Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 976–980, February 1991
Biochemistry

Sequence determination and modeling of structural motifs for the smallest monomeric aminoacyl-tRNA synthetase

(cysteine-tRNA synthetase/nucleotide-binding fold)

YA-MING HOU, KIYOTAKA SHIBA, CHRISTOPHER MOTTES, AND PAUL SCHIMMEL

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Paul Schimmel, November 19, 1990

ABSTRACT **Polypeptide chains of 19 previously studied *Escherichia coli* aminoacyl-tRNA synthetases are as large as 951 amino acids and, depending on the enzyme, have quaternary structures of α, α_2, $\alpha_2\beta_2$, and α_4. These enzymes have been organized into two classes which are defined by sequence motifs that are associated with specific three-dimensional structures. We isolated, cloned, and sequenced the previously uncharacterized gene for *E. coli* cysteine-tRNA synthetase (EC 6.1.1.16) and showed that it encodes a protein of 461 amino acids. Biochemical analysis established that the protein is a monomer, thus establishing this enzyme as the smallest known monomeric synthetase. The sequence shows that cysteine-tRNA synthetase is a class I enzyme that is most closely related to a subgroup that includes the much larger methionine-, isoleucine-, leucine-, and valine-tRNA synthetases, which range in size from 677 to 951 amino acids. The amino-terminal 293 amino acids of the cysteine enzyme can be modeled as a nucleotide-binding fold that is more compact than that of its closest relatives by virtue of truncations of two insertions that split the fold. This smaller nucleotide-binding fold accounts for much of the reduced size of the cysteine enzyme and establishes the limit to which the structure of this domain is contracted in the five members of this subgroup of class I enzymes.**

Aminoacyl-tRNA synthetases are among the oldest proteins, arising early in evolution to establish the rules of the genetic code through their interactions with amino acids and transfer RNAs (1, 2). These enzymes specifically attach an amino acid to its cognate tRNA in a reaction driven by the hydrolysis of ATP. The diversity of their polypeptide chain sizes, sequences, and quaternary structures is well characterized. However, nine of the enzymes have two sequence elements that suggest a common structural relationship (3, 4). One element is the signature sequence, which is an 11-amino acid segment that ends in the tetrapeptide His-Xaa-Gly-His (HXGH), where Xaa is a hydrophobic amino acid. The other element is the tetrapeptide Lys-Met-Ser-Lys (KMSK). In the x-ray structures of three enzymes [*Bacillus stearothermophilus* tyrosine- (5–7), *Escherichia coli* methionine- (8–10), and *E. coli* glutamine-tRNA synthetase (11)] that have these sequence motifs, both elements are found within an amino-terminal nucleotide-binding fold (Rossmann fold) of alternating β-sheets and α-helices (12). The signature sequence is located in the same part of each structure, and it starts at the carboxyl terminus of one strand of β-sheet, forms a loop, and terminates at the beginning of an α-helix with the HXGH motif. The KMSK tetrapeptide is near the carboxyl terminus of the nucleotide-binding fold domain. Both HXGH and KMSK sequence motifs interact with ATP (9, 11).

The 9 enzymes that share these sequence motifs are believed to have similar nucleotide-binding folds and include aminoacyl-tRNA synthetases for arginine, glutamine, glutamic, isoleucine, leucine, methionine, tryptophan, tyrosine, and valine. These are designated as class I synthetases (4). Each of 10 other synthetases lack the signature sequence and KMSK element, but 8 of these share three sequence motifs that are absent from the class I enzymes, while the remaining two enzymes (alanine- and glycine-tRNA synthetases) have one of the three motifs. The three-dimensional structure of the serine-tRNA synthetase has no nucleotide-binding fold and shows no similarity to any parts of the class I synthetases and thus establishes a structural basis for a distinct second class of aminoacyl-tRNA synthetases (13).

Although only four structures of synthetases have been solved, the use of structural modeling, together with point and deletion mutagenesis that tests predictions of the models, has suggested that the locations of specific elements of secondary structure can be predicted and modeled in some of the class I enzymes for which there is no structural information (14, 15). The differences in the sizes of the enzymes are explained in part by the insertion of variable lengths of polypeptides into the nucleotide-binding fold domain. A portion of one of these insertions has been demonstrated to be dispensable in isoleucine-tRNA synthetase (14) and, conversely, the analogous insertion in methionine-tRNA synthetase has been artificially expanded by the introduction of novel sequence cassettes (16).

