Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 2437–2441, May 1983
Biochemistry

Primary structure of the *Saccharomyces cerevisiae* gene for methionyl-tRNA synthetase

(DNA sequence analysis/protein sequence comparison)

PHILIPPE WALTER, JEAN GANGLOFF, JACQUES BONNET, YVES BOULANGER, JEAN-PIERRE EBEL, AND FRANCO FASIOLO

Institut de Biologie Moléculaire et Cellulaire du Centre National de la Recherche Scientifique—Laboratoire de Biochimie, 15, rue René Descartes 67084 Strasbourg Cedex, France

Communicated by Boris Magasanik, December 30, 1982

ABSTRACT The sequence of a 5-kilobase DNA insert containing the structural gene for yeast cytoplasmic methionyl-tRNA synthetase has been determined and a unique open reading frame of 2,253 nucleotides encoding a polypeptide chain of 751 amino acids (M_r, 85,500) has been characterized. The data obtained on the purified enzyme (subunit size, amino acid composition, and COOH-terminal sequence) are consistent with the gene structure. The protein sequence deduced from the nucleotide sequence reveals no obvious internal repeats. This protein sequence shows a high degree of homology with that of *Escherichia coli* methionyl-tRNA synthetase within a region that forms the putative methionyl adenylate binding site. This strongly suggests that both proteins derive from a common ancestor.

Aminoacyl-tRNA synthetases play a crucial role in protein synthesis because they catalyze the specific attachment of amino acids to their cognate tRNAs. Knowledge of their primary structure is one of the prerequisites for the complete understanding of the structure–function relationship. So far, the sequence of only one aminoacyl-tRNA synthetase has been completely determined by using classical protein sequence analysis techniques (1). However, the cloning of a number of aminoacyl-tRNA synthetases genes has allowed derivation of the primary structure of the corresponding enzyme from the DNA sequence—i.e., the alanyl- (2), tryptophanyl- (3), glutaminyl- (4), and methionyl- (5) tRNA synthetases from *Escherichia coli*.

In our laboratory, a yeast mutant strain lacking functional cytoplasmic methionyl-tRNA synthetase was available. It was complemented with a plasmid pool containing random fragments of wild-type yeast genomic DNA obtained by a partial *Sau*3A digestion (6). It was thus possible to isolate a 5.1-kilobase piece of DNA containing the methionyl-tRNA synthetase gene (*MES1*). The isolated gene product is a monomer (M_r, 80,000). However, in crude extracts from a wild-type strain the enzyme behaves as a dimer (M_r, 2 × 80,000). Because enzyme purification always led to a monomeric species with no detectable variation of M_r, the existence of a dimeric structure for native methionyl-tRNA synthetase is by no means proven (unpublished data).

In this paper we report the complete nucleotide sequence of the DNA insert containing the structural gene. Indeed, a unique open reading frame whose length corresponds exactly to that expected for the above protein size could be characterized.

MATERIALS AND METHODS

DNA Sequence Analysis. DNA sequence analysis was carried out by the chain termination method of Sanger *et al.* (7). Fragments of the cloned DNA were subcloned into the bacteriophage M13mp7 (8) and analyzed by using the primer synthesized by R. Crea and purchased from P-L Biochemicals.

Enzymes and Chemicals. Most restriction enzymes and T4 DNA ligase were obtained from New England BioLabs. The Klenow *E. coli* polymerase and nuclease S1 were from Boehringer Mannheim.

All chemicals were from Merck (Darmstadt, Federal Republic of Germany) and were analytical grade.

RESULTS AND DISCUSSION

Sequence of the *MES1* Gene. Fig. 1 shows the complete nucleotide sequence of a 3,215-base-pair-long DNA fragment bearing the structural gene for the yeast methionyl-tRNA synthetase in which a unique open reading frame was identified from nucleotide −33 to nucleotide 2,253 after which the next in-phase stop codon occurs. Nucleotide 1 is A of the first ATG encountered in this frame. As suggested for eukaryotic systems (9), we assume that this is the initiating codon because it is the first ATG after the start of transcription determined by nuclease S1 mapping which showed three major 5′-mRNA ends centered at positions −10, −18, and −17 (unpublished data). Therefore, this DNA section is large enough to encode a polypeptide chain of 751 amino acids (including the initiating methionine). The corresponding peptide sequence is also shown in Fig. 1. It yields a M_r of 85,500, in good agreement with that measured for the purified protein.

