Molecular cloning and sequence determination of the nuclear gene coding for mitochondrial elongation factor Tu of Saccharomyces cerevisiae

(protein synthesis/tuf gene probe/Southern hybridization/sequence conservation)

Shigekazu Nagata, Yasuko Tsunetsugu-Yokota, Ayako Naito, and Yoshito Kaziro

Institute of Medical Science, University of Tokyo, Minatoku, Takanawa, Tokyo 108, Japan

Communicated by Severo Ochoa, July 8, 1983

ABSTRACT A 3.1-kilobase Bgl II fragment of Saccharomyces cerevisiae carrying the nuclear gene encoding the mitochondrial polypeptide chain elongation factor (EF) Tu has been cloned on pBR327 to yield a chimeric plasmid pYYB. The identification of the gene designated as tufM was based on the cross-hybridization with the Escherichia coli tufB gene, under low stringency conditions. The complete nucleotide sequence of the yeast tufM gene was established together with its 5'- and 3'-flanking regions. The sequence contained 1,311 nucleotides coding for a protein of 437 amino acids with a calculated M_r of 47,980. The nucleotide sequence and the deduced amino acid sequence of tufM were 60% and 66% homologous, respectively, to the corresponding sequences of E. coli tufA, when aligned to obtain the maximal homology. Plasmid YRpYB was then constructed by cloning the 2.5kilobase EcoRI fragment of pYYB carrying tufM into a yeast cloning vector YRp-7. A mRNA hybridizable with tufM was isolated from the total mRNA of S. cerevisiae D13-1A transformed with YRpYB and translated in the reticulocyte lysate. The mRNA could direct the synthesis of a protein with M_r 48,000, which was immunoprecipitated with an anti-E. coli EF-Tu antibody but not with an antibody against yeast cytoplasmic EF-1 α . The results indicate that the tufM gene is a nuclear gene coding for the yeast mitochondrial EF-Tu.

The polypeptide chain elongation factor Tu (EF-Tu) promotes a GTP-dependent binding of an aminoacyl-tRNA to the A site of ribosomes (1). EF-Tu from *Escherichia coli* consists of a single polypeptide chain with M_r 43,000, and the primary structure comprised of 393 amino acid residues has been determined (2). The protein is encoded by two nearly identical genes on the *E. coli* chromosome (3), *tufA* at 73 min and *tufB* at 89 min (4). Both *tufA* (5) and *tufB* (6) have been cloned and their nucleotide sequences determined. The sequences of *tufA* (7) and *tufB* (8) are nearly homologous and differ only in 13 positions but the gene products, EF-TuA and EF-TuB, are identical except for the COOH-terminal amino acid (2).

In eukaryotes, the counterpart of prokaryotic EF-Tu, designated as EF-1 α , has been purified from various sources, including pig liver (9), rabbit reticulocytes (10), Artemia salina (11), wheat germ (12), and yeast (13), and was shown to consist of a single polypeptide chain, M_r 47,000–53,000. The partial amino acid sequence of EF-1 α from rabbit reticulocytes (14) and A. salina (15) was determined, and sequence conservation between A. salina EF-1 α and E. coli EF-Tu has been reported (15).

In addition to EF-1 α , which functions in the cytoplasmic fraction in conjunction with 80S ribosomes, eukaryotic cells possess mitochondrial EF-Tu (designated as mEF-Tu) that

functions in the mitochondrial translational apparatus (16, 17). Translational factors as well as ribosomal proteins in the mitochondria are encoded by nuclear genes, synthesized in cytoplasmic fractions, and transported into the mitochondria (18). The translational machineries of mitochondria have been thought to be closer to the prokaryotic ones than to the machinery present in eukaryotic cytoplasm (17, 19). However, the organization of the nuclear genes for the mitochondrial translational apparatus, their expression, and the incorporation of the products into the mitochondria are not well understood.

In this report, we describe the molecular cloning of the yeast mitochondrial EF-Tu gene (tufM) utilizing hybridization with the *E. coli tufB* gene under low stringency conditions. The nucleotide sequence as well as the deduced amino acid sequence of the tufM gene indicate that the structures of *E. coli* EF-Tu and yeast mitochondrial EF-Tu are remarkably conserved during evolution. The mRNA for tufM can be translated in the cell-free system from rabbit reticulocytes and the product has been identified as mEF-Tu by immunoprecipitation with an antibody against *E. coli* EF-Tu.

MATERIALS AND METHODS

Enzymes. Endonuclease *Eco*RI and T4 polynucleotide ligase were gifts from Takashi Yokota. Other restriction enzymes and T4 polynucleotide kinase were purchased from Takara Shuzo (Kyoto, Japan) or Bethesda Research Laboratories and were used essentially as recommended by the supplier (except for the use of 200 μ g of gelatin per ml instead of bovine serum albumin). *E. coli* DNA polymerase I (large fragment) was purchased from Boehringer Mannheim.