Although these and other experimental manipulations of synthetase structures have clarified some of the relationships between the class I enzymes, the natural variations of sequences and structures of individual enzymes are the most instructive. In this regard, it is noteworthy that no cysteine-tRNA synthetase (EC 6.1.1.16) from any source has previously been sequenced. Because, among other considerations, this enzyme provided the only remaining possibility to analyze a natural variation of one of the synthetase structural formats, efforts were made to clone, sequence,* characterize, and model the structure of cysteine-tRNA synthetase to test and refine concepts of synthetase architecture.

MATERIALS AND METHODS

Strains. *E. coli* strains UT171 and UT181 were provided by Asgeir Bjornsson (Uppsala University, Sweden). UT181 carries a temperature-sensitive allele of cysteine-tRNA synthetase [*cysS818* (ts), Tn*10*] (17) and is otherwise isogenic to UT171 (*cysS*$^+$, Tn*10*). The *cysS818* allele of UT181 and the *cysS*$^+$ allele of UT171 were separately transferred to wild-type strain JM103 by P1 transduction to obtain JM103 *cysS818* and JM103 *cysS*$^+$, respectively. Aminoacylation assays confirmed that extracts of JM103 *cysS818* have a temperature-sensitive defect in cysteine-tRNA synthetase activity. JM103 *cysS818* was used for screening a genomic library by complementation.

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.

Abbreviation: CP, connective peptide.

*The sequence reported in this paper has been deposited in the GenBank data base (accession no. M59381).

Cloning and Sequencing the Gene for Cysteine-tRNA Synthetase. Extraction of DNA from *E. coli* strain JM109 and partial digestion with *Sau*3A were carried out according to Silhavy *et al.* (18). Digested fragments were separated by electrophoresis on a 0.8% agarose gel. Fragments of sizes between 5 and 15 kilobase pairs (kbp) were excised, eluted, and ligated into the *Bam*HI site of plasmid pBR322. Competent JM103 *cysS818* cells were transformed with the ligation mixture and colonies that complemented the temperature-sensitive phenotype of the *cysS818* allele were selected at 42°C on ampicillin-containing Luria broth (LB) plates. From approximately 2000 ampicillin-resistant colonies that were screened, one colony complemented the temperature-sensitive phenotype of the *cysS818* allele. The plasmid DNA carried by this colony was designated pYM100 and had an insert of a 4.7-kbp fragment of genomic DNA.

A subclone of pYM100 (pYM101) complemented the *cysS818* allele and contained a 2.7-kbp fragment of the genomic DNA. This 2.7-kbp fragment was sequenced on both strands by the dideoxy termination method using the M13/pUC19 reverse primer (19).

Purification of Cysteine-tRNA Synthetase and Sequence Analysis. Aminoacylation assays were done by a variation of the procedures previously described (20). (Details are available upon request.) Cysteine-tRNA synthetase was purified from *E. coli* strain JM109 harboring plasmid pYM101. The preparation of crude extract and S100 fraction and chromatography on DEAE-cellulose were as described earlier (21). Activity for aminoacylation with cysteine eluted from the DEAE-cellulose column between 0.17 M and 0.2 M NaCl. The fractions that contained the highest activity were pooled, and an aliquot was electrophoresed on an SDS/10% polyacrylamide gel. The overproduced enzyme on the gel was blotted onto an Immobilon-P transfer membrane (22) for sequence analysis on an Applied Biosystems model 470A gas phase sequencer at the Whitehead Institute.

RESULTS

Molecular Cloning of the Gene for Cysteine-tRNA Synthetase. *In vitro* aminoacylation assays showed that extracts of strain JM103 *cysS818* contained a much-diminished activity of cysteine-tRNA synthetase at 42°C. While the cysteine tRNA charging activity in the cell extract of JM103 *cysS818* was only half of that of the wild-type JM103 *cysS*$^+$ at 30°C, the activity of the mutant was not detectable at 42°C (data not shown). Thus, the temperature-sensitive phenotype of JM103 *cysS818* correlates with the defective cysteine-tRNA synthetase activity measured *in vitro*. We expected that complementation of this temperature-sensitive defect should be achieved by introduction of a wild-type gene for cysteine-tRNA synthetase.

One clone of an *E. coli* genomic library was found to complement JM103 *cysS818* at 42°C. The chromosomal insert of this clone was subjected to sequence analysis on both strands. An open reading frame of 1386 nucleotides was identified, which starts with the ATG translation initiation codon and extends to the termination codon TAA (Fig. 1). The sequence encodes a polypeptide of 461 amino acids with a calculated molecular mass of 52,168 Da. An extract of *E. coli* that harbored the gene on plasmid YM101 contained an overexpressed protein species of an apparent molecular mass of 55 kDa on an SDS/polyacrylamide gel. The sequence starts with methionine, and the next 24 amino acids of this protein are identical to those encoded by the gene (data not shown). The partially purified protein was chromatographed through a Superose-12 column (Pharmacia) and the elution position of the aminoacylation activity was interpolated to a molecular mass of 53 kDa (data not shown). Thus, the cloned cysteine-tRNA synthetase behaves on chromatography as a monomer.