Unfortunately, NH_2-terminal sequence analyses of the enzyme purified from transformed yeast cells did not give any clear answer. Indeed, the negative results of dansylation, Edman–Chang degradation, and aminopeptidase digestion suggest that the NH_2-terminal residue is blocked. But three lines of evidence establish that the stretch of DNA shown in Fig. 1 does contain the *MES1* gene:

(*i*) The amino acid composition deduced from the DNA sequence fits well with that calculated from amino acid analysis of the protein purified from the transformed cells (Table 1).

(*ii*) The COOH terminus predicted from this DNA sequence is Gln-Gln-Val-stop which is in good agreement with that established by both hydrazinolysis and carboxypeptidase A and Y digestions of the protein: (Ser or amide)-(Ser or amide)-Val(OH).

(*iii*) There is a remarkable homology (see Fig. 2) between our protein sequence and that determined for the *E. coli* crystallized tryptic fragment of methionyl-tRNA synthetase (5) in the region corresponding to the nucleotide binding domain.

Because of the lack of an unambiguous NH_2-terminal sequence as well as small oligopeptides distributed along the DNA sequence, special care was taken to establish the latter. Indeed, 100% of it was read off both strands, thus increasing our con-

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.

$^{-321}$A AGCTTTCCCG AGAACGCGAG TTCGATGATA ATGGTGTATC CTATCGATCT GTCGGCTTGC TATGATTATC TTTGTTAGCC GCCGAAACGA TCATTTATGC

$^{-220}$ATTGCCCGAG TATAGTCACT TCTTCTCAAT ATCACGAGGC ACGAATGATT TCGGAGGCAG AAAAAATTTC TGTACGCAAC TTATCGAAAG AGAAAACGAA AAAAAAAAAA

** *

$^{-110}$AGACGATGAC AAATTTTCTT CGATGTCTAA TGTTAATGAT GATTTGCTTC CTGATATGCA TCTTTGTGGA GGAATAAAAC TTGAAAAAAA AGCGGAAATT TACAACAAGC

$^{+1}$ATG TCT TTC CTC ATT TCC TTT GAT AAA TCA AAG AAA CAT CCT GCC CAT TTG CAG TTA GCG AAC AAT TTG AAG ATT GCC CTA GCA CTT GAA
Met Ser Phe Leu Ile Ser Phe Asp Lys Ser Lys Lys His Pro Ala His Leu Gln Leu Ala Asn Asn Leu Lys Ile Ala Leu Ala Leu Glu$_{30}$

TAT GCA AGC AAA AAT TTA AAG CCC GAA GTT GAC AAT GAT AAT GCT GCC ATG GAA TTG CGC AAT ACA AAG GAA CCT TTC CTT TTG TTT GAT
Tyr Ala Ser Lys Asn Leu Lys Pro Glu Val Asp Asn Asp Asn Ala Ala Met Glu Leu Arg Asn Thr Lys Glu Pro Phe Leu leu Phe Asp$_{60}$

GCT AAC GCT ATT CTA AGA TAT GTC ATG GAT GAT TTT GAA GGT CAA ACT TCC GAC AAG TAT CAA TTT GCA TTG GCA TCT TTA CAA AAC CTG
Ala Asn Ala Ile Leu Arg Tyr Val Met Asp Asp Phe Glu Gly Gln Thr Ser Asp Lys Tyr Gln Phe Ala Leu Ala Ser Leu Gln Asn Leu$_{90}$

TTA TAT CAT AAA GAA TTA CCT CAA CAG CAT GTC GAG GTT TTG ACA AAT AAG GCC ATT GAA AAT TAC TTG GTC GAA TTA AAA GAA CCA TTG
Leu Tyr His Lys Glu Leu Pro Gln Gln His Val Glu Val Leu Thr Asn Lys Ala Ile Glu Asn Tyr Leu Val Glu Leu Lys Glu Pro Leu$_{120}$

