Structure of the Yeast Valyl-tRNA Synthetase Gene (VASI) and the Homology of Its Translated Amino Acid Sequence with Escherichia coli Isoleucyl-tRNA Synthetase*

Xavier Jordana, Bruno Chatton, Maria Paz-Weissbaar§, Jean-Marie Buhler¶, Fritz Cramer†, Jean Pierre Ebel, and Franco Fasiolo**

From the Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, 67084 Strasbourg, Cedex, France and the Service de Biochimie, Saclay, F-91191 Gif-sюр-Yvette, France

The VASI gene encoding the valyl-tRNA synthetase from yeast was isolated and sequenced. The gene-derived amino acid sequence of yeast valyl-tRNA synthetase was found to be 23% homologous to the Escherichia coli isoleucyl-tRNA synthetase. This is the highest level of homology reported so far between two distinct aminoacyl-tRNA synthetases and is indicative of an evolutionary relationship between these two molecules. Within these homologous sequences, two functional regions could be recognized: the HIGH region which forms part of the binding site of ATP and the KMSKS region which is recognized as the consensus sequence for the binding of the 3'-end of tRNA (Hountondji, C., Dessen, Ph., and Blanquet, S. (1986) Biochimie (Paris) 68, 1071–1078). Secondary structure predictions as well as the presence of both HIGH and KMSKS regions, delineating the nucleotide-binding domain and the COOH-terminal helical domain in aminoacyl-tRNA synthetases of known three-dimensional structure, suggest that the yeast valyl-tRNA synthetase could be folded into three domains: an NH₂-terminal α-helical region followed by a nucleotide-binding topology and a COOH-terminal domain composed of α-helices which probably carries major sites in tRNA binding.

The aminoacyl-tRNA synthetases are a vastly divergent family of enzymes differing in size and subunit structure but catalyzing the same reaction, the formation of an aminoacyl-tRNA, specific for both the amino acid and the tRNA. The mechanism of the aminoacylation involves the initial rapid formation of an aminoacyladenylate complex followed by the transfer of the aminoacyl moiety to the tRNA. Valyl-tRNA synthetase from yeast is a monomer of M₄, 120,000 (Kern et al., 1975) and belongs, together with leucyl- and isoleucyl-tRNA synthetases, to the class of enzymes having the largest polypeptide chain. Activation of a single amino acid by the aminoacyl-tRNA synthetase is, in most cases, very specific. However, valyl- and isoleucyl-tRNA synthetases do not discriminate between closely related amino acids in the adenylate formation step. In neither of these cases, however, is the misactivated amino acid used to form a stable aminoacyl-tRNA.

The mechanism of rejection is designated as a proof-reading or editing mechanism. The isoleucyl- and valyl-tRNA synthetases are known to hydrolyze the misacylated valyl and threonyl adenylates, respectively (Baldwin and Berg, 1966; Ferstl and Kaether, 1976; Igloi et al., 1977). Knowledge of their structure should be useful in defining structural elements involved in catalysis and/or specificity. The entire primary structure of Escherichia coli isoleucyl-tRNA synthetase has been reported (Webster et al., 1984). We present here the isolation and sequence of the VASI Saccharomyces cerevisiae gene coding for valyl-tRNA synthetase. Comparison of the translated amino acid sequence with that of isoleucyl-tRNA synthetase from E. coli shows the strongest homology ever reported for two distinct aminoacyl-tRNA synthetases.

MATERIALS AND METHODS

Yeast, Bacteria, Plasmids, Gene Libraries, and Growth Media—The yeast genomic bank from S. cerevisiae strain X 2180 in phage λgt11 and the host strain Y 1090 (Young and Davis 1983a, 1983b) were kindly provided by Dr. R. Young (Whitehead, MIT). The yeast genomic bank from S. cerevisiae strain FL100 in the plasmid vector pFL1 (Chevallier et al., 1980) was a gift from Dr. F. Lacroute (IBMC, Strasbourg, France). The strain FL11.1 (mes1,ura3) was the recipient for yeast transformation (Fasiolo et al., 1981). Parental and transformed yeast strains were grown on YNB (0.67% yeast nitrogen base without amino acids, 2% glucose) supplemented with 100 µl/ml methionine. Transformations of yeast and E. coli and preparation of nucleic acids were done using standard procedures.

Enzymes and Reagents—Restriction endonucleases, T4 DNA ligase, and E. coli DNA polymerase I (Klenow fragment) were purchased from Boehringer Mannheim. [α⁻³²P]dATP, α⁻²⁵S-labeled dATP, and 121 were purchased from New England Nuclear.

Antibody Preparation and Plaque Screening—Homogeneous yeast valyl-tRNA synthetase was prepared in our laboratory by Drs. D. Kern and R. Giege. Rabbits were immunized at 15-day intervals by three subcutaneous injections of 500 µg of enzyme dissolved in 500 µl of 10 mM potassium phosphate buffer (pH 7.4), 150 mM NaCl and emulsified in 500 µl of complete Freund’s adjuvant. One week after the last injection, the rabbits were bled, and the immunoglobulin fraction was purified from the serum by ammonium sulfate precipitation and DEAE-Sephadex chromatography. Purified antibodies were prepared by chromatography on valyl-tRNA synthetase bound to succinylaminomethyl-Sepharose 4B. Ten nmol of enzyme were coupled to 5 ml of packed gel with N-cyclohexyl-N'-(N-methylmorpholinophenyl)carbodiimide p-toluenesulfonate.

7189
Screening of the λgt11 genomic library was carried out essentially as described by Young and Davis (1983b) using affinity-purified antibodies at a concentration of 5-10 μg/ml and 35S-protein A (50 μCi/μg) at 1 μCi/ml. Positive plaques were purified by four additional cycles of screening.

Hybridization Procedures—DNA probes were purified by gel electrophoresis or sucrose gradient centrifugation from phage λgt11 or recombinant plasmids digested with the appropriate restriction enzymes. They were labeled by nick translation as described by Maniatis et al. (1982). DNA probes cloned in M13 phage were labeled by chain extension using the Klenow fragment of E. coli DNA polymerase I and [32P]dATP. The yeast genomic bank in vector pFL1 was screened by the high density colony-screening procedure described by Hanahan and Meselson (1983). Positives clones were purified by two additional cycles of screening. Southern blot hybridizations were carried out according to the procedures described by Maniatis et al. (1982).

Determination of Enzymatic Activities—Cytoplasmic valyl-tRNA synthetase was tested in crude extracts obtained by mechanical breakage with glass beads of exponentially growing cells. Protein concentration was estimated according to Bradford (1976).

The enzyme was tested using unfractionated yeast cytoplasmic tRNA isolated under the following conditions: 144 mM Tris-HCl (pH 7.8), 5 mM dithiothreitol, 10 mM ATP, 2 mM MgCl₂, 0.1 mM [35S]valine (25,000 cpm/nmol), 6 mg/ml yeast tRNA, and various amounts of crude extracts. The reaction mixture was 200 μl; and at various time intervals, 40-μl aliquots were spotted onto Whatman paper discs and quenched with 5% trichloroacetic. The precipitated aminoacylated tRNA was subjected to scintillation counting.

Western Blot—Protein samples were run on 10% polyacrylamide gels in the presence of 0.1% sodium dodecyl sulfate (Laemmli, 1970). Conditions for the transfer of proteins to nitrocellulose membranes were as described in the Schleicher & Schuell manual (No. 2). The protein band corresponding to valyl-tRNA synthetase was detected as described above using affinity-purified antibodies (5-10 μg/ml and 35S-protein A (0.1 μCi/ml).

DNA Sequence Analysis—The dideoxy-DNA sequencing method of Sanger et al. (1977) was used. EcoRI and SalI digests of pVASI recombinant generated fragments of 1.6, 1.2, and 1.3 kb, respectively. These DNA fragments were isolated and digested with AluI, HaeIII, TagI, and Sau3A. The resulting subfragments as well as the original fragments were cloned into suitable M13mp8 and M13mp9 vectors (Vieira and Messing, 1982).

Computer Analysis of Amino Acid Sequences—Amino acid sequences were analyzed with programs of the University of Wisconsin Genetics Computer Group edited by Devereux and Haebeli to locate sequences patterns: "Best fit" to align two sequences; "Gap" to find the optimal alignment for two sequences by adding gaps in either one to achieve the maximum number of matches; "Dotplot" and "Peptplot" to visualize the homology between two sequences; and "Choufas" to perform prediction of secondary structures.

RESULTS

Cloning of the VASI Gene—We have screened a yeast DNA library using the expression vector λgt11 which contains random genomic fragments in the unique EcoRI site (Young and Davis, 1983a, 1983b). Ten putative positive clones were obtained and further purified by three successive rounds of antibody screening at low plaque density after which only one clone remained positive. Yeast DNA inserted into the λgt11 recombinant is 2.5 kb, whereas the minimum expected length of the message for a protein of Mr 120,000 (Kern et al., 1975) is about 3.5 kb. In order to isolate the complete gene coding for valyl-tRNA synthetase, we have screened the pFL1 yeast DNA library (Chevallier et al., 1980) using the yeast EcoRI fragment from the λgt11 recombinant as hybridization probe. Only three clones (pVASI-1, -2, and -3) were purified, and their overlapping inserts were mapped with a number of restriction enzymes. Southern blot hybridization analysis of yeast nuclear DNA gave an identical genomic map for the two EcoRI and HindIII sites (Fig. 1).

