Volume 16 Number 8 1988 Nucleic Acids Research

Two genes encode related cytoplasmic elongation factors 1α (EF-1α) in *Drosophila melanogaster* with continuous and stage specific expression

Bernd Hovemann, Sabine Richter, Uwe Walldorf[1] and Celina Cziepluch

Zentrum für Molekulare Biologie Heidelberg (ZMBH), Universität Heidelberg, Im Neuenheimer Feld 282, D 6900 Heidelberg, FRG and [1]Biozentrum, Klingelbergstrasse 70, CH 4056 Basel, Switzerland

Received February 4, 1988; Accepted March 18, 1988 Accession nos: F1=X06869, F2=X06870

ABSTRACT

We have characterized two previously cloned genes, F1 and F2 (1)that code for elongation factor EF - 1α of Drosophila melanogaster. Genomic Southern blot hybridization revealed that they are the only gene copies present. We isolated cDNA clones of both transcripts from embryonal and pupal stage of development that cover the entire transcription unit. The 5' ends of both genes have been determined by primer extension and for F1 also by RNA sequencing. These start sites have been shown to be used consistently during development. Comparison of cDNA and genomic sequences revealed that EF - 1α,F1 consists of two and EF - 1α,F2 of five exons. The two described elongation factor genes exhibit several regions of strong sequence conservation when compared to five recently cloned eucaryotic elongation factors.

INTRODUCTION

The process of translation of genetic information from mRNA to protein follows a distinct pathway. One step, the elongation of the amino acid chain, involves a series of protein components that have been classified according to their function as elongation factor 1 (EF - 1) and elongation factor 2 (EF - 2). The cytoplasmic EF - 1 complex is thought to consist of three proteins: EF - 1α, ß and γ(2). EF - 1α facilitates the GTP dependent binding of charged tRNAs to the acceptor site of the ribosomes. This process includes the binding and hydrolysis of GTP, the binding of aminoacyl tRNA and the recognition and interaction with the 80S ribosome. The multifunctional demand on EF - 1α obviously leaves little room for an evolution of divergent protein structures. It is therefore not surprising that the eucaryotic cytoplasmic 1α elongation factors isolated to date (3-10) show a high degree of homology that is indicated by physical properties like pI and molecular weight. Their

close relationship also becomes evident by comparing the gene structure of various recently cloned and sequenced EF - 1α genes. Saccharomyces cerevisiae for example contains two genes for the cytoplasmic factor (11-13), either one is sufficient for cell viability. The fungus Mucor racemosus has three closely related EF - 1α genes (14), and the brine shrimp Artemia salina may even contain four copies of this gene (15). The number of genes detected in mouse (our unpublished observation) might exceed ten, including possible pseudogenes. The one human (16) and part of a mouse EF - 1α gene sequence known to date (17-19) all share a very high degree of homology within the protein coding portion with all cytoplasmic EF - 1α genes analysed so far.

Here, we present evidence that Drosophila melanogaster contains only two different gene copies of the cytoplasmic factor. We report the isolation and sequence of cDNA clones covering the entire mRNA for both EF - 1α,F1 and F2. Transcription start and processing sites are determined by comparison of cDNA and genomic sequences and primer extended RNA sequencing.

MATERIALS AND METHODS

DNA sequencing was performed according to the chemical cleavage method of Maxam and Gilbert (20) and the chain termination method of Sanger (21). For RNA sequencing we followed the protocol of Geliebter (22) using a 20mer oligonucleotide primer and AMV reverse transcriptase. 10 μg poly $(A)^+$ RNA and 0.5 pmol ^{32}P labelled oligonucleotide primer were coprecipitated and taken up in annealing buffer (250 mM KCl, 10 mM Tris/Cl pH 8.3). After heating to 80°C for 3 min the mixture was kept at 45°C for 30 min. The individual reactions were carried out at 43°C for 15 min.

For primer extended cDNA synthesis the reaction was set up as described previously (23). For Southern analysis of chromosomal DNA and for the screening of cDNA libraries we followed the standard protocols as summarized by Maniatis (24). Poly $(A)^+$ RNA was prepared as described by Hovemann et al. (23). Oligonucleotide primers were obtained from the service unit of the ZMBH. cDNA libraries were used that were originally

constructed by the laboratory of T. Kornberg. We screened batch E6 and Q4. Drosophila melanogaster strains Canton S (CS) and Oregon R (OR) were grown as mass population on standard cornmeal agar medium in fly cages. (25).

RESULTS AND DISCUSSION

In an approach to clone sex dependently regulated genes of Drosophila melanogaster we had previously described the isolation of two closely related genes that were arbitrarily named F1 and F2. According to Northern blot analysis, F1 RNA was found in all stages of development whereas F2 RNA was not present in early embryos and was peaking in pupae. We now identified the function of their putative protein products by comparison of their gene sequences with those of a series of recently published elongation factor (EF - 1α) genes. The genomic sequences of both genes (the F1 sequence has partly been published by us (1)) have now been determined in full length and are shown in Figure 1 and 2. In order to determine the length of the 5' untranslated leader sequences preceeding the translational reading frame of F1 and F2 we used two independent approaches. We determined the 5' end of the RNA indirectly by oligonucleotide primed cDNA synthesis and, in addition, we isolated and sequenced cDNA clones in an attempt to obtain full length RNA copies (Fig. 3, 4).

Gene structure of Dm EF - 1α,F1

For indirect RNA sequencing and primer extension we used a synthetic oligonucleotide complementary to positions -14 to +6 relative to the translational start site (Fig. 1 and 3). The sequence determined by reverse transcription of the F1 RNA deviated from the corresponding strand of the genomic DNA after the C at position -20 (Fig. 1). Indirect RNA sequencing as well as the primer extension experiments suggested that the 5' mRNA leader sequence was encoded by at least one additional miniexon of together 60 nucleotides (Fig.3). This was corroborated by the fact that the sequence difference started with an AG dinucleotide flanked by the splicing acceptor site consensus sequence observed in Drosophila (26). At the same time we

```
   1 GAATTCCAGA AGAAAAAGTA TTTAAACGTT ATATAAATTT AGTTATAAAT TTTTACAATA TTGTATAAAT TGATACAATA CCAAATTATT

  91 TCTAGTTTTA CTTGTAGCAT ATTTTAATTG TAGCAACGGA CGCAAAAAAA TGTATAAATA AATGAAACGT ATTACATAAA TTCAAACGAT

 181 CTCAAGCTTC CATTGTTATT TAAAGTTCTA TTACGTTAGG GTTCACATAC AAATTAAAGT GGCAGGTTCT ATCTCAAAAC ATTCGTTCAA

 271 AATGCGGACT ACTAATGCAA TTGTTATTGT TTTTACATAT TAAAAGATAT GTGTTCCAAT ATTACGTATA GAAATTATAG ACATCGTTTT

 361 GTAGAAAATA CTTTTGGAAT CACTGATTAT TTAGTTTTTC ATATAAAAAC AATGTCGAGC AAACAAGGTT TTTTAAATTC CTCAATCTTT

 451 AGGTTATTGT ATTTTGCCAC TTTCAATCAC TTAAATTTCA ATAAAATGAA GTGCTTCATT CGCGCGTAGT GGAAACACCG CAGTGGGAAC

 541 ACGGTTTCTG CTCTTTTGAC AGTTGCGTAG CTTCGGTCAC ACCATGTGTC AAACGAGGCT TCCTGTGCTG AGCTCTGCCG AACGCTCGTT

                                                  C
 631 CACTTTGTTC GAATCCGTCG CCGCTTAGAC TTCGTGATTT CTCATTCAGC TTATTAGAGA GTAAGTTTTA CCTGCGAGGC TATAATTAAG

 721 TGATTTCTGC AAAAAAACTG CAGGGGGGAA ACAATTTATA AACAAATATG CAGCTGAGAC GCCGAATTTG TGCATATTTC CAGTGTTTTT

 811 CCTGTGTGTG TGTAATAAAC CCGGAGATAA CCTCTAACTG CGGTTTTCCA AAGTGAAAGG TGGCCATAGA AGCAAACACG TGGCAAGTCT

 901 GCAAAGGCAA AAATTTTAAC TGGCGTTCCC AGTTAAAGTT CCCAGCATTC TCAAAATAAT TTTCCGGCTT TTCCGGCCGC ATTTTCGCCC

 991 TGCAATATGG TGCACTTAGC GTGTAATTAC TTTGCCACGC CCACGCCGGA CACAGAGGTC ATCCACCAGA TGTGCTCATT AACCGAGAAA

1081 AAAAAACGTG CTTTCTCTCT TGCCTTTGTC ATGGCCTATA GATATTCCTT ATTCTTTCTT TTTGCGGCAT GGAATTCTAA AATGGCGACC

1171 CAGTGGCGTG AGTCAAGTGG GCGAAAAAAT TCGCCTGGCA ACAAGCGAAA AAATGTGCTT TTTTGGGTTT CCAGCCCATT AGCATATCTG

1261 GTGTAATGGC ACTCGCATCA GCTATTTCGC CATTTCCAAC CGACTCAATA ATTGGTTTTG GTAAAATGGC TGCCGCTGCA CTACGTTCTT

1351 GATTAATTCG TTGTGTGCCC CTCTCTTTTT CATTTCTTTC CAATTACCAA TTGTGCCACC GCGGCGGAGA CGCTTGCATT TGTACAAGTC

1441 ACACACGCAC ACTAATGCAC ATCCGCCATT TTGGTCTCTC TCTCTTCCTC TCTTACTTTT TCCGGCCGGC AACAGCGTCA CACAAATACA

1531 CAGGCATAGA TATACACACG CATAGGCAGA TAAGCACATG TGTATTTGCG AATTAAATTT GCTGGAATTT TCCTTTGGAC TCTTCGATTT

1621 AACATGATGA TGATTTTTCA GTTCTGCTAC TGAAGAGAGT TGACAGAAAG CAAAAATACC AAAATCACTG AAACAAAATC GAGTTTCCAT

1711 ATGGAATTTT ATTTGCACGC TCTTTTCTGT AGTTGCGCCC CACTCGTTTT ACCCACACCC CTACATGCGG GCACTGGTCC TAACCTCAAA

1801 AAACACGTTT TGTACGGCTG CAAGAGTTTG AGGTTAGGTT GTGCTCGCGC ATGCAAACAA AAGTCGAACG TACGCTAGGG AAATGAGAAA

1891 GTGTTATACC CACTAATAAT TGTAGTTGTA ATCCCACCGA ATTGTTTTAC CCTTTGTTTA TTCCAACCTC TCTTGCTCGC CAACCCGCCG

1981 AACCCTGCAA CCTTCCAATG TTCCAACGTT CCGTTAATCC AACACTCGAA TACACACAAC AGCCATAGTG TAATCATCCA AC

2063 ATG GGC AAG GAA AAG ATT CAC ATT AAC ATT GTC GTG ATC GGA CAC GTC GAT TCC GGT AAG TCG ACC ACC ACC GGA
     Met Gly Lys Glu Lys Ile His Ile Asn Ile Val Val Ile Gly His Val Asp Ser Gly Lys Ser Thr Thr Thr Gly

2138 CAC TTG ATC TAC AAG TGC GGT GGT ATC GAC AAG CGT ACC ATC GAG AAG TTC GAG AAG GAG GCC CAG GAG ATG GGA
     His Leu Ile Tyr Lys Cys Gly Gly Ile Asp Lys Arg Thr Ile Glu Lys Phe Glu Lys Glu Ala Gln Glu Met Gly

2213 AAG GGA TCC TTC AAG TAC GCC TGG GTT TTG GAT AAG TTG AAG GCT GAG CGC GAG CGT GGT ATC ACC ATC GAT ATC
     Lys Gly Ser Phe Lys Tyr Ala Trp Val Leu Asp Lys Leu Lys Ala Glu Arg Glu Arg Gly Ile Thr Ile Asp Ile

2288 GCC CTG TGG AAG TTC GAA ACT GCC AAG TAC TAC GTG ACC ATC ATT GAT GCC CCC GGA CAC AGG GAT TTC ATC AAG
     Ala Leu Trp Lys Phe Glu Thr Ala Lys Tyr Tyr Val Thr Ile Ile Asp Ala Pro Gly His Arg Asp Phe Ile Lys

2363 AAC ATG ATC ACT GGT ACC TCG CAG GCC GAT TGC GCC GTG CAG ATT GAC GCC GCC GGA ACC GGA GAA TTC GAG GCC
     Asn Met Ile Thr Gly Thr Ser Gln Ala Asp Cys Ala Val Gln Ile Asp Ala Ala Gly Thr Gly Glu Phe Glu Ala

2438 GGT ATC TCG AAG AAC GAC CAG ACC CGC GAG CAC GCC CTG CTC GCC TTC ACC CTG GGT GTG AAG CAG CTG ATC GTT
     Gly Ile Ser Lys Asn Asp Gln Thr Arg Glu His Ala Leu Leu Ala Phe Thr Leu Gly Val Lys Gln Leu Ile Val

2513 GGT GTG AAC AAG ATG GAC TCC TCC GAG CCA CCA TAC AGC GAG GCC CGT TAT GAG GAA ATC AAG AAG GAA GTG TCC
     Gly Val Asn Lys Met Asp Ser Ser Glu Pro Pro Tyr Ser Glu Ala Arg Tyr Glu Glu Ile Lys Lys Glu Val Ser

2588 TCT TAC ATC AAG AAG GTC GGC TAC AAC CCA GCC GCC GTT GCC TTC GTG CCC ATT TCC GGA TGG CAC GGC GAC AAC
     Ser Tyr Ile Lys Lys Val Gly Tyr Asn Pro Ala Ala Val Ala Phe Val Pro Ile Ser Gly Trp His Gly Asp Asn
```

```
2663 ATG TTG GAA CCC TCT ACC AAC ATG CCC TGG TTC AAG GGA TGG GAA GTG GGA CGC AAG GAG GGT AAC GCT GAC GGC
     Met Leu Glu Pro Ser Thr Asn Met Pro Trp Phe Lys Gly Trp Glu Val Gly Arg Lys Glu Gly Asn Ala Asp Gly

2738 AAG ACC CTG GTC GAT GCC CTC GAT GCC ATC CTT CCC CCA GCC CGT CCC ACC GAC AAG GCC CTG CGT CTG CCC CTG
     Lys Thr Leu Val Asp Ala Leu Asp Ala Ile Leu Pro Pro Ala Arg Pro Thr Asp Lys Ala Leu Arg Leu Pro Leu

2813 CAG GAT GTG TAC AAA ATT GGC GGT ATT GGA ACA GTA CCC GTG GGT CGT GTG GAG ACT GGT GTG CTG AAG CCC GGT
     Gln Asp Val Tyr Lys Ile Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Val Leu Lys Pro Gly

2888 ACC GTT GTG GTC TTC GCC CCT GCT AAC ATC ACC ACT GAG GTC AAG TCC GTG GAG ATG CAC CAC GAG GCC CTG CAG
     Thr Val Val Val Phe Ala Pro Ala Asn Ile Thr Thr Glu Val Lys Ser Val Glu Met His His Glu Ala Leu Gln

2963 GAG GCC GTT CCC GGA GAC AAC GTT GGC TTC AAC GTC AAG AAC GTG TCC GTG AAG GAG CTG CGT CGT GGC TAC GTT
     Glu Ala Val Pro Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Glu Leu Arg Arg Gly Tyr Val

3038 GCC GGT GAC TCC AAG GCT AAC CCC CCC AAG GGA GCC GCC GAC TTC ACC GCC CAG GTC ATC GTG CTG AAC CAC CCC
     Ala Gly Asp Ser Lys Ala Asn Pro Pro Lys Gly Ala Ala Asp Phe Thr Ala Gln Val Ile Val Leu Asn His Pro

3113 GGT CAG ATT GCC AAC GGC TAC ACC CCA GTG TTG GAT TGC CAC ACC GCT CAC ATT GCT TGC AAG TTC GCT GAG ATC
     Gly Gln Ile Ala Asn Gly Tyr Thr Pro Val Leu Asp Cys His Thr Ala His Ile Ala Cys Lys Phe Ala Glu Ile

3188 TTG GAG AAG GTC GAC CGT CGT TCC GGC AAG ACC ACC GAG GAG AAC CCC AAG TTC ATC AAG TCT GGC GAT GCT GCC
     Leu Glu Lys Val Asp Arg Arg Ser Gly Lys Thr Thr Glu Glu Asn Pro Lys Phe Ile Lys Ser Gly Asp Ala Ala

3263 ATC GTC AAC CTG GTG CCC TCT AAG CCC CTG TGC GTG GAG GCC TTC CAG GAG TTC CCC CCT CTG GGT CGC TTC GCT
     Ile Val Asn Leu Val Pro Ser Lys Pro Leu Cys Val Glu Ala Phe Gln Glu Phe Pro Pro Leu Gly Arg Phe Ala

3338 GTG CGT GAC ATG AGG CAG ACC GTG GCT GTC GGT GTC ATT AAG GCT GTC AAC TTC AAG GAT GCC TCC GGT GGC AAG
     Val Arg Asp Met Arg Gln Thr Val Ala Val Gly Val Ile Lys Ala Val Asn Phe Lys Asp Ala Ser Gly Gly Lys

3413 GTC ACC AAG GCC GCC GAG AAG GCC ACC AAG GGC AAG AAG TAG CTGGTTTGCT TCCACTCAAC AACAACAACA ACACGCAGTA
     Val Thr Lys Ala Ala Glu Lys Ala Thr Lys Gly Lys Lys ***

3495 GTAGCAGCAA CAACAAGCAT ATAACCAACA TCATAATGCA GCCAACAACA CCACTCAATA ATACCAGCAA CAGCAGCAGC GAACACAATA

3585 GTAGTATAAC ACCAACACCT GTCCTGCGCA AGATGACCGA TAAGATGATG TTTCAGCAGA AGCATAAGTT TAATTTCTTC CATCGAAAGG

3675 AGTTTCGACG GATACGAATG CTAAATGCAG ACGAGGCCGC CTTCACTGGG AAATCGGTGG ATCCCAAGGA TAAGAGTGCA CACTGGGAAA

3765 ACACTTGCAT TTATGCATCC ACTCCTCATC CACTTCCCCG TCGATCTTTA GTTTACTAAA TATGGTATGA TGCACGCAGT TGACTTCGTT

3855 TTATCATATC ATATATAGGA ATCCTCTGTA GCATTTATGA TATCGTTTAA ATTAACCTTT ATACTTTGAT ATGTATCATT TATCTTACCC

3945 TACTTTTGCA CACACTACTT TGTACACAAG AAAAGAACCA GAATAGAAGC GATAAACTAT ATTTACAAAA AAAATAAAAA CCCTATTTTT

4035 GTATTTCTTT TGTTTTTACC ACCCAGCCCG TAAAAGAGCA CTCTCTTTTT GGTTGTTGCC TCCCGATTT
```

FIGURE 1
Sequence of the EF - 1α,F1 gene region. Throughout the coding portion aminoacids are written below the sequence. The intron area is boxed. Transcription start and polyadenylation site are indicated by an arrow and a dot, respectively. A sequence heterogeneity observed in cDNA cDm19 is indicated at position 686. The oligonucleotide primer is shown by an arrow (Pos. 2049-2068). Conserved sequence blocks are underlined.

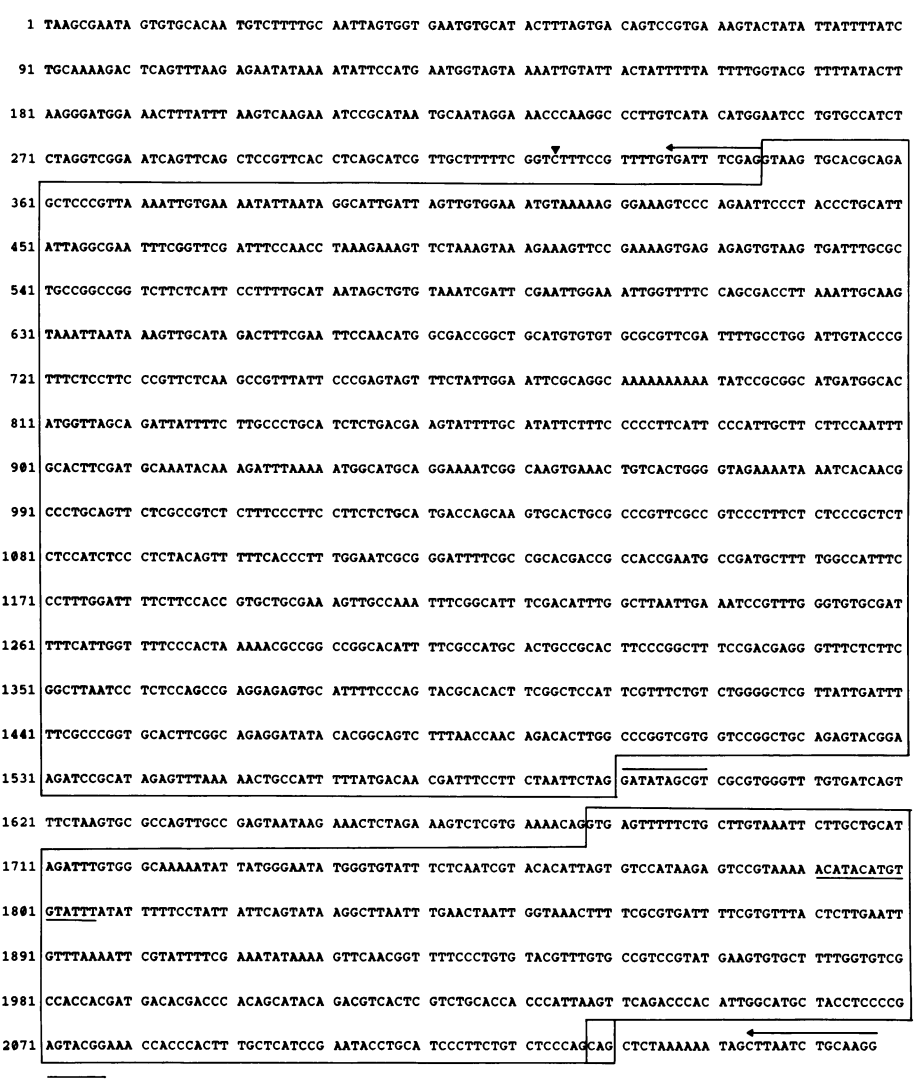
isolated several F1 cDNA clones from a λgt10 library which had been prepared from embryonal poly $(A)^+$ RNA. By comparison of their restriction site patterns with the genomic map some isolates seemed to contain a mRNA copy including the very 5'end. One such cDNA (cDm 19, Fig.3) was subsequently sequenced and showed the same 5' end portion when compared to the indirect RNA sequencing results. The heterogeneity observed at nucleotide 56 of the first exon may be attributed to a difference in the inbred strains Canton S and Oregon R that

were used for genomic and cDNA library construction, respectively (Fig. 1).

In order to locate the 5' end of EF - 1α,F1 mRNA within the genomic sequence, primer extended ^{32}P - labeled cDNA was hybridized to a blot of DNA containing the subcloned BamHI and two HindIII fragments preceeding about 10 kb upstream of the EF - 1α,F1 coding portion. Alternatively, cDNA clone cDm19 was hybridized to a blot of the F1 gene and its upstream sequences included in phage λCS1 (Fig. 3). Consistently, we observed hybridization to the 1.3 kb and the neighbouring 1.2 kb EcoRI fragments. In order to be able to precisely localize the 5' exon(s) in the genomic DNA we subsequently sequenced both EcoRI fragments (Fig. 1). The resulting data showed that the first exon is contained within the 1.2 kb EcoRI fragment. EF - 1α,F1 thus extents over 3.4 kb and consists of two exons that are separated by a 1.3 kb intron (Fig 3).

Structure and sequence of the EF - 1α,F2

An EF - 1α,F2 containing 4.0 kb EcoRI fragment of genomic DNA that had been identified by crosshybridization to the EF - 1α,F1 cDm49 probe (1) has been sequenced using both the methods of Maxam and Gilbert and of Sanger (Fig. 2). Comparison of this genomic sequence with those of a series of cDNA clones isolated from cDNA libraries of pupal RNA by hybridization with the same 4.0 kb EcoRI fragment revealed a single long open reading frame. It encoded a second, related elongation factor protein that is one amino acid shorter and shows 90% homology when compared to EF - 1α,F1 (Fig. 8). Two introns that interrupt the coding portion of the F2 gene had already been mapped (1). To determine the transcriptional start of the EF - 1α,F2 gene we proceeded the same way as described for F1 and used a synthetic oligonucleotide primer complementary to nucleotides -14 to +6 relative to the translation start. The major cDNA product primed with this oligonucleotide on pupal RNA was 142 nucleotides long (Fig. 4). In addition, cDNA clones that possibly reached the 5' end of the F2 mRNA were isolated from a library prepared from pupal poly $(A)^+$ RNA. cDNA pc3 was selected as the clone that extended farthest to the 5' end.


1 TAAGCGAATA GTGTGCACAA TGTCTTTTGC AATTAGTGGT GAATGTGCAT ACTTTAGTGA CAGTCCGTGA AAGTACTATA TTATTTTATC
91 TGCAAAAGAC TCAGTTTAAG AGAATATAAA ATATTCCATG AATGGTAGTA AAATTGTATT ACTATTTTTA TTTTGGTACG TTTTATACTT
181 AAGGGATGGA AACTTTATTT AAGTCAAGAA ATCCGCATAA TGCAATAGGA AACCCAAGGC CCTTGTCATA CATGGAATCC TGTGCCATCT
271 CTAGGTCGGA ATCAGTTCAG CTCCGTTCAC CTCAGCATCG TTGCTTTTTC GGTCTTTCCG TTTTGTGATT TCGAGGTAAG TGCACGCAGA
361 GCTCCCGTTA AAATTGTGAA AATATTAATA GGCATTGATT AGTTGTGGAA ATGTAAAAAG GGAAAGTCCC AGAATTCCCT ACCCTGCATT
451 ATTAGGCGAA TTTCGGTTCG ATTTCCAACC TAAAGAAAGT TCTAAAGTAA AGAAAGTTCC GAAAAGTGAG AGAGTGTAAG TGATTTGCGC
541 TGCCGGCCGG TCTTCTCATT CCTTTTGCAT AATAGCTGTG TAAATCGATT CGAATTGGAA ATTGGTTTTC CAGCGACCTT AAATTGCAAG
631 TAAATTAATA AAGTTGCATA GACTTTCGAA TTCCAACATG GCGACCGGCT GCATGTGTGT GCGCGTTCGA TTTTGCCTGG ATTGTACCCG
721 TTTCTCCTTC CCGTTCTCAA GCCGTTTATT CCCGAGTAGT TTCTATTGGA ATTCGCAGGC AAAAAAAAAA TATCCGCGGC ATGATGGCAC
811 ATGGTTAGCA GATTATTTTC TTGCCCTGCA TCTCTGACGA AGTATTTTGC ATATTCTTTC CCCCTTCATT CCCATTGCTT CTTCCAATTT
901 GCACTTCGAT GCAAATACAA AGATTTAAAA ATGGCATGCA GGAAAATCGG CAAGTGAAAC TGTCACTGGG GTAGAAAATA AATCACAACG
991 CCCTGCAGTT CTCGCCGTCT CTTTCCCTTC CTTCTCTGCA TGACCAGCAA GTGCACTGCG CCCGTTCGCC GTCCCTTTCT CTCCCGCTCT
1081 CTCCATCTCC CTCTACAGTT TTTCACCCTT TGGAATCGCG GGATTTTCGC CGCACGACCG CCACCGAATG CCGATGCTTT TGGCCATTTC
1171 CCTTTGGATT TTCTTCCACC GTGCTGCGAA AGTTGCCAAA TTTCGGCATT TCGACATTTG GCTTAATTGA AATCCGTTTG GGTGTGCGAT
1261 TTTCATTGGT TTTCCCACTA AAAACGCCGG CCGGCACATT TTCGCCATGC ACTGCCGCAC TTCCCGGCTT TCCGACGAGG GTTTCTCTTC
1351 GGCTTAATCC TCTCCAGCCG AGGAGAGTGC ATTTTCCCAG TACGCACACT TCGGCTCCAT TCGTTTCTGT CTGGGGCTCG TTATTGATTT
1441 TTCGCCCGGT GCACTTCGGC AGAGGATATA CACGGCAGTC TTTAACCAAC AGACACTTGG CCCGGTCGTG GTCCGGCTGC AGAGTACGGA
1531 AGATCCGCAT AGAGTTTAAA AACTGCCATT TTTATGACAA CGATTTCCTT CTAATTCTAG GATATAGCGT CGCGTGGGTT TGTGATCAGT
1621 TTCTAAGTGC GCCAGTTGCC GAGTAATAAG AAACTCTAGA AAGTCTCGTG AAAACAGGTG AGTTTTTCTG CTTGTAAATT CTTGCTGCAT
1711 AGATTTGTGG GCAAAAATAT TATGGGAATA TGGGTGTATT TCTCAATCGT ACACATTAGT GTCCATAAGA GTCCGTAAAA ACATACATGT
1801 GTATTTATAT TTTTCCTATT ATTCAGTATA AGGCTTAATT TGAACTAATT GGTAAACTTT TCGCGTGATT TTCGTGTTTA CTCTTGAATT
1891 GTTTAAAATT CGTATTTTCG AAATATAAAA GTTCAACGGT TTTCCCTGTG TACGTTTGTG CCGTCCGTAT GAAGTGTGCT TTTGGTGTCG
1981 CCACCACGAT GACACGACCC ACAGCATACA GACGTCACTC GTCTGCACCA CCCATTAAGT TCAGACCCAC ATTGGCATGC TACCTCCCCG
2071 AGTACGGAAA CCACCCACTT TGCTCATCCG AATACCTGCA TCCCTTCTGT CTCCCAGCAG CTCTAAAAAA TAGCTTAATC TGCAAGG
2158 ATG GGC AAG GAG AAG ATC CAT ATT AAC ATT GTG GTC ATT GGC CAT GTG GAC TCC GGC AAG TCG ACG ACC ACC GGC
Met Gly Lys Glu Lys Ile His Ile Asn Ile Val Val Ile Gly His Val Asp Ser Gly Lys Ser Thr Thr Thr Gly
2233 CAC TTG ATC TAC AAA TGC GGC GGC ATC GAC AAG CGT ACG ATT GAG AAG TTC GAG AAG GAG GCC CAG GAA ATG GGA
His Leu Ile Tyr Lys Cys Gly Gly Ile Asp Lys Arg Thr Ile Glu Lys Phe Glu Lys Glu Ala Gln Glu Met Gly
2308 AAA GGC TCC TTT AAG TAC GCT TGG GTA CTG GAC AAG CTG AAG GCA GAG CGG GAG CGG GGC ATC ACC ATC GAC ATT
Lys Gly Ser Phe Lys Tyr Ala Trp Val Leu Asp Lys Leu Lys Ala Glu Arg Glu Arg Gly Ile Thr Ile Asp Ile
2383 GCC CTA TGG AAG TTC GAG ACG TCC AAG TAC TAT GTG ACC ATC ATC GAT GCC CCT GGT CAC AGG GAT TTC ATC AAG
Ala Leu Trp Lys Phe Glu Thr Ser Lys Tyr Tyr Val Thr Ile Ile Asp Ala Pro Gly His Arg Asp Phe Ile Lys
2458 AAC ATG ATT ACC GGT ACC TCT CAG GCC GAT TGT GCG GTG CTG ATC GAC GCC GCC GGA ACT GGA GAG TTC GAG GCC
Asn Met Ile Thr Gly Thr Ser Gln Ala Asp Cys Ala Val Leu Ile Asp Ala Ala Gly Thr Gly Glu Phe Glu Ala
2533 GGG ATC TCG AAG AAC GGC CAG ACC CGC GAG CAC GCC CTT CTG GCA TTC ACG CTG GGC GTG AAG CAG CTT ATT GTG
Gly Ile Ser Lys Asn Gly Gln Thr Arg Glu His Ala Leu Leu Ala Phe Thr Leu Gly Val Lys Gln Leu Ile Val
2608 GGC GTC AAC AAG ATG GAC TCC ACT GAG CCG CCG TAC AGC GAG GCC CGC TAC GAG GAG ATC AAG AAG GAG GTG TCC
Gly Val Asn Lys Met Asp Ser Thr Glu Pro Pro Tyr Ser Glu Ala Arg Tyr Glu Glu Ile Lys Lys Glu Val Ser

```
2683 TCG TAC ATC AAG AAG ATC GGC TAC AAT CCG GCC TCG GTG GCC TTC GTG CCC ATC TCC GGA TGG CAC GGC GAC AAT
     Ser Tyr Ile Lys Lys Ile Gly Tyr Asn Pro Ala Ser Val Ala Phe Val Pro Ile Ser Gly Trp His Gly Asp Asn

2758 ATG CTG GAG CCG TCC GAG AAG ATG CCC TGG TTC AAG GGA TGG TCC GTG GAG CGC AAG GAA GGC AAG GCA GAG GGC
     Met Leu Glu Pro Ser Glu Lys Met Pro Trp Phe Lys Gly Trp Ser Val Glu Arg Lys Glu Gly Lys Ala Glu Gly

2833 AAG TGC TTG ATC GAC GCG CTG GAC GCG ATC CTT CCA CCC CAG CGT CCC ACC GAC AAG CCG CTG CGC CTG CCG CTC
     Lys Cys Leu Ile Asp Ala Leu Asp Ala Ile Leu Pro Pro Gln Arg Pro Thr Asp Lys Pro Leu Arg Leu Pro Leu

2908 CAG GAC GTC TAC AAG ATC GGA GGC ATC GGA ACC GTA CCA GTA GGT CGT GTG GAG ACT GGT CTC CTC AAG CCA G
     Gln Asp Val Tyr Lys Ile Gly Gly Ile Gly Thr Val Pro Val Gly Arg Val Glu Thr Gly Leu Leu Lys Pro

2981 GTAAGGCTCC GGGTTGATGA GGTCGGGTGT GGGCCCTCTT TTCTCTTTGG GCACTTCATA CATGTATTCT GCAAAATTTG GGTCGACAGT

3071 GGGCTGGCAT CCAACAGCCA CCGCCTCCAA AGCGGAGCCG CAACGAAGTC TTGCGCATGT ATGCATTATT GAGCGAACGT CTTCGTCGAG

3161 AGCGAGACCC TCCACCTCAT GCACTTGGTG AAATTCTCAC TCCGAAGAGC TTCCATTTTC AACATGAAAG TGAAAGGCCA TTAAAATAAA

3251 ATAACCCTAG CTAACATATT AATATATGTA GAGCTATTGA TTCAAATAAA AATAAATTGG AGTTAGTTCG AATAATATCG CTCCACGTTT

3341 CTCTCTCTGT ATGCACCCAC CCCCATCCAA ATGTCTACAC ATAACGTCCG GATATGTAAC TTCGTTTCGG TCGCTTCGTT TCCGGTTTCG

3431 TTTCAG    GC ATG GTC GTC AAC TTT GCG CCG GTC AAC CTG GTC ACC GAA GTA AAG TCT GTG GAG ATG CAC CAC
                Gly Met Val Val Asn Phe Ala Pro Val Asn Leu Val Thr Glu Val Lys Ser Val Glu Met His His

3502 GAG GCT CTC ACC GAA GCC ATG CCC GGC GAC AAC GTT GGC TTC AAC GTG AAG AAC GTG TCC GTG AAG GAG CTC CGT
     Glu Ala Leu Thr Glu Ala Met Pro Gly Asp Asn Val Gly Phe Asn Val Lys Asn Val Ser Val Lys Glu Leu Arg

3577 CGT GGC TAT GTG GCC GGC GAT TCC AAG AAC AAT CCT CCT AGG GGA GCA GCC GAC TTT ACC GCT CAG GTAGGGTAAC
     Arg Gly Tyr Val Ala Gly Asp Ser Lys Asn Asn Pro Pro Arg Gly Ala Ala Asp Phe Thr Ala Gln

3653 AAAGATGAGA AATCTTTGAT AGTTGAACTC ATCTTTGTTT GGTTTTTTTT TTTTCTTTTT GCCCACAG    GTG ATT GTG CTC AAC
                                                                                  Val Ile Val Leu Asn

3736 CAT CCG GGC CAG ATC GCC AAT GGG TAC ACT CCC GTC TTG GAT TGC CAC ACG GCG CAC ATT GCC TGC AAG TTT TCC
     His Pro Gly Gln Ile Ala Asn Gly Tyr Thr Pro Val Leu Asp Cys His Thr Ala His Ile Ala Cys Lys Phe Ser

3811 GAG ATC AAG GAG AAG TAC GAC CGC CGT ACG GGC GGA ACC ACC GAA GAC GGG CCG AAG GCT ATC AAG TCC GGG GAT
     Glu Ile Lys Glu Lys Tyr Asp Arg Arg Thr Gly Gly Thr Thr Glu Asp Gly Pro Lys Ala Ile Lys Ser Gly Asp

3886 GCG GCC ATC ATT GTG CTG GTG CCC AGC AAG CCG TTG TGC GTA GAG AGC TTC CAG GAG TTC CCA CCG CTG GGA CGG
     Ala Ala Ile Ile Val Leu Val Pro Ser Lys Pro Leu Cys Val Glu Ser Phe Gln Glu Phe Pro Pro Leu Gly Arg

3961 TTC GCT GTG CGC GAC ATG AGG CAG ACC GTG GCC GTG GGC GTC ATC AAG TCG GTG AAC TTT AAA GAG ACG ACC TCG
     Phe Ala Val Arg Asp Met Arg Gln Thr Val Ala Val Gly Val Ile Lys Ser Val Asn Phe Lys Glu Thr Thr Ser

4036 GGC AAG GTG ACA AAA GCC GCT GAG AAG GCA CAG AAG AAG AAA TAA CTAGGGTACC AGCAGAACAA CGTCATCACT
     Gly Lys Val Thr Lys Ala Ala Glu Lys Ala Gln Lys Lys Lys ***

4111 CGAACCCAAC AACAACAAAA ACAGACGGCT AGAGCAACAG CAGCAACAAC ACACAACAAC AATACACATG TCAAAATTAT AATACCCACT

4201 CGACGATCAA ATTCACACCT TGACTCCATG GCAAGAGAGA CACCAATTAC TACTATTACT AGCTGCTGGG AGAAGCGGCA GATATTAACC

4291 GAAATCGAGC AGATTATACC CTATATAATA ACCACACGTA CGATTAGCGA GGAGAGGAGC ATCAGGTGCA GCGAGGATGC GAAGGAGGAG

4381 CCCTTCCAGC CTCGCCGGGT CGGTTTTGGT CGCCTTCGCC GTGGTGGTCT ACTGCAGCTA TCTGAACATG TATCGTCACC GCAAGTCCTT

4471 TCGTAGGAAA CCACCCGCTA GCCACTCCGC AGAGTGGATA GGGGCCTCCG GAGCACTGCT GTAGCCCGCC CCTTCGATAT ATACTCATCT

4561 CTAAAACTAA CCTTACACTT GATTAGCAGC CACACATCCG GTCGCATCCA CCTGTTTCGA ATGGATTTTA AACACTTTTT ATACTTTTGA

4651 TAAGTCAAGT CGGAGGCATT CGATTTAAAA TCTATTGAAA TATGTAATTT CCGAATTTAG TTTTAAACCA CGTCCGCGCT CCCAAAAATC

4741 CCCCGAACCG AAAAGACTAC ATTCGCGATG AATTCAAAAT TTCTCTTGAA ACCAAAAAAA ACAAATGCTT AAGAAGTATT ACAAAAAAGA

4831 AATCAACATT ACACACATAA TCATGCGGTT TTTGAAAACA TTATAAATGT TTAATCGAGC CTCATTTGCA TTTGCATATT ACATAATATA

4921 CGTTAGCCAC ATGTCATCTC ATTGCCCATA ATAACCTGCA TCCTGCATAT TATACACGTT AATCTCACAC TCTGAATTTA TACAAACCGA

5011 AGACAATTGT AACCGACACC AGAACAATTC TTGGATACAG AACATGTTGG CTTGATAAAA GATCTTTTAA ATGATGAGAA AAATAAAGGA

5101 AGCTTAACCG TAAAATACCA CACACGAACG CCTTTTAATT GAAAAATACT TGAATATCTA TGAAGAAAAT GAATTC
```

FIGURE 2
Sequence of the EF - 1α,F2 gene region. The localization of transcription start, polyadenylation site, intron, exon, coding and noncoding portions is denoted as in Figure 1. According to cDNA pC3 and pC4 different splice acceptor sites at position 2129 and 2132, respectively, are used. Oligonucleotide primer sequences are indicated by an arrow above the sequence. Note that due to the intron the arrow representing the primer close to the 5'end is split. A conserved sequence motif is underlined.

Hybridization of the oligonucleotide primed cDNA of 142 nucleotides in length to an EcoRI XbaI digest of the subcloned fragment 1 (Fig.4) and of the 5'EcoRI - XbaI fragment of cDNA pc3 to a blot of λ17 DNA containing 10 kb sequence upstream of the F2 reading frame gave rise to hybridization signals in two different regions (Fig.4). Extending the sequence analysis of the genomic DNA to about 2.5 kb upstream from the translation start and determination of the cDNA pc3 sequence resulted in the localization of two 5' miniexons. They were separated from the coding portion of the gene by a 0.45 kb intron at position -27 and a 1.25 kb intron at position -114 relative to the translational start. When compared with the length of the oligonucleotide extension product of 142 nucleotides, cDNA pc3 was missing 1 nucleotide at the 5' end. In order to map the transcription start of F2 to the nucleotide precisely, we repeated the primer extension experiment with an oligonucleotide complementary to position
-124 to -104 relative to the translational start (position 336 in Fig: 2). An extension product of 12 nucleotides (data not shown) confirmed the result already obtained with the first primer. In summary: EF - 1α,F2 is organized in five exons separated by intron sequences of 1.24, 0.45, 0.45, and 0.08 kb in length and extends over 4.8kb.

Stable promoter elements

The use of alternative promoter elements has been shown to be a means of differential gene regulation. In order to find out whether alternative promoters were also used for the expression of EF - 1α genes during D. melanogaster development, we performed primer extension experiments using poly $(A)^+$ RNA from

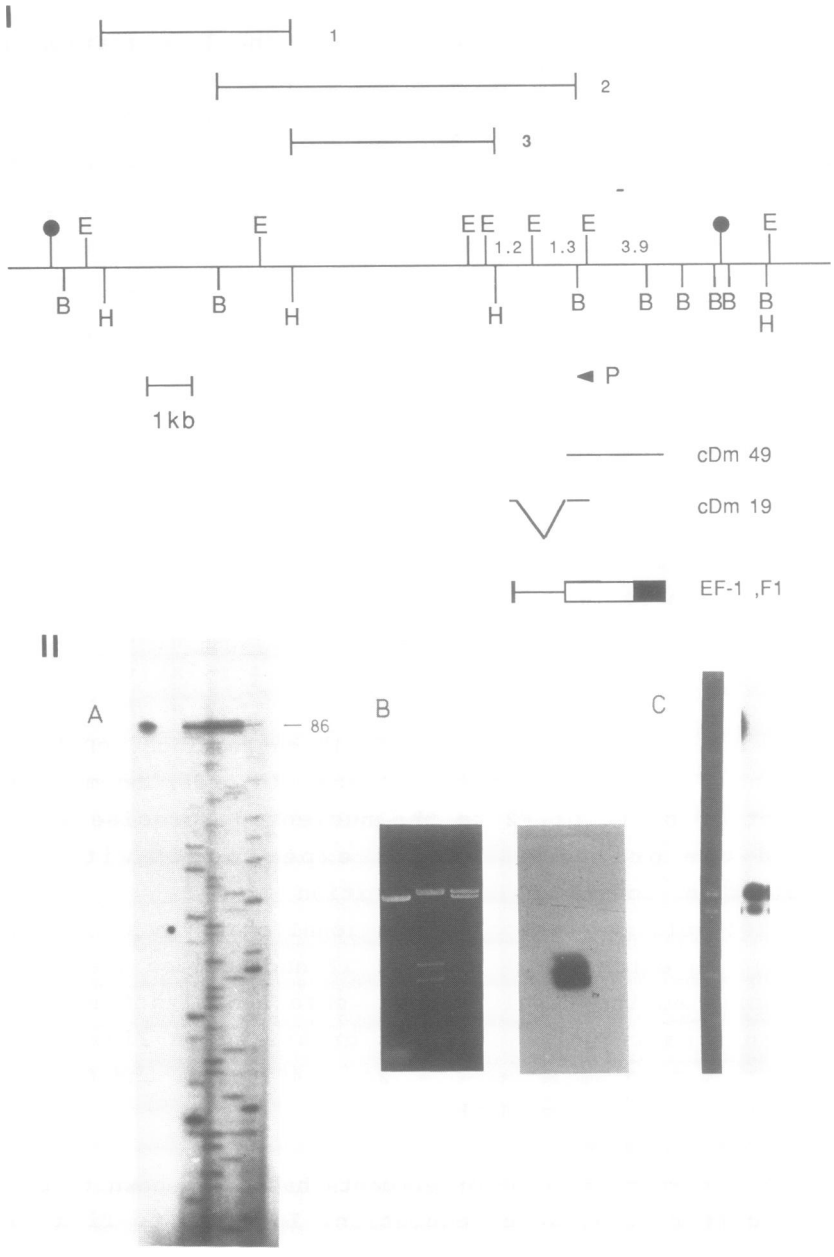


FIGURE 3
Structure of the EF - 1α,F1 gene locus.
Part I shows a restriction map of phage λ CS1 ( ) containing the F1 gene. Restriction sites are shown for EcoRI (E), HindIII

(H), and BamHI (B). The size of the EcoRI fragments harbouring the F1 gene is indicated in kilobases. Subcloned HindIII and BamHI fragments used to localize exon I are numbered 1-3. Below the restriction site map the localization of cDNAs 19 and 49 is denoted by horizontal lines. The complete F1 gene structure is given as thin line and boxed area for intron and exon sequences, respectively. The open box confines the coding part. P together with a filled triangel denotes the position from which the oligonucleotide primed cDNA synthesis is started.
Part II: Panel A shows the cDNA synthesis product and the indirect RNA sequencing result. Panel B shows an Ethbr stained agarose gel of the subcloned HindIII and BamHI fragments 1-3 digested with HindIII and EcoRI, respectively, together with an Southern blot of this gel hybridized with the ^{32}P - labelled oligonucleotide primed cDNA. Panel C: Agarose gel separation of the EcoRI fragments of the genomic DNA cloned in phage λ CS1. Hybridization of ^{32}P labelled cDNA cDM 19 gives rise to the pattern shown to the right.

0 - 16 hr embryos (E), climbing third instar larvae (L), 2 - 4 day old pupae (P), and of adult flies (A). Since both primers were complementary to sequences encompassing the beginning of the respective EF - 1α protein coding regions (Fig.1 and 2), a switch to an alternative promoter would have been reflected in the appearance of cDNAs exhibiting different lengths. As can be seen in Figure 5, the primer extension products remained of the same length throughout development therefore indicating, that a promoter switch is obviously not apparent during either F1 or F2 expression. Since the radioactive label incorporated into the cDNAs represents the relative amount of RNA available for primer extension at the respective stage of development, our results indicate that EF - 1α,F1 RNA is present in all stages of development, although to a variable amount. EF - 1α,F2 RNA, on the other hand, is highly expressed in the pupal stage but is also present to a lesser extend in third instar larvae and flies. In addition to the major cDNA product of 142 nucleotides in length a weak additional band of about 160 nucleotides appeared possibly indicating the existence of a minor transcription start site for EF - 1α,F2.

Only two gene copies for cytoplasmic EF - 1α exist in Drosophila melanogaster

E. coli or Saccharomyces cerevisiae have been reported to contain two nearly identical gene copies of EF - 1α. In

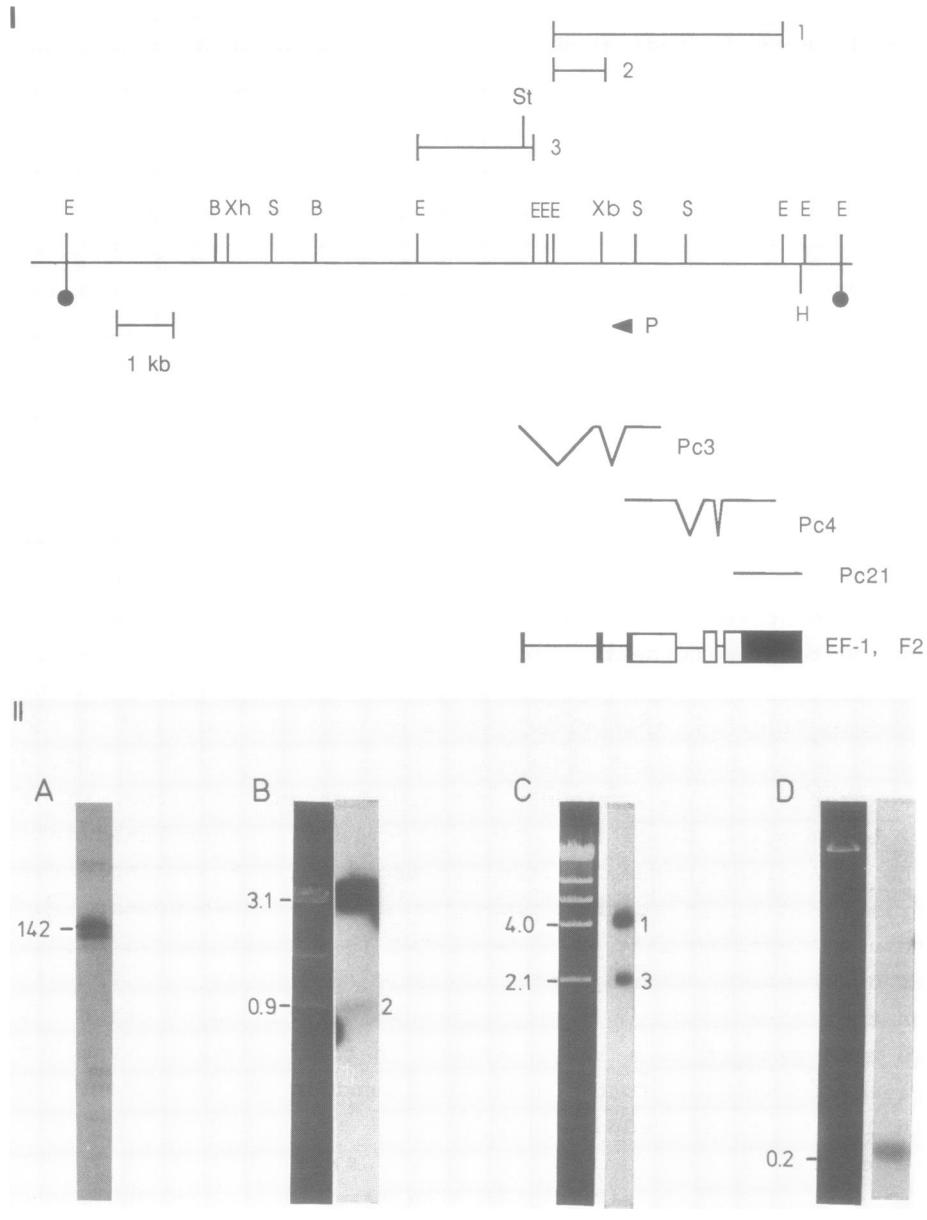


FIGURE 4
Structure of the EF - 1α,F2 gene locus. Part I shows a restriction map of the DNA cloned in phage λ17 ( ). Restriction sites are indicated for StyI (St), SalI (S), XhoI (Xh), XbaI (Xb) and as in Figure 3. 1-3 denotes subcloned

fragments used for localization of exon 1 and 2. Primer sequence, cDNA localization and resulting gene structure are shown as in Figure 3.
Part II: In panel A the cDNA synthesis products are shown. In panel B,C, and D Ethbr stained agarose gels are shown together with their Southern blot hybridization results: B subcloned fragment 1 digested with EcoRI and XbaI and hybridized with primer extended cDNA, C λ17 digested with EcoRI and hybridized with the 5'EcoRI-XbaI fragment of cDNA pC3. D subcloned fragment 3 digested with EcoRI and StyI hybridized as in in C. Only the StyI site that gives rise to the hybridizing fragment is indicated.

contrary, our results for Drosophila melanogster suggest the existence of two different elongation factor genes. In order to decide whether the two identified gene copies are the only ones that code for cytoplasmic elongation factor EF - 1α in Drosophila, we performed genomic Southern blot hybridization experiments using EF - 1α,F1 cDNA cDm49 (Fig. 3) as a probe. DNA was isolated from Canton S embryos, digested with either EcoRI or HindIII, blotted and hybridized. EcoRI digested DNA gave rise to a 4.0 kb hybridization signal of twice the intensity when compared to a second signal of 1.3 kb. Two bands appeared with HindIII digested genomic DNA (Fig 6). In each case, the hybridizing DNA fragments could be assigned to the map of the already cloned F1 and F2 DNA (Fig. 3 and 4). The part of the F1 gene, for example, that is contained in cDm49, spans a portion of the 1.3 kb and the 4.0 kb EcoRI fragments whereas the crosshybridizing part of the F2 gene is also

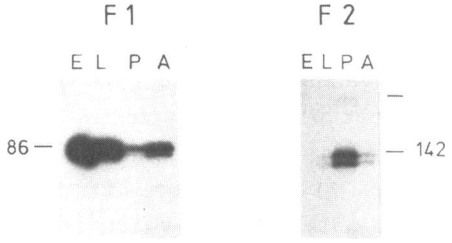


FIGURE 5
Primer extension products obtained using F1 and F2 specific oligonucleotides (see Fig.1 and 2). Poly (A)$^+$ RNA from 0-16 hour embryos (E), third instar larvae (L), two to four day old pupae (P) and flies (A) is used.

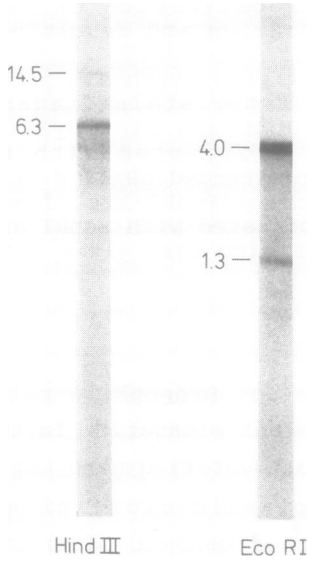


FIGURE 6
Genomic Southern hybridization with cDm49 (see Fig.3) DNA as hybridization probe. The size of the identified DNA fragments is indicated in kilobases.

located on a 4.0 kb EcoRI fragment. Titration experiments with gene specific hybridization probes revealed that both genes are present as single copy per haploid genome (34). Therefore, we can conclude that additional related gene copies encoding cytoplasmic EF - 1α do not exist in D. melanogaster.

Codon usage in the EF - 1α,F1 and F2 gene.

Since F1 (464 aminoacids) expression is generally markedly stronger when compared to F2 (463 aminoacids), we examined the codon usage of both coding regions in light of the assumption that highly expressed genes are subject to a more extreme codon bias (27-29). Consistent with the hypothesis, codon preference in F1 is restricted to 44 triplets whereas in F2 55 different codons are used (Fig. 7). Moreover, if one neglects those triplets that are used only once or twice in F1, codon usage in the residual 97% of the reading frame is biased to 34 triplets. This is different in the F2 gene where codon usage is random with a slight bias against A-T richness.

	T	F1	F2	C	F1	F2	A	F1	F2	G	F1	F2
	TTT PHE	0	5	TCT SER	4	2	TAT TYR	1	2	TGT CYS	0	1
T	TTC PHE	17	11	TCC SER	11	12	TAC TYR	10	10	TGC CYS	5	5
	TTA LEU	0	0	TCA SER	0	0	TAA OCH	0	1	TGA OPA	0	0
	TTG LEU	6	4	TCG SER	3	6	TAG AMB	1	0	TGG TRP	5	5
	CTT LEU	1	3	CCT PRO	2	3	CAT HIS	0	3	CGT ARG	11	6
C	CTC LEU	2	6	CCC PRO	17	7	CAC HIS	11	8	CGC ARG	4	6
	CTA LEU	0	1	CCA PRO	5	4	CAA GLN	0	0	CGA ARG	0	0
	CTG LEU	15	12	CCG PRO	0	11	CAG GLN	11	11	CGG ARG	0	3
	ATT ILE	11	10	ACT THR	4	4	AAT ASN	0	4	AGT SER	0	0
A	ATC ILE	20	23	ACC THR	23	16	AAC ASN	18	12	AGC SER	1	3
	ATA ILE	0	0	ACA THR	1	1	AAA LYS	1	5	AGA ARG	0	0
	ATG MET	8	10	ACG THR	0	7	AAG LYS	45	42	AGG ARG	2	3
	GTT VAL	7	1	GCT ALA	11	6	GAT ASP	12	6	GGT GLY	18	4
G	GTC VAL	14	9	GCC ALA	34	19	GAC ASP	12	16	GGC GLY	12	25
	GTA VAL	1	5	GCA ALA	0	5	GAA GLU	7	5	GGA GLY	13	10
	GTG VAL	22	25	GCG ALA	0	6	GAG GLU	25	29	GGG GLY	0	4

FIGURE 7
Codon usage as deduced from the nucleotide sequence of the F1 and F2 reading frames.

Evolutionary stability of EF - 1α genes

When we compared the amino acid sequences of EF - 1α,F1 and F2 with all eucaryotic cytoplasmic elongation factors (Fig.8) known to us, we observed a strong conservation of their primary structure. Some of the highly conserved regions have already been correlated with functional domains. The most strongly conserved NH_2 - terminal end sequence of EF - 1α had been assigned to GTP binding activity. EF - 1α also shares homology with several classes of nucleotide binding proteins and even the procaryotic elongation factor genes (37,38). Another highly conserved region further substantiated in our comparison is comprised in the sequences around Ala 92, Lys 244 and Lys 273. The corresponding amino acids in EF - Tu of *E. coli* were considered to be important for tRNA binding. This remarkable degree of sequence conservation is most easily explained if one considers the multifunctional nature of the EF - 1α protein. In the future, cloning, in combination with site directed mutagenesis, will open the way to design modified factors that can be tested for single functional steps in the elongation process.

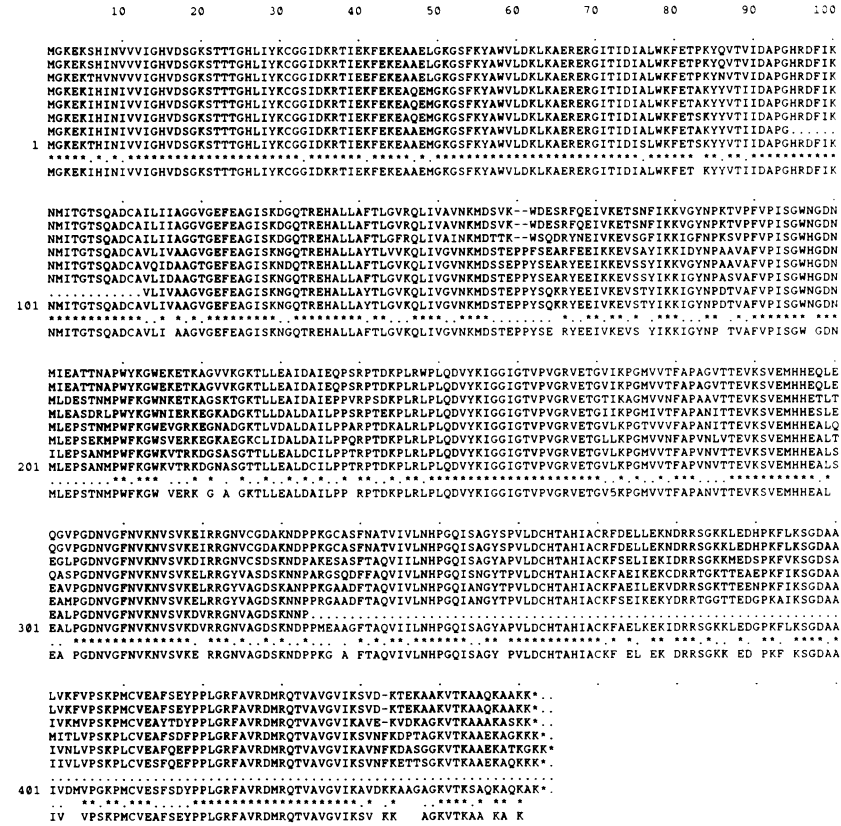


FIGURE 8
Comparison of the aminoacid sequences deduced from cloned EF - 1α gene sequences (11-19). For mouse only part of the sequence is known (17,18). Asterisks denote identical aminoacids, points indicate a conservation of at least 50% between the different proteins.

CONCLUDING REMARKS

cDNA clones were isolated that span the complete mRNA sequences of the two D. melanogaster elongation factors EF - 1α (F1 and F2). Using these clones in addition to oligonucleotide primer extended cDNA probes, we were able to characterize miniexons encoding the 5'end of the untranslated leader of both EF - 1α mRNAs. By comparison of cDNA and genomic sequences the gene structures could be determined. In contrast to the two nearly identical genes present in E. coli or Saccharomyces cerevisiae,

Drosophila contains two copies that are clearly different from each other with respect to sequence and structure. Even though it is not proven yet that both messages are actually translated in vivo, we detect both sequences in the high molecular weight fraction of polysomal gradients at their respective time of expression indicating active translation (Richter and Hovemann, unpublished result). The amount of cDNA synthesized using F1- and F2 specific oligonucleotide primers reflects the extent of RNA expression in the various developmental stages. According to this expression profile, F1 should represent the housekeeping gene that gives rise to the elongation factor needed in all growing cells. F2 transcription, peaking in pupal stage, represents on the other hand an elongation factor gene that is specifically expressed in certain developmental stages, possibly,even in a tissue specific manner. It is also conceivable that EF - 1α,F2 exerts a specific function by preferring its own pool of aminoacyl tRNAs. This would then mean that it would by itself be involved in some kind of translational control.

The complete gene structures allow us now to perform a comparison at the nucleotide level. The coding region of F1 and F2 differ in sequence between each other to a degree that is similar when compared with the Artemia salina or the human gene (15,16). We, therefore, conclude that the genetic separation of F1 and F2 is not a recent event. Since both genes established themselves with their unique expression profile, they could express independent functions. A stage specific EF - 1α like activity was also observed in Xenopus laevis previtellogenic oocytes (35,36). However, a molecular analysis of the oocyte specific activity is still missing.

Promoter regions of housekeeping function genes have been characterized of being devoid of the TATA motif (31). We did not recognize the TATA box sequence in either one of the two EF - 1α genes. Still, there is little sequence homology at all in front of the transcription start of F1 and F2. A comparable promoter motif, however, that is common to housekeeping genes is not yet established. Since EF - 1α belongs to the group of proteins that are engaged in protein synthesis, we searched for

common sequence motifs that might be indicative for the concerted regulation of genes encoding translation factors and ribosomal proteins. In the yeast Saccharomyces cerevisiae a general promoter enhancer motif, the HOMOL box (32), has been identified in front of the EF - 1α and several ribosomal protein genes. We find strong homology to this motif 373 nucleotides in front of the EF - 1α,F1 gene (position 258 in Fig.1). To our knowledge this is the first case that the yeast consensus sequence has been observed at the right distance in front of this group of genes in Drosophila. A sequence homology between F1 and F2 that might also be worth mentioning is located at the end of the intron that is preceeding the translational reading frame. The sequence of 12 nucleotides at position 1566 in Figure 1 and position 1797 in Figure 2 is completely conserved. In light of the highly divergent surrounding this consensus sequence is very unlikely to have evolved by accident.

EF - 1α,F1 transcription strength is comparable with that of the induced vitellogenin I gene (33). The usage of a limited number of codons in the F1 message, therefore, supports the hypothesis of a codon bias for such highly expressed genes. In Drosophila heat shock exerts a strong transcriptional and translational control. Translation of nonheatshock message is restrained by a slow down of initiation and elongation (34). The availability of both elongation factor genes will allow us to address questions regarding the possible involvement of one of the EF - 1α factors in this regulation phenomenon.

Acknowledgment

We are grateful to T. Kornberg for making cDNA libraries available to us. We would like to thank I. Sures for helpful criticism in the preparation of this manuscript. This work was supported by the Deutsche Forschungsgemeinschaft (SFB 229/C3). S. Richter is recipient of a G.F.G stipend of the Land Baden-Württemberg.

REFERENCES

1 Walldorf,U., Hovemann,B., and Bautz, E.K.F. (1985) Proc. Natl. Acad. Sci. USA 82, 5795-5799
2 Kaziro, Y. (1978) Biochim. Biophys. Acta 505, 95-127

3 Drews,J., Bednarik,K., and Grasmuk,H. (1974) Eur. J. Biochem. 41, 217-227
4 Kemper,W.M., Merrick,W.C., Redfield,B., Liu, C.-K.,and Weissbach,H. (1976) Arch. Biochem. Biophys. 174, 603-612
5 Nagata,S., Motoyoshi,K., and Iwasaki,K. (1978) J. Biochem. 83, 423-429
6 Slobin,L.I., and Möller,W. (1978) Eur. J. Biochem. 84, 69-77
7 Murakami,K., and Miyamoto,K. (1982) J. Neurochem. 38,1315-1322
8 Lauer,S.J., Burks,E., Irvin,J.D., and Ravel,J.M. (1984) J. Biol. Chem. 259, 1644-1648
9 Thiele,D., Cottrelle,P., Iborra,F., Buhler,J.-M., Sentenac,A., and Fromageot,P. (1985) J. Biol. Chem. 260, 3084-3089
10 Crechet,J.-B., Canceill,D., Bocchini,V., and Parmeggiani,A. (1986) Eur. J. Biochem. 161, 635-645
11 Nagata,S., Nagashima,K., Tsunetsugu-Yokota,Y., Fujimura,K., Miyazaki,M., and Kaziro,Y. (1984) EMBO J. 3, 1825-1830
12 Schirmaier,F. and Philippsen,P. (1984) EMBO J. 3, 3311-3315
13 Cottrelle,P., Thiele,D., Price,V.L., Memet,S., Micouin,J.-Y., Marck,C., Buhler,J.-M., Sentenac,A., and Fromageot,P. (1985) J. