Cloning, Nucleotide Sequence, and Expression of One of Two Genes Coding for Yeast Elongation Factor 1α*

(Received for publication, July 2, 1984)

Patrick Cottrelle, Danielle Thiele, Virginia L. Price‡, Sylvie Memet, Jean-Yves Micouin, Christian Marcq, Jean-Marie Buehler, André Sentenac, and Pierre Fromageot

From the Département de Biologie, Service de Biochimie, Centre d’Études Nucléaires de Saclay, 91191 Gif-sur-Yvette, Cedex, France

One gene coding for yeast cytoplasmic elongation factor 1α (EF-1α) was isolated by colony hybridization using a cDNA probe prepared from purified EF-1α mRNA. A recombinant plasmid, pLB1, with a 6-kilobase yeast DNA insert, was found by hybrid selection and selection experiments to carry the entire gene. The nucleotide sequence of the gene with its 5'– and 3'-flanking regions was determined. The 5' and 3' ends of EF-1α mRNA were localized by the S1 nuclease mapping technique. The cloned gene, called TEF1, encodes a protein of 458 amino acids (M, = 50,071) in a single, uninterrupted reading frame. The amino acid sequence shows a strong homology with several domains of Artemia salina EF-1α cytoplasmic factor, as evidenced by diagonal dot matrix analysis. Protein sequence homology is comparatively much lower with the yeast mitochondrial elongation factor.

S1 nuclease mapping of the mRNA, hybridization analysis of chromosomal DNA using intragenic or extragenic DNA probes, and gene disruption experiments demonstrated the existence of two genes coding for the cytoplasmic elongation factor EF-1α/haploid genome. The presence of an intact chromosomal TEF1 gene is not essential for growth of haploid yeast cells.

In Escherichia coli, two unlinked genes, designated tufA and tufB, code for the polypeptide chain elongation factor (EF2)-Tu (1–3). Both genes have been isolated and sequenced (4–7). Their nucleotide sequences differ only in 13 positions whereas the gene products EF-TuA and EF-TuB are identical, except for 1 C-terminal amino acid residue (6, 7). The two genes are distantly located on the E. coli chromosome (1–3).

The tufA gene is part of a large transcription unit with three other genes coding for two ribosomal proteins, S12 and S7, and for EF-G (1). The tufB gene is co-transcribed with a cluster of four tRNA genes (8). Duplication of tuf genes occurs in prokaryotes other than E. coli, but is not a general observation (9), and its functional significance is not clear.

Less information is available on the gene(s) encoding the eukaryotic cytoplasmic elongation factor. van Hemert et al. (10, 11) recently isolated several overlapping cDNA clones coding for EF-1α, the cytoplasmic factor from Artemia salina, and have determined by cDNA as well as by protein sequences the primary structure of the Artemia factor. Southern analysis has suggested the possibility of reiterated genes (12).

In addition to the cytoplasmic factor EF-1α, the eukaryotic cell contains a counterpart of EF-Tu in mitochondria and chloroplasts. Specific EF-Tu probes from E. coli DNA have been used to clone the gene for the yeast mitochondrial (13) and chloroplastic factors (14), both having retained a high level of sequence homology with the bacterial factor (11, 13).

Among eukaryotic cells, the yeast Saccharomyces cerevisiae offers the unique opportunity of combining the classical biochemical and genetic approaches with sophisticated recombinant DNA technology. Cloning the structural gene for yeast EF-1α would facilitate the structural and functional analysis of this presumed essential protein (see Ref. 22 and references therein). In the present work, we show that the haploid yeast genome contains two genes coding for the elongation factor 1α. One gene has been cloned and sequenced with its flanking regions. Although this gene is actively transcribed, it is not essential for growth. The amino acid sequences of EF-1α from yeast and Artemia are very similar, suggesting a strong conservation of this protein among eukaryotes.