ACT GCT ACA GAT TTG ATC CTG TTT GCT AAC GTT TAT GCT CTA AAT TCT TCT TTA GTT CAT TCT AAA TTC CCA GAA TTG CCA TCC AAA GTG
Thr Ala Thr Asp Leu Ile Leu Phe Ala Asn Val Tyr Ala Leu Asn Ser Ser Leu Val His Ser Lys Phe Pro Glu Leu Pro Ser Lys Val$_{150}$

CAT AAC GCT GTA GCA TTG GCT AAA AAG CAT GTT CCA CGT GAT TCT TCT TCT TTC AAA AAC ATC GGC GCA GTG AAA ATC CAA GCT GAC TTA
His Asn Ala Val Ala Leu Ala Lys Lys His Val Pro Arg Asp Ser Ser Ser Phe Lys Asn Ile Gly Ala Val Lys Ile Gln Ala Asp Leu$_{180}$

ACA GTT AAG CCA AAG GAT TCA GAA ATT TTG CCT AAG CCA AAC GAA AGA AAC ATC TTG ATC ACT TCG GCA TTA CCT TAT GTC AAC AAC GTT
Thr Val Lys Pro Lys Asp Ser Glu Ile Leu Pro Lys Pro Asn Glu Arg Asn Ile Leu Ile Thr Ser Ala Leu Pro Tyr Val Asn Asn Val$_{210}$

CCA CAC TTG GGT AAT ATC ATC GGT AGT GTT CTT TCA GCA GAC ATT TTT GCT CGT TAC TGT AAG GGA CGT AAT TAT AAT GCC TTG TTT ATT
Pro His Leu Gly Asn Ile Ile Gly Ser Val Leu Ser Ala Asp Ile Phe Ala Arg Tyr Cys Lys Gly Arg Asn Tyr Asn Ala Leu Phe Ile$_{240}$

TGT GGT ACT GAT GAA TAT GGT ACT GCC ACG GAA ACT AAA GCT TTG GAG GAA GGT GTG ACA CCA AGA CAA CTA TGT GAC AAA TAT CAC AAA
Cys Gly Thr Asp Glu Tyr Gly Thr Ala Thr Glu Thr Lys Ala Leu Glu Glu Gly Val Thr Pro Arg Gln Leu Cys Asp Lys Tyr His Lys$_{270}$

ATC CAC AGT GAC GTT TAC AAG TGG TTC CAA ATT GGA TTT GAT TAT TTC GGT AGA ACT ACG ACG GAT AAG CAA ACG GAG ATT GCT CAA CAT
Ile His Ser Asp Val Tyr Lys Trp Phe Gln Ile Gly Phe Asp Tyr Phe Gly Arg Thr Thr Thr Asp Lys Gln Thr Glu Ile Ala Gln His$_{300}$

ATT TTT ACA AAG CTA AAT TGC AAT GGT TAT CTA GAA GAA CAA TCT ATG AAG CAA TTG TAC TGT CCA GTT CAT AAT TCT TAT CTG GCT GAT
Ile Phe Thr Lys Leu Asn Ser Asn Gly Tyr Leu Glu Glu Gln Ser Met Lys Gln Leu Tyr Cys Pro Val His Asn Ser Tyr Leu Ala Asp$_{330}$

CGT TAT GTG GAA GGT GAA TGT CCA AAA TGT CAC TAC GAT GAT GCT CGT GGG GAT CAA TGT GAC AAA TGT GGT GCC CTG TTA GAT CCA TTT
Arg Tyr Val Glu Gly Glu Cys Pro Lys Cys His Tyr Asp Asp Ala Arg Gly Asp Gln Cys Asp Lys Cys Gly Ala Leu Leu Asp Pro Phe$_{360}$

GAA TTG ATC AAT CCA CGT TGT AAA TTA GAT GAT GCT TCT CCA GAA CCA AAA TAT TCT GAT CAT ATT TTC CTA TCG CTG GAT AAA TTA GAA
Glu Leu Ile Asn Pro Arg Cys Lys Leu Asp Asp Ala Ser Pro Glu Pro Lys Tyr Ser Asp His Ile Phe Leu Ser Leu Asp Lys Leu Glu$_{390}$