Modeling the Structural Motifs of Cysteine-tRNA Synthetase. Cysteine-tRNA synthetase has the HXGH and KMSK motifs that are characteristic of class I aminoacyl tRNA synthetases (3, 4). An alignment of these regions of the cysteine enzyme with those of other class I enzymes is shown in Fig. 2. The locations of elements of α-helix and β-sheet in the known three-dimensional structure of methionine-tRNA synthetase are indicated. Amino acid identities between cysteine and the other class I enzymes are indicated by shading. For all 10 enzymes the alignment of the HIGH

```
                M L K I F N T L T R Q K E E F K P I H A G E V G M Y V C G I T V Y D L   35
TCTAAACGGAATCAACGATGCTAAAAATCTTCAATACTCTGACACGCCAAAAAGAGGAATTTAAGCCTATTCACGCCGGGGAAGTCGGCATGTACGTGTGTGGAATCACCGTTTACGATC

C H I G H G R T F V A F D V V A R Y L R F L G Y K L K Y V R N I T D I D D K I I   75
TCTGTCATATCGGTCACGGGCGTACCTTTGTTGCTTTTGACGTGGTTGCGCGCTATCTGCGTTTCCTCGGCTATAAACTGAAGTATGTGCGCAACATTACCGATATCGACGACAAAATCA

K R A N E N G E S F V A M V D R M I A E M H K D F D A L N I L R P D M E P R A T  115
TCAAACGCGCCAATGAAAATGGCGAAAGCTTTGTGGCGATGGTGGATCGCATGATCGCCGAAATGCACAAAGATTTTGATGCTTTGAACATTCTGCGCCCGGATATGGAGCCGCGCGCGA

H H I A E I I E L T E Q L I A K G H A Y V A D N G D V M F D V P T D P T Y G V L  155
CGCACCATATCGCAGAAATTATTGAACTCACTGAACAACTGATCGCCAAAGGTCACGCTTATGTGGCGGACAACGGCGACGTGATGTTCGACGTCCCGACCGATCCAACTTATGGCGTGC

S R Q D L D Q L Q A G A R V D V V D D K R N P M D F V L W K M S K E G E P S W P  195
TGTCGCGTCAGGATCTCGACCAGCTGCAGGCAGGCGCGCGCGTTGACGTGGTCGACGACAAACGCAACCCAATGGACTTCGTTCTGTGGAAGATGTCGAAAGAGGGCGAACCGAGCTGGC

S P W G A G R P G W H I E C S A M N C K Q L G N H F D I H G G G S D L M F P H H  235
CGTCTCCGTGGGGCGCGGGTCGTCCTGGCTGGCACATTGAATGTTCGGCAATGAACTGCAAGCAGCTGGGTAACCACTTTGATATCCACGGCGGCGGTTCAGACCTGATGTTCCCGCACC

E N E I A Q S T C A H D G Q Y V N Y W M H S G M V M V D R E K M S K S L G N F F  275
ACGAAAACGAAATCGCGCAGTCCACCTGTGCCCATGATGGTCAGTATGTGAACTACTGGATGCACTCGGGGATGGTGATGGTTGACCGCGAGAAGATGTCCAAATCGCTGGGTAACTTCT

T V R D V L K Y Y D A E T V R Y F L M S G H Y R S Q L N Y S E E N L K Q A R A A  315
TTACCGTGCGCGATGTGCTGAAATACTACGACGCGGAAACCGTGCGTTACTTCCTGATGTCGGGCCACTATCGCAGCCAGTTGAACTACAGCGAAGAGAACCTGAAGCAGGCGCGTGCGG

L E R L Y T A L R G T D K T V A P A G G E A F E A R F I E A M D D D F N T P E A  355
CGCTGGAGCGTCTCTACACTGCGCTGCGCGGCACAGATAAAACCGTTGCGCCTGCCGGTGGCGAAGCGTTTGAAGCGCGCTTTATTGAAGCGATGGACGACGATTTCAACACCCCGGAAG

Y S V L F D M A R E V N R L K A E D M A A A N A M A S H L R K L S A V L G L L E  395
CCTATTCCGTACTGTTTGATATGGCGCGTGAAGTAAACCGTCTGAAAGCAGAAGATATGGCAGCGGCGAATGCAATGGCATCTCACCTGCGTAAACTTTCCGCTGTATTGGGCCTGCTGG

Q E P E A F L Q S G A Q A D D S E V A E I E A L I Q Q R L D A R K A K D W A A A  435
AGCAAGAACCGGAAGCGTTCCTGCAAAGCGGCGCGCAGGCAGACGACAGCGAAGTGGCTGAGATTGAAGCGTTAATTCAACAGCGTCTGGATGCCCGTAAAGCGAAAGACTGGGCGGCGG

D A A R D R L N E M G I V L E D G P Q G T T W R R K *                        461
CGGATGCGGCGCGTGATCGTCTTAACGAGATGGGGATCGTGCTGGAAGATGGCCCGCAAGGGACCACCTGGCGTCGTAAGTAATTGCGCTATTGCCGGATGCGAGTTTTCGCATCCGGTT
```

