Intron locations and functional deletions in relation to the design and evolution of a subgroup of class I tRNA synthetases

ARTICLE *in* PROTEIN SCIENCE · OCTOBER 1992

Impact Factor: 2.85 · DOI: 10.1002/pro.5560011018 · Source: PubMed

3 AUTHORS, INCLUDING:

Kiyotaka Shiba
Japanese Foundation for Cancer Research
174 PUBLICATIONS **4,934** CITATIONS

SEE PROFILE

Protein Science (1992), *1*, 1387-1391. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Intron locations and functional deletions in relation to the design and evolution of a subgroup of class I tRNA synthetases

PAUL SCHIMMEL, ALYSSA SHEPARD, AND KIYOTAKA SHIBA[1]
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(RECEIVED August 3, 1992; REVISED MANUSCRIPT RECEIVED August 12, 1992)

Ten aminoacyl tRNA synthetases belong to a distinct class, based on two characteristic sequence motifs that are part of an N-terminal nucleotide-binding fold held in common (Webster et al., 1984; Hountondji et al., 1986; Cusack et al., 1990; Eriani et al., 1990; Burbaum & Schimmel, 1991; Nagel & Doolittle, 1991; Moras, 1992). These 10 enzymes are further subdivided into a subclass of 5, which consists of Cys-, Ile-, Leu-, Met-, and Val-tRNA synthetases (Hou et al., 1991; Shiba & Schimmel, 1992). In addition to having the sequence motifs in the N-terminal halves of all class I enzymes, these five also have related C-terminal domains and, in general, a statistically more significant alignment of N-terminal sequences with each other than with the remaining five class I enzymes. Class II enzymes are unrelated to the class I synthetases and probably have a distinct origin.

Except for the dimeric alpha$_2$ *Escherichia coli* methionyl-tRNA synthetase, the other four class I *E. coli* enzymes in the subgroup are monomers. Removal of approximately 140 amino acids from the C-terminus of *E. coli* methionyl-tRNA synthetase converts it into an active monomer, and thus it also may be treated as monomeric (Cassio & Waller, 1971). Consistent with this is the observation that yeast MetRS is monomeric (Walter et al., 1983). For these five enzymes, polypeptide chain lengths vary in *E. coli* from 461 (Cys) to 959 (Val) amino acids and show a similar length variation in other organisms. The three-dimensional structure (Brunie et al., 1990) of the active N-terminal monomeric 547-amino acid polypeptide of methionyl-tRNA synthetase serves as a model for the structures of the other four related proteins. The N-terminal nucleotide-binding fold, with alternating beta-strands and alpha-helices, is fused to a C-terminal domain composed primarily of alpha-helices. Two insertions—connective polypeptide 1 (CP1) and CP2—split the nucleotide-binding fold after the third and fourth beta-strand, respectively. These connective elements bulge out from the structure, and much of the size differences between this subgroup of class I enzymes is accommodated by insertions into or deletions within CP1.

Twenty-one sequences of this subgroup of five class I enzymes are available (Fig. 1; Table 1). The sources of the four IleRS sequences, five ValRS sequences, six LeuRS sequences, five MetRS sequences, and the single CysRS sequence are listed in Table 1 along with the abbreviations used in Figure 1. Of these sequences, eight are from prokaryotes, three are mitochondrial, two are from thermophilic bacteria (one is from an archaebacterium), and eight are from eukaryotic cytoplasm.

One of the most recent sequences is that of the gene for *Tetrahymena thermophila* (*T. thermophila*) isoleucyl-tRNA synthetase, and this is the first to contain multiple introns. The possibility of introns separating structural or functional domains was mentioned by Csank and Martindale (1992).

Both the LeuRS sequences from *Neurospora crassa* cytoplasm (Nc-L) and mitochondria (Ncm-L) contain a single intron (Benarous et al., 1988; Chow et al., 1989). The locations of the eight introns in the structure of *T. thermophila* isoleucine tRNA synthetase (which is modeled after the *E. coli* methionine tRNA synthetase) are shown in the upper portion of Figure 1. The beta-strands and alpha-helices of the nucleotide-binding fold are represented as pentagons and rectangles, respectively. Six of the introns are located within or at the end of the N-terminal nucleotide-binding fold and with one exception (intron 2), all of these introns are between segments of secondary structure. The locations of introns 4, 5, and 6 are of particular interest.