AGC CAA ATT TCT GAA TGG GTT GAA AAG GCC TCT GAA GAA GGT AAC TGG TCA AAA AAT TCA AAA ACA ATT ACG CAA TCA TGG TTG AAG GAT
Ser Gln Ile Ser Glu Trp Val Glu Lys Ala Ser Glu Glu Gly Asn Trp Ser Lys Asn Ser Lys Thr Ile Thr Gln Ser Trp Leu Lys Asp$_{420}$

GGT TTG AAG CCA CGT TGT ATT ACA AGA GAT TTA GTT TGG GGT ACG CCA GTG CCT TTA GAA AAA TAT AAA GAC AAA GTC TTG TAC GTT TGG
Gly Leu Lys Pro Arg Cys Ile Thr Arg Asp Leu Val Trp Gly Thr Pro Val Pro Leu Glu Lys Tyr Lys Asp Lys Val Leu Tyr Val Trp$_{450}$

TTT GAC GCT ACA ATC GGC TAC GTT TCC ATC ACT TCC ATT TAC ACC AAA GAA TGG AAA CAA TGG TGG AAT AAT CCA GAG CAT GTT TCA TTG
Phe Asp Ala Thr Ile Gly Tyr Val Ser Ile Thr Ser Asn Tyr Thr Lys Glu Trp Lys Gln Trp Trp Asn Asn Pro Glu His Val Ser Leu$_{480}$

TAT CAA TTC ATG GGT AAG GAC AAT GTT CCT TTC CAT ACA GTT GTT TTC CCT GGT TCT CAA TTG GGT ACG GAA GAG AAC TGG ACT ATG TTG
Tyr Gln Phe Met Gly Lys Asp Asn Val Pro Phe His Thr Val Val Phe Pro Gly Ser Gln Leu Gly Thr Glu Glu Asn Trp Thr Met Leu$_{510}$

CAC CAT TTG AAT ACT ACA GAA TAC TTA CAA TAT GAG AAC GGT AAA TTT TCT AAA AGT AGG GGT GTT GGT GTT TTT GGT AAT AAC GCT CAA
His His Leu Asn Thr Thr Glu Tyr Leu Gln Tyr Glu Asn Gly Lys Phe Ser Lys Ser Arg Gly Val Gly Val Phe Gly Asn Asn Ala Gln$_{540}$