**Experimental Procedures**

Results

Cloning the Yeast EF-1α Gene—As observed in our previous work (22), EF-1α is a very abundant protein in actively growing yeast cells. With the assumption that the same situation applied for the EF-1α mRNA, we decided to clone the gene using a cDNA probe. In short, total yeast RNA was fractionated by electrophoresis on a polyacrylamide gel under denaturing conditions, and the different fractions were assayed in a reticulocyte *in vitro* translation system for EF-1α synthesis. The enriched EF-1α mRNA fraction was then used to direct the synthesis of a labeled cDNA probe which served to screen a yeast genomic library in plasmid pBR322. One hundred and thirty clones were selected by colony hybridization. When further screened by groups of 12, in hybrid selection experiments, one pool of 12 clones was found to contain DNA sequences complementary to EF-1α mRNA, as evidenced by immunoprecipitation of the *in vitro* translation products (results not shown). Finally, one positive clone was

---

*The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1754 solely to indicate this fact.

‡Present address: Immunex Corporation, Seattle, WA 98101.

The abbreviations used are: EF, elongation factor; kb, kilobase; Pipes, 1,4-piperazinediethanesulfonic acid; bp, base pairs.

* Portions of this paper (including "Experimental Procedures" and additional references) are presented in miniprint at the end of this paper. Miniprint is easily read with the aid of a standard magnifying glass. Full size photocopies are available from the Journal of Biological Chemistry, 9560 Rockville Pike, Bethesda, MD 20814. Request Document No. 84M-2002, cite the authors, and include a check or money order for $1.60 per set of photocopies. Full size photocopies are also included in the microfilm edition of the Journal that is available from Waverly Press.
isolated which contained a recombinant plasmid called pLB1. The restriction map of the yeast DNA insert is presented in Fig. 1. Unique sites were found in the insert for HpaI, PvuII, XbaI, and EcoRV enzymes. There was no restriction sites for KpnI, XhoI, SalI, PstI, SacI, or BclI. Coarse mapping of the EF-1α coding sequence was done by hybrid selection and translation experiments. One DNA fragment, the 0.87-kb EcoRI-HindIII fragment, was found to select by hybridization the EF-1α message in a total yeast RNA extract. Flanking restriction fragments on both sides were less efficient to select a translatable EF-1α mRNA (results not shown). This preliminary mapping suggested that the isolated recombinant plasmid pLB1 harbored the entire coding region (about 1.4 kb as inferred from the size of the protein) on a 2.97-kb AhaIII-AhaIII fragment. More precise mapping of the coding region was obtained by S1 mapping (see below).

**DNA Sequencing of the EF-1α Gene**—The final demonstration that the EF-1α structural gene had been cloned was obtained by DNA sequencing. One looked for the conserved sequence Gly-Ile-Thr-Ile and following residues corresponding to the trypsin-sensitive site which had been identified in our previous work (22). The strategy for DNA sequencing is presented in Fig. 1. When possible, double-restricted fragments were cloned into M13 mp10 and mp11 phages which allowed us to read the sequence from both ends using the same primer. A region of 2900 nucleotides has been sequenced. The 0.87-kb EcoRI-HindIII fragment contained, as expected, a large open reading frame. The conserved sequence Gly-Ile-Thr-Ile was found at the level of the ClaI site proximal to this EcoRI-HindIII fragment. The presence of this sequence and the following 10 amino acids (22) identified the EF-1α structural gene that we have called TEF1. Fig. 2 shows a sequence of 2112 nucleotides encompassing EF-1α TEF1 gene with its 5’- and 3’-flanking regions. There was only one AUG codon upstream from the Gly-Ile-Thr-Ile sequence (the ClaI site) and in the same open reading frame. It is located 218 nucleotides from this ClaI site. The 5’ end of the mRNA, positioned by S1 mapping (see below), is 35 nucleotides upstream from this AUG.

The gene encodes a protein of 458 amino acids in a single uninterrupted open reading frame. This corresponds to a protein molecular weight of 50,071, including the first methionine, in good agreement with the value of 49,000 estimated by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate (22).