Southern Hybridization. DNA was prepared from various sources including Saccharomyces cerevisiae 106A (mating type α , arginine requiring) as described by Cryer *et al.* (20) and completely digested with several restriction enzymes. Fragments were separated by electrophoresis on 1% agarose gel (21) and transferred to a nitrocellulose filter (Sartorius, 0.45-µm pore) as described by Southern (22). The 1.5-kilobase (kb) Hpa I fragment of pTUB1 (6) was subcloned at the Pvu II site of pBR322 to yield a hybrid plasmid pYT-1 which was used as a probe. This fragment covers the entire coding sequence of E. coli EF-TuB except for the NH₂-terminal 12 amino acids (see ref. 8). pYT-1 was labeled by nick-translation (23) with $[\alpha^{-32}P]dCTP$ (3,000 Ci/mmol; 1 Ci = 3.7×10^{10} Bq; Amersham Japan) to yield ≈ 1 $\times 10^8$ cpm/µg and was hybridized with yeast DNA fragments essentially as described by Wahl et al. (24), except that the hybridization temperature was lowered to 28°C, and the filter was washed at 44°C with 15 mM NaCl/1.5 mM sodium citrate, pH 7.0/0.1% NaDodSO₄.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EF, polypeptide chain elongation factor; bp, base pairs; kb, kilobase(s).

Biochemistry: Nagata et al.



FIG. 1. Southern hybridization analysis of *E. coli* and yeast DNA. *E. coli* DNA (100 ng) (lanes 1–4) or yeast DNA (10 μ g) (lanes 5–8) was digested with several restriction enzymes and electrophoresed on a 1% agarose gel. DNA was transferred to a nitrocellulose membrane filter and hybridized with nick-translated pYT-1 (1 × 10⁸ cpm/ μ g). Restriction enzymes used were *Eco*RI (lanes 1 and 5), *Bgl* II (lanes 2 and 6), *Eco*RI and *Bgl* II (lanes 3 and 7), and *Pst* I (lanes 4 and 8). A mixture of the *Sma* I single-digested and the *Sma* I and *Eco*RI double-digested pYT-1 (lane 9) and the *Eco*RI, *Pvu* II, and *Pst* I triple-digested pBR322 (lane 10) were run as size markers. Sizes of marker DNA fragments are given in kb.

Restriction Mapping and DNA Sequence Determination. Restriction sites were determined by a single or double digestion of the cloned DNA. In some instances the arrangement of the restriction fragments was determined by Smith-Birnstiel mapping (25) of the labeled fragments. DNA sequence analysis was carried out by the chemical degradation method (26) with restriction fragments labeled either at the 5' end with $[\gamma^{-32}P]$ -ATP (prepared as described in ref. 27) and T4 polynucleotide kinase or at the 3' end with $[\alpha^{-32}P]$ dCTP and *E. coli* DNA polymerase I (large fragment).

Hybridization-Translation. S. cerevisiae D13-1A (a his3-532 trp1 gal2) was transformed with YRpYB. This was constructed by cloning of the 2.5-kb EcoRI fragment containing the yeast nuclear gene for mitochondrial EF-Tu [cloned in pYYB (see Results)] into the yeast vector YRp7 (28). A transformant was grown on a minimal essential medium as described (29), except that

2% glucose was replaced by 2% lactate. Total RNA was prepared according to Struhl and Davis (30) and poly(A)-containing mRNA was selected by oligo(dT)-cellulose column chromatography (31).

The 2.5-kb EcoRI fragment of pYYB or EcoRI-digested pBR327 DNA was bound to a nitrocellulose filter (Schleicher & Schüll, 7×7 mm, 0.45- μ m pore) and the filters were hybridized with 20 μ g of mRNA prepared as above (32). The filter-hybridized mRNA was eluted and translated in the reticulocyte lysate system (33). Immunoprecipitation of the translational products with anti-sera and protein A adsorbent was carried out as described (34). The anti E. coli EF-Tu antibody (35) was a gift from A. Miyajima. The antibody against yeast cytoplasmic EF-1 α was a generous gift from M. Miyazaki (Nagoya University).

RESULTS

Presence of a Sequence Homologous to the Gene for E. coli EF-Tu in Eukaryotes. First, we searched for a sequence homologous to E. coli tufB in DNA from various eukaryotic sources, including yeast, mouse liver, and human leukocytes utilizing low stringency conditions (36). Because the original plasmid pTUB1 containing tufB possessed other E. coli sequences, such as a part of rrnB, genes for four tRNAs (thrT, glyT, tyrU, and thrU), a gene for an unidentified protein "U," and a part of rplK (6), we constructed plasmid pYT-1 by subcloning the 1.5-kb Hpa I fragment of pTUB1 into pBR322 and used this as a probe. The fragment covered almost the entire coding sequence of tufB as well as about 300 nucleotides of the 3'-flanking region.

As shown in Fig. 1, *E. coli* DNA (lanes 1-4) gave two bands hybridizing with the *tufB* probe due to the presence of two genes for EF-Tu (*tufA* and *tufB*) in *E. coli* (3). On the other hand, a single DNA band was detected with yeast DNA (lanes 5-8), indicating that at least one sequence homologous to the *E. coli* EF-Tu gene is present in yeast. Mouse liver and human leukocyte DNA did not give any detectable band hybridizing with pYT-1 under the same conditions (data not shown), suggesting that the sequence of the gene in mouse or human cells is more divergent.

Isolation of the Yeast Gene Homologous to the E. coli EF-Tu Gene. Two hundred micrograms of DNA from S. cerevisiae 106A was digested with Bgl II, and fragments (3.1 kb in length) hybridizing with pYT-1 were isolated by agarose gel electrophoresis. The fragments were ligated to pBR327 at the BamHI



FIG. 2. Restriction map of the yeast DNA fragment cloned on pYYB and the strategy for nucleotide sequence determination. Restriction map of the yeast Bgl II fragment cloned at the *Bam*HI site of pBR327 is shown in the upper part of the figure with major restriction sites. The lower part shows the expanded structure of the yeast tufM gene. The sequence corresponding to the coding region is indicated by the shadowed box; the direction of transcription is left to right. The circles represent labeled termini at 5' (\bullet) or 3' (\odot) ends, and the solid arrows show the sequence read off of the labeled fragments. bp, base pairs.