To demonstrate that the cloned gene codes for valyl-tRNA synthetase, we expressed the various clones in yeast to give catalytically active valyl-tRNA, synthetase. The activity in the crude extracts of the yeast transformants (pVASI-1 and -2) was approximatively 10 times higher than the basal level of enzyme in the recipient strain. In order to verify that the activity was associated with a full-length protein in the overproducing strains, proteins from a crude cytoplasmic extract were separated by electrophoresis on sodium dodecyl sulfate-polyacrylamide gels and transferred to nitrocellulose, and valyl-tRNA synthetase was detected using the specific cytoplasmic valyl-tRNA synthetase antibodies and 35S-labeled protein A. The results of the Western blot analyses are shown in Fig. 2. A protein band which co-migrated with the purified cytoplasmic valyl-tRNA synthetase was detected in the crude extract of the recipient strain (lane 2). The concentration of this protein was increased (lanes 3 and 4) in yeast transformants harboring the VASI gene on a multicopy plasmid (pVASI-1 and -2). The level of valyl-tRNA synthetase in the transformant corresponding to clone pVASI-3 was again similar to the basal level of the recipient strain and was probably due to lack of the 5'-upstream promoter sequences.

Determination of the Nucleotide Sequence of the VASI Gene—We have determined 80% of the entire sequence on both strands, and on one strand, the remaining 20%. All restriction endonuclease sites used for generating M13 clones were overlapped. This strategy enabled us to localize a 78-base pair EcoRI fragment between the large 1.6- and 1.2-kb EcoRI subfragments. A long open reading frame of 3,312 nucleotides was found only on one strand (Fig. 3). The trans-

---

1 Dereveux, J., and Haebeli, P. (1983) Program Library of the University of Wisconsin Genetics Computer Group, Madison, WI.

2 The abbreviation used is: kb, kilobase.
lated amino acid sequence from the first in-phase methionine codon includes 1,104 amino acid residues, yielding a protein of M, 125,000, in good agreement with the M, measured for the purified protein. Attempts to define the NH2-terminal peptide of the protein were unsuccessful due to a blocked NH2 terminus.

**DISCUSSION**

Sequence homologies among different aminoacyl-tRNA synthetases, with the exception of those specific for the same amino acid in different organisms, are rare or nonexistent. Similarities of the three-dimensional level of these enzymes, however, are expected to be much greater due to structural constraints imposed on the binding of tRNA which probably shares the same tertiary conformation (Moras et al., 1980) and to the necessity of bringing the adenylate site close to the terminal adenosine site of tRNA in order to achieve the chemical acylation step. Since the ATP and the 3'-CCA arm of tRNA are common to all aminoacyl-tRNA synthetases, it is reasonable to assume that identical or at least functionally...
equivalent residues are present in many aminoacyl-tRNA synthetases. Hence, a comparison of primary sequences can be useful to identify important binding and/or catalytic residues. A classical example derives from a structural comparison of E. coli methionyl-tRNA synthetase and Bacillus stea- 
thermophilus tyrosyl-tRNA synthetase (Blow et al., 1983). The three-dimensional structures of both enzymes indicate folding of the NH₂-terminal regions into similar and character-
istic nucleotide-binding domains, although there is only a short stretch of amino acid sequence homology. In particular, 1 cysteine and 2 histidine residues occupy identical positions in the two tertiary structures (Barker and Winter, 1982; Blow et al., 1983). These conserved residues are involved in the binding and catalysis of adenylate as demonstrated by site-directed mutagenesis experiments (Winter et al., 1982; Leatherbarrow et al., 1985).

The NH₂-terminal region of E. coli isoleucyl-tRNA synthetase shows a sequence homology of 11 consecutive amino acids with the corresponding region of E. coli methionyl-tRNA synthetase which allowed the authors (Webster et al., 1984) to conclude that isoleucyl-tRNA synthetase is similarly folded in an alternating β/α structure. The perfect peptide match includes the consensus HIGH region involved in ATP binding (see below).

We have compared the deduced amino acid sequences of yeast valyl-tRNA synthetase and E. coli isoleucyl-tRNA synthetase. Residues 177-726 of the yeast enzyme could be aligned with residues 50-618 of the bacterial enzyme (Fig. 4). Fig. 4 shows four short perfect matches of 5-13 conserved residues at the following peptide positions in the yeast se-
quen:

196-200, 431-435, 564-570, and 700-712. The overall homology is 23%.

Two functional regions can be recognized within this homology; one at the ATP-binding site and the other at the probable CCA-binding site of tRNA.

