Nucleic Acids Research, Vol. 19, No. 22 6215–6220

Cloning, sequencing and overexpression of the gene for prokaryotic factor EF-P involved in peptide bond synthesis

Hiroyuki Aoki, Sally-Lin Adams, Dae-Gyun Chung, Makoto Yaguchi[1], Shuang-En Chuang[2] and M.Clelia Ganoza*

Banting and Best Department of Medical Research, University of Toronto, Toronto, Ontario, [1]Department of Biology, National Research Council, Ottawa, Ontario, Canada and [2]Department of Genetics, University of Wisconsin, Madison, WI, USA

Received August 16, 1991; Revised and Accepted October 11, 1991 EMBL accession no. X61676

ABSTRACT

A soluble protein EF-P (elongation factor P) from *Escherichia coli* has been purified and shown to stimulate efficient translation and peptide-bond synthesis on native or reconstituted 70S ribosomes *in vitro*. Based on the partial amino acid sequence of EF-P, 18- and 24-nucleotide DNA probes were synthesized and used to screen λ phage clones from the Kohara Gene Bank. The entire EF-P gene was detected on λ clone #650 which contains sequences from the 94 minute region of the *E.coli* genome. Two DNA fragments, 3.0 and 0.78 kilobases in length encompassing the gene, were isolated and cloned into pUC18 and pUC19. Partially purified extracts from cells transformed with these plasmids overrepresented a protein which co-migrates with EF-P upon SDS polyacrylamide gel electrophoresis, and also exhibited increased EF-P mediated peptide-bond synthetic activity. Based on DNA sequence analysis of this gene, the EF-P protein consists of 187 amino acids with a calculated molecular weight of 20,447. The sequence and chromosomal location of EF-P establishes it as a unique gene product.

INTRODUCTION

Although the principal function of ribosomes is to transfer the information of the genetic transcript into proteins through the condensation of aminoacyl-tRNAs, the mechanism of peptide-bond synthesis remains incompletely understood (1, 2). In prokaryotic cells, this reaction is catalyzed by either 50S or 70S ribosomes and has been reported to occur in the absence of soluble factors or GTP (3). However, this activity of 50S particles is measurable only in the presence of exogenous methanol and elevated concentrations of substrate (4). In order for peptide-bond synthesis to occur efficiently under more physiological conditions, the 30S particle, which contributes to the overall environment of the native 70S ribosome, as well as a soluble protein, EF-P, which functionally couples the 50S peptidyl transferase to the 30S subunit, are required (5–8).

EF-P has been identified and purified based on its ability to stimulate the synthesis of N-formyl[^{35}S]Methionyl-puromycin from a complex of 70S ribosome·AUG·f[^{35}S]Met-tRNA and added puromycin (5, 6), and is active with other aminoacyl acceptors such as cytidyl(3′-5′)-[2′(3′)-*O*-L-aminoacyladenosines (CCAs) (9). In addition, purified EF-P stimulates translation programmed with MS2 RNA, or amber mutants of f_2 RNA which synthesize the N-terminal hexapeptide of the coat protein (10). The synthesis of poly-Lys programmed with poly(A) is also enhanced by EF-P (7). These experiments strongly suggest that EF-P functions during elongation.

Although its mechanism of action is still unknown, it is possible that EF-P functions indirectly by altering the affinity of the ribosome for aminoacyl-tRNAs, thus increasing their reactivity as acceptors for peptidyl transferase (5, 9). A model implicating EF-P in the kinetics of translocation has also been proposed (11). A third alternative is that EF-P, since it differentially stimulates incorporation of various amino acids, is involved in accommodating their side chains into the active center (9, 12). More generally, EF-P could be involved in the fine tuning of the overall conformation of the peptidyl transferase center (13). To ascribe a more detailed function to this protein in translational events, as well as its possible structural relationship to other established components of protein synthesis, we have cloned the gene for EF-P and determined its DNA sequence. The sequence and chromosomal location of the gene reported below establish that EF-P is coded by a previously undescribed gene.

MATERIALS AND METHODS

Materials

[λ-^{32}P]ATP (~4000 Ci/mmol) and α[^{35}S]dATP (>1000 Ci/mmol) were obtained from ICN Biomedicals and Amersham Corp., respectively. T_4 polynucleotide kinase, T_7 DNA polymerase and Nap-25 (Sephadex G-25) columns (2×8 cm) were from Pharmacia. T_4 ligase, pUC18, pUC19 and restriction endonucleases were from BRL. pBluescript and XL1-Blue cells *rec*A$^-$ (*lac*$^-$ λ$^-$ {F′ proAB, *lacI*^{Q}, *lacZ*ΔM15, Tn10}) were

* To whom correspondence should be addressed

from Stratagene. Zeta-Probe blotting membranes were from Bio-Rad. *E.coli* MRE600 mid-log cells were from the Grain Processing Corp. (Muscatine, IA, USA).