*The 5’- and 3’-Flanking Regions*—The general features of the 5’-flanking region upstream from the coding sequence are those previously reported for other highly expressed yeast genes (15, 16): the A + T content is very high (78% A + T over a 480-nucleotide stretch of DNA); a “CAAG” sequence is found at position –99/–96, about 40 nucleotides upstream from a C + T-rich block; a putative “TATA box” (TAATAA)

**Fig. 1.** Restriction map and DNA sequencing strategy for yeast EF-1α gene. Cleavage sites for restriction endonucleases with a 6-bp recognition sequence are shown in the schematic representation of the cloned yeast DNA insert, together with other sites used for DNA sequencing. The open box represents the coding region of the EF-1α gene with location of the ATG codon. Dots and arrows indicate the direction and extent of DNA sequence information obtained from the indicated restriction sites.

**Fig. 2.** DNA sequence of the EF-1α gene (TEF1) with 5’- and 3’-flanking regions. The sequence of the mRNA identical strand is shown together with the deduced amino acid sequence of the protein. The coding sequence of the EF-1α gene begins at position +1 (the A of the ATG codon) and extends to position 1374 in a single uninterrupted reading frame. The Gly-Ile-Thr-Ile conserved sequence and presumptive TATA box are underlined. A major transcription initiation and termination site identified by S1 mapping (see text).

is at position –121/–115; finally, a “CAAT box” (GGCAAT), similar to the conserved sequence GGGyAATCT, is present at position –210/–205.

At the 3’ side, the gene ends with a succession of three closely spaced stop codons. A stretch of 288 nucleotides was sequenced downstream from the first termination codon (see Fig. 2). The A + T content is high (73%) as for other yeast genes (17, 18), and one finds the tripartite sequence with the “TATG” block, tentatively proposed by Zaret and Sherman (15) as termination/polyadenylation signal, in some yeast genes.

**Mapping the 5’ and 3’ Termi of the EF-1α mRNA**—For 5’ end determination, total yeast RNA was hybridized with an Accl-SacI restriction fragment labeled at the 5’ end of the Accl site which is located at position +60 within the coding region. The resultant hybrid was treated at 37°C with varying concentrations of S1 nuclease and the size of the
S. cerevisiae Elongation Factor 1α (TEF1) Gene

Hybridized DNA determined by electrophoresis on a urea-polyacrylamide gel, alongside the G + A cleavage products of the same DNA probe (Fig. 3). Two groups of S1-resistant bands were obtained. The longest and major band, 95 nucleotides long, corresponded to RNA chains initiated at position -35. A short ladder of bands of smaller size was also found which indicated the existence of RNA chains complementary to the DNA probe from the AccI site up to position -5 ± 2 upstream from the ATG codon (the short ladder of bands formed at both sites resulted from a variable trimming of the hybrid by S1 nuclease). This finding of two distinct groups of S1-protected DNA fragments was similar to the earlier observation of Holland et al. (16) with the enolase and glyceraldehyde-3-phosphate dehydrogenase multicopy genes. In these cases, the long hybrid bands corresponded to mRNA transcribed from the gene cloned, while the short S1-resistant bands were hybridized with mRNA transcribed from the homologous genes. By analogy, our results suggested the possibility of two or more copies of the EF-1α gene with different 5' untranslated regions.

The location of the 3' terminus of the mRNA was determined similarly by S1 mapping, using a 1.4-kb HindIII-EcoRI probe labeled at the 3' end of the HindIII site (Fig. 4). Again, two major S1-resistant DNA bands were found. Hybrids length corresponded to RNA chains complementary to the DNA probe up to 5 ± 2 nucleotides and 115 ± 5 nucleotides downstream from the UAA termination codon. As for the 5' end, this result strongly suggested the existence of multicyclic genes with heterogeneous 3' noncoding regions, the longest RNA chain being transcribed from the cloned TEF1 gene. From the above 5' and 3' end mapping, the length of the nonpolyadenylated mRNA was calculated to be 1530 ± 10 nucleotides. The relative intensity of the hybridization signals suggests that the TEF1 gene and the homologous gene are transcribed in vivo with a similar efficiency.