FIG. 1. DNA and amino acid sequences of the coding region of the gene for *E. coli* cysteine-tRNA synthetase. An arbitrary number of nucleotides in the 5′ and 3′ untranslated region are also included. The HXGH and KMSK tetrapeptides are underlined.

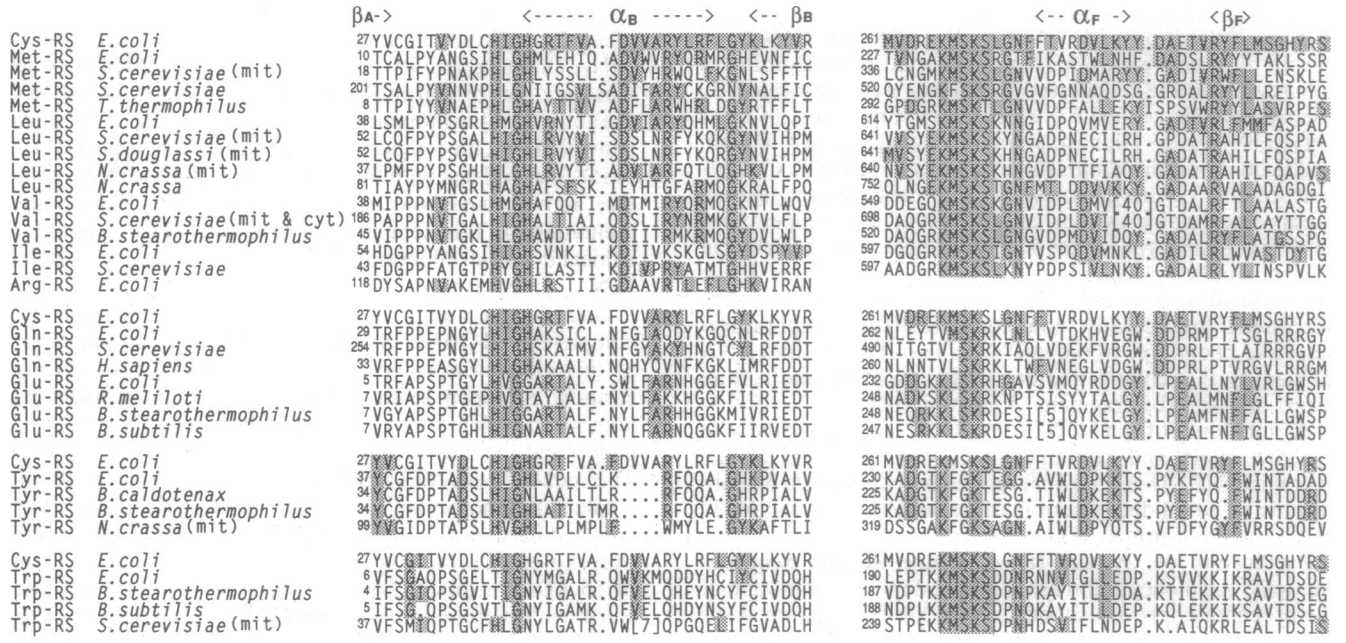

FIG. 2. Juxtaposition of regions containing HXGH and KMSK tetrapeptides of *E. coli* and other class I aminoacyl-tRNA synthetases. The locations of α-helix (α) and β-sheet (β) are based on the structure of methionine-tRNA synthetase (8–10). Amino acid identities between the cysteine enzyme and other class I synthetases are shaded. Numbers correspond to amino acid residues in the protein sequences. Dots indicate gaps and brackets indicate insertions in the alignment. The sequences of aminoacyl-tRNA synthetases (RSs) are from GenBank release 65.0 with the addition of those for methionine [*Saccharomyces cerevisiae* mitochondrial (mit) (23), *Thermus thermophilus* (24)], isoleucine [*E. coli* (25)], glutamine [*S. cerevisiae* (26)], glutamate [*Rhizobium meliloti* (27), *B. stearothermophilus* (28), *Bacillus subtilis* (28)], and tryptophan [*S. cerevisiae* (mit) (29)]. Species not named in full elsewhere are *Saccharomyces douglassi, Neurospora crassa, Homo sapiens*, and *Bacillus caldotenax*.

