Villafranca, J. E., Howell, E. E., Voet, D. H., Strobel, M. J., Ogden, R. C., Abelson, J. N., & Kraut, J. (1983) *Science 222*, 782–788.
Volz, K. W., Matthews, D. A., Alden, R. A., Freer, S. T., Hansch, C., Kaufman, B. T., & Kraut, J. (1982) *J. Biol. Chem. 257*, 2528–2536.
Wallace, C. J. A. (1987) *J. Biol. Chem. 262*, 16767–16770.
Williams, J. W., Morrison, J. F., & Duggleby, R. G. (1979) *Biochemistry 18*, 2567–2573.
Zakrzewski, S. F., & Sansone, A. M. (1971) *Methods Enzymol. 18B*, 728–731.
Zoller, M. J., & Smith (1972) *Nucleic Acids Res. 10*, 6487.

# A Mammalian Tryptophanyl-tRNA Synthetase Shows Little Homology to Prokaryotic Synthetases but Near Identity with Mammalian Peptide Chain Release Factor[†,‡]

Maurice Garret,[§] Bertrand Pajot,[§] Véronique Trézéguet,[§] Julie Labouesse,[§] Michel Merle,[§] Jean-Claude Gandar,[§] Jean-Pierre Benedetto,[§] Marie-Line Sallafranque,[§] Jeanine Alterio,[∥] Maryse Gueguen,[§] Claude Sarger,[§] Bernard Labouesse,[§] and Jacques Bonnet*[,§]

*Institut de Biochimie Cellulaire et Neurochimie du CNRS, Université de Bordeaux II, 1 rue Camille Saint-Saëns, 33077 Bordeaux Cedex, France, and INSERM U 118, 29 rue Wilhem, Paris, France*

*Received December 3, 1990; Revised Manuscript Received May 3, 1991*

ABSTRACT: Determination of the amino acid sequence of beef pancreas tryptophanyl-tRNA synthetase was undertaken through both cDNA and direct peptide sequencing. A full-length cDNA clone containing a 475 amino acid open reading frame was obtained. The molecular mass of the corresponding peptide chain, 53 728 Da, was in agreement with that of beef tryptophanyl-tRNA synthetase, as determined by physicochemical methods (54 kDa). Expression of this clone in *Escherichia coli* led to tryptophanyl-tRNA synthetase activity in cell extracts. The open reading frame included two sequences analogous to the consensus sequences, HIGH and KMSKS, found in class I aminoacyl-tRNA synthetases. The homology with prokaryotic and yeast mitochondrial tryptophanyl-tRNA synthetases was low and was limited to the regions of the consensus sequences. However, a 90% homology was observed with the recently described rabbit peptide chain release factor (eRF) [Lee et al. (1990) *Proc. Natl. Acad. Sci. 87*, 3508–3512]. Such a strong homology may reveal a new group of genes deriving from a common ancestor, the products of which could be involved in tRNA aminoacylation (tryptophanyl-tRNA synthetase) or translation termination (eRF).

Besides their primary role in tRNA aminoacylation, aminoacyl-tRNA synthetases exhibit in vivo complementary functions such as splicing of mitochondrial RNAs in *Neurospora crassa* or yeast (Atkins & Lambovitz, 1982; Herbert et al., 1986) and regulation of translation initiation (Clemens, 1990). They also exhibit in vitro several catalytic activities which are the consequence of their ability to synthesize a highly reactive aminoacyl adenylate (Weiss et al., 1959; Tada & Tada, 1975; Plateau et al., 1981; Lorber et al., 1982; Andrews et al., 1985). Complementary functions or properties may be expected in eukaryotes, because cytoplasmic aminoacyl-tRNA synthetases are larger in size than their prokaryotic counterparts (Schimmel & Söll, 1979; Schimmel, 1987). For example, eight to ten eukaryotic aminoacyl-tRNA synthetases can be purified as high molecular weight multienzyme complexes (Dang et al., 1982; Cirakoglu & Waller, 1986) while prokaryotic synthetases cannot. Beef tryptophanyl-tRNA synthetase (WRS), however, has never been found in such complexes (Bec et al., 1989), although its molecular weight is higher than that of prokaryotic ones [54 kDa as compared to 37 kDa for the *Bacillus stearothermophilus* protein (Winter & Hartley, 1977)]. Beef WRS presents another noticeable feature: its concentration is 2 orders of magnitude larger in pancreas than in other organs (Sallafranque et al., 1986; Favorova et al., 1989), amounting to around 1% of the total protein concentration of this gland. This amount by far exceeds what is necessary for protein biosynthesis. Furthermore, in beef pancreas, its activity is much higher than that of the other aminoacyl-tRNA synthetases (Davie et al., 1956).

The catalytic properties of beef WRS have been extensively studied (Akhverdyan et al., 1977; Trézéguet et al., 1986; Merle et al., 1988) but have not been related to its molecular structure, which to date has not been described. Up to now, very few aminoacyl-tRNA synthetases of higher eukaryotes have been sequenced and the three-dimensional structure of none has been solved. Therefore, further insight into the molecular properties of these enzymes requires better knowledge of their structures. To obtain such information, the determination of the amino acid sequence of beef WRS was undertaken through both cDNA cloning and direct peptide sequencing. Very unexpectedly, this sequence turned out to be nearly identical with that of the eukaryotic peptide release factor (Lee et al., 1990).

## Materials and Methods

### *Materials*

Radionucleotides were from Amersham Corp. Restriction and modifying enzymes and chemicals were from BRL, Boehringer, Pierce, and Sigma. They were of the highest

[†] Supported by grants from the CNRS, the Université de Bordeaux II, and the Région Aquitaine.
[‡] The full cDNA sequence in this paper has been submitted to the EMBL Data Bank under Accession Number X 53918 and to GenBank under Accession Number J05334.
[§] Université de Bordeaux II.
[∥] INSERM U 118.

0006-2960/91/0430-7809$02.50/0 

purity grade available. TPCK–trypsin, chymotrypsin, and clostripain were from Worthington, yeast tRNA was from Boehringer, and [$^{14}$C]-L-tryptophan was from Amersham.

*Methods*

*Tryptophanyl-tRNA Synthetase.* Tryptophanyl-tRNA synthetase was isolated from beef pancreas and purified as in Mêrault et al. (1978).

*Peptide Isolation and Sequencing.* Tryptophanyl-tRNA synthetase was reduced by incubation in 0.1 M Tris-HCl, pH 8.5, 6 M guanidium chloride, and 0.1 mM EDTA in the presence of DTT[1] (16-fold excess over cysteinyl residues). After a 3-h incubation, carboxymethylation was performed by adding iodoacetic acid (stoichiometric to thiol groups), followed after 90 min by the addition of 0.6 M $\beta$-mercaptoethanol. The protein was then dialyzed against 0.1 M ammonium bicarbonate, pH 7.6, and submitted to enzymatic digestion. Trypsin was added at a ratio of 1:50 over tryptophanyl-tRNA synthetase (50 $\mu$M), and clostripain and chymotrypsin were added at a ratio of 1:100, for 3–6 h at 37 °C. Chemical cleavage with cyanogen bromide was performed as described in Reinbolt et al. (1983).

Peptides were separated with an HPLC apparatus (Waters Associates) using either Micropack TSK gel columns (2000SW and 3000SW, Varian), a C18 5 ODS column (Ultratech Sphere), or an Aquapore RP300 $\mu$ column. The eluant was 0.1 M ammonium bicarbonate, pH 7.6, for the TSK columns. For the reverse-phase columns, peptides were eluted with various gradients from 10 mM ammonium formate to 80% methanol in water. Amino acid sequences were determined with an Applied Biosystems Model 470A protein sequencing system.

*Polyclonal Antibodies.* Polyclonal antibodies directed against beef tryptophanyl-tRNA synthetase were prepared and characterized as described in Sallafranque et al. (1986).