Reprint requests to: Paul Schimmel, Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.
[1] Present address: Cancer Institute, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan.

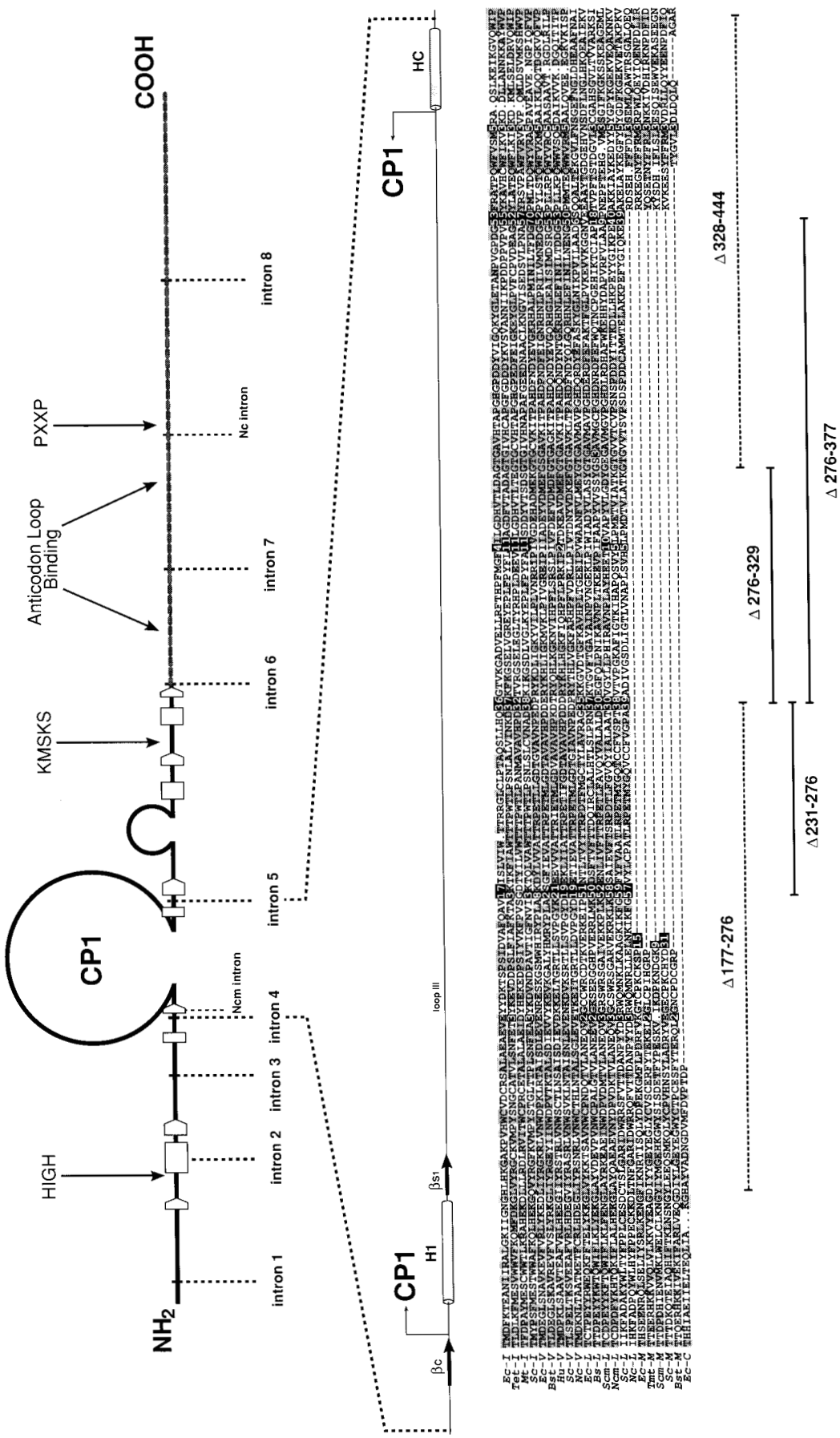
HIGH
KMSKS
Anticodon Loop Binding
PXXP
COOH
NH2
CP1
Ncm intron
Nc intron
intron 1
intron 2
intron 3
intron 4
intron 5
intron 6
intron 7
intron 8
βc
H1
βs1
loop III
CP1
HC
Δ177-276
Δ231-276
Δ276-329
Δ276-377
Δ328-444

