The EMBO Journal vol.3 no.8 pp.1825 – 1830, 1984

# Polypeptide chain elongation factor 1α (EF-1α) from yeast: nucleotide sequence of one of the two genes for EF-1α from *Saccharomyces cerevisiae*

**Shigekazu Nagata, Koji Nagashima, Yasuko Tsunetsugu-Yokota, Katsuya Fujimura[1], Masazumi Miyazaki[1] and Yoshito Kaziro**

Institute of Medical Science, University of Tokyo, Minatoku, Tokyo 108, and [1]Institute of Molecular Biology, School of Science, Nagoya University, Nagoya 464, Japan

Communicated by F.Chapeville

**Messenger RNA for yeast cytosolic polypeptide chain elongation factor 1α (EF-1α) was partially purified from *Saccharomyces cerevisiae*. Double-stranded complementary DNA (cDNA) was synthesized and cloned in *Escherichia coli* with pBR327 as a vector. Recombinant plasmid carrying yEF-1α cDNA was identified by cross-hybridization with the *E. coli tufB* gene and the yeast mitochondrial EF-Tu gene (*tufM*) under non-stringent conditions. A yeast gene library was then screened with the EF-1α cDNA and several clones containing the chromosomal gene for EF-1α were isolated. Restriction analysis of DNA fragments of these clones as well as the Southern hybridization of yeast genomic DNA with labelled EF-1α cDNA indicated that there are two EF-1α genes in *S. cerevisiae*. The nucleotide sequence of one of the two EF-1α genes (designated as *EF1αA*) was established together with its 5'- and 3'-flanking sequences. The sequence contained 1374 nucleotides coding for a protein of 458 amino acids with a calculated mol. wt. of 50 300. The derived amino acid sequence showed homologies of 31% and 32% with yeast mitochondrial EF-Tu and *E. coli* EF-Tu, respectively.**

*Key words:* elongation factor 1α/EF-1α gene sequence/*Saccharomyces cerevisiae*.

## Introduction

Prokaryotic polypeptide chain elongation factor Tu (EF-Tu) and its eukaryotic counterpart EF-1α promoting the GTP-dependent binding of aminoacyl-tRNA to the A site of ribosomes have been purified from various sources and their structure and function extensively studied (for a review, see Kaziro, 1978). EF-Tu from *Escherichia coli*, comprising 393 amino acid residues (Arai *et al.*, 1980) is encoded by two nearly identical but unlinked genes on the *E. coli* chromosome, designated as *tufA* (73 min) and *tufB* (89 min) (Jaskunas *et al.*, 1975). Both *tufA* (Shibuya *et al.*, 1979) and *tufB* (Miyajima *et al.*, 1979) have been cloned and their sequences determined (Yokota *et al.*, 1980; An and Friesen, 1980).

Eukaryotic cells possess two independent translational apparatuses; one functioning in the cytosolic fraction in conjunction with 80S ribosomes, and the other in the mitochondria with 70S ribosomes. The protein components of the mitochondrial translational apparatus including all the ribosomal proteins, aminoacyl-tRNA synthetases and soluble translational factors are encoded by nuclear genes, synthesized in the cytoplasmic fraction, and transported into the mitochondria.

In an attempt to study the genetic organization of eukaryotic translational machineries, we have screened yeast chromosomal DNA for homologous sequences with *E. coli tufB* and succeeded in cloning a gene for yeast mitochondrial EF-Tu (mtEF-Tu) (Nagata *et al.*, 1983). The gene designated as *tufM* contained a 1311 nucleotide long continuous reading frame coding for a protein of 437 amino acids. The derived amino acid sequence of *tufM* was 66% homologous to that of *E. coli tufA*.

Here, we describe the cloning of the yeast chromosomal gene coding for cytosolic EF-1α. Yeast cytosolic EF-1α has been purified to homogeneity, and the antibody against yeast EF-1α prepared (unpublished). Messenger RNA for EF-1α was partially purified, and cDNA for EF-1α, was prepared and cloned on plasmid pBR327. Two chromosomal genes for EF-1α were then isolated from the yeast gene library. The two genes for EF-1α are not closely linked to each other and the *tufM* gene was not found in the neighborhood of any of the *EF1α* genes. The sequence of one of the two *EF1α* genes was determined, and the deduced amino acid sequence was 70–80% homologous to the partial amino acid sequences reported for EF-1α from rabbit reticulocytes (Slobin and Möller, 1976) and *Artemia salina* (Amons *et al.*, 1983). The nucleotide sequence for yeast EF-1α was ~78% homologous to the partial nucleotide sequence of *A. salina* EF-1α cDNA (Van Hemert *et al.*, 1983).

## Results

### *Cloning of yeast EF-1α cDNA*

Since EF-1α is one of the most abundant proteins in eukaryotic cells (Slobin, 1980), it is expected that the mRNA for EF-1α is also an abundant species among the total cellular mRNAs. Therefore, we first purified the EF-1α mRNA by measuring its activity in an *in vitro* translation system using the antisera against yeast EF-1α. When 600 μg mRNA obtained from 10 g of wet cells was fractionated by sucrose density gradient centrifugation, the EF-1α mRNA activity was recovered as a single peak at the position corresponding to 18S rRNA. The active fractions were combined and re-chromatographed on an oligo(dT)-cellulose column and ~20 μg of the partially purified EF-1α mRNA was obtained. The purity of EF-1α mRNA at this stage was estimated to be ~3–4%, as judged from the amount of the $^{35}S$ radioactivity in the immunoprecipitates.