*Construction of λgt11 and λgt10 Libraries.* Fresh pancreas chunks were frozen in an ethanol–dry ice bath, lyophilized, and then used to prepare total RNA by phenol extraction. The yield was 4 mg of RNA/g of dried organ. Poly(A)+ RNA was purified by chromatography of total RNA (10 mg) through an oligo(dT)–cellulose column (Pharmacia) in 20 mM Tris-HCl, pH 7.5, 0.5 M NaCl, 0.1% SDS, and 1 mM EDTA. Elution was performed with the same buffer, but NaCl was omitted. Poly(A)+ RNA was then fractionated by sedimentation on a 10–32% sucrose gradient in 20 mM Tris-HCl, pH 7.4, 5 mM EDTA, and 0.1% SDS, for 12 h at 34 000 rpm and 20 °C. RNA migrating with a sedimentation coefficient of 18 s or more was selected. Complementary DNA synthesis was performed from 3 $\mu$g of poly(A)+ RNA (Maniatis et al., 1982). Half of the cDNA was inserted into the λgt11 vector, and the other half into the λgt10 vector, to yield $4 \times 10^5$ and $5 \times 10^5$ independent recombinants, respectively. A λgt11 library was also constructed from beef retina (Alterio et al., 1988).

Standard protocols for screening and preparation of positive clones were used (Maniatis et al., 1982; Grossberger, 1987).

*Production of a Fusion Protein.* Cloning in the Bluescribe plasmid [Stratagene, according to Maniatis et al. (1982)] allowed us to express a $\beta$-galactosidase–WRS fusion protein, using the promoter included in the plasmid upstream to the sequence. For this purpose, we used a clone that spanned the entire coding sequence except that corresponding to the first 17 amino acids of the N-terminal part of the protein (clone 7-1; see Results).

The fusion protein was produced as follows: *Escherichia coli*, strain DH5 $\alpha$, containing the 7-1 clone was grown to mid log phase ($OD_{600}$: 0.2) and expression was induced by 10 mM IPTG. Cells were further grown until the stationnary phase ($OD_{600}$: 1). They were then harvested and rinsed with 1 mL of 100 mM Tris-HCl, pH 8, containing 10 mM $MgCl_2$, 1 mM EDTA, 5 mM $\beta$-mercaptoethanol, and 10 mM DTT. They were resuspended in 1 mL of the same buffer, and the mixture was ground in a 1.5-mL Eppendorf tube with 300 $\mu$L of glass beads (0.45 mm in diameter) by shaking 6 × 30 s at maximum speed on a bench Vortex. The cells debris and glass beads were pelleted, and the supernatant was collected. It was dialyzed against 2 × 1000 volumes of 100 mM Tris-HCl, pH 8, containing 10 mM $MgCl_2$, 0.1 mM EDTA, and 1 mM PMSF and then concentrated by dialysis against the same solution containing 50% glycerol. Protein concentration of this crude extract was determined according to Ehreshmann et al. (1974).

*Determination of Enzymatic Activities.* tRNA aminoacylation was carried according to Dorizzi et al. (1977), at 25 °C, using 500 $\mu$g of protein of the crude extract in buffer A containing 100 mM Tris-HCl, pH 8, 1 mM EDTA, 15 mM magnesium acetate, and 0.05 mg/mL BSA, in the presence of 10 mM ATP, 0.1 mM [$^{14}$C]-L-tryptophan (54 mCi/mmol), and 7 mg/mL total yeast tRNA. Total yeast tRNA contained 2.5% of $tRNA^{Trp}$. At various times, 35-$\mu$L aliquots from a 200-$\mu$L reaction mixture were spotted on a Whatman disk and the aminoacylation reaction was quenched with 5% TCA and 0.5% cold tryptophan. The radioactivity retained on the filters after decontamination was counted by using an ACS scintillation cocktail (Amersham).

*Western Blotting.* Crude extract (100 $\mu$g) was submitted to PAGE on a 10% polyacrylamide gel in the presence of 0.1% SDS (Laemmli, 1970). Transfer of the protein to a nitrocellulose membrane (Schleicher & Schuell BA 86) was carried out according to the manufacturer's instruction by using an LKB Transphor system. The protein bands corresponding to WRS were detected by using DEAE-Sephadex-purified anti-WRS rabbit antibodies previously exhausted of cross-reacting material other than beef WRS (see Results). They were revealed with secondary anti-rabbit antibodies conjugated to alkaline phosphatase. The chromogen was NBT–BCIP (Sigma), used as described by Sambrook et al. (1989). The $M_r$ of the fusion protein was determined by using the following protein markers: phosphorylase, bovine serum albumin, ovalbumin, trypsin inhibitor, and lysozyme.

*cDNA Sequencing.* Clones of interest were subcloned into the Bluescript KS+ vector (Stratagene) and sequenced in both directions according to the Sanger dideoxy chain-termination technique (Sanger et al., 1977) using the T7 DNA polymerase sequencing kit from Pharmacia. Storage and analysis of DNA sequences were done by using PC-Gene sequence analysis software (Genofit-Intelligenetics). The full cDNA sequence was submitted to the EMBL Data Bank.

*Northern Analysis.* Thirty micrograms of total RNA isolated from beef brain was separated by electrophoresis on a 1% agarose gel in 6% formaldehyde. RNA was then transferred onto a nylon sheet (Hybond, Amersham) and hybridized as described by Fourney et al. (1987) with the $^{32}$P-labeled WRS cDNA probe. After hybridization, washings of the blot were done under moderate stringency: two 10-min washes in 1 SSC and 0.1% SDS at room temperature, followed by two

[1] Abbreviations: DTT, dithiothreitol; IPTG, isopropyl thio-$\beta$-D-galactoside; NBT–BCIP, nitro blue tetrazolium chloride–5-bromo-4-chloro-3-indolyl phosphate; PMSF, phenylmethanesulfonyl fluoride; SSC, sodium saline citrate; SDS, sodium dodecyl sulfate; TBS, Tris-buffered saline.

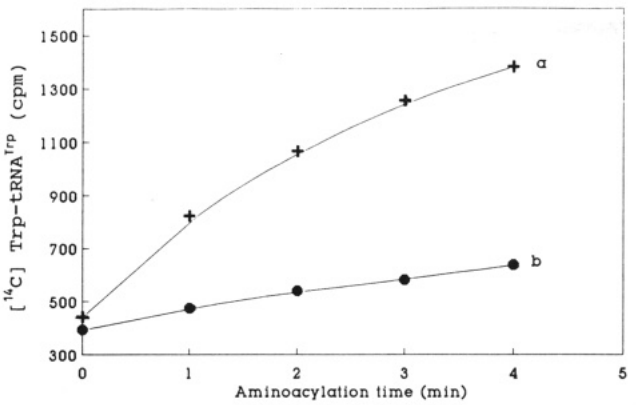


FIGURE 1: Aminoacylation activity of the expression product of the sequenced WRS clone. Clone 7-1 was expressed in *E. coli* either in the presence (curve a) or in the absence (curve b) of 10 mM IPTG. The cell extract was preincubated for 1 min in buffer A with 0.1 mM [$^{14}$C]-L-tryptophan and 7 mg/mL total yeast tRNA (corresponding to 7 $\mu$M tRNA$^{Trp}$ in the final assay). The aminoacylation reaction was then started by addition of ATP (10 mM in the final assay).

20-min washes in 0.1 SSC and 0.1% SDS at 60 °C.