sequence that ends at the amino terminus of α_B is evident, and scattered amino acid identities continue throughout this helix and on into β_B. The identities with cysteine-tRNA synthetase are somewhat more frequent with the subgroup that includes methionine-, leucine-, valine-, isoleucine-, and arginine-tRNA synthetases.

Around the KMSK sequence, which lies between β_E and α_F in the structure of the methionine enzyme, amino acid identities with cysteine-tRNA synthetase are also evident. [Because arginine-tRNA synthetase has no KMSK sequence (30), it is not listed in the alignment of this region.] The frequency of identities is greatest with the synthetases for methionine and for the hydrophobic amino acids, and it extends to the end of the nucleotide fold of methionine-tRNA synthetase at β_F, where there is an RXF tripeptide that is shared by several of these enzymes.

The alignment of the sequence of the cysteine-tRNA synthetase with all of the elements of secondary structure in the nucleotide-binding fold of *E. coli* methionine-tRNA synthetase is shown in Fig. 3. Sequences of the *T. thermophilus* and *S. cerevisiae* cytoplasmic and mitochondrial methionine enzymes have been included for comparison. The fold is split by two insertions (14): connective polypeptide 1 (CP1) and connective polypeptide 2 (CP2). These are between β_C and β_D (CP1) and β_D and α_E (CP2). Sequence identities between cysteine-tRNA synthetase and at least one of the methionine enzymes are present throughout the nucleotide-binding fold, including the segment between β_B and β_C and the two ends of CP2.

The alignment of Fig. 3 makes possible the identification of those parts of the cysteine-tRNA synthetase sequence that can be assigned to CP1 and CP2. These extend from His-116 to Trp-184 and Ser-193 to His-235, respectively. A schematic diagram of the locations of the secondary structure elements and the CP1 and CP2 segments for cysteine-, methionine-, leucine-, valine-, and isoleucine-tRNA synthetases is shown

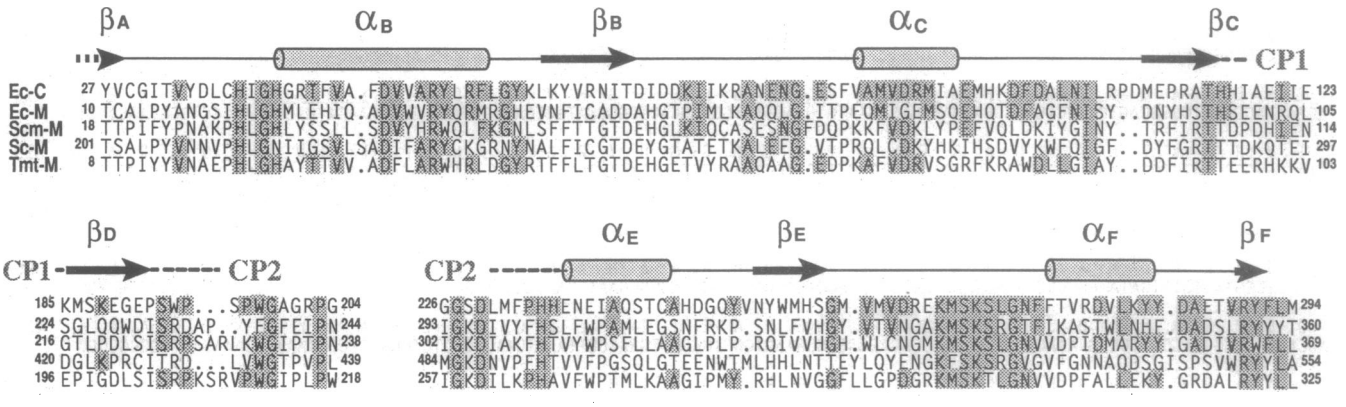

FIG. 3. Alignment of secondary structural elements in the nucleotide-binding fold of *E. coli* cysteine-tRNA synthetase (Ec-C) with methionine-tRNA synthetases from *E. coli* (Ec-M), *S. cerevisiae* mitochondria (Scm-M), *S. cerevisiae* (Sc-M), and *T. thermophilus* (Tmt-M). Regions of connective polypeptides CP1 and CP2 that are omitted are indicated by the numbering of the amino acid residues.