Double-stranded cDNA was prepared using the above partially purified mRNA, and joined to plasmid pBR327. After transformation of *E. coli* SK1592 with the hybrid DNA, 2800 tetracycline-resistant colonies were obtained. Identification of the clone harboring EF-1α cDNA was carried out as follows. Firstly, 132 clones strongly hybridizing with the labelled cDNA prepared by reverse transcription of the partially purified preparation of EF-1α mRNA were selected out of 1152 transformants by *in situ* colony hybridization (Hanahan and Meselson, 1980). Then plasmid DNA from 48 of the above clones were prepared (Holmes and Quigley, 1981),

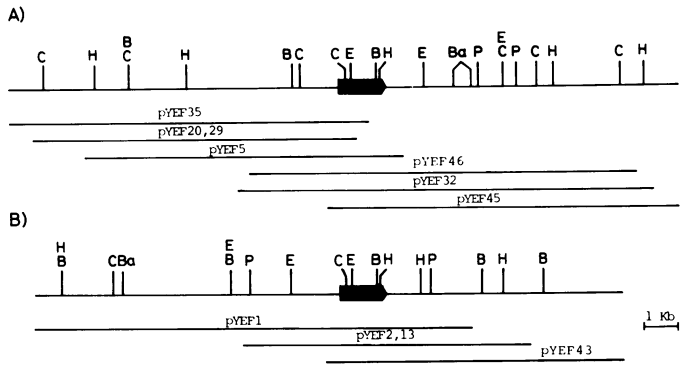


**Fig. 1.** Restriction maps of recombinant plasmids carrying the *EF-1α* genes. The solid bars indicate the *EF-1α* genes and the direction of transcription is from left to right. Restriction sites are abbreviated as: B, *Bgl*II site; Ba, *Bam*HI site; C, *Cla*I site; E, *Eco*RI site; H, *Hind*III site; P, *Pst*I site. kb denotes kilobases. **(A)** Maps of yeast DNA segments carrying the *EF1αA* gene. An additional site was present for each *Cla*I and *Hind*III, but was not mapped. **(B)** Maps of yeast DNA segments carrying the *EF1αB* gene.

electrophoresed on a 1% agarose gel, and transferred to a nitrocellulose membrane. Southern hybridization was carried out under non-stringent conditions (50% formamide, at 20°C) using a mixture of the 1.5-kb *Hpa*I fragment of *E. coli tufB* (Yokota *et al.*, 1980) and the 1.0-kb *Pvu*II fragment of yeast *tufM* (Nagata *et al.*, 1983), which were labelled by nick-translation. One plasmid which gave a positive response was designated as pKN31.

The 700-nucleotide long cDNA in pKN31 contained a single *Cla*I and *Eco*RI site, and DNA sequence analysis was carried out using the unique *Cla*I site. The nucleotide sequence corresponded to the position from 100 to 492 in Figure 4. The amino acid sequence derived from that nucleotide sequence revealed an extensive coincidence with the partial amino acid sequence of EF-1α from *S. carlsbergensis* which was determined with a large fragment obtained by limited proteolysis (data not shown). Twenty eight out of 30 amino acid residues were identical; at two positions (67 and 69 in Figure 4) arginine was replaced by lysine. The difference in amino acid sequences may be due to the difference of the yeast strains used for cDNA cloning (*S. cerevisiae*) and for preparation of EF-1α (*S. carlsbergensis*). From this result, we conclude that pKN31 is a clone carrying the 5′-half of the cDNA for yeast EF-1α.

*Cloning of the chromosomal gene for EF-1α*

Carlson and Botstein (1982) have prepared a recombinant plasmid library by insertion of the partial *Sau*3A digests of yeast genomic DNA into the *Bam*HI site of the yeast plasmid YEp24. This gene library was screened by colony hybridization (Hanahan and Meselson, 1980) using $^{32}$P-labelled EF-1α cDNA excised from pKN31. Forty six positive clones were identified by screening 4 x $10^4$ clones, and purified by repeated colony hybridization. Plasmid DNAs were prepared (Wilkie *et al.*, 1979) from 11 clones, and restriction sites for *Eco*RI, *Cla*I, *Hind*III, *Bgl*II, *Pst*I and *Bam*HI were determined by the conventional procedure. The gene for EF-1α was then located by Southern hybridization with the *Pst*I fragment of pKN31 cDNA.

As shown in Figure 1, the recombinant plasmids contained ~8−15 kb of the yeast genomic fragment. Plasmids pYEF20 and pYEF29, as well as pYEF2 and pYEF3 were independent isolates of the same clone. By comparing the

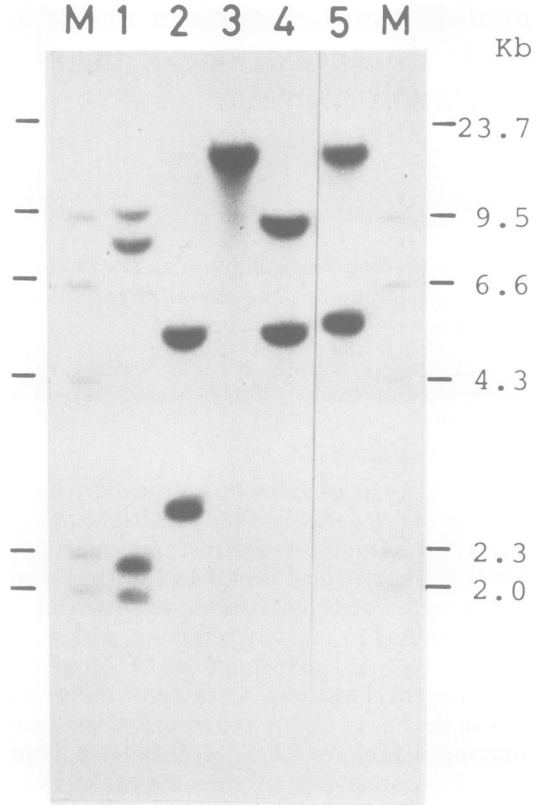


**Fig. 2.** Southern hybridization analysis of yeast DNA. 20 μg of yeast DNA was digested with several restriction enzymes and electrophoresed on 0.8% agarose gel. The DNA was transferred to a nitrocellulose membrane filter and hybridized with the nick-translated *Pst*I fragment of pKN31. Restriction enzymes used were *Eco*RI **(lane 1)**, *Bgl*II **(lane 2)**, *Bam*HI **(lane 3)**, *Hind*III **(lane 4)** and *Pst*I **(lane 5)**. 5′-$^{32}$P-labelled *Hind*III-digested lambda DNA was used as a size marker **(lane M)**.

restriction map, these clones could be classified into two groups. In the first group, the DNA inserts from the six clones pYEF35, 20, 5, 46, 32 and 45 overlapped each other and covered the range of 19.5 kb of the chromosomal DNA. The EF-1α gene in this group was designated as *EF1αA*. The second group consisting of pYEF1, 2, and 43, covered 17 kb of the contiguous chromosomal DNA on which the second EF-1α gene designated as *EF1αB* was present.