## RESULTS

*Isolation of Beef Tryptophanyl-tRNA Synthetase cDNA.* A λgt11 library was constructed from beef pancreas and probed with antibodies directed against beef tryptophanyl-tRNA synthetase. These antibodies have been previously characterized (Sallafranque et al., 1986). In particular, it was shown that WRS preparations did not contain any contaminant that could be detected by these antibodies, having an $M_r$ different from that of WRS. From a total of $2 \times 10^5$ cDNA clones screened with these antibodies, seven immunoreactive clones were isolated. One of these coded for a 116-kDa fusion protein, the sequence of which revealed that it was a false positive, when compared to the WRS peptide sequences obtained independently. The other clones were shown to contain cDNAs coding for the same protein. Hence, this fusion protein was used to exhaust the anti-WRS polyclonal antibodies of the immunoglobulins directed against the foreign material. A total amount of $2 \times 10^5$ plaque-forming units of the first false positive clone were plated and transferred onto nitrocellulose filters. The filters were then incubated at room temperature in a solution containing 200 mg/mL antibodies and 0.1% ovalbumin in TBS (25 mM Tris-HCl, pH 8.0, 140 mM NaCl, 2.5 mM KCl) at room temperature. The antibodies were recovered after overnight incubation and were used for a second screening of the λgt11 library. From $4 \times 10^5$ clones probed with the new antibody preparation, six immunoreactive clones were isolated, five of which were false positive. The sixth one contained a short open reading frame coding for MTPRKLSYDFQ-STOP. This sequence included both the known C-terminus of WRS, [K,S,Y]DFQ-COOH (Gros et al., 1972), and the end of one of the peptides isolated from WRS, -MTPR. Considering these results, a λgt10 library was probed with this latter clone, and seven clones were identified. The sequence of the largest one (clone 7-1, 1.6 kb) exhibited a truncated open reading frame encoding a peptide chain of 458 amino acids. Its sequence was consistent with that of all of the peptides isolated from beef WRS. To obtain the full-length cDNA, a cDNA library constructed from beef retina was screened with clone 7-1 and a 17-mer oligonucleotide complementary to the 5′ end of 7-1. Clones positive with both probes were plaque-purified. Nine clones were obtained. After restriction analysis, the largest one (Trs2) was subcloned into the Bluescript plasmid KS+ and sequenced.

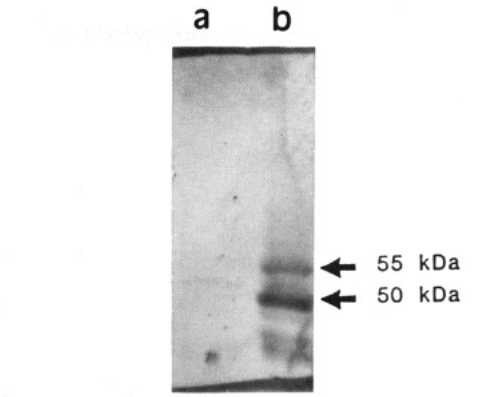


FIGURE 2: Western blot of the protein extracts from bacterial cells expressing clone 7-1. After gel electrophoresis and transfer to a nitrocellulose sheet, the proteins were revealed with purified rabbit anti-WRS antibodies and anti-rabbit immunoglobulin antibodies conjugated to alkaline phosphatase. Lane a: Cells incubated in the absence of IPTG. Lane b: Cells incubated in the presence of IPTG.

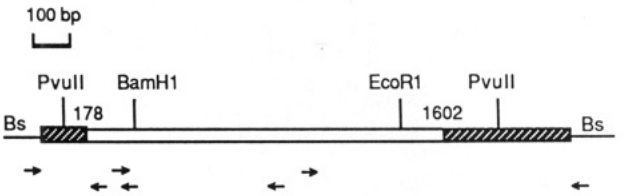


FIGURE 3: Sequencing strategy of clone Trs2. The arrows represent the oligonucleotides used for sequencing. The open box corresponds to the open reading frame coding for beef WRS.

RNA analysis by Northern blotting detected a single transcript of around 3 kb using Trs2 DNA as a probe (data not shown). WRS gene size was estimated to be around 23 kbp by Southern blotting of genomic DNA from total beef blood (data not shown).

*Expression of the cDNA.* After expression of the 7-1 clone in *E. coli*, the aminoacylation activity of the bacterial extract was measured with [$^{14}$C]-L-tryptophan, using total yeast tRNA. The tRNA$^{Trp}$ accepting concentration brought into the incubation mixture by total yeast tRNA was 7 $\mu$M. Figure 1, curve a, shows that tRNA was tryptophanylated when the cells were grown in the presence of IPTG. Control experiments with cells grown in the absence of IPTG showed much slower incorporation of [$^{14}$C]-L-tryptophan (Figure 1, curve b). It is known that the rate of *E. coli* tRNA$^{Trp}$ aminoacylation by beef WRS is at least 4 orders of magnitude slower than that of yeast tRNA$^{Trp}$ (Dorizzi et al., 1977). The IPTG-induced expression of the yeast tRNA aminoacylating activity could be therefore unambiguously assigned to the fusion protein. The *E. coli* extract transfected with clone 7.1 and expressed in the presence or in the absence of IPTG was submitted to electrophoresis under denaturing conditions and revealed by Western blotting with anti-WRS antibodies. These antibodies were exhausted of cross-reacting material by incubation, first, with the false positive 116-kDa fusion protein resulting from the first cloning and then with an *E. coli* cell extract. Two bands were observed, of mass 55 and 50 kDa (Figure 2). The expected molecular mass for the fusion protein expressed by the Bluescribe vector would be 56 kDa.