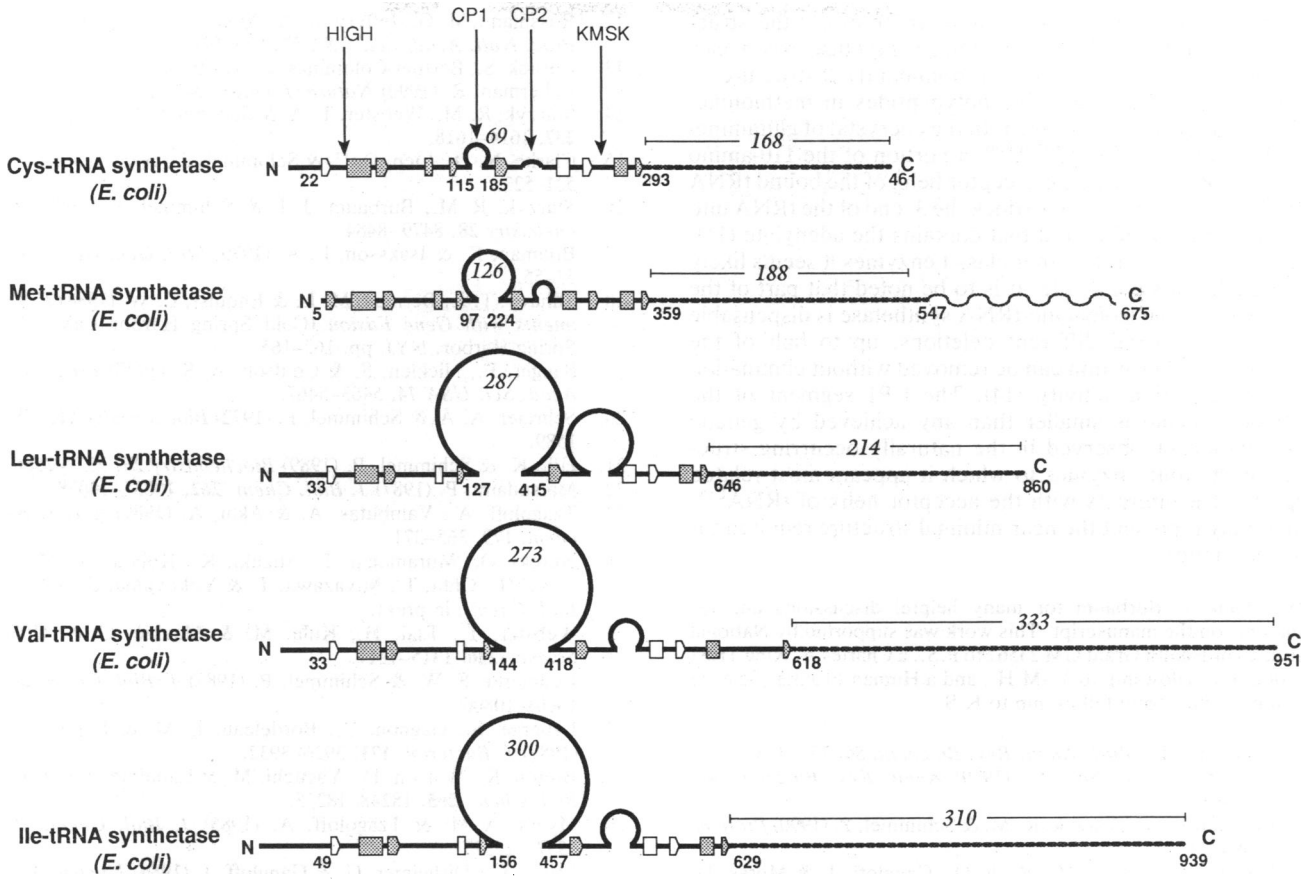

FIG. 4. Structural organization of *E. coli* cysteine-, methionine-, leucine-, valine-, and isoleucine-tRNA synthetases. Rectangles indicate α-helices, pentagons indicate β-sheets, and loops indicate connective polypeptides. Italicized numbers represent lengths of polypeptide sequences in the proteins. Shaded regions of secondary structure of methionine-tRNA synthetase are confirmed by the crystal structure (8–10) and, for the other enzymes, are highly probable secondary structure elements based on sequence relationships to the methionine enzyme. The open symbols are regions of secondary structure that have a weak sequence relationship to the analogous elements in methionine-tRNA synthetase. The domain responsible for dimerization of methionine-tRNA synthetase is shown as a wavy region at the carboxyl terminus between amino acids 547 and 675. The four other synthetases are monomers.

in Fig. 4. Except for methionine-tRNA synthetase, all of these assignments are based on analyses similar to those outlined here and are supported in the case of *E. coli* isoleucine-tRNA synthetase by experimental tests (14).