Two Genes Encode EF-1α—The existence of reiterated genes with different 5' and 3' DNA sequences was evidenced by Southern analysis of genomic yeast DNA. Yeast DNA fragments restricted with EcoRI or HindIII nucleases were separated by agarose gel electrophoresis, transferred to nitrocellulose membranes, and hybridized with three different probes originating from the cloned TEF1 gene. The 2.2-kb EcoRI-EcoRI-labeled probe, which encompasses most of the

**Fig. 3.** S1 nuclease mapping of the 5' end of EF-1α mRNA.

Total yeast RNA was hybridized to a radioactive double-stranded probe 5' end-labeled with 32P at the AccI site (position +60) and extending to the Sau3A site (position -425, see Fig. 1). After treatment with varying S1 nuclease concentrations, the nuclease-resistant hybrids were analyzed on an 8% polyacrylamide gel containing 7 M urea. Lane a, 1000 units/ml S1 nuclease; and lane b, 3000 units/ml S1 nuclease. Lane c, shows the G + A chemical degradation products of the probe used for hybridization. Partial gene sequences (mRNA complementary) are indicated. Numbering is as shown in Fig. 2.

**Fig. 4.** S1 nuclease mapping of the 3' end of EF-1α mRNA.

Total yeast RNA was hybridized to the 1.4-kb HindIII-EcoRI fragment 32P-labeled at the 3' end of the HindIII site (see Fig. 1). Hybrids resistant to S1 nuclease (1000 or 3000 units/ml) were analyzed on a 5% polyacrylamide sequencing gel. Lane a, molecular weight DNA markers of known sequence. Their size (in bases) is indicated in the margin; lane b, DNA probe incubated with yeast RNA, without S1 treatment; lane c, DNA probe incubated with 200 μg of E. coli RNA, treated with S1 nuclease (1000 units/ml); lanes d and e, DNA probe incubated with yeast RNA, treated with S1 nuclease at 1000 units/ml and 3000 units/ml, respectively. S1-resistant bands are indicated by arrows numbered 1 and 2. Some residual full-length probe is seen on top of the gel in lanes d and e. When compared to the amount of DNA probe loaded in lane b, this S1-resistant DNA probe corresponded to about 2–5% of the input.
gene and 1.3 kb on its 3' side, hybridized with two genomic EcoRI fragments, one of an expected length of 2.2 kb and an additional one of about 9 kb (Fig. 5, panel 2, lanes a and b). The same probe hybridized with four HindIII genomic DNA fragments, strongly with three bands of 10, 5.9, and 3.5 kb and weakly with a 1.5-kb fragment (Fig. 5, panel 2, lanes d and e). Since a HindIII site is located within the coding region, this result was compatible with the existence of two genes with different 3'-flanking regions. The weak signal on the 1.5-kb band could therefore belong to the second gene hybridizing to the probe only over 142 nucleotides of homologous coding region, starting from the HindIII site.

This conclusion was confirmed by the observation that the 3' probe EcoRI-BamHI (contiguous to the previous EcoRI probe) hybridized only to the HindIII genomic fragment of 3.5 kb (Fig. 5, panel 3, lanes d and e). Similarly, this 3' probe hybridized to a single EcoRI genomic fragment of 2.4 kb (panel 3, lanes a and b). The third probe was a ClaI-ClaI 1.57-kb fragment overlapping the 5' coding region over 217 bp. This 5' probe hybridized strongly with the 5.9-kb HindIII fragment already detected and weakly with a 10-kb HindIII band (Fig. 5, panel 1, lanes d and e). Similarly, with the same 5' probe, two signals of different intensity were found using EcoRI-restricted DNA of 11 and 2 kb (Fig. 5, panel 1, lanes a and b). These two signals again argue in favor of two genes with different 5' surroundings.