*cDNA and Amino Acid Sequences of Beef Tryptophanyl-tRNA Synthetase.* Figure 3 shows the strategy followed to sequence clone Trs2, and Figure 4 gives the resulting nucleotide sequence with the corresponding amino acids. The peptide chain would have a molecular mass of 53 728 Da, in close agreement with that of WRS subunits, as determined by physicochemical methods (54 kDa; Gros et al., 1972). Besides

```
GAGGTGAGTAGGCTGCTCTTTTATTATTGAGACATGAATCGTGTCTCAACAGAAACTAAA     60
TGAACATTTCAGAATAAATTATTGCTGTATGTAAACTGTTACTGAAAAAAAAAAAAAAAA   120
AGGACCACTGAGCAAATCTGCCAGCACCCCAACCAGCTGTGAACATCTTCCCAGAAGATG   180
                                                          M
GCAGACATGTCCAATGGCGAGCAGGGCTGTGGGTCCCCCCTGGAGCTGTTCCACAGCATC   240
 A  D  M  S  N  G  E  Q  G  C  G  S  P  L  E  L  F  H  S  I
GCTGCTCAGGGGGAGCTCGTAAGGGACCTCAAAGCCAGAAACGCAGCCAAGGATGAAATT   300
 A  A  Q  G  E  L  V  R  D  L  K  A  R  N  A  A  K  D  E  I
GATTCTGCAGTGAAGATGTTGTTGTCCTTAAAAACGAGCTACAAAGCTGCCACAGGGGAG   360
 D  S  A  V  K  M  L  L  S  L  K  T  S  Y  K  A  A  T  G  E
GATTACAAGGTGGACTGTCCTCCAGGGGACCCGGCGCCTGAGAGTGGTGAGGGCCTGGAT   420
 D  Y  K  V  D  C  P  P  G  D  P  A  P  E  S  G  E  G  L  D
GCCACCGAAGCAGACGAGGACTTTGTGGATCCCTGGACAGTGCAGACAAGCAGCGCCAAA   480
 A  T  E  A  D  E  D  F  V  D  P  W  T  V  Q  T  S  S  A  K
GGCATTGACTACGACAAGCTCATTGTTCGATTTGGAAGCAGTAAAATTGACAAGGAGCTG   540
 G  I  D  Y  D  K  L  I  V  R  F  G  S  S  K  I  D  K  E  L
GTAAACCGAATAGAGAGAGCCACGGGCCAGAGACCACACCGCTTCCTGCGCAGGGGAATC   600
 V  N  R  I  E  R  A  T  G  Q  R  P  H  R  F  L  R  R  G  I
TTCTTCTCACACAGAGACATGCATCAGATTCTGGATGCCTATGAAAACAAGAAGCCGTTC   660
 F  F  S  H  R  D  M  H  Q  I  L  D  A  Y  E  N  K  K  P  F
TATCTCTACACGGGCAGGGGCCCCTCTTCTGAAGCCATGCATGTGGGTCACCTCATCCCA   720
 Y  L  Y  T  G  R  G  P  S  S  E  A  M  H  V  G  H  L  I  P
TTCATCTTCACCAAGTGGCTGCAGGATGTGTTCAACGTCCCCTTGGTCATCCAGATGACT   780
 F  I  F  T  K  W  L  Q  D  V  F  N  V  P  L  V  I  Q  M  T
GATGACGAGAAGTACCTGTGGAAGGACCTGACCCTGGATCAGGCCTATGGCTACGCTGTG   840
 D  D  E  K  Y  L  W  K  D  L  T  L  D  Q  A  Y  G  Y  A  V
GAGAACGCCAAGGATATCACGTGCGGCTTTGACATCAACAAAACGTTCATCTTCTCTGAC   900
 E  N  A  K  D  I  T  C  G  F  D  I  N  K  T  F  I  F  S  D
CTTGACTACATGGGGATGAGCCCTGGCTTCTACAAGAACGTGGTGAAGATCCAGAAGCAC   960
 L  D  Y  M  G  M  S  P  G  F  Y  K  N  V  V  K  I  Q  K  H
GTCACCTTCAACCAAGTGAAAGGCATTTTCGGCTTCACTGACAGCGACTGCATTGGGAAG  1020
 V  T  F  N  Q  V  K  G  I  F  G  F  T  D  S  D  C  I  G  K
ATCAGTTTTCCTGCCATCCAGGCCGCTCCCTCCTTCAGCAACTCATTCCCTCAGATCTTC  1080
 I  S  F  P  A  I  Q  A  A  P  S  F  S  N  S  F  P  Q  I  F
CGAGACCGGACGGACGTCCAGTGCCTCATCCCGTGCGCCATCGACCAGGACCCGTACTTC  1140
 R  D  R  T  D  V  Q  C  L  I  P  C  A  I  D  Q  D  P  Y  F
AGGATGACCAGGGACGTGGCCCCCAGGATCGGCTACCCCAAGCCAGCCCTCTTGCACTCG  1200
 R  M  T  R  D  V  A  P  R  I  G  Y  P  K  P  A  L  L  H  S
ACCTTCTTCCCTGCCCTGCAGGGGGCCCAGACCAAGATGAGTGCCAGTGACCCCAACTCT  1260
 T  F  F  P  A  A  L  G  A  Q  T  K  M  S  A  S  D  P  N  S
TCCATCTTCCTCACGGACACGGCCAAGCAGATCAAGACCAAGGTCAACAAGCACGCCTTC  1320
 S  I  F  L  T  D  T  A  K  Q  I  K  T  K  V  N  K  H  A  F
TCCGGAGGCAGGGACACCGTTGAGGAGCACCGGCAGTTTGGGGGCAACTGTGACGTGGAC  1380
 S  G  G  R  D  T  V  E  E  H  R  Q  F  G  G  N  C  D  V  D
GTGTCCTTCATGTACCTGACCTTCTTCCTGGAGGATGATGACAAGCTGGAGCAGATCAGG  1440
 V  S  F  M  Y  L  T  F  F  L  E  D  D  D  K  L  E  Q  I  R
AGGGACTACACGAGCGGGGCCATGCTCACCGGCGAGCTCAAGAAGGAGCTCATCGAAGTC  1500
 R  D  Y  T  S  G  A  M  L  T  G  E  L  K  K  E  L  I  E  V
CTGCAGCCCTTGATTGCCGAGCACCAGGCCCGGCGCAAGGAGGTCACCGACGAGATCGTG  1560
 L  Q  P  L  I  A  E  H  Q  A  R  R  K  E  V  T  D  E  I  V
AAAGAATTCATGACCCCCCGGAAGCTGTCCTACGACTTTCAGTAACGCTCACTTCCTCTA  1620
 K  E  F  M  T  P  R  K  L  S  Y  D  F  Q  STOP
TGCTTACTCAGAGATGTAATTTACCCATAATCCCAACTCAGTCAAATCAGCACCCCCCTG  1680
CTGCTGGCCCTAGGCTCTATCAAATGGTAATTATTGGCCCAGGTCTGTAAGTTCTGTGTT  1740
TCTCTCAGCGCTGTTTCTTCCCTTGAATGTCATTCCGGCCCCTGTGGTAACTGGGTGCCG  1800
GGTAGTATCGTGTGGTCACACAGACCAAGTCCAGCTGGAGTTCCTCCCACAAAGCAGCCC  1860
CTGGGTGATGGTCCACCAACCAGCCTGGTTGTGGTCCACAGGCAACTCCCTGTTTGAAAG  1920
CATCCTGCGTCCAAACACATTCTTCCACCTGGCTGTATAAAAGGATATATAGTTCCATCC  1980
TATATGTACCTGGAATCAACATGGTGTTGCCATAAAACTTTTTTTTAAGTGCTTTGAAAC  2040
AGTAATGTAAGAAAGCTTAATGGGCACAAGAGGCGTTTCAGCCCATGCCCCCCTCCCATC  2160
TTATTGAAGACGAAATGATGCCAAAGGCTTTGCTAAACTTTAAATTCCAAGATCTCACAG  2160
```

FIGURE 4: Nucleotide and amino acid sequences of beef tryptophanyl-tRNA synthetase. Nucleotide sequence of clone Trs2 is given with the corresponding deduced amino acids of the largest open reading frame (the single-letter code is used). Underlined stretches of amino acids represent the sequences determined from peptides isolated after chemical or proteolytic cleavage of WRS (see Table I).