As can be seen in Figure 1, the restriction maps of the two genes are very similar within the structural genes for EF-1α, while the restriction maps of the 5′- and 3′-flanking regions are different. These results suggest that the two EF-1α genes are located in the different chromosomal loci of *S. cerevisiae* not closely linked to each other. Also, when the above cloned chromosomal DNA segments were analyzed by Southern hybridization under stringent conditions with *tufM* as a probe, no hybridizing fragment was observed (data not shown). Therefore, the *tufM* gene coding for mtEF-Tu does not seem to be linked to either of the two EF-1α genes.

*Southern hybridization of yeast genomic DNA with EF-1α cDNA*

To confirm the existence of two EF-1α genes, yeast genomic DNA was analyzed by Southern hybridization with labelled EF-1α cDNA. Yeast DNA was prepared from *S. cerevisiae* 106A (Cryer *et al.*, 1975) and cleaved completely with several restriction endonucleases. Figure 2 shows the results of Southern hybridization under stringent conditions (Wahl *et al.*, 1979). A single band was observed when the DNA was

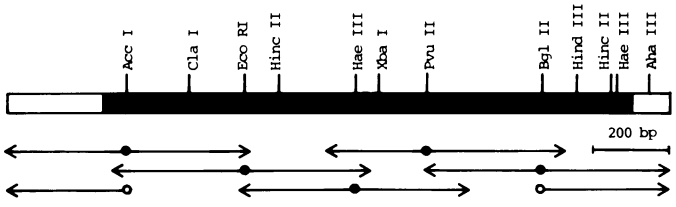


**Fig. 3.** Restriction map of the *EF1αA* gene and the strategy for nucleotide sequence analysis. The restriction map of the *EF1αA* gene on pYEF46 is shown in the figure. The sequence corresponding to the coding region is indicated by the shadowed box; the direction of transcription is from left to right. The circles represent labelled termini at 5′ (●) or 3′ (○) ends, and the solid arrows show the sequence read off the labelled fragments.

digested with *Bam*HI. Two bands were observed in DNA digested with *Hind*III, *Bgl*II or *Pst*I, while *Eco*RI digests gave four bands. Since cDNA of pKN31 had a single site for *Eco*RI but no site for *Bam*HI, *Hind*III, *Bgl*II and *Pst*I, the results indicate either that there are two EF-1α genes in the yeast genome or that the EF-1α gene is split by several intervening sequences. However, the latter possibility was excluded from the subsequent nucleotide sequence analysis of the cloned EF-1α gene (see Figure 4) which revealed no intervening sequences at least in the coding region of EF-1α gene.

The DNA fragments observed in Southern hybridization of yeast genomic DNA (Figure 2) could be assigned on the cloned DNA fragments. For example, the 2.2-kb *Eco*RI, 5.3-kb *Hind*III, 2.7-kb *Bgl*II and 15-kb *Pst*I fragments hybridizing with EF-1α cDNA were derived from the *EF1αA* gene while the 1.9-kb *Eco*RI, 9.0-kb *Hind*III, 5.0-kb *Bgl*II and 5.5-kb *Pst*I fragments were from the *EF1αB* gene. The 15-kb of *Bam*HI fragment is a double band consisting of two different fragments each containing the *EF1αA* or *EF1αB* gene, since two EF-1α genes cannot be on the single 15-kb *Bam*HI fragment judged from the restriction map shown in Figure 1.

*Nucleotide sequence analysis*

The restriction map of the *EF1αA* gene was constructed using pYEF46 and the nucleotide sequence was determined with fragments labelled either at the 5′ or the 3′ end by the strategy shown in Figure 3.

Figure 4 shows the nucleotide sequence of a 1713-bp long *EF1αA* gene. The part of the sequence was identical to the sequence of EF-1α cDNA originally cloned in pKN31 and S1 mapping experiments showed that EF-1α mRNA is co-linear with the *EF1αA* gene (data not shown). The 1374-nucleotide long open reading frame starts from the methionine codon ATG at positions 1−3 which was preceded by the termination codon UAG at positions −30 to −28 in the same reading frame. The termination codon for translation was assigned to UAA at 1375−1377. The *EF1αA* gene thus codes for a protein of 458 amino acid residues including the $NH_2$-terminal methionine with a calculated mol. wt. of 50 300. This value is in good agreement with the previous estimations of mol. wts. of 47 000−49 000 for yeast EF-1α (Dasmahapatra *et al.*, 1981 and unpublished).

It is likely that the $NH_2$-terminal amino acid is modified, since Edman degradation of the EF-1α from *S. carlsbergensis* and *S. pombe* did not release the terminal amino acid (unpublished). Modification of the N-terminal amino acid was also observed for EF-1α from rabbit reticulocytes (Slobin *et al.*, 1981) and pig liver (unpublished) as in the case of *E. coli* EF-Tu (Arai *et al.*, 1980).