Table 1. *Class I subgroup aminoacyl-tRNA synthetases*

Enzyme	Organism	Abbreviation	Classification
IleRS	*Escherichia coli*	Ec-I	Prokaryote
IleRS	*Tetrahymena thermophila*	Tet-I	Eukaryote
IleRS	*Methanobacterium thermoautotrophicum*	Mt-I	Archaebacterium
IleRS	*Saccharomyces cerevisiae*	Sc-I	Eukaryote
ValRS	*E. coli*	Ec-V	Prokaryote
ValRS	*Bacillus stearothermophilus*	Bst-V	Prokaryote
ValRS	*Homo sapiens*	H-V	Eukaryote
ValRS	*S. cerevisiae*	Sc-V	Eukaryote
ValRS	*Neurospora crassa*	Nc-V	Eukaryote
LeuRS	*E. coli*	Ec-L	Prokaryote
LeuRS	*Bacillus subtilis*	Bs-L	Prokaryote
LeuRS	*S. cerevisiae*	Scm-L	Mitochondria
LeuRS	*N. crassa*	Ncm-L	Mitochondria
LeuRS	*S. cerevisiae*	Sc-L	Eukaryote
LeuRS	*N. crassa*	Nc-L	Eukaryote
MetRS	*E. coli*	Ec-M	Prokaryote
MetRS	*Thermus thermophilus*	Tmt-M	Prokaryote
MetRS	*S. cerevisiae*	Scm-M	Mitochondria
MetRS	*S. cerevisiae*	Sc-M	Eukaryote
MetRS	*B. stearothermophilus*	Bst-M	Prokaryote
CysRS	*E. coli*	Ec-C	Prokaryote

The CP1 insertion is flanked by introns 4 and 5, suggesting that a distinct exon encoding the CP1 polypeptide was recruited into the nucleotide-binding fold. In the related class I glutaminyl-tRNA synthetase, the CP1 insertion interacts with the tRNA acceptor stem and facilitates docking of the 3′-end of the tRNA near the bound adenylate (Rould et al., 1989). Mutations that alter tRNA recognition by methionyl-tRNA synthetase occur in CP1 (Meinnel et al., 1991), and these mutations have been proposed to alter acceptor helix interactions (Schimmel, 1991). The tRNA acceptor helices can be modeled as 7-bp microhelix substrates for specific aminoacylation by several synthetases (Francklyn & Schimmel, 1989, 1990; Francklyn et al., 1992; Frugier et al., 1992; Martinis & Schimmel, 1992). The recruitment of an exon-encoded polypeptide insertion next to the site of adenylate synthesis may have been a step in the evolution of class I enzymes, which provided a structural format for achieving aminoacylation of RNA oligonucleotides that became tRNA acceptor stems.

In the alignment of the 21 sequences in the region of CP1 of the subgroup of class I enzymes shown in the lower portion of Figure 1, the relatively enlarged sizes of the CP1 insertions of Ile-, Leu-, and Val-tRNA synthetases are apparent. Earlier work showed that, in *E. coli* isoleucyl-tRNA synthetase, substantial parts of CP1 are dispensable for aminoacylation (Starzyk et al., 1987). These deletions are represented at the bottom of Figure 1. At the time, only the sequence of *E. coli* methionyl-tRNA synthetase was used for comparison. This larger data base shows that these functional CP1 deletions in isoleucyl-tRNA synthetase remove segments that are mostly absent in methionyl- and cysteinyl-tRNA synthetases (six lower sequences in alignment, Fig. 1; Shiba & Schimmel, 1992). The CP1 sequences common to both the active isoleucyl-tRNA synthetase deletion proteins and to methionyl- and cysteinyl-tRNA synthetases may form the