this cDNA sequence determination, 26 peptides obtained by chemical or proteolytic cleavages of WRS were isolated and submitted to Edman degradation (see Table I). Altogether, they represented more than 50% of the peptide chain, and their sequences were found to be in almost perfect agreement with those deduced from the cDNA (Figure 4). The protein sequence exhibited two stretches analogous to the two consensus sequences HIGH and KMSKS found in class I aminoacyl-tRNA synthetases (Burbaum et al., 1990): HVGH (176–179) and KMSAS (354–358). The presence of an alanine in the

```
Bf  MADMSNGEQGCGSPLELFHSIAAQGELVRDLKARNAAKDEIDSAVKMLLSLKTSYKAATGEDYK

Bf  VDCPPGDPAPESGEGLDATEADEDFVDPWTVQTSSAKGIDYDKLIVRFGSSKIDKELVNRIERA

                                        =   ---=  ==    -=--  =      =
Bf  TGQRPHRFLRRGIFFSHRDMHQILDAYENKKPFYLYTGRGPSSEAMHVGHLIPFIFTKW....
Ym  MSNWISTVQRA......DFKLNSSEALHSNAT.TVFSMIQPTGLCFHLGNYLGA.TRVWTDLC
Ec                                MTKP.IVFSGAQPSGE.LTIGNYMGALRQ.W....
Bst                               MTK...IFSGIQPSGV.ITIGNYIGARLQ.F....
Bsu                                MKQT.IFSGIQPSGS.VTLGNYIGAMKQ.F....

                 - =  -  =                 =     =  = =   = =   =   -
Bf  ............LQD..VFNVP..LVIQMTDDEKYLWKDLTLDQAYGYAVENAKD.IT..CGFD
Ym  ELKQPGQELIFGVADLHAITVPKPDGEMFRKFRHEAVA.SIL..AVGVDPEKASV.IY.QSAIP
Ec  VKMQDDYHCIYCIVDQHAITVRQDAQKLRKATLDIL.A...LYLACGIDPEK.S..IFVQSHVP
Bst VELQHEYNCYFCIVDQHAITV.WQDPHELRQNIRRLAA...LYLAVGVDPEKASV.IY.QSEV.
Bsu VELQHDYNSYFCIVDQHAITV.PQDRLELRKNIRNLAA...LYLAVGLDPEKATLFI..QSEVP

          -  =       =   -   =  =                              =-=
Bf  INKTFIFSDLDYMGMSPGFY.KNV.VKIQKHVTFNQVK.................GIFGFTDSD
Ym  QHSELHWL.LSTL.ASMGLLNRMTQWKS.KSNIKQST.N.GDVYLVNDSDGKVRLGLFS.....
Ec  EHAQLGWA.LNCY.TYFGELSRMTQFKD.KSARYAENINA..............GLFD.....
Bst AHAQAAWM.LQCI.VYIGELERMTQFKE.KSAGKEAV.SA..............GLLT.....
Bsu AHAQAGWM.MQCV.AYIGELERMTQFKD.KSKGNEAVVS..............GLLT.....

          -= - ==                          -=   ==       -=--= =     =
Bf  CIGKISFPAIQAAPSFSN.SFPQIFRDRTDVQCLIPCAIDQDPYFRMTRDVAPR.I..GY....
Ym  ......YPVLQAADILLYKSTH............VPVGDDQSQHL..TRHLAEK.FNKMYGEIF
Ec  ......YPVLMAADILLYQTNL............VPVGEDQKQHL..SRDIAQR.FNALYGEIF
Bst ......YPPLMAADILLYNTDI............VPVGEDQKQ....TRDLAER.FNKRYGELF
Bsu ......YPPLMAADILLYGTDL............VPVGEDQKQHLELTRNLAER.FNKKYNDIF

            ==              =     ===  ====    = -= =  = -  =  = = -
Bf  .......PKPAL.LHSTFFP..ALQGAQTKMSASDPN..SSI.FLTDTAKQIKTKVNKHAFSGG
Ym  KV.....PKPVTMLAQTKKVLSSLSTPEKKMSKSDPNNHDSVIFLNDEPKAIQ.KRL.EALT..
Ec  KVPEPFIPK.SGARVN......SLLEPTKKMSKSDDNRNNVIGLLEDP.KSVV.KKIKRAVT..
Bst TIPEARIPK.VGARIM......SLVDPTKKMSKSDPNPKAYITLLDDA.KTIE.KKIKSAVT..
Bsu TIPEVKIPK.VGARIM......SLNDPLKKMSKSDPNQKAYITLL.DEPKQLE.KKIKSAVT..

     =-  =               = -  - -          = -   =    =    = -=  -= =
Bf  RDTVEEHRQFGGNCDVDV..SFMY.L.TFFLEDDDKLEQIRRDYTS..GAMLTGELKK.ELIE
Ym  .DSISDRFYYDPVERPGV..SNLINIVSGIQRKSI..EDVVEDVSRFN.RDF....KDY.VSE
Ec  .DS.DEPPVVRYDVQNKAGVSNLLDILSAVTGQSIP.E.LEKQFE...GKMY.GHLK.GEVAD
Bst .DS..EGT.IRYDKEAKPGISNLLNIYSTLSGQSIE.E.LERQYE...GKGY.GYFK.ADLAQ
Bsu .DS..EG.IVKFDKENKPGVSNLLTIYSILGNTTIE.E.LEAKYE...GKGY.GEFK.GDLAE

    =-   =    = =        =              =
Bf  VLQP.LIAEHQARRKEVTD..EIVKEFMTP....RKLSYDFQ
Ym  VIIEELKGP...RTEFEKYINEPTYLHSVVESGMRKAREKAAKNLPTFI
Ec  AVSGMLTEL.QERYHRFRN.DEAFLQQVMKD.GAEKASAHAS
Bst VVIETLPI..QERYHHWMESEELDVRL...DEGAEKANRVASEMVRKMEQAMGLGRR
Bsu VVVNALKPI.QDRYYELIESEELDRIL...DEGAERANRTANKMLKKMENAMGLGRKRR
```

FIGURE 5: Comparison of the predicted amino sequence of beef WRS with the published sequences of prokaryotic and yeast mitochondrial WRS sequences. Bf, beef pancreas (this work); Ym, yeast mitochondrial (Myers & Tzagoloff, 1985); Ec, *E. coli* (Hall et al., 1982); Bst, *B. stearothermophilus* (Barstow et al., 1986); Bsu, *B. subtilis* (Chow & Wong, 1988). The two consensus sequences, HIGH and KMSKS, are boxed. The possible homologous residues in beef WRS and in at least two other WRS are marked with (=). The isologous residues in beef WRS and in the prokaryotic WRS are marked with (-). Gaps in the sequences are indicated by dots.

KMSAS stretch was confirmed by direct peptide sequencing and indicates that the second lysine of the KMSKS sequence is not critical for tryptophanyl-tRNA synthetase activity. Comparison of all the known WRS sequences (Barstow et al., 1986; Hall et al., 1982; Myers & Tzagoloff, 1985; Chow & Wong, 1988) showed that, including these two consensus sequences, beef WRS shared around 20% homology with prokaryotic and yeast mitochondrial WRS, depending on the chosen alignment (Figure 5).

A more surprising result is shown in Figure 6, where beef WRS and rabbit peptide chain release factor (Lee et al., 1990)] sequences are compared. Both proteins have the same amino acid length (475) and share 86% of strict homology, which can be extended to 92% when considering a few isologous replacements such as D/E, K/R, F/Y, S/T, and V/I. The high homology in the cDNA sequences (83%) would suggest that they originate from the same ancestor gene.

## Discussion

Considering the striking homology between the amino acid sequences of the rabbit peptide chain release factor (Lee et al., 1990) and that of beef tryptophanyl-tRNA synthetase, the

```
                                                                    50
WRS MADMSNGEQGCGSPLELFHSIAAQGELVRDLKARNAAKDEIDSAVKMLLS
eRF ---VT---R.-A--Q---S---------KS----K-P-E----------
                                                                   100
WRS LKTSYKAATGEDYKVDCPPGDPAPESGEGLDATEADEDFVDPWTVQTSSA
eRF ------E-M-----A-----NST-D-HGDPE--DDK---------R----
                                                                   150
WRS KGIDYDKLIVRFGSSKIDKELVNRIERATGQRPHRFLRRGIFFSHRDMHQ
eRF ----------Q-------------------------------------N-
                                                                   200
WRS ILDAYENKKPFYLYTGRGPSSEAMHVGHLIPFIFTKWLQDVFNVPLVIQM
eRF V-----------------LLKQCN---------------D----V--
                                                                   250
WRS TDDEKYLWKDLTLDQAYGYAVENAKDIT.CGFDINKTFIFSDLDYMGMSPG
eRF S-------------E-V---TL---Q--MP----V----------------
                                                                   300
WRS FYKNVVKIQKHVTFNQVKGIFGFTDSDCIGKISFPAIQAAPSFSNSFPQI
eRF --------------------------------------------------

                                                                   350
WRS FRDRTDVQCLIPCAIDQDPYFRMTRDVAPRIGYPKPALLHSTFFPALQGA
eRF --HGQA-------------------------------------------D
                                                                   400
WRS QTKMSASDPNSSIFLTDTAKQIKTKVNKHAFSGGRDTVEEHRQFGGNCDV
eRF --------------------------------I-----------
                                                                   450
WRS DVSFMYLTFFLEDDDKLEQIRRDYTSGAMLTGELKKELIEVLQPLIAEHQ
eRF ----------------------K--S--------------D-----V---
                                    475
WRS ARRKEVTDEIVKEFMTPRKLSYDFQ-COOH
eRF ---------M--------Q-CFHY--COOH
```

