbon framework at the center represents the bound ligand tryptophanyl-5′-AMP.

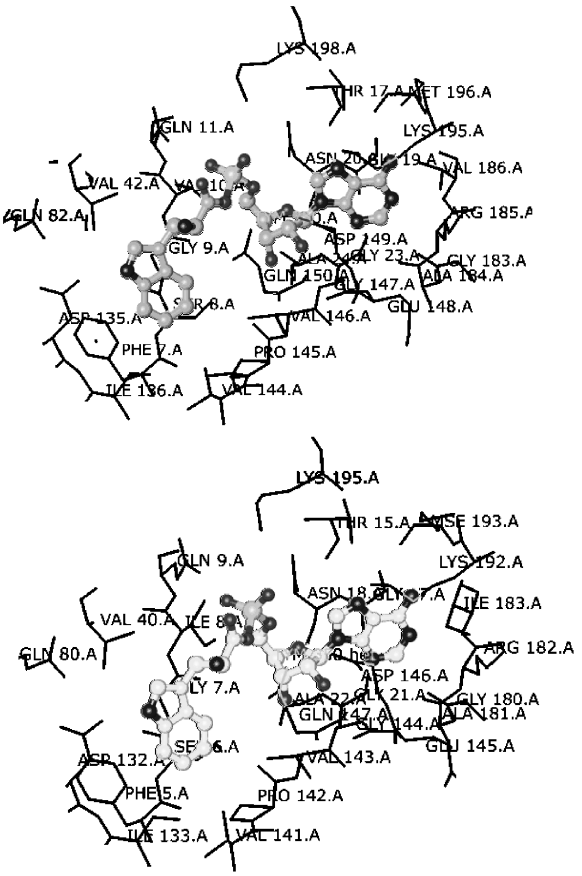

Fig. 6. A closer look at the binding site of Trp-RS of *A. actinomycetemcomitans*. Guided docking (SYBYL 7.0) of the ligand tryptophanyl-5′-AMP revealed a highly identical active site in comparison with the bacterial template. The Trp-RS *A. actinomycetemcomitans* is shown at the top, and the template-1I6K is shown as the bottom figure. The binding site residues are rendered as a wire frame model and the ligand as a ball and stick model.

Trp-RS template of *A. actinomycetemcomitans* and a quantitative conservation of the Trp binding site (Fig. 6).

Conclusion

In conclusion, the prevalence of tryptophanyl-tRNA synthetase conserved motifs of *Actinobacillus actinomycetemcomitans* and its molecular phylogenetic analysis have been carried out. Comparison of *A. actinomycetemcomitans* with other oral bacterial genomes revealed that Trp-RS has glutamic acid residues in its adenylation domain that are not present in other related oral bacterial species. The constructed Trp-RS phylogenetic tree revealed the relationship with a group of oral species. Our oral microbiota have diversified bacterial inhabitants and this molecular phylogenetic study on oral bacteria could support future studies to construct a comprehensive phylogenetic tree on Trp-RS among the available oral bacterial species. Since the aminoacyl-tRNA synthetase family of enzymes is universal and can be found in all living organisms, their universality will allow a future construction of a detailed phylogenetic tree of AARSs in all extant organisms.

Acknowledgements

This work was supported by Grant # P20 RR-16481 from IDeA Networks of Biomedical Research Excellence (INBRE) program of the National Center for Research Resources, USA and partially by Grant #MCB-0639356 from National Science Foundation (NSF), USA to Narayanan Rajendran. We thank Deanna James for laboratory assistance and Bruce Griffis for correcting the manuscript.

Blight S. K., Larue R. C., Mahapatra A., Longstaff D., Chang E., Zhao G., Kang P. T., Green-Church K. B., Chan M. K., and Krzycki J. A. (2004), Direct charging of tRNA(CUA) with pyrrolysine *in vitro* and *in vivo*. Nature **431**, 333–335.

Brown J. R., Weiss R., and Doolittle W. F. (1997), Evidence for the early divergence of trptophanyl-tRNA and tyrosyl-tRNA synthetases. J. Mol. Evol. **45**, 9–16.

Buddha R. M. and Crane B. R. (2005), Structures of tryptophanyl-tRNA synthetase II from *Deinococcus radiodurans* bound to ATP and tryptophan: Insight into subunit cooperativity and domain motions linked to catalysis. J. Biol. Chem. **280**, 31965–31973.

Carnio M. C., Stachelhaus T., Kevin P. F., and Scherer S. (2001), Pyridinyl polythiazole class peptide antibiotic micrococcin P1, secreted by foodborne *Staphylococcus equorum* WS2733, is biosynthesized nonribosomally. Eur. J. Biochem. **268**, 6390–6400.

Doekel S. and Marahiel M. A. (2001), Biosynthesis of natural products on modular peptide synthetases. Metabolic Eng. **3**, 64–77.