																					(<u>tu</u>	<u>£M)</u> 1	rric/	TAT	FTTT	-220 AAOG	ATTT	IGTT	ITAG	CACO	-200 CATO	CGA
-180 CCTCAGTCAATA	TATOCTT	70007	-160 GACCA	VGGC1	ma	nca	- 	-140 IGCT	CTA	CTO	TTA	CAGAT	-120 TTTR	CTAT	7777	OGTO	-	100 TTAT	стт	GAAA	CTGA	TTAA	-80 CTG	~~~~	AATT	IGAG	CTTC	-60 FTTG	TIGT	AAAC	TATT	TTG
-40 TGCTTTCAGITT	таттста	-20 GCTCC) Gaca <i>i</i>	AGGI	raac <i>i</i>	GAC/	-1 ••••	l ATG Met	TCA Ser	GCT Ala	TTA Leu	TTA Leu	CCA Pro	AGA Arg	TTA Leu	CTC Leu	30 ACA Thr	AGA Arg	ACA Thr	GCT Ala	TTT . Phe i	AAA (Lys i	CT 1	ICT (Ser (GOG /	AAA Lys	60 CTT (Leu)	CTG /	AGG (Arg)	CTC -	TCT ' Ser :	ICA Ser
GTA ATT TCT	90 AGG ACC	TTT	тст	CAA	ACT	ACT	ACT	1 TCC	ŤAŤ	ĢCĂ	120 GCT	сст	iii	ĠĂT	ĊĊŤ	тċс	10 ÅÅÅ	ĊĊĠ	ĊĂT	ĠŤA_	150 ÅÅT	ATA (ict i	icc i	ĂŤĊ (ööč	20 ČÅT (iii (ĜĂT (ĊĂT (180 366 /	AAG -
Val Ile Ser	Arg Thr 30	Phe	Ser	Gin	Thr	Thr	Thr	Ser	Tyr 1 Ser	Ala Lys	Ala 40 Glu	Ala Lys	Phe. Phe	Asp Glu	Arg Arg	Ser Thr	Lys Lys	Pro 10 Pro	His His	Val Val	Asn 50 Asn	Ile Val	Gly (fhr fhr	Ile (Gly Gly_	His His	Val / 20 Val /	Asp.	His (Gly 1 60 Gly 1	Lys Lys
in the min		***	:	:*~	210	(<u>ti</u>	<u>IIA</u>)	1	ici		GAA	AAA	TTT im	GAA	240	AÇA	30	ççç	çac	ĢŢT	AAC /	GTT. (ACT .	270		60	ят (iei	GAC ·	çạc ·	QGT	••••
The The Leu	Thr Ala	Ala 30	Ile	Thr	Lys 70	Thr	Leu	Ala	Ala	Lys	Gly	Gly 40	Ala	AAC	Phe 80	Leu	Asp	TAT	Ala	Ala	Ile	GAT Asp 50	Lys	Ala	Pro 90	Glu	Glu	Arg	Ala	Arg	Gly	ATT Ile 60
ACT ACT CTG		ĢÇĂ	ATC.	ÂÇC	ÂCC	GTA	cic	ĢÇT		ACC	TAC 120		ĢCT	GCT	GCT	CGT	ĢCA	TIC	ASP GAC	CAG	ATC 150	Asp Çat	ASN (GAA GAA	GAA GAA	AAA		Arg ÇÇT		ATC
ACA ATT TCT Thr Ile Ser 100	<u>ĂĈŤ</u> GĈA Thr Ala	<u>ČĂĊ</u> His	<u>Ġi</u> G Val	<mark>ĠĂĂ</mark> Glu	ŤĂĊ Tyr	GÅA Glu 70	ĂĊG Thr	GČC Ala	AAG Lys 110	AĜA Arg	CĂT His	ŤÅT Tyr	TĊT Ser	ĊĂĊ His	ĠĨC Val	GĂĊ Asp 80	İĞT Cys	ĊĊA Pro	GGC G1y 120	CAC His	ĠĊŢ Ala	ĠĂT Asp	iic Tyr	ATC . 11e	ÅÅG . Lys	ÅÅT Asn 90	ÅİĞ Met	ÅŤT Ile	ACC Thr 130	<u>ởởi</u> Gly	ĠĊŤ Ala	<u>ČĊŤ</u> Ala
Thr Ile Asn	Thr Ser AÇT TÇT	His ÇAÇ	Val ÇTT	Glu GAA	Tyr TFC 210	Asp GAC	Thr ACC	Pro CÇG	Thr ACC	Arg CÇT	His ÇAC	Tyr Tạc	Ala OÇA	His ÇAÇ	Val GTA 240	Asp ÇAÇ	Cys TÇC	Pro ÇÇG	Gly ççç	His ÇAÇ	Ala CC	Asp GAC	Tyr Tật	Val GTT	Lys AAA 270	Asn AAC	Met ATÇ	Ile ATC	<u>Thr</u> AÇÇ	<u>Gly</u> Çî	Ala ÇÇT	Ala 997
<u>ČĂA ĂÎĞ ĞĂT</u> Gln Met Asp	<mark>ởềT ổ</mark> CT Gly Ala	ĂŤĊ 11e	420 ATT Ile	ĠŤT Val.	ĠŤA Val	<u>ĠĊŤ</u> Ala	<u>ĠĊT</u> Ala	ACC Thr	ĠÅT Asp	ĠĠA Gly	ĊAA Gln	ATG Met	450 CCC Pro	ĊÅA Gln	ÁĊŤ Thr	AĞA	ĠĂA Glu	ĊĂT His	TŤA Leu	ĊŤT Leu	TİĞ Leu	åcc Ala	480 AGĂ Arg	<u>ČÅA</u> Gln	<u> ČŤT</u> Val	<mark>ĠĠT</mark> Gly	<u>ởic</u> Val	Ċ AA Gln	CÅT His	ÅİT Ile	GİĊ Val	ĞİT Val