FIGURE 6: Comparison of the predicted amino acid sequences of beef WRS and rabbit eRF. The amino acids of eukaryotic release factor (Lee et al., 1990) are represented by dashes (lower sequence) when they are identical with those of WRS (upper sequence) and specified when they are different. The two consensus sequences of the HIGH and KMSKS series are boxed (positions 176–179 and 354–358). The peptide sequences obtained by direct Edman degradation are underlined.

question arises of having mistakenly cloned and sequenced beef eRF instead of beef WRS. There are strong arguments against this hypothesis: (i) The sequenced clone has been expressed as active WRS in *E. coli*. (ii) The sequenced cDNA corresponds to the protein characterized as WRS, because there is an almost perfect match between the translated cDNA sequence and the sequences of the peptide stretches obtained by direct amino acid sequencing and distributed along the major part of the protein chain. (iii) The WRS preparation used for peptide isolation and polyclonal antibodies preparation could not be a mixture of beef eRF and beef WRS. Indeed, the protein isolated as WRS is an $\alpha_2$ dimer (Gros et al., 1982) and possesses two binding sites for tryptophan (Graves et al., 1980; Mazat et al., 1982) and two binding sites for $tRNA^{Trp}$ (Akhverdyan et al., 1977). It catalyzes both the formation of 2 mol of tryptophanyl adenylate/mol of enzyme (Merle et al., 1984) and the transfer of the two tryptophanyl moieties from tryptophanyl adenylate to $tRNA^{Trp}$ (Trézéguet et al., 1986). Such stoichiometries imply that the sequenced protein is really WRS and is not significantly contaminated by a protein species that could have a similar sequence and no binding sites for the different WRS ligands. These data suggest that WRS and eRF are two different but extremely close proteins.

A tryptophanyl-tRNA synthetase activity associated with the rabbit eRF molecule itself is unlikely. Such an activity has not been mentioned by Lee et al. (1990), although these authors were aware of this possibility since they pointed out that rabbit eRF and prokaryotic WRS present stretches of sequence similarity. One very limited but specific sequence difference between beef WRS and rabbit eRF does exist, however, and can account for the absence of synthetase activity in eRF. This specific difference lies upstream to the HIGH sequence. Using site-directed mutagenesis techniques, Leatherbarrow et al. (1985) investigated the role of several amino acid residues of *E. coli* tyrosyl-tRNA synthetase (YRS) that could stabilize the transition state of the aminoacyl adenylate·enzyme complex, the structure of which was established from crystallographic data (Bhat et al., 1982). Thr 40 and His 45 of YRS were suggested to contribute largely to the proper binding of the $\gamma$-phosphate group of ATP by hydrogen bonding with the side chains of these amino acids. His 45 in YRS is the first histidine residue of the HIGH signature sequence present in class I synthetases (Burbaum et al., 1990; Eriani et al., 1990). Replacement in YRS mutants of Thr 40 and His 45 by residues unable to give rise to hydrogen bonds decreases the enzyme activity by more than 5 orders of magnitude, while the single replacement of Thr 40 by a glycine residue reduces it 6000-fold (Leatherbarrow et al., 1985). The HIGH signature sequence is present in WRS as HVGH at positions 176–179, where His 176 corresponds to His 45 of *E. coli* YRS. Upstream to this sequence, there are two serine residues (Ser 170 and 171) which could act as hydrogen donors, like Thr 40 in YRS. It is noticeable that, in the aminoacyl-tRNA synthetase family featuring the HIGH signature sequence, all of the members have hydrogen donor

Table I: Isolated and Sequenced Peptides from Beef WRS

| | | |
|---|---|---|
| | Trypsin | |
| $K_{56}$ | AATGEDYKVDCPPGDPAPESGE | |
| $K_{64}$ | VDCPPGDPAPESGEGLDA | |
| $K_{119}$ | ELVNR | |
| $R_{146}$ | DMHQILDAYENK | |
| $F_{161}$ | YLYTGR | |
| $K_{186}$ | WLQDV | |
| $K_{205}$ | YLWK | |
| $K_{209}$ | DLTLDQAYGYAVENAK | |
| $K_{225}$ | DIIACGFDINK | ($T_{228}$) |
| $K_{234}$ | TFIFCLDYMGMSPGFYK | ($S_{240}D_{241}$) |
| $K_{268}$ | GIFGFTDSDCIGK | |
| $R_{302}$ | DRTDVQCLIPCAIDQDPYF | |
| $R_{330}$ | IGYPK | |
| $T_{409}$ | FLEDDQKLEQI | ($D_{415}$) |
| $S_{426}$ | GAMLT | |
| | Clostripain | |
| $R_{111}$ | FGSSKIDKELVNR | |
| $R_{330}$ | IGYPKPALLHSTFFPALQGAQTK | |
| $R_{453}$ | KEVTDEIDVKEFMTPR | ($V_{461}$) |
| | Chymotrypsin | |
| $A_{23}$ | QGELVRKARN | |
| $L_{32}$ | KARNAAKDEIDSAVKMLLSLK | |
| $Y_{55}$ | KAATGEDYKVDCPPGDPAPESGEGLDAT | |
| | CNBr | |
| $M_{148}$ | HQILDAYENKK | |
| $M_{174}$ | HVGHLIPFIFTKWLQDVFNVPVLVIQM | |
| $M_{200}$ | TDDEKYLWKDLTLDQAYGYAVENAK | |
| $M_{354}$ | SASDPN | |
| $M_{405}$ | YLTFFLE | |

[a] The peptides were obtained after cleavage of WRS by TPCK-trypsin, clostripain, chymotrypsin, and cyanogen bromide, respectively. The amino acids and their positions, appearing in front of the peptides, refer to the residues according to the cDNA sequence immediately before the N-termini of the sequenced peptides (Figure 4). The underlined residues are those for which a difference with the cDNA sequence was found: on the right of the peptide is (are) given the residue(s) assigned from the cDNA.

residues four to six amino acids upstream to this sequence (Table II). In contrast to this observation, eRF, though it possesses an NVGH sequence (176–179) comparable to HVGH of WRS, has two leucine residues in positions 170 and 171, which cannot act as hydrogen donors and probably repulse the γ-phosphate group of ATP. The release factor is probably unable to activate any amino acid, unless Lys 172 is equivalent to Thr 40 of YRS. Therefore, it may be that eRF and WRS do not present the same catalytic activities although they have nearly identical amino acid sequences.

These two proteins, however, probably share common properties. The sequence homology between eRF and WRS is high enough to imply that eRF possesses a tertiary structure similar to that of WRS and is able to bind a tRNA-like molecule. The second consensus sequence KMSKS, found in several aminoacyl-tRNA synthetases (Hountondji et al., 1986a), is present in both eRF and WRS, but as KMSAS. This is consistent with the fact that reaction of WRS with periodate-oxidized ATP does not interfere with tryptophanyl adenylate formation (Fournier et al., 1987), in contrast with what is observed for aminoacyl-tRNA synthetases featuring the KMSKS sequence (Mechulam et al., 1991). From chemical and crystallographic data (Bhat et al., 1982; Hountondji et al., 1986a,b; Burbaum et al., 1990), this region appears to be involved in the binding of the tRNA 3′ end and in the transfer of the aminoacyl moiety from the adenylate to tRNA. Owing to the presence of this KMSAS sequence, eRF may be able to bind the 3′ end of a tRNA. The question then arises whether eRF interacts with stop codons of the mRNA by the intermediate of a non-aminoacylatable tRNA-like molecule, not yet identified, or directly with the stop codon.