(Fig. 1 continues on next page.)

```
GAC TCT GGA ATT TCT CCA AGT GTT TGG AGA TAC TAC CTG GCA TCT GTT AGA CCA GAA TCT AGT GAT TCT CAT TTC TCA TGG GAT GAC TTT
Asp Ser Gly Ile Ser Pro Ser Val Trp Arg Tyr Tyr Leu Ala Ser Val Arg Pro Glu Ser Ser Asp Ser His Phe Ser Trp Asp Asp Phe 570

GTT GCT AGA AAC AAC AGT GAA TTG TTG GCT AAC TTG GGT AAC TTT GTT AAC AGA TTA ATT AAG TTT GTT AAT GCC AAA TAT AAT GGT GTT
Val Ala Arg Asn Asn Ser Glu Leu Leu Ala Asn Leu Gly Asn Phe Val Asn Arg Leu Ile Lys Phe Val Asn Ala Lys Tyr Asn Gly Val 600

GTT CCA AAA TTC GAC CCC AAG AAG GTT TCC AAT TAT GAT GGT TTA GTT AAA GAT ATC AAC GAA ATT TTA TCA AAT TAC GTC AAG GAA ATG
Val Pro Lys Phe Asp Pro Lys Lys Val Ser Asn Tyr Asp Gly Leu Val Lys Asp Ile Asn Glu Ile Leu Ser Asn Tyr Val Lys Glu Met 630

GAA CTT GGA CAT GAA AGA CGT GGT CTA GAA ATT GCC ATG TCG TTA AGT GCT CGT GGT AAC CAG TTT TTG CAA GAA AAT AAG TTG GAL AAT
Glu Leu Gly His Glu Arg Arg Gly Leu Glu Ile Ala Met Ser Leu Ser Ala Arg Gly Asn Gln Phe Leu Gln Glu Asn Lys Leu Asp Asn 660

ACC TTG TTT TCA CAG TCC CCA GAA AAG TCA GAT GCT GTT GTC GCA GTT GGT TTG AAT ATT ATT TAC GCT GTT AGC TCT ATT ATC ACA CCA
Thr Leu Phe Ser Gln Ser Pro Glu Lys Ser Asp Ala Val Val Ala Val Gly Leu Asn Ile Ile Tyr Ala Val Ser Ser Ile Ile Thr Pro 690

TAT ATG CCA GAA ATA GGT GAG AAA ATA AAC AAG ATG TTA AAT GCA CCA GCT TTA AAA ATT GAT GAT AGA TTC CAT TTG GCA ATC CTA GAA
Tyr Met Pro Glu Ile GLy Glu Lys Ile Asn Lys Met Leu Asn Ala Pro Ala Leu Lys Ile Asp Asp Arg Phe His Leu Ala Ile Leu Glu 720

GGA CAT AAT ATA AAC AAG GCA GAA TAC TTG TTC CAA CGT ATT GAT GAA AAG AAA ATC GAT GAA TGG AGA GCC AAA TAT GGT GGT CAA CAA
Gly His Asn Ile Asn Lys Ala Glu Tyr Leu Phe Gln Arg Ile Asp Glu Lys Lys Ile Asp Glu Trp Arg Ala Lys Tyr Gly Gly Gln Gln 750

GTG TAA   GGTTCATCGG AAGCATTTTT TTTCCTCTCT TTTCCCTTTC AATCTTCACA TGATAAAAGT GTATCAATTA ATATATGAAT TTTACTCTGA TATAAAAACA +100
Val Stop

AGTTTAAAAA ACAATTCTCC CGCTAACACG TTTATACCCT TTCTTTTAGT TGAGTAATTG TGTAGTGTTA AGCTGAGAAT ATTCATGATT TGATTTCTGA TAAATAACTT +210

CTGCTTTATA ATAACAAATG TTTAACGTGT TAGCGCCATA ATTATATAGA ATTGAGGGAA AAAGAAACTG AAAACCAAAA ATACCCTGCG TCGCTTGAAA TGAGAACAGC +320
```

FIG. 1. Sequence of the *MES1* gene and adjacent regions. This is the complete sequence of a region encompassing the *MES* structural gene and 320 nucleotides of flanking DNA in each direction. Only the mRNA identical strand (i.e., plus strand) of the gene is shown. The second base of stop codons is marked by an asterisk, thus indicating the reading frame. Numbers above the sequence are those of nucleotides (position +1 is assigned to A of the first ATG); those below the sequence correspond to amino acid positions (starting from the first putative translated methionine). Dyad symmetries are indicated by lines above the sequence.

Table 1. Amino acid compositions of methionyl-tRNA synthetase as calculated from amino acid analyses and deduced from DNA sequence

Amino acid	Composition*	
	From analyses	From DNA sequence
Asx	89	99
Thr	29	31
Ser	46	54
Glx	87	78
Pro	34	34
Gly	37	37
Ala	45	48
Cys†	9	10
Val	37	46
Met†	9	10
Ile	38	41
Leu	72	75
Tyr	28	35
Phe	28	32
His	20	22
Lys	56	62
Trp‡	13	13
Arg	24	24
Total	701	751

* Numbers correspond to residues per mol on the basis of $M_r = 80,000$.

† Cysteine and methionine were determined as cysteic acid and methionine sulfone, respectively, after performic acid oxidation of the protein.

‡ Tryptophan was determined after hydrolysis with methanesulfonic acid.

fidence in the predicted polypeptide primary structure and minimizing the risks of error.

Sequence Homology Between the *E. Coli* and Yeast Methionyl-tRNA Synthetases. Comparison of the two protein sequences, shown in Fig. 2, reveals an extensive homology (44%) between residue block 192–594 of the yeast enzyme and the NH_2-terminal 401 amino acids of the bacterial methionyl-tRNA synthetase. The alignment of similar residues has been maximized by introducing small deletions (between one and six amino acids) into both sequences, and the homology includes 30% amino acid identities and 14% conservative replacements. There appears to be no further significant homology outside this region, although the COOH-terminal 150 amino acids of the *E. coli* sequence are not yet available for comparison. Homology at the level of nucleotide sequence is much less apparent and this presumably reflects codon degeneracy.