Gin Met Asp CAG ATC CAC 300	100 Gly Ala GGC GGG	Ile ATC	Leu CŢG	Val ÇTA	Val ÇT	Ala QCT	Ala ÇÇG	Thr ACT	Asp GAC 330	110 Gly ÇÇC	Pro çcc	Met ATG	Pro QCG	Gln Ç A G	Thr ACT	Arg CÇT	<u>Glu</u> Ç A G	His ÇAC	Ile ATC 360	120 Leu Ç T G	Leu CŢÇ	Gly ÇCT	Arg OÇT	<u>Gln</u> Ç A G	Val ÇTA	<u>Gly</u>	Val ÇTT	Pro çCC	Tyr TAC 390	130 11e ATC	Ile AŢÇ	Val ĢļG
510 TTT GTT AAC Phe Val Asn	<u>ÅÅG</u> GTT Lys Val	ĜÅT Asp	ÂCC Thr	AŤŤ Ile	ĞÅŤ Asp	ĠÅT Asp	540 CCA Pro	ĠĂA Glu	AŤĠ	<u>TŤA</u> Leu	ĠÅG Glu	TŤA Leu	<u>Ĝîc</u> Val	ĠĂĂ Glu	ÅŤĠ Met	ĠĂĂ Glu	570 AľG Met	AĠA Arg	ĠĂĂ Glu	ĊŤŤ Leu	TTA	AAC Asn	GÅA Glu	İÅT Tyr	ÖGG Gly	iir Phe	600 GAC Asp	<mark>ост</mark> Gly	<u>ĠĂT</u> Asp	AÅT Asn	GĈŤ Ala	ČČA Pro
170 Phe Leu Asn TTC CTG AAC	Lys Cys	Asp	Met ATG	140 Val GTT	Asp GAT	Asp GAC	180 Glu GAA	Glu ÇAG	Leu CŢĢ	Leu CIG	Glu GAA	Leu CIG	Val	150 Glu GAA	Met ATG	Glu	190 Val GTT	Arg CÇT	Glu GAA	Leu CTT	Leu	Ser TCT	Ģln CAG	160 Тут ТАС	Asp ÇAC	Phe TIC	200 Pro CCG	Gly ÇÇC	Asp GAC	Asp GAC	Thr AÇŢ	Pro ÇÇG
ATT ATC ATG		<u></u>		TOC	ÖČT	TİĞ	ĠĂĂ Glu	<u>ÖĞT</u>		CAÁ	660 CCT	ĠÅA G I J	ATT	GÖĞ Gly	ČÅG	CAĜ	GCC	ÅŤĊ	AŤĜ	AÅÅ LVS	690 CTT	TIG	ĜAT Aso	GCA Ala	GÎĜ Val	ĠÅŤ IAsti	GAG Glu	<u>ŤÅŤ</u>	<u>ĂŤŤ</u>	<u>ČČŤ</u> Prod	720 ACA	
170 11e Val Arg	210 Gly Ser	Ala	Leu	Lys	Ala	Leu	Glu	180 Gly	Asp	Ala	220	Glu		Trp	Glu	Ala	Lys	Ile	Leu	190 Glu GAA	230 Leu CTG	Ala	Gly	Phe	Leu CTG	Asp	Ser TCT	TYT	Ile ATT	200 Pro	240 Glu GAA	Pro
510 Gầa aga gat	tỉg mắc	ÄÄG	йт йт	Ħċ	750 TTG	AŤĠ	548 <u>cc</u>	GİT	ĠĂĂ	ĠÅT	AŤC	 Ħċ	îĊT	ĂŤĊ	780 780	ĠĠŦ	AĠA	Ğ	578 ÅCT	čic	<u>Gic</u>	ÄĊT	661	<u>čči</u>	810 GTG	ĠĂĂ	NĠG	<u>àci</u>	600 Åat	 ТТА	Ä	
Glu Arg Asp Glu Arg Ala	Leu Asn Ile Asp	Lys	Pro Pro	Phe 210 Phe	Leu 250 Leu	Met Leu	Pro Pro	Val Ile	Glu	Asp	Ile Val	Phe Phe	Ser Ser	Ile 220 Ile	Ser 260 Ser	Gly	Arg	Gly Gly	Thr Thr	Val Val	Val Val	Thr Thr	Gly Gly	Arg 230 Arg	Val 270 Val	Glu Glu	Arg Arg	Gly Gly	Asn Ile	Leu Ile	Lys Lys	Lys Val
<u>ÇŅG CĢT</u> ĢOG	AŢT GŅÇ	ANG 	000 630	TIC	cić	cić		AŢC	670 870	Ģ	. GŢA	TTÇ	. TCC 660	ATC	TX:	997	्र	997	, vốc 300	ĢŢT	ĢŢT	NCC.	990 990	 	ĢŢA	ĊŴ	ogc	997 1779	ATC 930	AŢC		GTT
GCT GAG GAA Gly Glu Glu 240 280	TTG GAA Leu Glu	Ile	Val	GCT	CAC His	AAC Asn	TCC Ser	ACC Thr 250	CCA Pro 290	Leu	Lys	ACA Thr	Thr	Val	Thr	Gly	Ile	GAA	Met 300	Phe	Arg	Lys	GAA Glu	Leu	Asp	Ser	Ala	Met	Ala 270	Gly	Asp	Ast