The residues that can be suggested to be homologous between eukaryotic and prokaryotic WRS are concentrated in the regions presenting the two consensus sequences (Figure 5). One of these homologous peptide stretches, around the KMSAS sequence, has been already pointed out by Lee et al. (1990), who found 20% sequence similarity between amino acid positions 280 and 375 of eRF. It is possible to extend the homology between bovine WRS (or eRF) and prokaryotic

Table II: Sequence Analogies Upstream to HIGH Region of Class I Aminoacyl-tRNA Synthetases (AARS)[a]

| AARS | Source | Sequence | Ref |
|---|---|---|---|
| ArgRS | *E. coli* | **NVAKEMHVGH** | Eriani *et al.*, 1989 |
| CysRS | *E. coli* | **YDLCHIGH** | Eriani *et al.*, 1989 |
| GlnRS | *E. coli* | **NGYLHIGH** | Hoben *et al.*, 1982 |
| GluRS | *E. coli* | **TGYLHVGG** | Breton *et al.*, 1986 |
| IleRS | *E. coli* | **NGSIHIGH** | Webster *et al.*, 1984 |
| LeuRS | *E. coli* | **SGRLHMGH** | Haertlein & Madern, 1987 |
| MetRS | *E. coli* | **NGSIHLGH** | Dardel *et al.*, 1984 |
| ThrRS | *E. coli* | **HSCAHLLGH** | Mayaux *et al.*, 1983 |
| TrpRS | *B. stearo* | **SGELTIGN** | Winter *et al.*, 1977 |
| TrpRS | Beef | **SSEAMHVGH** | *This work* |
| TyrRS | *B. stearo* | **TADSLHIGH** | Winter *et al.*, 1983 |
| ValRS | *B. stearo* | **TGKLHIGH** | Borgford *et al.*, 1987 |
| eRF | Rabbit | PLLKQC**NVGH** | Lee *et al.*, 1990 |

[a] Only one example is given in this table for each aminoacyl-tRNA synthetase, except for TrpRS. Nonquoted aminoacyl-tRNA synthetases specific for a given amino acid have sequences similar to the listed ones. The potential hydrogen donor residue(s) upstream to the HIGH region, and the HIGH region itself, is (are) in boldface. Though threonyl-tRNA synthetase belongs to class II synthetases (Eriani et al., 1990), it is mentioned because it contains an HLLGH sequence, close to the HIGH sequence, preceded by a serine and a histidine, potential hydrogen donors. *stearo* = *stearothermophilus*.

WRS, when taking into account the consensus sequences of the HIGH family, present as TIGN or TLGN in these prokaryotic enzymes (Burbaum et al., 1990). Around 20% homologous amino acids may then be tentatively aligned along the entire chains of the prokaryotic (*E. coli*, *B. stearothermophilus*, and *Bacillus subtilis*) and the eukaryotic WRS (of the order of 70–80 out of 334 residues of *E. coli* WRS, including the HIGH and the KMSKS regions) (Figure 5). When considering that more than 25% identity should be observed to suggest a homology between two proteins (Sander & Schneider, 1991), it is questionable whether the eukaryotic WRS and the prokaryotic WRS share many significant structure similarities, besides their two clear signature sequence regions. This absence of homology is consistent with the observation previously mentioned, that beef WRS aminoacylates very poorly $tRNA^{Trp}$ from *E. coli*.

Part of the N-terminal region of beef WRS is dispensable to the catalytic activity of the enzyme. Lemaire et al. (1975) and Sheinker et al. (1979) have observed that the enzyme can undergo limited proteolysis, reducing the molecular mass of its subunit from 54 to 51 and then to 41 kDa. This proteolysis splits off only the N-terminal part of the enzyme, without affecting enzyme activity. This is consistent with the fact that the IPTG-induced expression of WRS could be obtained with a cDNA lacking the 5′-terminal part corresponding to the first 17 N-terminal amino acids of the protein (clone 7-1). In the 41-kDa enzyme, about 120 amino acids have been removed, leaving an N-terminal aspartic acid (Lemaire et al., 1975). A slightly deeper proteolysis has also been obtained through the action of chymotrypsin, to yield a protein of 40-kDa subunits. This protein has lost its catalytic properties (Epely et al., 1976). Hence, considering that the only Asp residue in the vicinity of amino acid 120 is Asp 118, the minimum chain length essential to beef tryptophanyl-tRNA synthetase for catalyzing tRNA aminoacylation is of the order of 358 residues (475 minus 117). This chain length is close to that of the homologous prokaryotic enzyme (328 for *B. stearothermophilus*, 330 for *B. subtilis*, and 334 for *E. coli*). The first 117 amino acids on the N-terminal side of beef tryptophanyl-tRNA synthetase are therefore more probably necessary for other functions than tryptophan activation and tRNA aminoacylation.

## ACKNOWLEDGMENTS

We thank Dr. J. M. Jeltsch, Laboratory of Prof. P. Chambon, Strasbourg, for his help in the initial use of the λgt 10 vector by M. Garret and Prof. K. Muench, University of Miami, in whose laboratory the WRS peptide sequencing was initiated.

## REFERENCES

Akhverdyan, V. L., Kisselev, L. L., Knorre, D. G., Lavrik, O. I., & Nevinsky, G. A. (1977) *J. Mol. Biol. 113* 475–501.

Alterio, J., Halley, C., Brou, C., Soussi, T., Courtois, Y., & Laurent, H. (1988) *FEBS Lett. 242*, 41–46.

Andrews, D., Trêzêguet, V., Merle, M., Graves, P. V., Muench, K. H., & Labouesse, B. (1985) *Eur. J. Biochem. 146*, 201–209.

Atkins, R. A., & Lambovitz, A. M. (1982) *Cell 50*, 331–345.

Barker, D. G., & Winter, G. (1982) *FEBS Lett. 145*, 191–193.

Barstow, D. A., Sharman, A. F., Atkinson, T., & Minton, N. P. (1986) *Gene 46*, 37–45.

Bec, G., Kerjan, P., Zha, X. D., & Waller, J. P. (1989) *J. Biol. Chem. 264*, 21131–21137.

Bhat, T. M., Blow, D. M., Brick, P., & Nyborg, J. (1982) *J. Mol. Biol. 155*, 63–81.

Borgford, T. J., Brand, N. J., Gray, T. E., & Fersht, A. R. (1987) *Biochemistry 26*, 2480–2486.

Breton, R., Sanfaçon, H., Papyannopoulos, I., Biemann, K., & Lapointe, J. (1986) *J. Biol. Chem. 261*, 10610–10617.

Burbaum, J., Starzyk, R. M., & Schimmel, P. (1990) *Proteins 7*, 99–111.

Chow, C. M., & Wong, J. T. F. (1988) *Gene 73*, 537–543.

Cirakoglu, B., & Waller, J.-P. (1986) *Eur. J. Biochem. 149*, 353–361.

Clemens, M. J. (1990) *Trends Biochem. Sci. 173*, 172–175.

Dang, C. V., Johnson, D. L., & Yang, D. C. H. (1982) *FEBS Lett. 142*, 1–6.

Dardel, F., Fayat, G., & Blanquet, S. (1984) *J. Bacteriol. 160*, 1115–1122.

Davie, E. W., Koningsberger, V. V., & Lipman, F. (1956) *Arch. Biochem. Biophys. 65*, 21–38.

Dorizzi, M., Mêrault, G., Fournier, M., Labouesse, J., Keith, G., Dirheimer, G., & Buckingham, R. (1977) *Nucleic Acids Res. 4*, 31–42.

Ehresmann, B., Imbault, P., & Weil, J. H. (1974) *Biochimie 56*, 1341–1345.

Eldred, E. W., & Schimmel, P. (1972) *J. Biol. Chem. 247*, 2961–2964.

Epely, S., Gros, C., Labouesse, J., & Lemaire, G. (1976) *Eur. J. Biochem. 61*, 139–146.

Eriani, G., Dirheimer, G., & Gangloff, J. (1989) *Nucleic Acids Res. 17*, 5725–5736.

Eriani, G., Delarue, M., Poch, O., Gangloff, J., & Moras, D. (1990) *Nature 347*, 203–206.