The NH_2-terminal 300 residues of the *E. coli* enzyme fold into a tertiary conformation of alternating β-sheet, loop, and α-helix with five parallel β-strands in a Rossman fold topology (10). This structure is also found within the dehydrogenase and kinase enzyme families, where it functions as a nucleotide-binding domain (11). All enzymes that are topologically equivalent to liver alcohol dehydrogenase have a glycine residue at the end of the loop between the first β-strand and the subsequent α-helix (12). This residue (glycine-23) is also present in the *E. coli* methionyl-tRNA synthetase and is conserved (glycine-214) in the yeast sequence. Coupled with the fact that the homologous region encompasses the entire Rossmann fold structure, these findings provide strong evidence that the yeast methionyl-tRNA

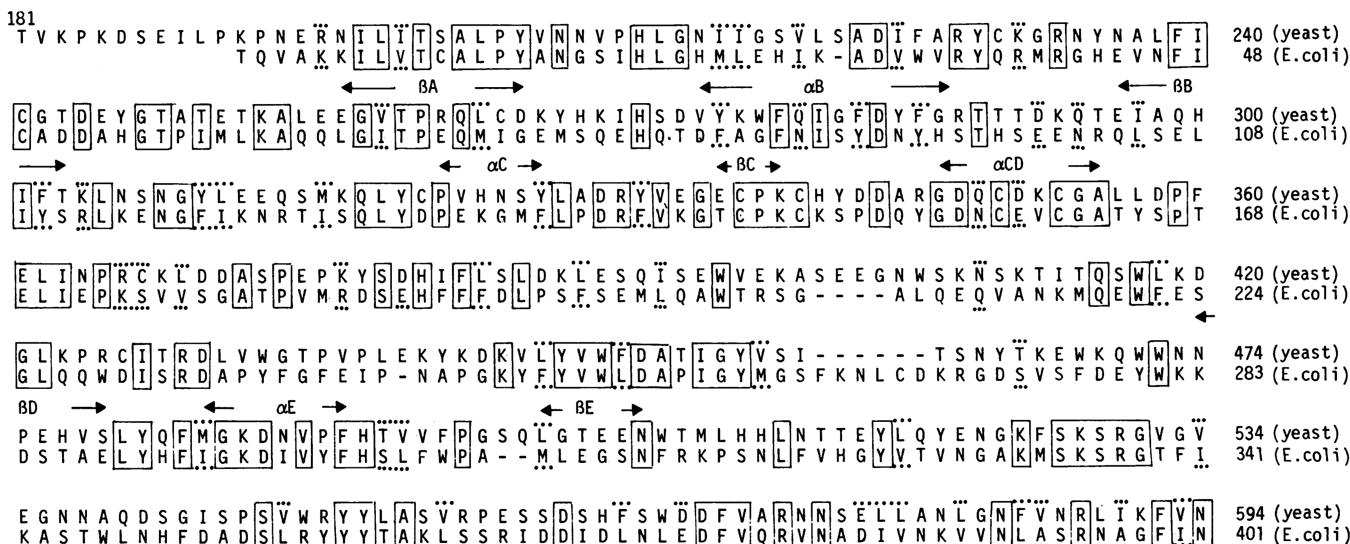

FIG. 2. Sequence homologies between an active tryptic fragment of the *E. coli* methionyl-tRNA synthetase and the entire yeast cytoplasmic enzyme. In both cases the amino acid sequences were deduced from the respective cloned genes with corroborative protein chemical data available for the bacterial methionyl-tRNA synthetase only. Deletions have been introduced at positions 31, 209–212, 244, and 308–309 in the bacterial enzyme sequence and 461–466 in the yeast enzyme sequence in order to maximize the alignment. Amino acids that are identical in both sequences are boxed; those that are similar are underlined. The secondary structural features of the Rossman fold (β-sheets and α-helices) are also indicated on the *E. coli* sequence and were kindly provided by J. L. Risler and C. Zelwer. For clarity, the yeast sequence is limited to residues 181–594 and the *E. coli* sequence, to residues 1–401.

synthetase has many tertiary structural homologies with its prokaryotic counterpart.