GIY GIU GIU GGT GAA GAA 720	VAL GIU GIT GNA	Â		<u>60</u>	ATÇ	Lys , AAA	GAG 750	ACT.	ÇAG	-	AAG AAG	ser TÇT	ACC 990	TGŢ	ACT		GŢŢ	GAA 780	AŢĢ	TTC	â	ÂÂA	CTG	CŢÇ	ĢļÇ	GAA	çoc	007 810	Ģ	901	Ģ	ŅĊ
GCC GCT CTT Ala Gly Val	TTA ĈTT Leu Leu	AĞA Arg	GLY 320	<u>ĂİĊ</u> Ile	ÅGG Arg	AĜA Arg	ĞÅT Asp	C ÀÀ Gln	TIC Lev	AÅ G Lys	AĞA Arg	Gly	ATG Met	GİC Val	TŤA Leu		AAG Lys	Pro	Gly	ÂCC Thr	GŤT Val	ääa Lys	GOC Ala 340	CĂT His	ACA Thr	Lys	ĂŤT Ile	CTĂ Leu	où: Ala	TCI Ser	Leu	Tyr
Val Gly Val GIA GGT GTT	Leu Leu CIG ÇIG	Arg COT 840	Gly GÇT	Ile AÇ	Lys ÀAA	Arg CÇT	Glu GAA	Glu	Ile AŢC	Glu GŅA	Arg	G1y GC1 870	Gln C A G	Val ÇTA	CIG	Ala ÇŢ	Lys	Pro QQC	Gly GOC	Thr ACC	Ile AŢC	Lys AAG 900	Pro CÇG	His CAC	Thr ACC	Lys AMG	Phe	Glu GAA	Ser TÇI	Glu GAM	Val GŢĢ	Tyr TAC 930
1050 ATT TTA TOC Ile Leu Ser	ĂĂĂ ĠĂG Lys Glu	<u>ĠĂĂ</u> Glu	<u>ост</u> Gly	<u>ÖĞT</u> Gly	AGA Arg	ĊĂŤ His	1080 TCT Ser	GCG Gly	iii Phe	GGT Gly	GÅÅ Glu	AAČ Asn	ŤÅĊ Tyr	AĜA Arg	ČĊA Pro	ĊÅA Gln	1110 ATG Met	ŤTT Phe	AŤA Ile	AĜA Arg	ÁĊA Thr	GČŤ Ala	<u>ĠĂT</u> Asp	<u>ĠŤT</u> Val	ÁĊA Thr	ĠTĪ Val	1140 GTG Val	ÅTG Met	AGĂ Arg	TT Phe	CT Pro	AÅG Lys
310 Ile Leu Ser ATT CIG TOC	Lys Asp AAA GAT	Glu ÇAA	Gly GC	Gly ÇÇC	Arg CCT	His CAT 960	Thr ACT	Pro CCÇ	Phe	Phe	Lys	Gly ccç	Tyr TAC	Arg	Pro	Gln CAG 990	330 Phe TTC	Tyr TAC	Phe TIC	Arg CÇT	Thr	Thr AÇŢ	Asp GAC	Val ÇŢG	Thr	Gly GCT 1020	Thr	Ile ATC	Glu GAA	: Leu CŢG	Pro	Glu G ŅA
ĠAG ĠĨT ĠĂA Glu Val Glu	1170 GAT CAT Asp His	TCT Ser	ÁİĞ Met	C AA Gln	GİT Val	ÅİĞ Met	ČČA Pro	ĠĠŢ Gly	ĠĂĊ Asp	ÅÅT Asn	1200 GTT Val	GÅÅ Glu	<u>ÅŤĠ</u> Met	GAA Glu	тст Суз	GAT	TİĞ	<u>ÅŤĊ</u> Ile	CĂT His	ČČT Pro	1230 ACC Thr	CĊA Pro	TŤA Leu	ĠÅA Glu	GTT Val	G Gly	Ċ AA Gln	C Arg	<u>tič</u> Phe	AA1 Asn	1260 AIC	AGA Arg
Gly Val Glu GGC GTA GAG	390		Met		350 Val GTA	Met ATG	Pro	Gly	Asp	Asn	400 Ile ATC	Lys AAA	Met	Val	360 Val GTT	Thr	Leu CTG	Ile ATC	His	Pro	410 Ile ATC	Ala GÇG	Met AŢG	Asp GAC	370 Asp GAC	Gly	Leu ÇTG	Arg	Phe	Ala GCA	420 11e	Arg Cçî
čie (*** ***	1. I.	ر مىيە	.050	ر مرد	1290	~	***	×-	~~~	,		Å	+1-	1080	320		C 3 3 C		(7 73 m	9 1111		135	0	1110		~~~		TA 1.] •••••••	.380		
Glu Gly Gly	Arg Thr	Val	Gly	Thr	G1y 430	Leu	Ile	ACA Thr	Arg	Ile	Ile	Glu	IAG	ACT	TAT	GAIG	CAAC	ruca	STAT	ATT	CIAL	AIAI.	iCIG	11CA		uu	CICA		AIA]	.AU11	G117	Jort
Glu Gly Gly GAA GGC GGC 1140	Arg Thr CÇT ÂÇC	Val GTT	Gly çç	Ala GÇG	Gly ççC	Val GTT	Val GTT	Ala GCT 1170	Lys	Val GTT	Leu CŢG	Gly çcc	ŢŅA	i -																		