The amino acid composition of EF-1α deduced from the nucleotide sequence analysis of the *EF1αA* gene agrees with the amino acid composition which was determined with EF-1α from *S. cerevisiae* (Dasmahapatra *et al.*, 1981) and *S. carlsbergensis* (unpublished). The number of the total basic amino acid residues greatly exceeded that of the total acidic amino acid residues; 49 lysine and 18 arginine residues being present as compared with 24 aspartic acid and 31 glutamic acid residues. The basic character of the EF-1α protein had previously been observed for EF-1α from yeast (Dasmahapatra *et al.*, 1981), pig liver (Nagata *et al.*, 1977), and *A. salina* (Slobin and Möller, 1976).

The codon usage of yeast genes has recently been tabulated by Bennetzen and Hall (1982). According to them, the codon bias could be correlated with the level of expression, that is, the codon usage for highly expressed genes is restricted, while that for genes expressed at low level is less biased. As shown in Table I, the *EF1αA* gene uses only 35 codons of 61 possible triplets, and a very strong preference (9.45%) was observed for 25 selective codons pointed out by Bennetzen and Hall (1982). On the other hand, the nuclear gene coding for yeast mitochondrial EF-Tu (*tufM*) uses 55 codons and only 65% of the total amino acid residues are coded by 25 selected codons. In agreement with this result, the *EF1αA* gene was found to be expressed at >10 times higher level than *tufM* (unpublished).

*Comparison of nucleotide and amino acid sequences of yeast EF1αA, tufM and E. coli tufB*

The nucleotide sequence of the yeast *EF1αA* gene is homologous to yeast *tufM* and *E. coli tufB* by 44.1% and 44.8%, respectively. Figure 5 shows the comparison of the amino acid sequence of yeast EF-1αA, yeast mtEF-Tu, and *E. coli* EF-Tu. To obtain the maximal homology, three gaps in EF-1αA sequence, seven gaps in yeast mtEF-Tu and 12 gaps in *E. coli* EF-TuA were introduced. Without counting the gaps, the protein sequence of yeast EF-1αA shows homologies of 31.2% with yeast mtEF-Tu and 32.2% with *E. coli* EF-Tu, while mtEF-Tu and *E. coli* EF-Tu have a homology of 64%. As seen in Figure 5, successive five to six amino acids residues match perfectly in four blocks. In addition, changes among yeast EF-1αA, mtEF-Tu and *E. coli* EF-Tu are conservative with respect to the chemical properties of the amino acid side chains and to the observed evolutionary amino acid replacements. If the replacement with those amino acids were considered as homologous, the homologies amount to 55% between EF-1αA and mtEF-Tu, and 52% between EF-1αA and *E. coli* EF-Tu.