The five enzymes vary in size from 461 to 951 amino acids. Most striking is the variation in size of CP1 and, to a lesser extent, of CP2. CP1 is approximately 300 amino acids in isoleucine-tRNA synthetase, is about 120 amino acids in methionine-tRNA synthetase, and is about 68 amino acids in cysteine-tRNA synthetase. CP2 is condensed from approximately 70 amino acids in methionine-tRNA synthetase to just 42 amino acids in cysteine-tRNA synthetase. Thus, the reduction in the size of cysteine-tRNA synthetase appears to conserve the secondary structure elements of the nucleotide-binding fold in the methionine-like enzymes and to achieve compaction solely through the elimination of large portions of the connective polypeptides.

DISCUSSION

Compared to cysteine-tRNA synthetase, the next-smallest monomeric synthetase is the *E. coli* glutamate enzyme, which has 471 amino acids (31). While the subunits of *E. coli* histidine-, tyrosine-, and tryptophan-tRNA synthetases are smaller (334 to 424 amino acids) than the cysteine-tRNA synthetase polypeptide, the former three enzymes are α_2 dimers (1) that are, therefore, larger than the monomeric cysteine enzyme. In the case of tyrosine-tRNA synthetase, it is known that the dimeric structure is required for aminoacylation activity, particularly because the tRNA is believed to bind across the subunit interface (32–34). While the dimeric methionine-tRNA synthetase can be truncated at Lys-547 to yield an amino-terminal monomeric fragment that has near full activity (35), further carboxyl-terminal deletions beyond Ile-534 result in complete loss of activity (36). Thus, to our knowledge, cysteine-tRNA synthetase is the smallest natural or artificially truncated aminoacyl-tRNA synthetase with full activity and all of the catalytic sites confined to one polypeptide chain.

Some similarities between methionine-tRNA synthetase and the synthetases for leucine, isoleucine, and valine have been noted previously (14, 37), and the results reported here suggest that the cysteine enzyme should be included in this group. These five enzymes may have a stronger historical relationship with each other than with the five other class I enzymes. In this connection, it is worth noting that, even though features such as the region encompassing the HIGH tetrapeptide are closely similar, the overall organization of the nucleotide-binding fold (six β-strands) of these five enzymes differs from that of the fold (five β-strands) of the less related class I tyrosine enzyme or of the glutamine enzyme. Also, the carboxyl-terminal domain of the methionine enzyme is predicted (unpublished results) and observed (8) to be predominantly α-helical. The carboxyl-terminal domains of the cysteine, leucine, isoleucine, and valine enzymes are also strongly predicted (38) to be α-helical (unpublished results). Thus, the relationship between the cysteine, methionine, isoleucine, leucine, and valine enzymes may extend

beyond the nucleotide-binding fold. In contrast, the structurally less related glutamine-tRNA synthetase has a carboxyl-terminal domain that is predominantly β-structure.

The role of the connective polypeptides in methionine-tRNA synthetase is not known. In the cocrystal of glutamine-tRNA synthetase with $tRNA^{Gln}$, a portion of the 110-amino acid CP1 interacts with the acceptor helix of the bound tRNA and, therefore, is required to dock the 3′ end of the tRNA into the nucleotide-binding fold that contains the adenylate (11). For at least some of the other class I enzymes it seems likely that CP1 has a similar role. It is to be noted that part of the CP1 region of the isoleucine-tRNA synthetase is dispensable because, in several different deletions, up to half of the 300-amino acid insertion can be removed without elimination of aminoacylation activity (14). The CP1 segment of the cysteine enzyme is smaller than any achieved by genetic manipulation or observed in the naturally occurring structures of the four enzymes to which it appears most related (Fig. 4). If it interacts with the acceptor helix of $tRNA^{Cys}$, then it may represent the near-minimal structure required for that interaction.

We thank J. Burbaum for many helpful discussions and for comments on the manuscript. This work was supported by National Institutes of Health Grant GM 23562 to P.S., a Charles A. King Trust postdoctoral fellowship to Y.-M.H., and a Human Frontier Science Program postdoctoral fellowship to K.S.