Fig. 1. Upper panel: Predicted secondary structure of IleRS and locations of *Tetrahymena thermophila* IleRS intron positions. The proposed organization of secondary structural elements in *Escherichia coli* IleRS are based on modeling using sequence alignments (Starzyk et al., 1987; Hou et al., 1991; Shiba & Schimmel, 1992) and the *E. coli* MetRS structure (Brunie et al., 1990). The HIGH, KMSKS, and PXXP (where X is predominantly hydrophobic) conserved sequences are indicated, as well as the proposed anticodon loop binding sites (Ghosh et al., 1990, 1991; Despons et al., 1991; Meinnel et al., 1991). The *T. thermophila* intron positions (Csank & Martindale, 1992) were based on an alignment between the *E. coli* and *T. thermophila* IleRS sequences. The positions of the single introns in the *N. crassa* mitochondria (Ncm) and cytoplasmic (Nc) LeuRS genes are also indicated (Benarous et al., 1988; Chow et al., 1989). **Lower panel:** The alignment for the 21 sequences of the class I subgroup are shown for the CP1 region. The secondary structure of *E. coli* MetRS is shown above the alignment; nomenclature for the alpha-helices, beta-sheets, and loop are based on this structure (Brunie et al., 1990). The alignment was determined from the Genetics Computer Group (Madison, Wisconsin) software GAP, PILEUP, PROFILE, and PRETTY (Needleman & Wunsch, 1970; Feng & Doolittle, 1987; Gribskov et al., 1990; see also Shiba & Schimmel, 1992). The shaded residues identify semiconserved sequences. Four IleRS sequences are from *E. coli* (Ec-I; Webster et al., 1984), *T. thermophila* (Tet-I; Csank & Martindale, 1992), *Methanobacterium thermoautotrophicum* (Mt-I; Jenal et al., 1991), and *Saccharomyces cerevisiae* (Sc-I; Englisch et al., 1987; Martindale et al., 1989). Also given are ValRS sequences: *E. coli* (Ec-V; Härtlein et al., 1987), *Bacillus stearothermophilus* (Bst-V; Borgford et al., 1987), human (Hu-V; Hsieh & Campbell, 1991), *S. cerevisiae* (Sc-V; Jordana et al., 1987), *Neurospora crassa* (Nc-V; Kubelik et al., 1991); LeuRS sequences: *E. coli* (Ec-L; Härtlein & Madern, 1987), *Bacillus subtilis* (Bs-L; Horn & Zahler, 1992), *S. cerevisiae* mitochondria (Scm-L; Tzagoloff et al., 1988), *N. crassa* mitochondria (Ncm-L; Chow et al., 1989), *S. cerevisiae* (Sc-L; Hohmann, 1991), *N. crassa* (Nc-L; Benarous et al., 1988); MetRS sequences: *E. coli* (Ec-M; Dardel et al., 1984), *Thermus thermophilus* (Tmt-M; Nureki et al., 1991), *S. cerevisiae* mitochondria (Scm-M; Tzagoloff et al., 1989), *S. cerevisiae* cytoplasm (Sc-M; Walter et al., 1983), *B. stearothermophilus* (Bst-M; Mechulam et al., 1991), and the single CysRS sequence from *E. coli* (Ec-C; Avalos et al., 1991; Eriani et al., 1991; Hou et al., 1991). The previously generated deletions in *E. coli* IleRS (Starzyk et al., 1987) are shown below the alignment, where the deletions depicted with dotted lines indicate those that resulted in abolished aminoacylation and adenylate formation, whereas those represented by a solid line were active. Ec, *E. coli*; Tet, *T. thermophila*; Mt, *M. thermoautotrophicum*; Sc, *S. cerevisiae*; Bst, *B. stearothermophilus*; Hu, human; Nc, *N. crassa*; Bs, *B. subtilis*; Scm, *S. cerevisiae* mitochondria; Ncm, *N. crassa* mitochondria; Tmt, *T. thermophilus*; Bst, *B. stearothermophilus*.

structures that interact with the acceptor stem microhelices.