 $\frac{1410}{1440} \frac{1440}{1470} \frac{1470}{1470}$

FIG. 3. Comparison of the nucleotide sequences of yeast tufM and E. coli tufA gene and the deduced amino acid sequences. The sequence of the E. coli tufA gene is from Yokota *et al.* (7). The sequences of tufM and tufA are aligned to give maximal homology by introducing several gaps in tufA. Identical amino acids are framed and identical nucleotides are marked by a dot.

site, and the hybrid plasmids were used to transform *E. coli* SK1592 (F gal thi T1 endA sbc15 hsdR⁴ hsdM) by the method of Hoeijmakers et al. (37), to yield 768 ampicillin-resistant and tetracycline-sensitive clones.

Clones containing the yeast tufM gene were identified by Southern hybridization of plasmid DNA with a tufB probe under low stringency conditions. Plasmid DNA was prepared (38) from pools of 24 bacterial clones, and $0.5-1.0 \mu g$ of DNA from each pool was electrophoresed on a 1% agarose gel, to separate host DNA, and transferred to a nitrocellulose membrane filter. As a probe, we used the 1.5-kb Hpa I fragment containing *E. coli tufB* prepared from plasmid pTUB311 (unpublished data). Because pTUB311 is an R-derivative plasmid containing tufB, any vector DNA that might contaminate the Hpa I fragment would not hybridize with the sequence of pBR327 (a colE1 derivative) (39) in the hybrid plasmids.

One of 32 groups of 24 clones gave a positive hybridization with the 32 P-labeled *Hpa* I fragment, whereas DNA from other groups or pBR327 gave no hybridization, even at the low stringency conditions used for the present experiments. Plasmid DNA from each individual clone of this group was then prepared and analyzed as above, and a clone that gave a positive result was designated as pYYB.

Nucleotide Sequence Analysis. Fig. 2 shows the physical restriction map of pYYB constructed as described in Materials and Methods and the strategy for the DNA sequence analysis. Southern hybridization analysis of pYYB with E. coli tufB showed that most of the yeast sequence homologous to tufB resides within the 1.0-kb Pvu II fragment (data not shown). Therefore we determined the sequence of this region as well as the regions flanking it, as indicated in Fig. 2. The nucleotide sequence (1,712 bp) thus obtained is shown in Fig. 3 together with the sequence of the E. coli tufA gene (7). The translation initiation site for yeast tufM was assigned to the methionine codon AUG at nucleotide positions 1-3, because this was the first AUG triplet downstream of the nonsense codon UAA (positions -12 to -10) found in the open reading frame of the nucleotide sequence aligned for maximal homology with E. coli tufA (e.g., positions 124-149). The termination codon for translation was assigned to UAG at positions 1,318-1,320 in the same reading frame. Because we could not find any intron in the coding sequence of tuf M (unpublished data), the gene codes for a protein of 437 amino acids including the NH2-terminal methionine with a calculated M_r of 47,800, in a single open reading frame of 1,311 nucleotides.

A comparison of the sequence of yeast tufM with that of E. coli tufA reveals that the homology is 60% and 66% for the nucleotide and amino acid sequences, respectively. To obtain maximum homology, the initiator methionine for tufA was aligned to Ser-37 of tufM, and a limited number of gaps were introduced into the sequence of tufA. As seen in Fig. 3, several distinct regions are highly conserved between tufM and tufA. A most remarkable homology was found in amino acid residues 74-118 of tufA and 111-155 of tufM, where 41 of 44 amino acid residues were identical (91% homology). Because this region is supposed to be an active site for interaction with aminoacyl-tRNAs (reviewed in ref. 1; see also ref. 40), the sequence conservation of this region might be due to a functional requirement. Other homologies were found in amino acid sequences 5-32, 208-234, and 371-386 of tufA with the corresponding sequences of tufM, where 25 of 28, 24 of 27, and 13 of 16 amino acid residues are identical, respectively. Whether or not these conserved regions constitute functional domains of the protein remains to be determined.

The Yeast tufM Gene Codes for the Mitochondrial Factor. To determine whether the yeast tufM gene homologous to bacterial tufA and tufB codes for the cytoplasmic or the mitochondrial factor, the mRNA specified by tufM was isolated by hybridization to the cloned tufM gene and translated in the reticulocyte cell-free system (33). As shown in Fig. 4, the tufMmRNA markedly stimulated the incorporation of [³⁵S]methionine into a protein of M_r 48,000 (lane 3), whereas little increase of incorporation was observed with the mRNA hybridized to pBR322 (lane 2) as compared to the incubation without added mRNA (lane 1). The product synthesized under the direction of the tufM mRNA was immunoprecipitable with antibody against *E. coli* EF-Tu (lane 5) but not with antibody against yeast cytoplasmic EF-1 α (lane 4). Furthermore, as will be described elsewhere, the M_r 48,000 protein was transported into yeast mitochondria on incubation in a cell-free system (unpublished data). From these results, it was concluded that the yeast tufMgene codes for the mitochondrial EF-Tu.

DISCUSSION

In the present study, a hybrid plasmid containing the yeast mitochondrial EF-Tu gene was isolated by subcloning and identified by Southern hybridization by using the *E. coli tufB* sequence as a probe. This procedure was found to be relatively simple and very effective. It may be applicable for isolation of other eukaryotic genes having a homology with *E. coli* genes for example, the eukaryotic genes encoding for the mitochondrial translation factors and ribosomes.