1. Schimmel, P. (1987) *Annu. Rev. Biochem.* **56**, 125–158.
2. Schimmel, P. & Soll, D. (1979) *Annu. Rev. Biochem.* **48**, 601–648.
3. Burbaum, J. J., Starzyk, R. M. & Schimmel, P. (1990) *Proteins* **7**, 99–111.
4. Eriani, G., Delarue, M., Poch, O., Gangloff, J. & Moras, D. (1990) *Nature (London)* **347**, 203–206.
5. Bhat, T. N., Blow, D. M., Brick, P. & Nyborg, J. (1982) *J. Mol. Biol.* **158**, 699–709.
6. Blow, D. M. & Brick, P. (1985) in *Biological Macromolecules and Assemblies*, eds. Jurnak, F. A. & MacPherson, A. (Wiley, New York), Vol. 2, pp. 442–467.
7. Brick, P., Bhat, T. N. & Blow, D. M. (1988) *J. Mol. Biol.* **208**, 83–98.
8. Brunie, S., Mellot, P., Zelwer, C., Risler, J.-L., Blanquet, S. & Fayat, G. (1987) *J. Mol. Graphics* **5**, 18–21.
9. Zelwer, C., Risler, J.-L. & Brunie, S. (1982) *J. Mol. Biol.* **155**, 63–81.
10. Risler, J.-L., Zelwer, C. & Brunie, S. (1981) *Nature (London)* **292**, 384–386.
11. Rould, M. A., Perona, J. J., Soll, D. & Steitz, T. A. (1989) *Science* **246**, 1135–1142.
12. Rossmann, M. G., Jeffery, B. A., Main, P. & Warren, S. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 515–524.
13. Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N. & Leberman, R. (1990) *Nature (London)* **347**, 249–255.
14. Starzyk, R. M., Webster, T. A. & Schimmel, P. (1987) *Science* **237**, 1614–1618.
15. Clarke, N. D., Lien, D. C. & Schimmel, P. (1988) *Science* **240**, 521–523.
16. Starzyk, R. M., Burbaum, J. J. & Schimmel, P. (1989) *Biochemistry* **28**, 8479–8484.
17. Bohman, K. & Isaksson, L. A. (1979) *Mol. Gen. Genet.* **176**, 53–55.
18. Silhavy, T. J., Berman, M. L. & Enquist, L. W. (1984) *Experiments with Gene Fusion* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 162–165.
19. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
20. Schreier, A. A. & Schimmel, P. (1972) *Biochemistry* **11**, 1582–1589.
21. Hill, K. & Schimmel, P. (1989) *Biochemistry* **28**, 2577–2586.
22. Matsudaira, P. (1987) *J. Biol. Chem.* **262**, 10035–10038.
23. Tzagoloff, A., Vambutas, A. & Akai, A. (1989) *Eur. J. Biochem.* **179**, 365–371.
24. Nureki, O., Muramatsu, T., Suzuki, K., Kohda, D., Matsuzawa, H., Ohta, T., Miyazawa, T. & Yokoyama, S. (1991) *J. Biol. Chem.*, in press.
25. Webster, T., Tsai, H., Kula, M. & Mackie, G. A. (1984) *Science* **226**, 1315–1317.
26. Ludmerer, S. W. & Schimmel, P. (1987) *J. Biol. Chem.* **262**, 10801–10806.
27. Laberge, S., Gagnon, Y., Bordeleau, L. M. & Lapointe, J. (1989) *J. Bacteriol.* **171**, 3926–3932.
28. Breton, R., Watson, D., Yaguchi, M. & Lapointe, J. (1990) *J. Biol. Chem.* **265**, 18248–18255.
29. Myers, A. M. & Tzagoloff, A. (1985) *J. Biol. Chem.* **260**, 15371–15377.
30. Eriani, G., Dirheimer, G. & Gangloff, J. (1989) *Nucleic Acids Res.* **17**, 5725–5736.
31. Breton, R., Sanfacon, H., Papayannopoulos, I., Biemann, K. & Lapointe, J. (1986) *J. Biol. Chem.* **261**, 10610–10617.
32. Waye, M. M. Y., Winter, G., Wilkinson, A. J. & Fersht, A. R. (1983) *EMBO J.* **2**, 1827–1829.
33. Bedouelle, H. & Winter, G. (1986) *Nature (London)* **320**, 371–373.
34. Carter, P., Bedouelle, H. & Winter, G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1189–1193.
35. Cassio, D. & Waller, J. P. (1971) *Eur. J. Biochem.* **20**, 283–300.
36. Mellot, P., Mechulam, Y., Le Corre, D., Blanquet, S. & Fayat, G. (1989) *J. Mol. Biol.* **208**, 429–443.
37. Heck, J. D. & Hatfield, G. W. (1988) *J. Biol. Chem.* **263**, 868–877.
38. Ralph, W. W., Webster, T. A. & Smith, T. F. (1987) *Comput. Appl. Biosci.* **3**, 211–216.