Intron 6 delineates the boundary between the end of the nucleotide-binding fold and the beginning of the helical C-terminal domain. In the class I methionyl- and glutaminyl-tRNA synthetases, the C-terminal domain has the motif for interaction with the anticodon trinucleotide (Rould et al., 1989, 1991; Brunie et al., 1990; Ghosh et al., 1990, 1991; Despons et al., 1991; Meinnel et al., 1991). The anticodon is distal to the amino acid attachment site and is part of a second stem-loop domain found in tRNA structures. The recruitment of one or more exons joined to the C-terminal side of the coding sequence for the nucleotide-binding fold, as evident in the gene for the *T. thermophila* isoleucyl-tRNA synthetase (Csank & Martindale, 1992), may have provided the framework for binding tRNA elements distal to the acceptor helix.

In methionyl- and glutaminyl-tRNA synthetases, the sequences and structures of the C-terminal domains are unrelated (Rould et al., 1989; Brunie et al., 1990), even though the two enzymes belong to the same class by virtue of the N-terminal nucleotide-binding fold held in common. Thus, although the sequences of CP1 polypeptides and C-terminal domains for the subgroup of enzymes considered here are related, these sequences are also highly diversified amongst all of the class I enzymes when considered as a group. The gene structure of *T. thermophila* isoleucyl-tRNA synthetase suggests that the diversification of C-terminal domains and CP1 polypeptides may have arisen from a variety of exons for RNA-binding interactions, which were joined to a relatively fixed adenylate synthesis domain.

Recently, we demonstrated the assembly of randomly cleaved isoleucine tRNA synthetase, including the generation of active enzyme from fragments that were generated by interruptions of conserved sequences and secondary structure elements (Shiba & Schimmel, 1992). These results indicated that the assembly of a protein from fragments is not dependent on having fragments that approximately begin and end at the boundaries of exons. With this in mind, although the gene structure of *T. thermophila* IleRS and related data can be viewed as reflecting the assembly of at least a portion of the synthetase from defined exon-like domains, the available data on this and other systems do not rule out additional possibilities, such as the build-up of synthetase structure from fragments that do not correspond to exon-like units. Some clarification will be provided when more sequences of genes of synthetases become available.

Acknowledgments

This work was supported by grant GM23562 from the National Institutes of Health, American Cancer Society Postdoctoral Fellowship grant PF-3664 to A.S., and a grant from the Human Frontiers Science Organization. We thank Drs. Karin Musier-Forsyth, Douglas Buechter, and Sunghoon Kim for their suggestions and comments throughout this work.

References

Avalos, J., Corrochano, L.M., & Brenner, S. (1991). Cysteinyl-tRNA synthetase is a direct descendent of the first aminoacyl-tRNA synthetase. *FEBS Lett. 286*, 176–180.

Benarous, R., Chow, C.M., & RajBhandary, U.L. (1988). Cytoplasmic leucyl-tRNA synthetase of *Neurospora crassa* is not specified by the leu-5 locus. *Genetics 119*, 805–814.

Borgford, T.J., Brand, N.J., Gray, T.E., & Fersht, A.R. (1987). The valyl-tRNA synthetase from *Bacillus stearothermophilus* has considerable sequence homology with the isoleucyl-tRNA synthetase from *Escherichia coli*. *Biochemistry 26*, 2480–2486.

Brunie, S., Zelwer, C., & Risler, J.-L. (1990). Crystallographic study at 2.5 Å resolution of the interaction of methionyl-tRNA synthetase from *Escherichia coli* with ATP. *J. Mol. Biol. 216*, 411–424.

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

Cassio, D. & Waller, J.P. (1971). Modification of methionyl-tRNA synthetase by proteolytic cleavage and properties of the trypsin-modified enzyme. *Eur. J. Biochem. 20*, 283–300.

Chow, C.M., Metzenberg, R.L., & RajBhandary, U.L. (1989). Nuclear gene for mitochondrial leucyl-tRNA synthetase of *Neurospora crassa*: Isolation, sequence, chromosomal mapping and evidence that the leu-5 locus specifies structural information. *Mol. Cell. Biol. 9*, 4631–4644.

Csank, C. & Martindale, D.W. (1992). Isoleucyl-tRNA synthetase from the ciliated protozoan *Tetrahymena thermophila*. *J. Biol. Chem. 267*, 4592–4599.