Assay of EF-P activity and purification of EF-P

EF-P activity was assayed as described (6, 13) by monitoring the stimulation of N-formyl[^{35}S]Methionyl-puromycin synthesis from ribosome·AUG·f[^{35}S]Methionyl-tRNA complexes and added puromycin. EF-P was purified to homogeneity from an S100 extract of *E.coli* K12 cells as described by Chung *et al.* (13).

Partial amino acid sequencing and analysis

Automated gas-phase sequencing was performed on an Applied Biosystems 475A Protein Sequencing System (Foster City, CA, USA) incorporating a model 470A gas-phase sequencer equipped with an on-line 120A PTH analyzer under the control of a model 900A control/data analysis module. Cyanogen bromide (CNBr) fragments of EF-P (500 pmole) on a polyvinylidine difluoride membrane were analyzed as described by Matsudaira (14). Amino acid analysis on 0.5−1.5 μg of EF-P was performed with an Applied Biosystems 420H-03 Amino Acid Analyzer (model 420A hydrolyzer and derivatizer, model 130A PTC analyzer and model 920A data module).

Oligonucleotide probes and screening of the *Escherichia coli* genomic library

Two mixed probes, consisting of 18 and 24 nucleotides, were prepared commercially using an Applied Biosystems 380A Synthesizer. The sequence of the 18-mer was 5'GAC(or T)-TTC(or T)-CGC(or T)-GC(N)-GGC(or T)-CTG3'. The sequence of the 24-mer was 5'ATG-AAT(or C)-AAT(or C)-GAA(or G)-AC(N)-TTC(or T)-GAA(or G)-CAA(or G)3'. The oligonucleotides were 5'-end labelled as described in (15). Preliminary screening of the Kohara Gene Bank (16) was performed with the 18-mer mixed oligomer. The DNA from each λ clone was isolated, dot blotted onto nitrocellulose, and probed as in (15) at 52°C. The positive λ clones were subsequently screened with the 24-mer mixed oligomer. The λ DNA from each clone was isolated and subjected to 0.7% agarose gel electrophoresis (15). After Southern blotting onto a Zeta-Probe membrane as in (17), the DNA was prehybridized, hybridized and washed as in (15) at 48°C.

Cloning of the EF-P gene

Stock bacteriophage harbouring λ clone #649 or 650 were purified using CsCl gradients as described in (15). Restriction enzyme analysis of the DNA from these λ clones was carried out by digestion with the enzymes described in (16) and electrophoresis on 0.7% agarose gels. Southern blot hybridization with the 18- and 24-oligonucleotide probes were performed using the conditions described above except that probing with the 18-mer probe was at 45°C. Fragments to be cloned were isolated by restriction enzyme digestion of λ clone #649 or 650 DNA with 40 units of the appropriate enzyme(s). Incubation was for 2 hrs at 37°C. Following 1.2% agarose gel electrophoresis (15), the portions of the gel containing the fragments of interest were excised and electroeluted (Analytical Electroeluter, IBI). The isolated fragments were ligated to 0.5 μg of pBluescript SK/KS (+/−), pUC18 and pUC19. The vector-insert mixtures were used to transform 200 μl of *E.coli* XL1-Blue cells made competent with $CaCl_2$ (15). The cells were cultured on SOB media containing 100 μg/ml ampicillin, 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and 40 μg/ml 5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (15). White colonies were selected and cultured. Colonies were confirmed to contain the correct insert by Southern blot hybridization.

Nucleotide sequence determination and analysis

The DNA sequence was determined by the double stranded dideoxy chain termination method of Sanger *et al.* (18) using modified T_7 DNA polymerase and α[^{35}S]dATP. The primers used were the M13mp18 universal primer supplied in the sequencing kit, the reverse primer, SK and KS primers (only when the vector was pBluescript) and the two synthetic 18 and 24 oligomer probes. A computer search for nucleotide and amino acid homology was carried out using the FASTA program (19). Open reading frames, amino acid composition of the predicted polypeptide and restriction enzyme mapping were performed using the DNA Strider 1.0 Program (20).