Effect of Gene Disruption on Cell Viability—The Southern analysis together with the S1 mapping experiments strongly suggest the existence of two genes encoding EF-la and that both are transcribed. A gene disruption experiment was performed to see whether the expression of both genes is required for cell viability. The intragenic 0.87-kb EcoRI-HindIII fragment was subcloned into the integrative plasmid YIp5. This plasmid, which carries the URA3 gene as marker, can only be propagated by integration into chromosomal DNA (19). The recombinant plasmid was linearized with BglII which cuts at a unique site present in the EcoRI-HindIII insert, to direct its integration within EF-la structural gene, and used to transform ura3- haploid yeast cells. URA3+ transformants were analyzed for the integration of the plasmid DNA within EF-la gene by Southern analysis (Fig. 6). BamHI-restricted genomic DNA fragments from several transformants were hybridized with two probes: the intragenic 0.87-kb EcoRI-HindIII fragment and the 3' probe EcoRI-BamHI which is specific of the cloned gene. Three or four DNA fragments were revealed in the transformants instead of one large ~12-kb fragment in the parental haploid cells (Fig. 6, panel 1). (The two genes are included within large BamHI fragments of very similar size which are not resolved in this experiment.) The three-band pattern corresponded to the disruption of one of the two EF-la genes which generated two fragments of ~10 kb and ~8 kb and left intact the 12-kb fragment which harbored the second gene. In the four-band pattern, the extra band of 6 kb corresponded to the full-length recombinant plasmid which had been integrated in one or more copies and was released by BamHI digestion. To decide which of the two EF-la genes had been disrupted in the transformants, we used the extragenic 3' specific probe from the cloned TEF1 gene to hybridize with the same genomic fragments. In the

Fig. 6. Disruption of the TEF1 gene; hybridization analysis of genomic DNA from transformed haploid cells. For gene disruption, ura3- haploid yeast cells X2180 were transformed with the integrative plasmid YIp5 carrying the 0.87-kb EcoRI-HindIII intragenic fragment of TEF1 gene, cloned into the plasmid EcoRI-HindIII sites, plus the auxotrophic URA3 marker. The recombinant plasmid was cleaved with BglII, which cuts at a unique site within the EcoRI-HindIII insert, to promote its preferential integration within EF-la structural gene. From several independent URA3+ transformants, the DNA was isolated, restricted with BamHI, and analyzed by electrophoresis and filter transfer hybridization with two different probes: the 0.87-kb EcoRI-HindIII fragment which contains a large part of the TEF1 gene (panel 1) and the 1.25-kb EcoRI-BamHI fragment which lies 3' outside the TEF1 gene (panel 2). a, control parental DNA; b-d, DNA from three transformants. The size of genomic fragments is given in kilobases.

For gene disruption, ura3- haploid yeast cells X2180 were transformed with the integrative plasmid YIp5 carrying the 0.87-kb EcoRI-HindIII intragenic fragment of TEF1 gene, cloned into the plasmid EcoRI-HindIII sites, plus the auxotrophic URA3 marker. The recombinant plasmid was cleaved with BglII, which cuts at a unique site within the EcoRI-HindIII insert, to promote its preferential integration within EF-la structural gene. From several independent URA3+ transformants, the DNA was isolated, restricted with BamHI, and analyzed by electrophoresis and filter transfer hybridization with two different probes: the 0.87-kb EcoRI-HindIII fragment which contains a large part of the TEF1 gene (panel 1) and the 1.25-kb EcoRI-BamHI fragment which lies 3' outside the TEF1 gene (panel 2). a, control parental DNA; b-d, DNA from three transformants. The size of genomic fragments is given in kilobases.

Fig. 5. Hybridization analysis of yeast chromosomal DNA. Genomic DNA (0.5 µg) from strain 4094 B was digested with 3 or 10 units of EcoRI endonuclease (lanes a and b, respectively), or with 3 or 10 units of HindIII (lanes d and e, respectively), electrophoresed on a 0.7% agarose gel, transferred to a nitrocellulose filter, and hybridized to three different DNA probes as depicted in the diagram. The probes were 32P-labeled, nick-translated restriction fragments ClaI-ClaI (1.57 kb, probe 1), EcoRI-EcoRI (2.2 kb, probe 2), and EcoRI-BamHI (1.25 kb, probe 3). The autoradiograms of the filters 1, 2, and 3 correspond to hybridization with probes 1, 2, and 3, respectively. Lane c, in the three panels, contains size markers derived from plasmid pLB1.
three transformants, the 8-kb fragment strongly hybridized with the TEF1-specific 3' probe (Fig. 6, panel 2)\(^3\). Note that this probe hybridized to the 12-kb band from parental DNA but not from the transformants, as expected if the remaining 12-kb fragment harbored the second gene. The same hybridization results were found with four additional independent transformants. Therefore, the gene interrupted was always the one which had been cloned. It is clear that the activity of TEF1 gene is not essential for cell growth since haploid cells had been transformed.