Donoghue P. O. and Luthey-Schulten Z. (2003), On the evolution of structure in amino-acyl-tRNA synthetases. Microbiol. Mol. Biol. Rev. **67**, 550–573.

Duncan M. J. (2003), Genomics of oral bacteria. Crit. Rev. Oral Biol. Med. **14**, 175–187.

Farahi K., Pusch G. D., Overbeek R., and Whitman W. B. (2004), Detection of lateral gene transfer events in the prokaryotic tRNA synthetases by ratios of evolutionary distances method. J. Mol. Evol. **58**, 615–631.

Guo L. T., Chen X. L., Zhao B. T., Shi Y., Li W., Xue H., and Jin Y. X. (2007), Human tryptophanyl-tRNA synthetase is switched to a tRNA-dependent mode for tryptophan activation by mutations at V85 and I311. Nucleic Acids Res. **35**, 5934–5943.

Harper-Owen R., Dymock D., Booth V., Weightman A. J., and Wade W. G. (1999), Detection of unculturable bacteria in periodontal health and disease by PCR. J. Clinical Microbiol. **37**, 1469–1473.

Healy F. G., Gibson D. M., and Loria R. (2000), The txt AB genes of the plant pathogen *Streptomyces acidiscabies* encode a peptide required for phytotoxin thaxtomin A production and pathogenicity. Mol. Microbiol. **38**, 794–804.

Jia J., Feng X., Jin Y., and Wang D. (2002), Two essential regions for tRNA recognition in *Bacillus subtilis* Trp-tRNA synthetase. Biochemistry **365**, 749–756.

Kim S. and Choi S. Y (2003), Aminoacyl-tRNA synthetases and inhibitors as a novel family of antibiotics. Appl. Microbiol. Biotechnol. **61**, 278–288.

Kitabatake M., Demain A., Sakamota K., Yokoyama S., and Soll E. (2002), Indolmycin resistance of *Streptomyces coelicolar* by induced expression of one of its two tryptophanyl-tRNA synthetases. J. Biol. Chem. **277**, 23882–23887.

Marahiel M. A., Stachelhaus T., and Mootz H. D. (1997), Modular peptide synthetases involved in nonribosomal peptide synthesis. Chem. Rev. **97**, 2651–2673.

Marti-Renom M. A., Stuart A., Fiser A., Sánchez R., Melo F., and Sali A. (2000), Comparative protein structure modeling of genes and genomes. Annu. Rev. Biophys. Biomol. Struct. **29**, 291–325.

Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., and Ferrin T. E. (2004), UCSF chimera – A visualization system for exploratory research and analysis. J. Comput. Chem. **25**, 1605–1612.

Polycarpo C., Ambrogelly A., Bérubé A., Winbush S. M., McCloskey J. A., Crain P. F., Wood J. L., and Söll D. (2004), An aminoacyl-tRNA synthetase that specifically activates pyrrolysine. PNAS **101**, 12450–12454.

Rajendran N. (1999), Identification and cloning of a gene locus encoding peptide synthetase of *Pseudomonas fluorescens*. Z. Naturforsch. **54c**, 105–109.

Rajendran N., Stachelhaus T., and Marahiel M. A. (1999), Multifunctional peptide synthetase required for nonribosomal peptide synthesis. In: Comprehensive Natural Products Chemistry, Vol. 4 (Barton D., Nakanishi K., and Kelly J. W., eds.). Elsevier Science Ltd, Oxford, England, pp. 198–220.

Ribas P. L., Fugier M., Quinn C. L., and Schimmel P. (1996), Evidence that two present-day components needed for the genetic code appeared after nucleated cells separated from Eubacteria. PNAS **93**, 166–170.

Schulz M. E., Vukov N., Shaheen R., Andersson M., Martlbauer E., and Scherer S. (2005), Identification and partial characterization of the nonribosomal peptide synthetase gene responsible for cereulide production in emetic *Bacillus cereus*. Appl. Environ. Microbiol. **71**, 105–113.

Srinivasan G. and Krzycki J. (2002). Pyrolysine encoded by IAG in Archaea: charging of a UAG-decoding specialized tRNA. Science **296**, 1459–1462.

Thompson J. D., Gibson T. J., Plewniak F., Jeanmougin F., and Higgins D. G. (1997), The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24**, 4876–4882.

Tumbula D. L., Becker H. D., Chang W. Z., and Soll D. (2000), Domain-specific recruitment of amide amino acids for protein synthesis. Nature **407**, 106–110.

Turgay K. and Marahiel M. A. (1994), A general approach for identifying and cloning peptide synthetase genes. Peptide Res. **7**, 238–241.

Woese C. R., Olsen G. J., Ibba M., and Soll D. (2000), Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiol. Mol. Biol. Rev. **64**, 202–236.

Wolf Y. I., Aravind L., Grishin N. V., and Koonin E. V. (1999), Evolution of aminoacyl tRNA synthetases analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. Genome Res. **9**, 689–710.