Detection by SDS-PAGE of EF-P in *E.coli* extracts containing EF-P clones

XL1-Blue cells transformed with pUC18 or EF-P clones, bearing the entire gene, were grown at 37°C in LB broth with 100 μg/ml ampicillin until the optical density at 650 nm was 0.7. The cells were then harvested and washed with 10 mM Tris-HCl (pH 7.4), 50 mM NH_4Cl, 10 mM $MgCl_2$ and 1 mM DTT. The cells (3.0 g) were ground with an equal weight of alumina, followed by addition of 6 mls of the above buffer and 2 μg/ml of DNase. Supernatant extracts were obtained upon centrifugation at 10,000×g for 15 min. and 30,000×g for 30 min. After removal of ribosomes by centrifugation at 100,000×g for 2.5 hrs, the S100s were subjected to 30 to 60% ammonium sulfate precipitation. The pellets were resuspended in a minimum volume of 10 mM Tris-HCl (pH 7.4) and 1 mM DTT, and dialyzed against the same buffer. The dialyzed fractions were analyzed by 10% SDS-PAGE analysis as described in (21). The protein bands were examined quantitatively by the GS 300 Transmittance/Reflectance Scanning Densitometer (Hoeffer Scientific Instruments). Protein concentrations were determined by the Bio-Rad Protein Microassay using bovine serum albumin as the standard.

RESULTS AND DISCUSSION

The process of translation requires the codon-dependent binding of aminoacyl-tRNAs to prokaryotic 70S ribosomes followed by peptide-bond formation. The central task of condensing the aminoacyl ester of the resulting nascent peptidyl-tRNA with the amino group of the incoming aminoacyl-tRNA is accomplished by the activity of the peptidyl transferase, which is thought to be an integral part of the prokaryotic 50S ribosomal subunit. The peptidyl transferase has been extensively studied in a 'fragment' reaction where peptide bonds are formed between an acceptor, e.g. puromycin, which is an analogue of the 3' terminus of aminoacyl-tRNA, and donor fragments of aminoacyl-tRNA (1, 2). This model reaction has allowed the identification of the minimum 50S components required (22, 23) and has led to numerous other biochemical and genetic studies aimed at determining the reaction mechanism.

In contrast, little is known about the requirements for peptide-bond synthesis on 70S ribosomes which are the native site of synthesis. Such studies have revealed the requirement for a new

protein EF-P (5–8). The protein is scored in the absence of methanol using low acceptor concentrations. Under these conditions, the 50S subunit exhibits low activity, as does the 70S particle. However, EF-P greatly enhances peptide-bond synthesis by 70S ribosomes or reassembled subunits containing bound fMet-tRNA, suggesting that it is not involved in associating the subunits.

Peptidyl transferase is thought to be composed of the 23S rRNA and the ribosomal proteins L2, L3, L4, L16 and L20 (22, 23). However, many other helper proteins also stimulate the reaction (22, 24). To determine if the activity of EF-P depends on the 50S ribosomal proteins which are essential to the proper functioning of the peptidyl transferase, 50S core particles lacking L16, L6, L11 and L7/L12 were reconstituted by adding back each protein and assaying for peptide-bond formation. These results underscored the specificity of L16 in this 70S ribosome-dependent reaction and demonstrated the requirement for EF-P (8).

Many of the factors that actuate translation such as EF-Tu, EF-G and IF-2, involve ribosome-dependent hydrolysis of GTP which depends on the 50S protein complex L7/L12 (3). Unlike these factors, the EF-P-mediated reaction does not require L7/L12 (25). Interestingly, the consensus sequences that constitute the characteristic G-protein domains inscribed in the elongation factors, EF-Tu and EF-G, as well as in IF-2 (26), are missing or very weak in the EF-P sequence described here (Figure 3). Thus, EF-P may regulate a unique translation function which mediates protein chain propagation.

The inability to find that a single protein of the peptidyl transferase domain effects catalysis has led to speculation that the entire ribosome may act as an inert matrix to align the substrates. Many lines of evidence suggest that such a catalytic center may exist (1, 2). Some emphasis has been given to the proposal that a histidine residue in an essential protein forms a carbonyl intermediate analogous to that found in other acid/base catalyzed reactions (1, 2, 23). However, such an intermediate has not yet been observed in the model reaction (1, 2).

The conserved region of the 23S rRNA, domain V, has also been implicated as the catalytic center by antibiotic binding and mutational studies (27, 28). However, other 23S rRNA domains, *e.g.* IV, may also be involved (29). Thus, it is possible that the 50S particle has a structural and catalytic locus whose conformation is defined by RNA and several neighboring proteins. The regions surrounding such a site must be flexible in order to accommodate different amino acid side chains.

The activity of acceptors in the peptidyl transferase reaction depends largely on the nature of the aminoacyl moiety (1, 2, 30, 31). For example, 70S ribosomes cannot efficiently synthesize certain dipeptides from CA amino acids. EF-P stimulates the synthesis of fMet-Gly and fMet-Leu, for example, but not that of fMet-Phe (9). In this regard, several antibiotics (ie. anisomycin and chloramphenicol) inhibit peptide-bond synthesis with some aminoacyl-tRNAs, but not with others (30, 31). For example, chloramphenicol inhibits the synthesis of polyphenylalanine more than that of polylysine or polyproline (28, 30, 31). Interestingly, EF-P stimulates polylysine synthesis programmed with poly(A) more than polyphenylalanine synthesis directed by poly(U) (7). It is possible, therefore, that the protein helps to create a hydrophobic locus on the native 70S ribosome which is essential to the reaction and specificity of peptide-bond synthesis.