Cusack, S., Berthet-Colominas, C., Härtlein, M., Nassar, N., & Leberman, R. (1990). A second class of synthetase structure revealed by X-ray analysis of *Escherichia coli* seryl-tRNA synthetase at 2.4 Å. *Nature 347*, 249–255.

Dardel, F., Fayat, G., & Blanquet, S. (1984). Molecular cloning and primary structure of the *Escherichia coli* methionyl-tRNA synthetase gene. *J. Bacteriol. 160*, 1115–1122.

Despons, L., Walter, P., Senger, B., Ebel, J.-P., & Fasiolo, F. (1991). Identification of potential amino acid residues supporting anticodon recognition in yeast methionyl-tRNA synthetase. *FEBS Lett. 289*, 217–220.

Englisch, U., Englisch, S., Markmeyer, P., Schischkoff, J., Sternbach, H., Kratzin, H., & Cramer, F. (1987). Structure of the yeast isoleucyl-tRNA synthetase gene (ILS1): DNA-sequence, amino-acid sequence of proteolytic peptides of the enzyme and comparison of the structure to those of other known aminoacyl-tRNA synthetases. *Biol. Chem. Hoppe-Seyler 368*, 971–979.

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

Eriani, G., Dirheimer, G., & Gangloff, J. (1991). Cysteinyl-tRNA synthetase: Determination of the last *E. coli* aminoacyl-tRNA synthetase primary structure. *Nucleic Acids Res. 19*, 265–269.

Feng, D.-F. & Doolittle, R.F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol. 25*, 351–360.

Francklyn, C. & Schimmel, P. (1989). Aminoacylation of RNA minihelices with alanine. *Nature 337*, 478–481.

Francklyn, C. & Schimmel, P. (1990). Enzymatic aminoacylation of an eight-base-pair microhelix with histidine. *Proc. Natl. Acad. Sci. USA 87*, 8655–8659.

Francklyn, C., Shi, J.-P., & Schimmel, P. (1992). Overlapping nucleotide determinants for specific aminoacylation of RNA microhelices. *Science 255*, 1121–1125.

Frugier, M., Florentz, C., & Giegé, R. (1992). Anticodon-independent aminoacylation of an RNA minihelix with valine. *Proc. Natl. Acad. Sci. USA 89*, 3990–3994.

Ghosh, G., Kim, H.Y., Demaret, J.-P., Brunie, S., & Schulman, L.H. (1991). Arginine-395 is required for efficient in vivo and in vitro aminoacylation of tRNAs by *Escherichia coli* methionyl-tRNA synthetase. *Biochemistry 30*, 11767–11774.

Ghosh, G., Pelka, H., & Schulman, L.H. (1990). Identification of the tRNA anticodon recognition site of *Escherichia coli* methionyl-tRNA synthetase. *Biochemistry 29*, 2220–2225.

Gribskov, M., Luthy, R., & Eisenberg, D. (1990). Profile analysis. *Methods Enzymol. 183*, 146–159.

Härtlein, M., Frank, R., & Madern, D. (1987). Nucleotide sequence of *Escherichia coli* valyl-tRNA synthetase gene valS. *Nucleic Acids Res. 15*, 9081–9082.

Härtlein, M. & Madern, D. (1987). Molecular cloning and nucleotide sequence of the gene for *Escherichia coli* leucyl-tRNA synthetase. *Nucleic Acids Res. 15*, 10199–10210.

Hohmann, S. (1991). *Genbank Release 71.0.* Accession #X62878.

Horn, P.B.V. & Zahler, S.A. (1992). Cloning and nucleotide sequence of the leucyl-tRNA synthetase gene of *Bacillus subtilis. J. Bacteriol. 174*, 3928–3935.

Hou, Y.-M., Shiba, K., Mottes, C., & Schimmel, P. (1991). Sequence determination and modeling of structural motifs for the smallest monomeric aminoacyl-tRNA synthetase. *Proc. Natl. Acad. Sci. USA 88*, 976–980.

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

Hsieh, H.L. & Campbell, D.R. (1991). Evidence that gene G7a in the human major histocompatibility complex encodes valyl-tRNA synthetase. *Biochem. J. 278*, 809–816.