Homology at the ATP-binding Site—Fig. 5 compares the homologies centered around the HIGH region of tyrosyl-
tRNA synthetase from B. stea- 
thermophilus, methionyl- and isoleucyl-tRNA synthetases from E. coli, and the methionyl- and valyl-tRNA synthetases from S. cerevisiae. The importance of the HIGH region in ATP binding and catalysis has become apparent from the studies of Fersht et al. (1984). This region is in the NH₂-terminal portion of the bacterial enzymes mentioned above, as is the case for the majority of prokaryotic tRNA synthetases; whereas we located the HIGH sequence in both yeast methionyl- and valyl-tRNA synthetases to approximately 200 amino acid residues from the NH₂-terminus. That this region corresponds to the ATP-binding site in yeast valyl-tRNA synthetase can be deduced by analogy with similar positions of the folded α/β topology in yeast methionyl-
tRNA synthetase (Walter et al., 1983). Thus, the two yeast enzymes bear an NH₂-terminal chain extension with respect to the mononucleotide binding fold. In yeast valyl-tRNA synthetase, this NH₂-terminal extension is mainly an α-helical region as deduced from predicted secondary structures.

Homology at the CCA-binding Site of tRNA—Covalent labeling of methionyl-tRNA synthetase from E. coli with 2,3'-diadialdehyde tRNA has led to the identification of a peptide encompassing Lys-335 (Hountondji and Blanquet, 1985). Although the exact position of this lysine residue in the crystal structure has not yet been located, it is part of the COOH-terminal helical domain of the synthetase (see Brumie et al.)

The sequence of the E. coli gene coding for valyl-tRNA synthetase was sent to us before publication by Dr. R. Leberman (LEBM, Grenoble, France) and co-workers. It turned out that the protein sequence was 45% homologous to the yeast enzyme and 23% homol-
ogous to the E. coli isoleucyl-tRNA synthetase.

---

\[\text{(Footnote: 2)}\]

The sequence of the E. coli gene coding for valyl-tRNA synthetase was sent to us before publication by Dr. R. Leberman (LEBM, Grenoble, France) and co-workers. It turned out that the protein sequence was 45% homologous to the yeast enzyme and 23% homologous to the E. coli isoleucyl-tRNA synthetase.

---

**Fig. 4. Homology between the amino acid sequences of yeast valyl-tRNA synthetase and E. coli isoleucyl-tRNA synthetase.** The comparisons shown in both A and B use programs from the University of Wisconsin Genetics Computer Group. The E. coli sequence is from Webster et al. (1984). The comparison in A uses the Dot Matrix program. Average score values were calculated for pairs of 25-amino acid segments using the mutation matrix of Staden (1982). If the average score value was equal to or greater than 25, a dot was printed at the corresponding position of the matrix. In B, the
Homology of Yeast Valyl- and E. coli Isoleucyl-tRNA Synthetases

Fig. 5. Alignment of the amino acid sequences from the HIGH regions. The numbering indicates the distance from the NH₂ terminus. The letters in parentheses indicate the reference of the sequence: a, Winter et al., 1983; b, Walter et al., 1983; c, Barker and Winter, 1982; d, Webster et al., 1984; and e, this work.

(a) TyrRSbs 33  LYGFDPTADSLHIGHLATI 52
(b) MetRS 9  VCALPYANGISHLGMEH 28
(c) MetRSsc 200  ITSAFLPYVNVPHGLNIGS 219
(d) IleRS 53  LHDPPYANGISHGSMNK 72
(e) ValRSsc 183  IAPPNTYGALHIGHATI 204

Fig. 6. Alignment of the amino acid sequences around the KMSKS regions. The origins of the sequences are indicated by the same nomenclature used in Fig. 5. The numbering indicates the distance from the NH₂ terminus. The references are as follows: a, Barker et al., 1982a; b, Barker et al., 1982b; c, Walter et al., 1983; d, Webster et al., 1984; and e, this work.

(a) TyrRS 223  TVPITKAGDKGFKGL - T 238
(b) MetRS 329  NGAKMSKSRGT - FIKAS 344
(c) MetRSsc 522  EGNKFSKSRGV 532
(d) IleRS 599  QGRKMSSKSLGTNYSPQD 615
(e) ValRSsc 709  QGRKMSKSLENVIDPLD 716

Acknowledgments—We thank Professor Y. Boulanger and Dr. P. Remy for critically reviewing this work and helpful comments and Dr. Ph. Walter for his help in the computer analysis. We acknowledge the skillful technical assistance of G. Nussbaum.
REFERENCES

Hountondji, C., Dessen, Ph., and Blanquet, S. (1986b) Biochimie (Paris) 68, 1071–1078