An increasing number of primary structures of aminoacyl-tRNA synthetases are now available for a search of sequence homologies (1–5) but we will limit our comparison (Fig. 3) to those for which similar tertiary structures have been obtained. In this respect, x-ray crystallographic studies of tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* have shown that the NH_2-terminal part of the molecule folds into a Rossman configuration (14). Barker and Winter (13) found little sequence homology between the two bacterial enzymes (*E. coli* methionyl- and *B. stearothermophilus* tyrosyl-tRNA synthetases), although they could identify a cysteine and two histidine residues that occupy homologous positions in the two tertiary structures. These residues lie close to the COOH terminus of the β-sheet that has been shown to bind the tyrosyl adenylate (15), and we can deduce by analogy that the equivalent domain of the *E. coli* methionyl-tRNA synthetase is responsible for methionine activation. In the yeast sequence, of the three potentially catalytic side chains only the histidine at position 212 is conserved.

Therefore, although the topology of this region of the active site is probably essential for adenylate formation, we can deduce little about the identity of the key catalytic residues. However, recent studies (16) have emphasized the essential role of cysteine-35 in the binding of ATP by tyrosyl-tRNA synthetase.

Finally, we would like to point out that Putney *et al.* (2) have proposed a structural model for *E. coli* alanyl-tRNA synthetase in which the NH_2-terminal domain (residues 1–400) contains the catalytic site where the alanyl-adenylate is formed whereas the COOH-terminal part is involved in the transfer step and subunit association. In addition, the enzyme contains clusters of arginine residues and mixed-charge residues distributed along the sequence. Such charged areas could interact with the cognate tRNA. This model does not fit entirely with the yeast methionyl-tRNA synthetase in which the putative nucleotide-binding domain corresponds to the middle of the molecule (between residues 192 and 594). But it is worth noting that mixed-charge clusters can be found in the NH_2- and COOH-terminal parts (residues 8–12, 183–186, 603–608, 635–637, 709–713, and 733–745).

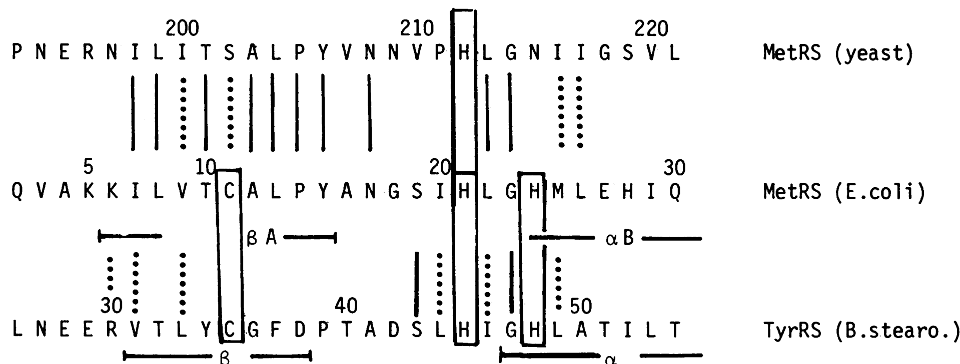

FIG. 3. Sequence homologies between yeast and *E. coli* methionyl-tRNA synthetases and *Bacillus stearothermophilus* tyrosyl-tRNA synthetase. The homologous cysteine and histidine residues previously detected by Barker and Winter (13) in the two bacterial enzymes are boxed. Vertical lines indicate homologies between the three enzymes (solid = identity; dotted = conservative replacement). Secondary structure assignment for tyrosyl-tRNA synthetase is taken from Bhat *et al.* (14).