Jenal, U., Rechsteiner, T., Tan, P.-Y., Buhlmann, E., Meile, L., & Leisinger, T. (1991). Isoleucyl-tRNA synthetase of *Methanobacterium thermoautotrophicum* Marburg: Cloning of the gene, nucleotide sequence, and localization of a base change conferring resistance to pseudomonic acid. *J. Biol. Chem. 266*, 10570–10577.

Jordana, X., Chatton, B., Paz-Weisshaar, M., Buhler, J.-M., Cramer, F., Ebel, J.P., & Fasiolo, F. (1987). Structure of the yeast valyl-tRNA synthetase gene (VAS1). and the homology of its translated amino acid sequence with *Escherichia coli* isoleucyl-tRNA synthetase. *J. Biol. Chem. 262*, 7189–7194.

Kubelik, A.R., Turcq, B., & Lambowitz, A.M. (1991). *Genbank Release 70.0.* Accession #M64703.

Martindale, D.W., Gu, Z.M., & Csank, C. (1989). Isolation and complete sequence of the yeast isoleucyl-tRNA synthetase gene (ILS1). *Curr. Genet. 15*, 99–106.

Martinis, S.A. & Schimmel, P. (1992). Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine. *Proc. Natl. Acad. Sci. USA 89*, 65–69.

Mechulam, Y., Schmitt, E., Panvert, M., Schmitter, J.-M., Lapadat-Tapolsky, M., Meinnel, T., Dessen, P., Blanquet, S., & Fayat, G. (1991). Methionyl-tRNA synthetase from *Bacillus stearothermophilus*: Structural and functional identities with the *Escherichia coli* enzyme. *Nucleic Acids Res. 19*, 3673–3681.

Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S., & Fayat, G. (1991). Selection of suppressor methionyl-tRNA synthetase: Mapping the tRNA anticodon binding site. *Proc. Natl. Acad. Sci. USA 88*, 291–295.

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

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

Needleman, S.B. & Wunsch, C.D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol. 48*, 443–453.

Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T., & Yokoyama, S. (1991). The methionyl-tRNA synthetase gene from an extreme thermophile, *Thermus thermophilus* HB8. *J. Biol. Chem. 266*, 3268–3277.

Rould, M.A., Perona, J.J., Söll, D., & Steitz, T.A. (1989). Structure of *E. coli* glutaminyl-tRNA synthetase complexed with $tRNA^{Gln}$ and ATP at 2.8 Å resolution. *Science 246*, 1135–1142.

Rould, M.A., Perona, J.J., Söll, D., & Steitz, T.A. (1991). Structural basis of anticodon loop recognition by glutaminyl-tRNA synthetase. *Nature 352*, 213–218.

Schimmel, P. (1991). Mutant enzymes and dissected tRNAs that elucidate motifs for protein-RNA recognition. *Current Opin. Struct. Biol. 1*, 811–816.

Shiba, K. & Schimmel, P. (1992). Functional assembly of a randomly cleaved protein. *Proc. Natl. Acad. Sci. USA 89*, 1880–1884.

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

Tzagoloff, A., Akai, A., Kurkulos, M., & Repetto, B. (1988). Homology of yeast mitochondrial leucyl-tRNA synthetase and isoleucyl- and methionyl-tRNA synthetase of *Escherichia coli. J. Biol. Chem. 263*, 850–856.

Tzagoloff, A., Vambutas, A., & Akai, A. (1989). Characterization of MSM1, the structural gene for yeast mitochondrial methionyl-tRNA synthetase. *Eur. J. Biochem. 179*, 365–371.

Walter, P., Gangloff, J., Bonnet, J., Boulanger, Y., Ebel, J.P., & Fasiolo, F. (1983). Primary structure of the *Saccharomyces cerevisiae* gene for methionyl-tRNA synthetase. *Proc. Natl. Acad. Sci. USA 80*, 2437–2441.

Webster, T.A., Tsai, H., Kula, M., Mackie, G.A., & Schimmel, P. (1984). Specific sequence homology and three-dimensional structure of an aminoacyl transfer RNA synthetase. *Science 226*, 1315–1317.