**EF-1α, a Highly Conserved Protein**—The sequence of yeast EF-1α was analyzed in comparison with the homologous yeast mitochondrial (13) and Artemia (11) genes. Codon usage is shown in Table I. The number of codons used is much more restricted in the case of the yeast cytoplasmic gene (34 codons used instead of 54 and 55 for coding the mitochondrial and Artemia protein, respectively). A small set of 25 codons accounts for 96% of yeast EF-1α sequence instead of 50 and 46 for the mitochondrial and Artemia proteins, respectively.

Conservation of essential protein domains was explored by diagonal dot matrix analysis of the amino acid sequence of yeast cytoplasmic and mitochondrial elongation factors, Artemia EF-1α and E. coli EF-Tu (Fig. 7). A striking homology was found between the yeast cytoplasmic and Artemia factor as evidenced by the perfect diagonal obtained at moderate stringency (≥7 matches over 21 contiguous amino acid residues). At high stringency (≥15 matches over 21 residues, which corresponded to ≥70% homology), the small background disappeared while two long domains of sequence homology and three small ones remained visible on the diagonal. The first N-terminal third of the protein was the most highly conserved, more so than the C-terminal part. A central region of divergence separated the two large domains of homology. A much smaller degree of sequence conservation was observed when comparing yeast EF-1α and yeast mitochondrial factor. Nevertheless, stretches of homology were found at low stringency (7/21 residues). Interestingly, the protein regions which showed some sequence conservation or divergence were the same as those noted above when comparing yeast and Artemia EF-1α: protein sequence homology was more maintained in the N-terminal than C-terminal region, and a region of divergence was confined within the middle of the protein. A similar pattern was observed with bacterial EF-Tu which reflected the high homology between the mitochondrial and bacterial factors (6, 7, 13).

**DISCUSSION**

The TEF1 Gene—This is the first report on the cloning and sequencing of a gene with its flanking DNA regions for cytoplasmic elongation factor from an eukaryotic cell (EF-1α). cDNA clones for A. salina factor have been isolated and sequenced recently (11). Cloning the yeast EF-1α gene by the mRNA strategy was probably facilitated by the abundance of the mRNA, as suggested by the higher level of EF-1α in yeast cells (22). The gene, called TEF1, was identified unambiguously by sequencing. The N-terminal sequence of proteolyzed EF-1α determined previously (22) was found in the deduced amino acid sequence in the region of the ClaI site. This sequence contained the apparently universal tetrypeptide found in EF-Tu and EF-1α factors analyzed so far. Furthermore, as discussed below, a detailed sequence comparison of the yeast and A. salina factors revealed a strong homology (≥70%) over 68% of their total sequence.

The complete nucleotide sequence of the yeast TEF1 gene was established, together with its 5’- and 3’-flanking regions. The gene sequence codes for a protein of 468 amino acids (calculated M, = 50,071), in close agreement with the size of the purified protein (apparent M, = 49,000). Additional evidence indicating the absence of introns are the S1 mapping data and the lack of the essential splice site (TACTAAXCA) found in the introns of yeast genes (20). Nevertheless, although unlikely, the presence of a very small size intron could not be totally ruled out. In fact, the chloroplast gene for elongation factor Tu does contain two small introns (14).

The codon usage in TEF1 was highly nonrandom as for all highly expressed yeast genes (21). There were large differences in the codon usage of the two nuclear genes coding for cytoplasmic and mitochondrial homologous proteins. It very likely reflects the relative abundance of the two proteins in the cell (21).