Southern hybridization analysis of the yeast restriction fragments with labeled DNA containing the *E. coli tufB* gene revealed only a single band that hybridized with the probe, suggesting that the mitochondrial EF-Tu is coded by a single nuclear



FIG. 4. In vitro translation of the hybridization-selected mRNA and immunoprecipitation of the product. The 2.5-kb EcoRI fragment of pYYB containing tufM gene or EcoRI-digested pBR322 was bound to nitrocellulose membrane filters and hybridized with mRNA obtained from yeast transformed with YRpYB. The hybridized mRNA was eluted and translated in 20 μ l of reticulocyte lysate in the presence of 12 μ Ci of [³²S]methionine (1,490 Ci/mmol, Amersham Japan). One third of the product was electrophoresed on 15% polyacrylamide gels in the presence of NaDodSO₄ (41) and visualized by fluorography (42). Translation products without added mRNA (lane 1) or with mRNA hybridized to pBR322 (lane 2) or tuf M (lane 3). One third each of the product used for the experiment of lane 3 was immunoprecipitated by antibodies against either yeast cytoplasmic EF-1 α (lane 4) or E. coli EF-Tu (lane 5), and the precipitates were analyzed as above. Lane M, ¹⁴C-labeled molecular weight standards (Amersham Japan) shown as $M_r \times 10^-$ (from top to bottom): myosin, phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, and lysozyme.

cytochrome c peroxidase	Met	-Thr-Thr-	Ala	-Val-Arg-	Leu-Leu-Pro	-Ser-	Leu	Gly-	Arg-Thr-Ala	His-	Lys	-Arg-	Ser	-Leu-Tyr-	-Leu	┟
mEF-Tu :	Met	Ser-	Ala		Leu-Leu-Pro	-Arg-	Leu	-Leu-Thr-	Arg-Thr-Ala	-Phe-	Lys	-Ala-	Ser	-Gly-Lys-	-Leu	ł

FIG. 5. Comparison of the NH2-terminal sequence of yeast mitochondrial EF-Tu and the signal sequence of yeast cytochrome c peroxidase. The amino acid sequence of yeast mitochondrial EF-Tu is from Fig. 3 and the signal sequence of yeast cytochrome c peroxidase is from Kaput et al. (45). The boxed amino acids indicate identical amino acids; dashes show gaps introduced to obtain the maximal homology.

gene and that the gene for the cytoplasmic factor may be more divergent. This agrees with the previous finding that yeast mitochondrial EF-Tu is functionally interchangeable with E. coli EF-Tu but not with yeast cytoplasmic EF-1 α (16). More recently, Piechulla and Küntzel (17) have shown that an antibody against yeast mitochondrial EF-Tu crossreacts with E. coli EF-Tu. As we have shown, the deduced amino acid sequence of yeast *tufM* is remarkably homologous to that of *E. coli* EF-Tu, especially in some regions. Amons et al. (15) have recently determined the partial amino acid sequence of EF-1 α from A. salina and pointed out the homology with E. coli EF-Tu. However, the homology of E. coli EF-Tu with A. salina EF-1 α is far less pronounced than that with yeast mitochondrial EF-Tu (compare figure 4 of ref. 15 with Fig. 3 of the present paper).

In general, nuclear-coded mitochondrial proteins are synthesized as precursors in the cytoplasm and transported into mitochondria (43). Although it is not known whether yeast mitochondrial EF-Tu is synthesized as a precursor that is processed to the mature protein during transport into the mitochondria, it is noteworthy that the tuf M gene codes for a protein 37 amino acids longer than E. coli EF-Tu at the NH2-terminal end. The 37-amino acid peptide is strongly basic, having 4 arginine and 2 lysine residues but no acidic amino acids. It is also rich in threonine (six residues) and serine (seven residues).

Recently, the structures of the proteolipid subunit of mitochondrial ATP synthase (44) and cytochrome c peroxidase (45)have been published. The signal sequences of these proteins are 68 and 66 amino acids long, respectively, and possess properties similar to the above NH₂-terminal sequences. Because the precursor of cytochrome c peroxidase is transported into the inner membrane space of the mitochondria, with the first 18 amino acid residues of the signal sequence facing the mitochondrial matrix (45), we have compared the structure of the signal sequence of cytochrome c peroxidase with that of the mitochondrial EF-Tu. As shown in Fig. 5, the first 20 amino acid residues of mitochondrial EF-Tu have a distinct homology with that of cytochrome c peroxidase, which may suggest that a homologous sequence is important for proteins made in the cytoplasm that are to be imported into the mitochondria.

The sequences that are conserved in many eukaryotic promoters are found at -225 nucleotides upstream from the first ATG codon of the tufM gene. A consensus sequence for poly(A) addition T-A-A-A-T-A-A-A-G in yeast is also present at 96-103 nucleotides downstream from the TAG termination codon. More detailed characteristics of the sequence as well as the analysis of the transcriptional and translational products will be discussed elsewhere.

We thank Dr. M. Miyazaki (Nagoya University) for his generous gift of the antibody against yeast cytoplasmic EF-1 α . Thanks are also due to Dr. R. W. Davis for providing us with vector plasmid YRp7 and strain D13-1A.