Internal Sequence Homology. Extensive sequence repeats were found by tryptic mapping in *E. coli* methionyl-tRNA synthetase, and similar experiments performed on other aminoacyl-tRNA synthetases from various sources led to the widespread view that those enzymes with subunit sizes ranging from M_r 62,000 to 130,000 could contain such long stretches of repeated sequences in polypeptide chain (for a review, see ref. 17). No obvious internal redundancy was found in the predicted sequence of yeast methionyl-tRNA synthetase and the same is true in *E. coli* alanyl-, glutaminyl-, and methionyl-tRNA synthetase (2, 4, 5). Therefore, all these results cast serious doubts on the validity of the above hypothesis.

To conclude, we emphasize that transformation to prototrophy of our yeast K_m mutant could be due to the cloning of either cytoplasmic or mitochondrial methionyl-tRNA synthetase genes because both belong to the nuclear genome and both enzymes aminoacylate cytoplasmic tRNA (18). However, the two enzymes can be distinguished on the basis of two criteria (18): (*i*) their chromatographic behavior on hydroxyapatite columns, the mitochondrial enzyme being eluted at 0.1 M phosphate and the cytoplasmic one at 0.2 M; and (*ii*) the specificity of aminoacylation. Indeed, mitochondrial methionyl-tRNA synthetase aminoacylates both yeast cytoplasmic and *E. coli* tRNA at similar rates whereas its cytoplasmic counterpart aminoacylates *E. coli* tRNA at a rate 1/7th that of yeast cytoplasmic tRNA (18). We recently have found that antibodies raised against the purified gene product do not inhibit the mitochondrial methionyl-tRNA synthetase. These differences show that our cloned gene is indeed that of the cytoplasmic enzyme.

The authors thank G. de Marcillac for his help with sequence computer analysis, Dr. M. Yaniv who introduced F.F. to Sanger's sequence analysis method, and Dr. D. Barker for preprints of his articles before publication. The skillful technical assistance of G. Nussbaum is gratefully acknowledged. This work was supported by grants from the Centre National de la Recherche Scientifique. P.W. was the recipient of a fellowship from the Délégation Générale à la Recherche Scientifique et Technique.

1. Winter, G. P. & Hartley, B. S. (1977) *FEBS Lett.* **80,** 340–342.
2. Putney, S. D., Royal, N. J., Neuman de Vegvar, H., Herlihy, W. C., Biemann, K. & Schimmel, P. (1981) *Science* **213,** 1497–1501.
3. Hall, C. V., Van Cleemput, M., Muench, K. H. & Yanofsky, C. (1982) *J. Biol. Chem.* **257,** 6132–6136.
4. Hoben, P., Royal, N., Cheung, A., Yamao, F., Biemann, K. & Söll, D. (1982) *J. Biol. Chem.* **257,** 11644–11650.
5. Barker, D. G., Ebel, J. P., Jakes, R. & Bruton, C. J. (1982) *Eur. J. Biochem.* **127,** 449–457.
6. Fasiolo, F., Bonnet, J. & Lacroute, F. (1980) *J. Biol. Chem.* **256,** 2324–2328.
7. Sanger, F., Nickless, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
8. Messing, J., Crea, R. & Seeburg, P. H. (1981) *Nucleic Acids Res.* **9,** 309–321.
9. Kozak, M. (1978) *Cell* **15,** 1109–1123.
10. Risler, J. L., Zelwer, C. & Brunie, S. (1981) *Nature (London)* **292,** 384–386.
11. Rossman, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. (1975) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), Vol. 11, 3rd Ed., pp. 61–102.
12. Rossman, M. G. & Argos, P. (1977) *J. Mol. Biol.* **109,** 99–129.
13. Barker, D. G. & Winter, G. P. (1982) *FEBS Lett.* **145,** 191–193.
14. Bhat, T. N., Blow, D. M., Brich, P. & Nyborg, J. (1982) *J. Mol. Biol.* **158,** 699–709.
15. Rubin, J. & Blow, D. M. (1981) *J. Mol. Biol.* **145,** 489–500.
16. Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M. & Smith, M. (1982) *Nature (London)* **299,** 756–758.
17. Ofengand, J. (1978) in *Molecular Mechanisms of Protein Biosynthesis*, eds. Weissbach, H. & Petska, S. (Academic, New York), pp. 8–79.
18. Schneller, J. M., Schneider, C. & Stahl, A. J. C. (1978) *Biochem. Biophys. Res. Commun.* **85,** 1392–1399.