The sequence of the 5’ upstream region was characteristic of highly expressed yeast genes (15) with its high A + T content, putative "TATA box," and "CAAT box" (see "Results"). A curious feature of the 5’ region is the 14-bp repeated

Table I

<table>
<thead>
<tr>
<th>Y'</th>
<th>A'</th>
<th>Y'</th>
<th>Y'</th>
<th>A'</th>
<th>Y'</th>
<th>Y'</th>
<th>A'</th>
<th>Y'</th>
<th>Y'</th>
<th>A'</th>
<th>Y'</th>
<th>Y'</th>
<th>A'</th>
<th>Y'</th>
<th>Y'</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUU Phe</td>
<td>2; 8; 9</td>
<td>UCU Ser</td>
<td>14; 8; 12</td>
<td>UAU Tyr</td>
<td>0; 2; 5</td>
<td>UGU Cys</td>
<td>6; 4; 2</td>
<td>UUC Phe</td>
<td>15; 10; 5</td>
<td>UCC Ser</td>
<td>6; 8; 4</td>
<td>UAC Tyr</td>
<td>8; 11; 4</td>
<td>UGC Cys</td>
<td>1; 2; 1</td>
</tr>
<tr>
<td>UUG Leu</td>
<td>21; 7; 12</td>
<td>UGC Ser</td>
<td>0; 1; 0</td>
<td>UAG AMB</td>
<td>0; 4; 2</td>
<td>UGG Trp</td>
<td>6; 5; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUC Leu</td>
<td>0; 12; 6</td>
<td>CCA Pro</td>
<td>0; 7; 6</td>
<td>CAU His</td>
<td>5; 4; 10</td>
<td>CGU Arg</td>
<td>0; 11; 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUA Leu</td>
<td>0; 1; 2</td>
<td>CCC Pro</td>
<td>0; 3; 2</td>
<td>CAC His</td>
<td>6; 7; 4</td>
<td>CGG Arg</td>
<td>0; 1; 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CUG Leu</td>
<td>0; 0; 5</td>
<td>CCA Pro</td>
<td>23; 12; 9</td>
<td>CAA Gin</td>
<td>12; 4; 11</td>
<td>CGA Gin</td>
<td>0; 6; 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC Ile</td>
<td>17; 20; 17</td>
<td>ACU Thr</td>
<td>14; 11; 10</td>
<td>AAU Asn</td>
<td>0; 3; 7</td>
<td>AGU Ser</td>
<td>1; 2; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUA Ile</td>
<td>13; 12; 12</td>
<td>ACC Thr</td>
<td>14; 6; 9</td>
<td>AAC Asn</td>
<td>16; 13; 6</td>
<td>AGG Ser</td>
<td>0; 0; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUG Met</td>
<td>8; 9; 18</td>
<td>ACG Thr</td>
<td>0; 4; 4</td>
<td>AAG Lys</td>
<td>46; 27; 12</td>
<td>AGG Arg</td>
<td>0; 0; 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGU Val</td>
<td>26; 12; 18</td>
<td>GCC Ala</td>
<td>30; 23; 18</td>
<td>GAA Asp</td>
<td>8; 13; 18</td>
<td>GGU Gly</td>
<td>42; 22; 30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAC Val</td>
<td>19; 3; 7</td>
<td>CCA Ala</td>
<td>7; 15; 11</td>
<td>GAC Arg</td>
<td>16; 11; 6</td>
<td>GGC Gly</td>
<td>0; 10; 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUA Val</td>
<td>0; 5; 3</td>
<td>GCA Ala</td>
<td>0; 2; 6</td>
<td>GAA Glu</td>
<td>30; 22; 26</td>
<td>GGA Gly</td>
<td>0; 7; 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUG Val</td>
<td>1; 1; 5</td>
<td>GCC Ala</td>
<td>0; 1; 0</td>
<td>GAG Glu</td>
<td>1; 10; 7</td>
<td>GGG Gly</td>
<td>0; 5; 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
sequence (CGTTTCTTTTTCTT) at positions -234 and -190, (i.e. 113 and 69 bp upstream from the TATA box). This DNA region could be involved in regulation of transcription.