To study the structure and mechanism of EF-P action, the gene for the protein was cloned, mapped and sequenced as described below and in Materials and Methods.

Partial amino acid sequence analysis of EF-P

Purified EF-P was subjected to N-terminal amino acid analysis as well as CNBr cleavage and subsequent peptide sequencing, as described in Materials and Methods. The results indicated that the N-terminal methionine is cleaved from the protein. Given the amino acid sequence information, two DNA probes (18 and 24 nucleotides in length) were synthesized corresponding to the least degenerate and most commonly used *E.coli* codons (32).

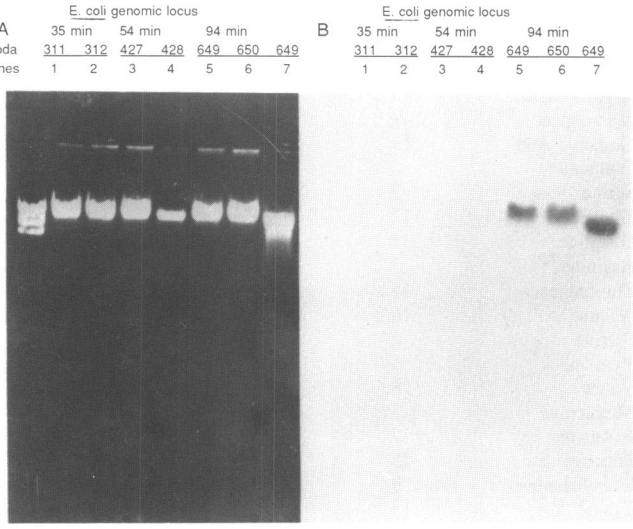

Figure 1. Southern blot hybridization of restriction fragments from bacteriophage λ clones of the Kohara *E.coli* genomic library bearing the 35, 54 and 94 minute regions of the *E.coli* chromosome. A: λ clones #311, 312 (located at 35 minutes), #427, 428 (at 54 minutes) and #649, 650 (at 94 minutes) were digested with BamHI (lanes 1–6) and HindIII (lane 7) and electrophoresed on an agarose gel as described in Materials and Methods. B: Autoradiogram of the Southern blot hybridized with the 24-nucleotide mixed probe corresponding to the downstream sequence of the EF-P gene.

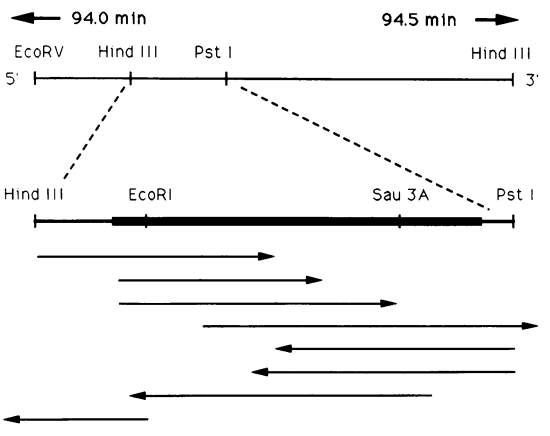

Figure 2. Restriction map and sequencing strategy for the EF-P gene. The top line is part of the *E.coli* genome encompassing the 94.3 minute region. The closed box indicates the coding region of the EF-P gene. The arrowed lines indicate the direction and extent of sequencing.

Screening of the Kohara Gene Bank

Screening with the 18-nucleotide probe (which corresponds to the N-terminal CNBr peptide) allowed the identification of several λ clones containing BamHI inserts of the *E.coli* chromosome corresponding to 35, 54 and 94 minutes of the Kohara Gene Bank (16). Agarose gel electrophoresis of λ clones #649 and #650 (Figure 1A), followed by Southern blot hybridization with the 24-nucleotide probe (Figure 1B) suggested that the EF-P gene maps to 94.3 minutes on the *E.coli* chromosome. Further, evidence that this is the correct map location is that the EF-P protein was identified in the same position in a gene-protein index by an independent electrophoretic method (33).

Cloning of the EF-P gene

The two positive clones were used to isolate and clone five fragments containing part or all of the EF-P gene. A 0.34 kb (EcoRI) fragment was obtained from 9B1 (λ clone #649) by hybridization with the 24-mer mixed probe. A 1.1 kb (EcoRI-EcoRV) fragment and a 0.51 (EcoRI-PstI) fragment were isolated from 1D4 (λ clone #650) by hybridization with the 18- and 24-mer probes, respectively. The entire EF-P gene was located within the 3.0 kb (HindIII) and the 0.78 kb (HindIII-PstI) fragments by hybridization of λ clone #650 with both probes. Both clones have the 5′ EF-P promoter region with the longer fragment containing an additional 2 kb of unsequenced region at the 3′ end. The location of the EF-P gene on the *E.coli*

```
-252 ACGCATCTGCGAGTAGTACTGACGCAAAGAGATCGTCGATGCCTTCTCCATGCGATCGATAAA

-189 GCTTTTTGGCGCTGCGTCCGGCTAACAGTTTTTCCTCCGCGTCTATATTCAAAAGACGCAGAA

-126 GTTCATCAGGATCGGTCACAACATCGGCAAGTTGCGTTAACCAATCTTCTCTGGATGGGGTAT

 -63 TTAGGGTTACAATATGCGCCATTTTGTGGCTTAGCTACCAATTAACAAATTTCAGAGGGCCTT

   1 ATG GCA ACG TAC TAT AGC AAC GAT TTT CGT GCT GGT CTT AAA ATC ATG

   1 M   A   T   Y   Y   S   N   D   F   R   A   G   L   K   I   M

  49 TTA GAC GGC GAA CCT TAC GCG GTT GAA GCG AGT GAA TTC GTA AAA CCG
  17 L   D   G   E   P   Y   A   V   E   A   S   E   F   V   K   P

  97 GGT AAA GGC CAG GCA TTT GCT CGC GTT AAA CTG CGT CGT CTG CTG ACC
  33 G   K   G   Q   A   F   A   R   V   K   L   R   R   L   L   T

 145 GGT ACT CGC GTA GAA AAA ACC TTC AAA TCT ACT GAT TCC GCT GAA GGC
  49 G   T   R   V   E   K   T   F   K   S   T   D   S   A   E   G

 193 GCT GAT GTT GTC GAT ATG AAC CTG ACT TAC CTG TAC AAC GAC GGT GAG
  65 A   D   V   V   D   M   N   L   T   Y   L   Y   N   D   G   E

 241 TTC TGG CAC TTC ATG AAC AAC GAA ACT TTC GAG CAG CTG TCT GCT GAT
  81 F   W   H   F   M   N   N   E   T   F   E   Q   L   S   A   D

 290 GCA AAA GCA ATT GGT GAC AAC GCT AAA TGG CTG CTG GAT CAG GCA GAG
  97 A   K   A   I   G   D   N   A   K   W   L   L   D   Q   A   E

 337 TGT ATC GTA ACT CTG TGG AAT GGT CAG CCG ATC TCC GTT ACT CCG CCG
 113 C   I   V   T   L   W   N   G   Q   P   I   S   V   T   P   P

 385 AAC TTC GTT GAA CTG GAA ATC GTT GAT ACC GAT CCG GGC CTG AAA GGT
 129 N   F   V   E   L   E   I   V   D   T   D   P   G   L   K   G

 433 GAT ACC GCA GGT ACT GGT GGC AAA CCG GCT ACC CTG TCT ACT GGC GCT
 145 D   T   A   G   T   G   G   K   P   A   T   L   S   T   G   A

 481 GTG GTT AAA GTT CCG CTG TTT GTA CAA ATC GGC GAA GTC ATC AAA GTG
 161 V   V   K   V   P   L   F   V   Q   I   G   E   V   I   K   V

 529 GAT ACC CGC TCT GGT GAA TAC GTC TCT CGC GTG AAG TAA TGCGGTTGTGG
 177 D   T   R   S   G   E   Y   V   S   R   V   K   Z

 579 TGCGCCTGCAGGCTGCACATCACTTATTCAGGTCAGAGATGATGAACGCTATCGTCTGTTGCT

 642 GCAGCACGCTGCTCACGGCTGTACACGTCGCGTTCGCATGCACACTGCACTCATTGCTGCAGC

 705 TATT
```

Figure 3. Nucleotide and deduced amino acid sequence of the EF-P gene. In the 5′ flanking region, the putative Pribnow box, the transcription start site and the Shine-Dalgarno sequence are underlined. In the coding region, the amino acids corresponding to the CNBr peptides are underlined. The amino acids that are double underlined indicate the position of the 18- and 24-oligonucleotide probes.

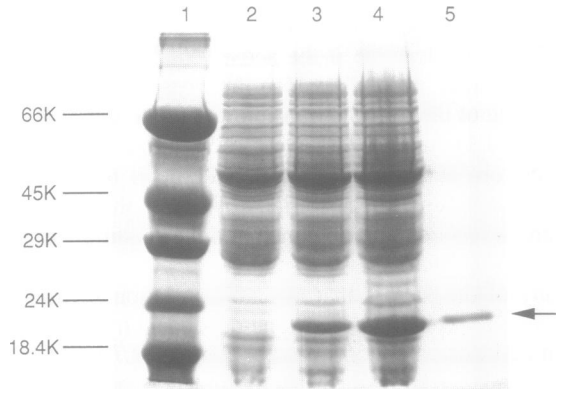

Figure 4. Detection of EF-P in *E.coli* extracts transformed with EF-P clones. SDS polyacrylamide gel electrophoretic analysis of ribosome-free XL1-Blue extracts. Lane 1 contains the molecular weight markers; bovine serum albumin, 66 kDa; ovalbumin, 45 kDa; carbonic anhydrase, 29 kDa; trypsinogen, 24 kDa; β-lactoglobulin, 18.4 kDa. Lane 2 contains extracts of pUC18 transformed cells. Lane 3 contains extracts transformed with pUC19 bearing the 0.78 kb EF-P insert. Lane 4 contains extracts transformed with pUC18 bearing the 3.0 kb EF-P insert. The amount of protein in lanes 2–4 is 58.2 μg. Lane 5 contains 2.6 μg of purified EF-P.

Table 1. Comparison of the amino acid composition of purified EF-P and the putative protein derived from the gene sequence.

Amino acid	DNA sequence	Amino acid analysis*
Aspartic acid	13	22.2 ± 0.7(Asp+Asn)
Asparagine	8	
Glutamic acid	13	16.1 ± 0.8(Glu+Gln)
Glutamine	5	
Serine	9	8.4 ± 0.3
Glycine	17	16.6 ± 0.0
Histidine	1	0.8 ± 0.0
Arginine	7	6.4 ± 0.2
Threonine	15	14.7 ± 0.3
Alanine	16	16.0†
Proline	8	8.2 ± 0.4
Tyrosine	6	5.2 ± 0.3
Valine	18	13.4 ± 0.8+
Methionine	3	3.3 ± 0.7
Isoleucine	7	5.3 ± 0.5+
Leucine	15	14.0 ± 1.0
Phenylalanine	9	8.6 ± 0.4
Lysine	13	11.1 ± 0.2
Cysteine	1	–
Tryptophan	3	–
Total	**187**	

*This column is an average of three analyses ± standard deviations.
†16 residues of alanine per mol is assumed.
+The significant low recovery of valine and isoleucine is due to the presence of acid hydrolysis resistant peptide bonds in EF-P (Val67-Val68, Ile114-Val115, Ile135-Val136, Val161-Val162, Val 173-Ile174).
– not determined.

chromosome, as well as the direction and extent of sequencing are shown in Figure 2.

Nucleotide sequence of the gene encoding EF-P

Sequencing was accomplished using all five fragments such that confirmation was achieved in both directions. The nucleotide and deduced amino acid sequence of the EF-P gene are given in Figure 3.

The EF-P gene is 567 base pairs long with a putative protein length of 187 amino acids and a predicted molecular weight of 20,447. This molecular weight was confirmed upon SDS-PAGE analysis of the purified protein (Figure 4). All the CNBr peptides are located within the deduced amino acid sequence. As well, the 18- and 24-oligonucleotide probes are located at nucleotides 22−39 and 253−276 from the initiation site of translation, respectively. These lines of evidence confirm the correspondence between the gene and the protein product.

A previously reported amino acid composition for EF-P determined by an independent method (8) agrees with the data presented in Table 1 except that Ser, Gly, and His were found to be much higher in the earlier determination. The His peak co-eluted with the Tris buffer used for EF-P isolation and Ser and Gly were probably non-protein contaminants. The hydrolyzer used in this study eliminated the manual multi-step hydrolysis procedure which frequently overestimates Ser and Gly when μg amounts of protein are hydrolyzed.

Residue 28 was assumed as Cys because PTH-Cys was not recovered and all other common PTH were identified in the system. The DNA sequence predicts Lys at position #31, but the amino acid sequence failed to confirm this position. Preliminary data suggests that Lys-31 is modified. The DNA sequencing was therefore repeated in both directions. This verified that Cys indeed occurs at position 82 and Lys at position 31 of the EF-P sequence.

The amino acid composition of the purified EF-P (Table 1) agrees well with that derived from the DNA sequence. This confirms again that the correct gene was cloned and sequenced. The deduced amino acid sequence of EF-P has 33% polar, 42% non-polar and 25% neutral amino acids. The isoelectric point calculated from the sequence is 5.8, which is very close to the estimated pI of 5.9 for the pure protein (38). The amino acid composition and sequence indicate that EF-P is a hydrophilic protein different from known ribosomal proteins (3). This is consistent with the purification procedures previously published (6, 13).

In Figure 3, the unique open reading frame starts with AUG and ends with UAA. The nucleotides neighbouring the initiation site are UUAUGG, which is a preferred sequence for initiation of translation in prokaryotes (34, 35). The region also contains a polypurine tract (GAGG) which corresponds to the 16S rRNA complementary region (36). A putative Pribnow box and an RNA polymerase start site are located 5′ to the Shine-Dalgarno sequence and may be the promoter region of the gene (37). A −35 region was not found, however.

Overproduction of EF-P from clones containing the EF-P gene

To confirm that the putative gene encodes the EF-P protein, *E.coli* cells containing clones of the whole EF-P gene or control cells containing only the vector were compared. SDS-PAGE analysis of *E.coli* XL1-Blue cell extracts containing pUC18 or the two EF-P clones with the whole gene (ie. having either the 3.0 or 0.78 kb insert) revealed that there was a 14-fold increase of protein that co-migrated with purified EF-P in the sample containing the 0.78 kb EF-P insert clone, compared to the sample containing only the vector (as determined by densitometry) (Figure 4). The fraction with the 3.0 kb EF-P insert had a 24-fold increase of this protein compared to the control.

Thus, the 0.78 and 3.0 kb fragments, when cloned into the multi-copy vectors pUC18 or pUC19, direct the synthesis and overexpression of a protein which co-migrates electrophoretically with pure EF-P. Partially purified cloned EF-P also showed the stimulation of the peptidyl transferase activity (unpublished data). As mentioned above, the overexpressed protein may not be modified as is naturally occurring EF-P. Thus it may be of interest to compare the kinetics of the overexpressed protein and of the naturally occurring EF-P.

The frequency of optimal codon use (F_{op}) for EF-P is 0.89, which reflects the preferential codon usage of *E.coli* genes expressed at high levels such as those encoding ribosomal proteins and *tuf*A coding for EF-Tu. The F_{op} of these genes is 0.90 and 0.93, respectively (39). As a comparison, the F_{op} of a gene expressed at mid-levels (*lac*Y coding for lactose permease) is 0.61, whereas an example of low level expression (*ara*C coding for a regulatory protein controlling transcription of the *ara*BAD operon) is 0.54 (39). The estimated cellular amount of EF-P in *E.coli* is one per 10 ribosomes (38) which is consistent with the observed high F_{op} value. Thus, EF-P may function with a turnover roughly analogous to that of initiation and termination factors, although not as high as that of EF-Tu and ribosomal proteins (3, 33).

Both the nucleotide and amino acid sequences are not represented in any known gene or protein databases as searched in the GenBank, release 67 and the SWISS-PROT, release 18, respectively. The predicted EF-P sequence differs from that of known prokaryotic ribosomal proteins, activating enzymes or translation factors. An interesting, but weak, homology to residues 47 through 183 of the eukaryotic IF-4E (*Saccharomyces cerevisiae*) cap binding protein does exist (40). Although these proteins appear to act in an altogether different manner, these regions may depict a conserved nucleic acid-protein recognition site or a conserved ribosome binding region. A homology to eIF-4D was carefully searched for since eIF-4D stimulates fMet-puromycin synthesis in reticulocyte cell-free systems after the 80S initiation complex is formed (41). However, no significant homology to this protein was detected. As well, EF-P is not homologous to yeast eIF-3 which may have a role in translation (42).

The unique sequence and chromosomal location identify EF-P as a novel translation factor. The cloning and sequencing of the gene for this protein should, therefore, facilitate a detailed study of the active center and the mechanism whereby it modulates synthesis. Several lines of evidence indicate that reactive histidine groups in proteins of the peptidyl transferase center participate in the peptide-bond synthetic process (23, 43). Interestingly, the stimulation of peptide-bond formation and transesterification by EF-P are also blocked by specific histidine modification (8, 13). Since EF-P has a unique histidine residue (His82), the nucleotides that correspond to this amino acid will be among the first targets for site-directed mutagenesis of the cloned gene. The cloned gene will also be useful in future studies in creating an EF-P mutant by gene replacement. This will enable experiments of the gene product in a background that contains no endogenous protein.

ACKNOWLEDGEMENTS

We thank Dr F. Blattner for his collaboration. We are grateful to P. Yaworsky, I. Kozieradzki and D.C. Watson for expert technical assistance. This work was funded by a grant from Medical Research Council of Canada to MCG.

REFERENCES

1. Nierhaus,K.H., Schulze,H. and Cooperman,B.S. (1980) *Biochem. Int.* **1**, 185–192.
2. Rychlik,I. and Cerna,J. (1980) *Biochem. Int.* **1**, 193–200.
3. Lucas-Lenard,J. and Beres,L. (1974) in *The Enzymes*, 3rd edn, ed. Boyer,P. (Academic Press, New York), Vol 10, pp. 53–86.
4. Traut,R.R. and Monro,R.E. (1964) *J. Mol. Biol.* **10**, 63–72.
5. Glick,B.R. and Ganoza,M.C. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 4257–4260.
6. Glick,B.R., Green,R.M. and Ganoza,M.C. (1979) *Can. J. Biochem.* **57**, 749–757.
7. Glick,B.R. and Ganoza,M.C. (1976) *Eur. J. Biochem.* **71**, 483–491.
8. Ganoza,M.C., Zahid,N.D. and Baxter,R.M. (1985) *Eur. J. Biochem.* **146**, 287–294.
9. Glick,B.R., Chladek,S. and Ganoza,M.C. (1979) *Eur. J. Biochem.* **97**, 23–28.
10. Green,R.H., Glick,B.R. and Ganoza,M.C. (1985) *Biochem. Biophys. Res. Com.* **126**, 792–798.
11. Glick,B.R. (1990) *J. Theor. Biol.* **82**, 149–153.
12. Ganoza,M.C., Baxter,R.M. and Fox,J.L. (1988) in *The Roots of Modern Biochemistry*, eds. Kleinkauf,H., von Dohren,H. and Jaenicke,L. (Walter de Gruyter and Co., Berlin-New York), pp. 551–554.
13. Chung,D.-G., Zahid,N.D., Baxter,R.M. and Ganoza,M.C. (1990) in *Ribosomes and Protein Synthesis; A Practical Approach*, ed. Spedding,G. (IRL Press, Oxford Publishing Co., Oxford), pp. 69–80.
14. Matsudaira,P. (1987) *J. Biol. Chem.* **262**, 10035–10038.
15. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
16. Kohara,Y., Akiyama,K. and Isono,K. (1987) *Cell* **50**, 495–508.
17. Meinkoth,J. and Wahl,G. (1984) *Anal. Biochem.* **138**, 267–284.
18. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
19. Pearson,W.R. and Lipman,D.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
20. Marck,C. (1988) *Nucl. Acids Res.* **16**, 1829–1836.
21. Maizel,J.V. (1969) in *Methods in Virology*, eds. Maramorosch,K. and Koprowski,H. (Academic Press, New York), Vol. 5, pp. 179–249.
22. Schulze,H. and Nierhaus,K.H. (1982) *EMBO J.* **1**, 609–613.
23. Sumpter,V.G., Tate,W.P., Nowotny,P. and Nierhaus,K.H. (1991) *Eur. J. Biochem.* **196**, 255–260.
24. Wittmann,H.G. (1983) *Ann. Rev. Biochem.* **52**, 35–65.
25. Glick,B.R. (1977) *FEBS Lett.* **73**, 1–5.
26. Bourne,H.R., Sanders,D.A. and McCormick,F. (1991) *Nature* **349**, 117–126.
27. Noller,H.F. (1984) *Ann. Rev. Biochem.* **53**, 119–162.
28. Vester,B. and Garrett,R.A. (1988) *EMBO J.* **7**, 3577–3587.
29. Beauclerk,A.A.D. and Cundliffe,E. (1988) *EMBO J.* **7**, 3589–3594.
30. Rychlik,I., Cerna,J., Chladek,S., Pulkrabek,P. and Zemlicka,J. (1970) *Eur. J. Biochem.* **16**, 136–142.
31. Kucan,Z. and Lipmann,F. (1964) *J. Biol. Chem.* **239**, 516–520.
32. Grantham,R., Gautier,C., Gouy,M., Jacobzone,M. and Mercier,R. (1981) *Nucl. Acids Res.* **9**, r43–r73.
33. VanBogelen,R.A., Hutton,M.E. and Neidhardt,F.C. (1990) *Electrophoresis* **11**, 1131–1166.
34. Ganoza,M.C., Sullivan,P., Cunningham,C., Hader,P., Kofoid,E.C. and Neilson,T. (1982) *J. Biol. Chem.* **257**, 8228–8232.
35. Eckhardt,H. and Luhrmann,R. (1981) *Biochemistry* **20**, 2075–2080.
36. Shine,J. and Dalgarno,L. (1975) *Nature* **254**, 34–38.
37. Hawley,D.K. and McClure,W.R. (1983) *Nucl. Acids Res.* **11**, 2237–2255.
38. An,G., Glick,B.R., Friesen,J.D. and Ganoza,M.C. (1980) *Can. J. Biochem.* **58**, 1312–1314.
39. Ikemura,T. (1981) *J. Mol. Biol.* **151**, 389–409.
40. Brenner,C., Nakayama,N., Goebl,M., Tanaka,K., Toh-E,A. and Matsumoto,K. (1988) *Mol. Cell. Biol.* **8**, 3556–3559.
41. Smit-McBride,Z., Dever,T.E., Merrick,W.C. and Hershey,J.W.B. (1989) *J. Biol. Chem.* **264**, 1578–1583.
42. Qin,S., Xie,A., Bonato,M.C.M. and McLaughlin,C.S. (1990) *J. Biol. Chem.* **265**, 1903–1912.
43. Baxter,R.M. and Zahid,N.D. (1986) *Eur. J. Biochem.* **155**, 273–277.