Eriani, G., Dirheimer, G., & Gangloff, J. (1991) *Nucleic Acids Res. 19*, 265–269.

Favorova, O. O., Zargarova, T. A., Rukosuyev, V. S., Beresten, S. F., & Kisselev, L. L. (1989) *Eur. J. Biochem. 184*, 583–588.

Fourney, R. M., Miyakoshi, J., Day, R. S., III, & Paterson, M. C. (1987) *Focus 10*, 5–7.

Fournier, M., Labouesse, B., & Labouesse, J. (1987) *Eur. J. Biochem. 167*, 533–540.

Graves, P.-V., de Bony, J., Mazat, J.-P., & Labouesse, B. (1980) *Biochimie 62*, 33–41.

Gros, C., Lemaire, G., Van Rappenbusch, R., & Labouesse, B. (1972) *J. Biol. Chem. 247*, 2921–2943.

Grossberger, D. (1987) *Nucleic Acids Res. 15*, 6537.

Haertlein, M., & Madern, D. (1987) *Nucleic Acids Res. 15*, 10199–10210.

Hall, C. V., Van Cleemput, M., Muench, K. H., & Yanofsky, C. (1982) *J. Biol. Chem. 257*, 6132–6136.

Herbert, C. J., Labouesse, M., Dujardin, G., & Slonimski, P. P. (1986) *EMBO J. 7*, 473–483.

Hoben, P., Royal, N., Cheung, A., Yamao, F., Biemann, K., & Söll, D. (1982) *J. Biol. Chem. 257*, 11644–11650.

Hountondji, C., Dessen, P., & Blanquet, S. (1986a) *Biochimie 68*, 1071–1078.

Hountondji, C., Lederer, F., Dessen, P., & Blanquet, S. (1986b) *Biochemistry 25*, 16–21.

Laemli, V. K. (1970) *Nature 227*, 680–685.

Leatherbarrow, R. J., Fersht, A. R., & Winter, G. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 7840–7844.

Lee, C. C., Craigen, W. J., Muzny, D. M., Harlow, E., & Caskey, C. T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3508–3512.

Lemaire, G., Gros, C., Epely, S., Kaminsky, M., & Labouesse, B. (1975) *Eur. J. Biochem. 51*, 237–252.

Lorber, B., Kern, D., Giegê, R., & Ebel, J. P. (1982) *FEBS Lett. 146*, 59–64.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Mayaux, J.-F., Fayat, G., Fromant, M., Springer, M., Grunberg-Manago, M., & Blanquet, S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6152–6156.
Mazat, J.-P., Merle, M., Graves, P.-V., Mêrault, G., Gandar, J.-C., & Labouesse, B. (1982) *Eur. J. Biochem. 128*, 389–398.
Mechulam, Y., Dardel, F., Lecorre, D., Blanquet, S., & Fayat, G. (1991) *J. Mol. Biol. 217*, 465–476.
Mêrault, G., Graves, P. V., Labouesse, B., & Labouesse, J. (1978) *Eur. J. Biochem. 87*, 541–550.
Merle, M., Graves, P.-V., & Labouesse, B. (1984) *Biochemistry 23*, 1716–1723.
Merle, M., Trêzêguet, V., Gandar, J.-C., & Labouesse, B. (1988) *Biochemistry 27*, 2246–2252.
Myers, A. A., & Tzagoloff, A. (1985) *J. Biol. Chem. 260*, 15371–15377.
Plateau, P., Mayaux, J. F., & Blanquet, S. (1981) *Biochemistry 20*, 4654–4662.
Reinbolt, J., Hounwanou, N., Boulanger, Y., Wittmann-Liebold, B., & Bosserhoff, A. (1983) *J. Chromatogr. 259*, 121–130.
Sallafranque, M. L., Garret, M., Benedetto, J. P., Fournier, M., Labouesse, B., & Bonnet, J. (1986) *Biochim. Biophys. Acta 882*, 192–199.
Sambrook, J., Fritsch, F. F., & Maniatis, T. (1989) *Molecular Cloning; A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sander, C., & Schneider, R. (1991) *Proteins 9*, 625–637.
Sanger, F., Niclen, S., & Coulson, A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5468.
Scheinker, V., Beresten, S. F., Degtyarev, S., & Kisselev, L. L. (1979) *Nucleic Acids Res. 7*, 625–637.
Schimmel, P. (1987) *Annu. Rev. Biochem. 56*, 125–158.
Schimmel, P., & Söll, D. (1979) *Annu. Rev. Biochem. 48*, 601–648.
Tada, M., & Tada, M. (1975) *Nature 255*, 510–512.
Trêzêguet, V., Merle, M., Gandar, J.-C., & Labouesse, B. (1986) *Biochemistry 25*, 7125–7135.
Webster, T., Tsai, H., Kula, M., Mackie, G. A., & Schimmel, P. (1984) *Science 226*, 1315–1317.
Weiss, S. B., Zachau, H. G., & Lipmann, F. (1959) *Arch. Biochem. Biophys. 83*, 101–105.
Winter, G. P., & Hartley, B. S. (1977) *FEBS Lett. 80*, 340–342.
Winter, G. P., Koch, G. L., Hartley, B. S., & Barker, D. G. (1983) *Eur. J. Biochem. 132*, 383–387.

# Energetics of Subunit Dimerization in Bacteriophage λ *c*I Repressor: Linkage to Protons, Temperature, and KCl†

Kenneth S. Koblan and Gary K. Ackers*
*Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110*
*Received February 7, 1991; Revised Manuscript Received April 25, 1991*

ABSTRACT: A common feature of gene regulatory systems is the linkage between reversible protein oligomerization and DNA binding. Experimental dissection using temperature dependence of the subunit–subunit energetics and their linkage to processes such as ion binding and release is necessary for characterization of the chemical forces that contribute to cooperativity and site specificity. We have therefore studied the effects of temperature, proton activity, and monovalent salt on monomer–dimer assembly of the λ *c*I repressor using a recently developed gel chromatographic procedure. This technique has made possible studies in the previously inaccessible picomolar concentration ranges where the assembly reactions occur. Upon formation of the dimer interface in the range pH 5–9, we find an overall absorption of protons which is temperature-dependent. The dimerization reaction displays a large negative enthalpy of association at all conditions studied (pH 5, 7, and 9). The reaction is also dependent on monovalent salt concentration: subunit association is weaker at low-salt conditions. The results suggest that a repulsive interaction between negatively charged side chains (i.e., aspartates and glutamates) on each monomer surface is attenuated by increasing concentrations of KCl. Formation of the dimer interface may be mediated by absorption of cations which stabilize the complex.

The regulation of transcriptional initiation for many prokaryotic and eukaryotic genes is governed by the interaction of regulatory proteins and specific DNA sequences. The functional energetics of each macromolecular species are also linked to any reactions of conformational change, subunit polymerization, and ion association/dissociation that it may undergo. Experimental determination of functionally relevant protein–protein and protein–DNA energetics is necessary for characterization of the noncovalent forces which contribute to site specificity and strength of interaction. A research program of this laboratory has been focused on the energetics of interactions between λ *c*I repressor and the right operator region ($O_R$)[1] (Ackers et al., 1982; Shea & Ackers, 1985; Brenowitz

†Supported by National Institutes of Health Grants GM39343 and R37-GM24486.
* Correspondence should be addressed to this author.

[1] Abbreviations: BSA, bovine serum albumin; $NaDodSO_4$, sodium dodecyl sulfate; $O_R$, λ right operator; $C_T$, total *c*I plateau concentration (monomer units); $V_e$, centroid elution volume; $\sigma_w$, weight-average partition coefficient; Tris, tris(hydroxymethyl)aminomethane; Bistris, [bis-(2-hydroxyethyl)amino]tris(hydroxymethyl)methane; NBS, National Bureau of Standards.

0006-2960/91/0430-7817$02.50/0 