## Discussion

We have isolated the yeast nuclear gene coding for mtEF-Tu (*tufM*) utilizing the cross-hybridization with *E. coli tufB* under non-stringent conditions (Nagata *et al.*, 1983). This simple and effective procedure has again been used successfully to identify yeast EF-1α cDNA. EF-1α mRNA was first enriched to ~3−4% and bacterial clones harboring the cDNA for the abundant mRNA species was then screened by colony hybridization. A plasmid containing EF-1α cDNA was finally identified by Southern hybridization with a mixture of *E. coli tufB* and *S. cerevisiae tufM* as probes. The hybridization condition was extremely non-stringent but even under this condition non-specific hybridization was not observed. The nucleotide sequence analysis of the cDNA revealed only ~40% homologies with *E. coli tufB* and yeast

```
        *       *       *     -200       *       *       *       *     -150       *       *       *
CTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTTGAAAATTTTTTTTTTTGATTTTTTTCTCTTTCGATGACCTCCCATTGATATTTAAGTTA

        *     -100       *       *       *       *      -50       *       *       *       *
ATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGCTCATTAGAAAGAAAGCATAGCAATCTAATCTAAGTTTTAATTACAAA

                                                                50
ATG GGT AAA GAG AAG TCT CAC ATT AAC GTT GTC GTT ATC GGT CAT GTC GAT TCT GGT AAG TCT ACC ACT ACC GGT CAT TTG ATT TAC AAG
Met Gly Lys Glu Lys Ser His Ile Asn Val Val Val Ile Gly His Val Asp Ser Gly Lys Ser Thr Thr Thr Gly His Leu Ile Tyr Lys
  1                                10                                  20                                  30

          100                                                   150
TGT GGT GGT ATT GAC AAG AGA ACC ATC GAA AAG TTC GAA AAG GAA GCC GCT GAA TTA GGT AAG GGT TCT TTC AAG TAC GCT TGG GTT TTG
Cys Gly Gly Ile Asp Lys Arg Thr Ile Glu Lys Phe Glu Lys Glu Ala Ala Glu Leu Gly Lys Gly Ser Phe Lys Tyr Ala Trp Val Leu
                                 40                                  50                                  60

                        200                                                   250
GAC AAG TTA AAG GCT GAA AGA GAA AGA GGT ATC ACT ATC GAT ATT GCT TTG TGG AAG TTC GAA ACT CCA AAG TAC CAA GTT ACC GTT ATT
Asp Lys Leu Lys Ala Glu Arg Glu Arg Gly Ile Thr Ile Asp Ile Ala Leu Trp Lys Phe Glu Thr Pro Lys Tyr Gln Val Thr Val Ile
                                 70                                  80                                  90

                                300                                                               350
GAT GCT CCA GGT CAC AGA GAT TTC ATC AAG AAC ATG ATT ACT GGT ACT TCT CAA GCT GAC TGT GCT ATC TTG ATT ATT GCT GGT GGT GTC
Asp Ala Pro Gly His Arg Asp Phe Ile Lys Asn Met Ile Thr Gly Thr Ser Gln Ala Asp Cys Ala Ile Leu Ile Ile Ala Gly Gly Val
                                100                                 110                                 120

                                                    400                                                           450
GGT GAA TTC GAA GCC GGT ATC TCT AAG GAT GGT CAA ACC AGA GAA CAC GCT TTG TTG GCT TTC ACC TTG GGT GTT AGA CAA TTG ATT GTT
Gly Glu Phe Glu Ala Gly Ile Ser Lys Asp Gly Gln Thr Arg Glu His Ala Leu Leu Ala Phe Thr Leu Gly Val Arg Gln Leu Ile Val
                                130                                 140                                 150

                                                        500
GCT GTC AAC AAG ATG GAC TCC GTC AAA TGG GAC GAA TCC AGA TTC CAA GAA ATT GTC AAG GAA ACC TCC AAC TTT ATC AAG AAG GTT GGT
Ala Val Asn Lys Met Asp Ser Val Lys Trp Asp Glu Ser Arg Phe Gln Glu Ile Val Lys Glu Thr Ser Asn Phe Ile Lys Lys Val Gly
                                160                                 170                                 180

          550                                                   600
TAC AAC CCA AAG ACT GTT CCA TTC GTC CCA ATC TCT GGT TGG AAC GGT GAC AAC ATG ATT GAA GCT ACC ACC AAC GCT CCA TGG TAC AAG
Tyr Asn Pro Lys Thr Val Pro Phe Val Pro Ile Ser Gly Trp Asn Gly Asp Asn Met Ile Glu Ala Thr Thr Asn Ala Pro Trp Tyr Lys
                                190                                 200                                 210

                        650                                                   700
GGT TGG GAA AAG GAA ACC AAG GCC GGT GTC GTC AAG GGT AAG ACT TTG TTG GAA GCC ATT GAC GCC ATT GAA CAA CCA TCT AGA CCA ACT
Gly Trp Glu Lys Glu Thr Lys Ala Gly Val Val Lys Gly Lys Thr Leu Leu Glu Ala Ile Asp Ala Ile Glu Gln Pro Ser Arg Pro Thr
                                220                                 230                                 240

                                750                                                               800
GAC AAG CCA TTG AGA TTG CCA TTG CAA GAT GTT TAC AAG ATT GGT GGT ATT GGT ACT GTG CCA GTC GGT AGA GTT GAA ACC GGT GTC ATC
Asp Lys Pro Leu Arg Leu Pro Leu Gln Asp Val Tyr Lys Ile Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Val Ile
                                250                                 260                                 270

                                                    850                                                           900
AAG CCA GGT ATG GTT GTT ACT TTT GCC CCA GCT GGT GTT ACC ACT GAA GTC AAG TCC GTT GAA ATG CAT CAC GAA CAA TTG GAA CAA GGT
Lys Pro Gly Met Val Val Thr Phe Ala Pro Ala Gly Val Thr Thr Glu Val Lys Ser Val Glu Met His His Glu Gln Leu Glu Gln Gly
                                280                                 290                                 300

                                                        950
GTT CCA GGT GAC AAC GTT GGT TTC AAC GTC AAG AAC GTT TCC GTT AAG GAA ATC AGA AGA GGT AAC GTC TGT GGT GAC GCT AAG AAC GAT
Val Pro Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Glu Ile Arg Arg Gly Asn Val Cys Gly Asp Ala Lys Asn Asp
                                310                                 320                                 330

          1000                                                  1050
CCA CCA AAG GGT TGC GCT TCT TTC AAC GCT ACC GTC ATT GTT TTG AAC CAT CCA GGT CAA ATC TCT GCT GGT TAC TCT CCA GTT TTG GAT
Pro Pro Lys Gly Cys Ala Ser Phe Asn Ala Thr Val Ile Val Leu Asn His Pro Gly Gln Ile Ser Ala Gly Tyr Ser Pro Val Leu Asp
                                340                                 350                                 360

                        1100                                                  1150
TGT CAC ACT GCT CAC ATT GCT TGT AGA TTC GAC GAA TTG TTG GAA AAG AAC GAC AGA AGA TCT GGT AAG AAG TTG GAA GAC CAT CCA AAG
Cys His Thr Ala His Ile Ala Cys Arg Phe Asp Glu Leu Leu Glu Lys Asn Asp Arg Arg Ser Gly Lys Lys Leu Glu Asp His Pro Lys
                                370                                 380                                 390

                                1200                                                              1250
TTC TTG AAG TCC GGT GAC GCT GCT TTG GTC AAG TTC GTT CCA TCT AAG CCA ATG TGT GTT GAA GCT TTC AGT GAA TAC CCA CCA TTA GGT
Phe Leu Lys Ser Gly Asp Ala Ala Leu Val Lys Phe Val Pro Ser Lys Pro Met Cys Val Glu Ala Phe Ser Glu Tyr Pro Pro Leu Gly
                                400                                 410                                 420

                                                    1300                                                          1350
AGA TTC GCT GTC AGA GAC ATG AGA CAA ACT GTC GCT GTC GGT GTT ATC AAG TCT GTT GAC AAG ACT GAA AAG GCC GCT AAG GTT ACC AAG
Arg Phe Ala Val Arg Asp Met Arg Gln Thr Val Ala Val Gly Val Ile Lys Ser Val Asp Lys Thr Glu Lys Ala Ala Lys Val Thr Lys
                                430                                 440                                 450

                                                        1400        *       *       *       *   1450        *
GCT GCT CAA AAG GCT GCT AAG AAA TAAGGAGATTGATAAGACTTTTCTAGTTGCATATCTTTTATATTTAAATCTTATCTATTAGTTAATTTTTTGTAATTTATCCTTATA
Ala Ala Gln Lys Ala Ala Lys Lys
                            458

        *       *
TATAGTCTGGTTATTCTAAAA
```

**Fig. 4.** The nucleotide sequence of the yeast *EF1αA* genes. The sequence of the yeast *EF1αA* gene is presented along with the deduced amino acid sequence.

**Table I.** Codon usage in the *EF1αA* gene and *tufM* gene

| | *EF1αA* | *tufM* | | *EF1αA* | *tufM* | | *EF1αA* | *tufM* | | *EF1αA* | *tufM* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Phe UUU | 2 | 9 | Ser UCU° | 14 | 12 | Tyr UAU | 0 | 3 | Cys UGU° | 6 | 1 |
| UUC° | 15 | 4 | UCC° | 6 | 5 | UAC° | 8 | 6 | UGC | 1 | 2 |
| Leu UUA | 3 | 14 | UCA | 0 | 2 | Ter UAA | 1 | 0 | Ter UGA | 0 | 0 |
| UUG° | 21 | 12 | UCG | 0 | 0 | UAG | 0 | 1 | Trp UGG° | 6 | 0 |
| Leu CUU | 0 | 6 | Pro CCU | 0 | 6 | His CAU | 5 | 10 | Arg CGU | 0 | 5 |
| CUC | 0 | 2 | CCC | 0 | 2 | CAC° | 6 | 4 | CGC | 0 | 1 |
| CUA | 0 | 2 | CCA° | 23 | 9 | Gln CAA° | 12 | 11 | CGA | 0 | 0 |
| CUG | 0 | 1 | CCG | 0 | 2 | CAG | 0 | 1 | CGG | 0 | 0 |
| Ile AUU° | 17 | 17 | Thr ACU° | 14 | 10 | Asn AAU | 0 | 6 | Ser AGU | 1 | 0 |
| AUC° | 13 | 12 | ACC° | 14 | 10 | AAC° | 16 | 7 | AGC | 0 | 0 |
| AUA | 0 | 2 | ACA | 0 | 11 | Lys AAA | 3 | 10 | Arg AGA° | 18 | 19 |
| Met AUG° | 8 | 18 | ACG | 0 | 4 | AAG° | 46 | 12 | AGG | 0 | 4 |
| Val GUU° | 26 | 19 | Ala GCU° | 30 | 19 | Asp GAU | 8 | 19 | Gly GGU° | 42 | 28 |
| GUC° | 19 | 6 | GUU° | 7 | 10 | GAC° | 16 | 5 | GGC | 0 | 2 |
| GUA | 0 | 3 | GCA | 0 | 6 | Glu GAA° | 30 | 26 | GGA | 0 | 3 |
| GUG | 1 | 5 | GCG | 0 | 0 | GAG | 1 | 7 | GGG | 0 | 5 |

°Codons preferentially used in yeast (Bennetzen and Hall, 1982).

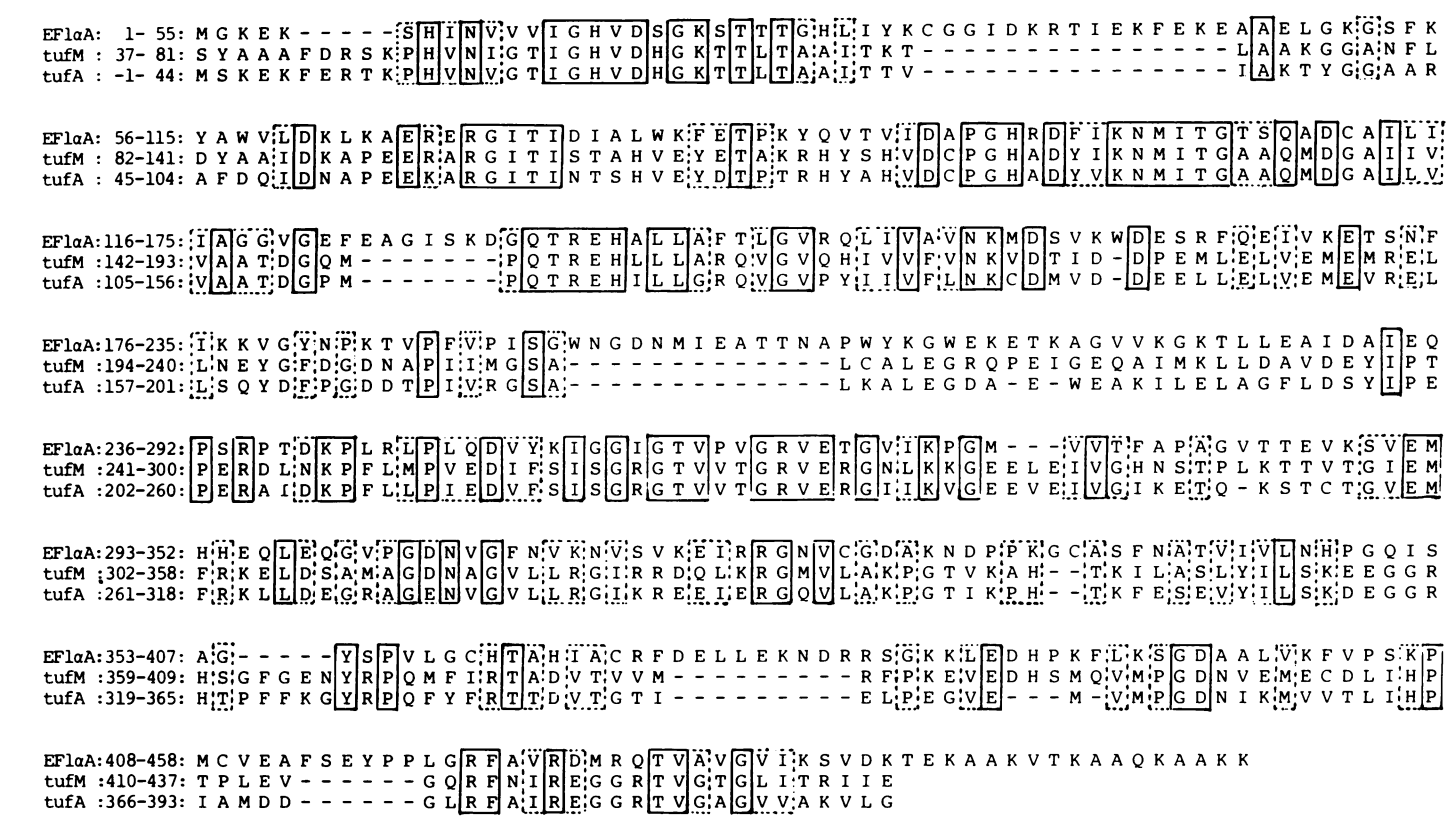

**Fig. 5.** Comparison of the amino acid sequence of yeast EF-1αA, mtEF-Tu and *E. coli* EF-TuA. The amino acid sequences of yeast EF-1αA, mtEF-Tu (Nagata *et al.*, 1983) and *E. coli* EF-TuA (Yokota *et al.*, 1980) are aligned to give maximal homology by introducing several gaps (−). The one-letter amino acid notation is used.The amino acid residue number 1 in yeast EF-1αA and mtEF-Tu are tentatively assigned to methionine at the initiator codon ATG while that for *E. coli* EF-TuA to serine which was found at the N-terminal of *E. coli* EF-Tu (Arai *et al.*, 1980). Sets of three identical amino acid residues at one aligned position are enclosed within solid lines, and sets of three residues regarded as favored substitutions are enclosed within dotted lines. Favored amino acid substitutions are defined as pairs of residues belonging to one of the following groups: S, T, P, A and G; N, D, E and Q; H, R and K; M, I, L and V; F, Y and W (Dayhoff *et al.*, 1978).

*tufM*. This procedure for isolation of EF-1α cDNA may be applicable to isolation of other yeast genes possessing the evolutionary homology with *E. coli* genes.

By comparison of the restriction maps of the chromosomal DNA segments containing the EF-1α gene (Figure 1) and by Southern hybridization analysis of yeast genomic DNA with EF-1α cDNA (Figure 2), two genes for yeast EF-1α, *EF1αA* and *EF1αB*, were detected. It has been known that the EF-1α proteins purified from pig liver (Nagata *et al.*, 1977), *A. salina* (Slobin and Möller, 1976) and yeast (Dasmahapatra *et al.*, 1981) are heterogeneous with respect to their isoelectric points. The molecular heterogeneity of the protein may be explained, at least in part, by the multiplicity of the EF-1α gene.

The two EF-1α genes in *S. cerevisiae* do not seem to be closely linked to each other (see Figure 1), neither was the *tufM* gene located in the neighborhood of either of the two EF-1α genes. In general, it has not yet been definitively settled whether or not two independent genes exist for the enzymes catalyzing the same activity in cytoplasmic fraction and mitochondria. In the present study, the three genes *EF1αA*, *EF1αB* and *tufM* are apparently unlinked. However, in the case of a tRNA modification enzyme, Hopper *et al.* (1982) have shown that a single mutation in yeast abolished the synthesis of $N^2$,$N^2$-dimethyl-guanosine on the cytoplasmic and mitochondrial tRNAs. On the other hand, Beauchamp *et al.* (1977) proposed a model in which cytosolic and mitochondrial leucyl-tRNA synthetases are encoded by the two closely linked genes since a single mutation in promoter affected the expression of both enzymes.

EF-1α and EF-Tu interact with aminoacyl-tRNA and GDP or GTP. Chemical modification (Arai *et al.*, 1974) and photooxidation studies (Nakamura and Kaziro, 1981) with *E. coli* EF-Tu suggested that Cys-81 and either His-66, 75, 78 or 84, may be involved in interaction with aminoacyl-tRNA. Although those amino acid residues except His-84 were replaced by other non-conservative amino acid residues in the *EF1αA* gene (Figure 5), the region from Tyr-69 to Gly-110 of *E. coli* EF-Tu was well conserved (70% homology considering conservative amino acid replacement as homologous) among yeast *tufM* and *EF1αA* (Figure 5). Recently, Leberman and Egner (1984) deduced the structure of the nucleotide binding site of *E. coli* EF-Tu by comparing the amino acid sequence of the six GTP-binding proteins. The sequence from Pro-113 to Val-140 of *E. coli* EF-Tu showed remarkable homology (75%) with the yeast *tufM* and *EF1αA* genes (Figure 5).

Partial amino acid sequence (Amons *et al.*, 1983) of EF-1α from *A. salina* and its cDNA sequence (Van Hemert *et al.*, 1983) were reported, although the cDNA was not full length and contained a termination codon in the middle of the coding sequence. A comparison of amino acid sequence of yeast *EF1αA* with that of *A. salina* EF-1α showed 225 identical amino acids out of 287 amino acids (78.4%), while the nucleotide sequence was homologous by 77.4% (425 identical out of 522).

## Materials and methods

*Preparation of double-stranded cDNA and transformation*

Total RNA was prepared from *S. cerevisiae* 106 A (α Arg) according to Struhl and Davis (1981) and poly(A) mRNA was enriched by oligo(dT)-cellulose column chromatography (Aviv and Leder, 1972). EF-1α mRNA was partially purified by centrifugation at 35 000 r.p.m. for 12 h on a 5–23% sucrose density gradient as described (Nagata *et al.*, 1980). Each fraction was assayed for EF-1α mRNA activity in an *in vitro* translation system using the antiserum against yeast EF-1α as described (Nagata *et al.*, 1983). The active fractions were pooled and the poly(A) mRNA was recovered by oligo(dT)-cellulose column chromatography.

Double-stranded cDNA was prepared essentially as described (Hoeijmaker *et al.*, 1980) using sucrose density gradient-purified poly(A) mRNA (15 μg). The DNA was size-fractionated by electrophoresis on a 1.0% agarose gel, and the material longer than 750 bp was recovered from the gel (Vogelstein and Gillespie, 1979). The cDNA was elongated with dCMP residues, annealed to dGMP-elongated, *Pst*I-cleaved pBR327 as described (Hoeijmakers *et al.*, 1980), and used to transform *E. coli* SK1592 (F *gal thi* T1 *endA sbc15 hsdR*$^4$ *hsdM*$^+$) cells. The efficiency of transformation was 4 x $10^6$ colonies/μg DNA.

*Hybridization, restriction mapping and nucleotide sequence determination*

Colony hybridization was carried out as described by Hanahan and Meselson (1980). Southern hybridization was done according to Wahl *et al.* (1979). When non-stringent conditions were to be used, the hybridization temperature was lowered to 20°C, and the filter was washed at 30°C with 1.5 mM sodium citrate (pH 6.45) containing 15 mM NaCl and 0.1% SDS.

Highly labelled cDNA (> 5 x $10^6$ c.p.m./μg) was prepared by reverse transcription of the sucrose density gradient purified EF-1α mRNA using reverse transcriptase and [α-$^{32}$P]dCTP (3000 Ci/mmol, Amersham, Japan) in the presence of calf thymus DNA primer (Shank *et al.*, 1978). Nick-translation of the DNA fragments was carried out with [α-$^{32}$P]dCTP as described (Nagata *et al.*, 1983) to yield the labelled DNA (1.0 x $10^8$ c.p.m./μg). Restriction mapping and DNA sequencing were carried out as previously described (Nagata *et al.*, 1983).

## Acknowledgements

We are grateful to Dr D.Botstein for his kind supply of the yeast gene library.

## References

Amons,R., Pluijms,W., Roobol,K. and Möller,W. (1983) *FEBS Lett.*, **153**, 37-42.
An,G. and Friesen,J.D. (1980) *Gene*, **12**, 33-39.
Arai,K., Kawakita,M., Nakamura,S., Ishikawa,I. and Kaziro,Y. (1974) *J. Biochem. (Tokyo)*, **76**, 523-534.
Arai,K., Clark,B.F.C., Duffy,L., Jones,M.D., Kaziro,Y., Laursen,R.A., L'Italien,J., Miller,D.L., Nagarkatti,S., Nielsen,K.M., Petersen,T.E., Takahashi,K. and Wade,M. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 1326-1330.
Aviv,H. and Leder,P. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1408-1412.
Beauchamp,P.M., Horn,E.W. and Gross,S.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 1172-1176.
Bennetzen,J.L. and Hall,B.D. (1982) *J. Biol. Chem.*, **257**, 3026-3031.
Carlson,M. and Botstein,D. (1982) *Cell*, **28**, 145-154.
Cryer,D.R., Ecdeshall,R. and Marmur,J. (1975) *Methods Cell Biol.*, **12**, 39-44.
Dasmahapatra,B., Skogerson,L. and Chakraburtty,K. (1981) *J. Biol. Chem.*, **256**, 10005-10011.
Dayhoff,M.O., Schwartz,R.M. and Oroutt,B.C. (1978) in *Atlas of Protein Sequence and Structure*, Vol. **5**, Suppl. 3, pp. 345-352.
Hanahan,D. and Meselson,M. (1980) *Gene*, **10**, 63-67.
Hoeijmakers,J.H.J., Borst,P., Van Den Burg,J., Weissmann,C. and Cross, G.A.M. (1980) *Gene*, **8**, 391-417.
Holmes,D.S. and Quigley,M. (1981) *Anal. Biochem.*, **114**, 193-197.
Hopper,A.K., Furukawa,A., Pham,H.D. and Martin,N.C. (1982) *Cell*, **28**, 543-550.
Jaskunas,S.R., Lindahl,L., Nomura,M. and Burgess,R.R. (1975) *Nature*, **257**, 458-462.
Kaziro,Y. (1978) *Biochim. Biophys. Acta*, **505**, 95-127.
Leberman,R. and Enger,U. (1984) *EMBO J.*, **3**, 339-341.
Miyajima,A., Shibuya,M. and Kaziro,Y. (1979) *FEBS Lett.*, **102**, 207-210.
Nagata,S., Iwasaki,K. and Kaziro,Y. (1977) *J. Biochem.*, **82**, 1633-1646.
Nagata,S., Taira,H., Hall,A., Johnsrud,L., Streuli,M., Ecsödi,J., Boll,W., Cantell,K. and Weissmann,C. (1980) *Nature*, **284**, 316-320.
Nagata,S., Tsunetsugu-Yokota,Y., Naito,A. and Kaziro,Y. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 6192-6196.
Nakamura,S. and Kaziro,Y. (1981) *J. Biochem. (Tokyo)*, **90**, 1117-1124.
Shank,P.R., Cohen,J.C., Varmus,H.E., Yamamoto,K.R. and Ringold,G.M. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 2112-2116.
Shibuya,M., Nashimoto,H. and Kaziro,Y. (1979) *Mol. Gen. Genet.*, **170**, 231-234.
Slobin,L.I. (1980) *Eur. J. Biochem.*, **110**, 555-563.
Slobin,L.I. and Möller,W. (1976) *Eur. J. Biochem.*, **69**, 351-366.
Slobin,L.I., Clark,R.V. and Olson,M.O.J. (1981) *Biochemistry (Wash.)*, **20**, 5761-5767.
Struhl,K. and Davis,R.W. (1981) *J. Mol. Biol.*, **152**, 535-552.
Van Hemert,F.J., Vanormondt,H. and Möller,W. (1983) *FEBS Lett.*, **157**, 289-293.
Vogelstein,B. and Gillespie,D. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 615-619.
Wahl,G.M., Stern,M. and Stark,G.R. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 3683-3687.
Wilkie,N.M., Clements,J.B., Boll,W.M., Mantei,N., Lonsdale,D. and Weissmann,C. (1979) *Nucleic Acids Res.*, **7**, 859-877.
Yokota,T., Sugisaki,H., Takanami,M. and Kaziro,Y. (1980) *Gene*, **12**, 25-31.

*Received on 11 May 1984*

**Note added in Proof**

Recently, the primary structure of EF-1α from *A. salina* has been published (van Hemart *et al.* (1984) *EMBO J.*, **3**, 1109-1113). We thank the authors for sending their preprint before publication.