Two Genes Code for EF-1α—The existence in the haploid yeast genome of multicopy genes coding for EF-1α was first suggested by S1 nuclease analysis, showing one group of protected DNA bands mapping at the boundaries of the coding sequence. A similar result has been reported by Holland et al. (16) for the enolase and glyceraldehyde-3-phosphate dehydrogenase multicopy genes. Southern hybridization to genomic DNA of intragenic or extragenic probes originating from the cloned TEF1 gene suggests the existence of two very similar genes with different surroundings. Taken together with the S1 mapping results, it is likely that the two genes have very homologous coding sequences, but differ in their 5′ and 3′ noncoding regions. The second gene was cloned recently and sequenced on its 5′ end: the 5′ coding region was identical to that of TEF1 over 384 bp, but its sequence rapidly diverged in the 5′ noncoding region, confirming our interpretation of the S1 analysis.4 We have called this second gene TEF2. Both genes are transcribed to the same extent as evidenced by S1 analysis.

The gene disruption experiment in the haploid cell further confirms the existence of an active gene distinct from TEF1 and shows that TEF1 is not essential for growth when a nondisrupted copy of TEF2 is present in the cell. Whether TEF2 gene is essential or not will be answered by performing gene disruption in a diploid cell in the presence or absence of one active TEF1 gene. It should be recalled that, in E. coli, two genes code for EF-Tu and that tufA could never be totally inactivated (3). The reasons for the presence of two genes in E. coli are still unknown. Sequencing of the TEF2 gene will possibly reveal slight variations in protein sequence which may have a functional significance. TufB in E. coli is included in a transcription unit with four tRNA genes (8). We found no evidence for tRNA genes in the yeast DNA inserts containing TEF1 or TEF2 genes, using a yeast reconstituted transcription system.5

EF-1α, a Highly Conserved Protein—The amino acid sequence derived from the TEF1 gene sequence was compared by diagonal dot matrix analysis to the sequence of EF-1α recently determined from A. salina (11). A striking homology was found. Using a low stringency criterion (≥35% homology), the two sequences were indistinguishable. At high stringency (≥70% homology), highly conserved protein domains became apparent, with divergent sequences in between. Hence, starting from the N-terminal end, there were two large, highly conserved domains separated by a central region of lesser homology. Blocks of high homology were also found in the C-terminal part. It is tempting to speculate that these conserved domains are essential for the functional integrity of the protein, while the regions in between may be less directly involved in protein function.

Acknowledgments—We thank W. Möller for communication of A. salina EF-1α sequence prior to publication and F. Sor for introducing one of us (P. C.) to DNA sequencing.

Note Added in Proof—Both TEF genes of S. cerevisiae were also cloned by Nagata et al. (23) and by Schirmer andPhilipsen (1984) EMBO J. 3, 3311–3315. The sequence of TEF2 was established by the latter.

REFERENCES


Additional references are found on p. 3096.

4 M. Cool, unpublished results.

5 S. Camier, unpublished observations.
S. cerevisiae Elongation Factor 1a TEF1 Gene

Hybridization analysis of yeast chromosomal DNA. Yeast genomic DNA from strain 11.5B was digested with 1, 2 or 10 units of EcoRI or HindIII restriction enzymes in 500 µl of HindIII restriction buffer and HindIII restriction fragments were subjected to electrophoresis on a 0.7% agarose gel. The TEF1 restriction fragments were transferred to nitrocellulose filter paper. The filters were hybridized with [32P]dCTP-labeled cDNA probes derived from the yeast Elongation Factor 1a TEF1 gene. Hybridization and autoradiography were done essentially as described in the legend.

Sequence data analysis. Editing and analysis of DNA sequences, base pair determination and protein sequence comparison by diapositive dot matrix analysis were performed on the UNIGRAF line printer with programs written partially on machine language.

REFERENCES