- Kaziro, Y. (1978) Biochim. Biophys. Acta 505, 95-127. 1.
- Arai, K., Clark, B. F. C., Duffy, L., Jones, M. D., Kaziro, Y., 2. Laursen, R. A., L'Italien, J., Miller, D. L., Nagarkatti, S., Nak-amura, S., Nielsen, K. M., Petersen, T. E., Takahasi, K. & Wade, M. (1980) Proc. Natl. Acad. Sci. USA 77, 1326–1330.
- 3. Jaskunas, S. R., Lindahl, L., Nomura, M. & Burgess, R. R. (1975) Nature (London) 257, 458-462.

- Bachmann, B. J. & Low, K. B. (1980) Microbiol. Rev. 44, 1-56.
- Shibuya, M., Nashimoto, H. & Kaziro, Y. (1979) Mol. Gen. Genet. 170. 231-234.
- 6. Miyajima, A., Shibuya, M. & Kaziro, Y. (1979) FEBS Lett. 102, 207 - 210.
- Yokota, T., Sugisaki, H., Takanami, M. & Kaziro, Y. (1980) Gene 7. 12. 25-31
- An, G. & Friesen, J. D. (1980) Gene 12, 33-39. 8.
- Nagata, S., Iwasaki, K. & Kaziro, Y. (1977) J. Biochem. (Tokyo) 82, 1633-1646.
- 10. Slobin, L. I. (1980) Eur. J. Biochem. 110, 555-563.
- Slobin, L. I. & Möller, W. (1976) Eur. J. Biochem. 69, 351-366. 11.
- Bollini, R., Soffientini, A., Bertani, A. & Lanzani, G. A. (1974) 12. Biochemistry 13, 5421-5425.
- Dasmahapatra, B., Skogerson, L. & Chakraburtty, K. (1981) J. Biol. Chem. 256, 10005-10011. 13
- Slobin, L. I., Clark, R. V. & Olson, M. O. J. (1981) Biochemistry 14. 20, 5761-5767.
- Amons, R., Pluijms, W., Roobol, K. & Möller, W. (1983) FEBS 15. Lett. 153, 37-42.
- 16. Richter, D. & Lipmann, F. (1970) Biochemistry 9, 5065-5070.
- Piechulla, B. & Küntzel, H. (1983) Eur. J. Biochem. 132, 235-240. 17.
- Richter, D. (1971) Biochemistry 10, 4422-4425. 18.
- 19. Borst, P. (1972) Annu. Rev. Biochem. 41, 333-376.
- 20. Cryer, D. R., Eccleshall, R. & Marmur, J. (1975) Methods Cell Biol. 12, 39-44.
- Nagata, S., Mantei, N. & Weissmann, C. (1980) Nature (London) 21. 287, 1-7.
- 22.
- Southern, E. M. (1975) J. Mol. Biol. 98, 503-517. Jeffreys, A. J. & Flavell, R. A. (1977) Cell 12, 429-439. 23.
- 24. Wahl, G. M., Stern, M. & Stark, G. R. (1979) Proc. Natl. Acad. Sci. USA 76, 3683-3687
- 25. Smith, H. D. & Birnstiel, M. L. (1976) Nucleic Acids Res. 3, 2387-2398.
- 26. Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. 65, 499-559
- 27. Walseth, T. F. & Johnson, R. A. (1979) Biochim. Biophys. Acta 562, 11-31
- 28. Struhl, K., Stinchcomb, D. T., Scherer, S. & Davis, R. W. (1979) Proc. Natl. Acad. Sci. USA 76, 1035–1039. Hitzeman, R. A., Hagie, F. E., Hayflick, J. S., Chen, C. Y., See-
- 29. burg, P. H. & Derynck, R. (1982) Nucleic Acids Res. 10, 7791-7808.
- Struhl, K. & Davis, R. W. (1981) J. Mol. Biol. 152, 535-552. 30.
- Aviv, H. & Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-31. 1412
- 32. Harpold, M. M., Dobner, P. R., Evans, R. M. & Bancroft, F. C. (1978) Nucleic Acids Res. 5, 2039–2053.
- Pelham, H. R. B. & Jackson, R. J. (1976) Eur. J. Biochem. 67, 247-33. 256.
- Kessler, S. W. (1975) J. Immunol. 115, 1617-1624. 34.
- 35.
- Miyajima, A. & Kaziro, Y. (1978) J. Biochem. (Tokyo) 83, 453-462. Shaw, G. D., Boll, W., Taira, H., Mantei, N., Lengyel, P. & Weissmann, C. (1983) Nucleic Acids Res. 11, 555-573. 36.
- 37. Hoeijmakers, J. H. J., Borst, P., Van Den Burg, J., Weissmann, C. & Cross, G. A. M. (1980) Gene 8, 391–417
- Wilkie, N. M., Clements, J. B., Boll, W., Mantei, N., Lonsdale, D. & Weissmann, C. (1979) Nucleic Acids Res. 7, 859-877. 38.
- Soberon, X., Covarreubias, L. & Bolivar, F. (1980) Gene 9, 287-39. 305
- **40**. Nakamura, S. & Kaziro, Y. (1981) J. Biochem. (Tokyo) 90, 1117-1124
- 41. Laemmli, U. K. (1970) Nature (London) 227, 680-685.
- Bonner, W. M. & Laskey, R. A. (1974) Eur. J. Biochem. 46, 83-42. 88.
- Schatz, G. & Butow, R. A. (1983) Cell 32, 316-318. 43.
- Viebrock, A., Perz, A. & Sebald, W. (1982) EMBO J. 1, 565-571. 44. 45. Kaput, J., Goltz, S. & Blobel, G. (1982) J. Biol. Chem. 257, 15054-
- 15058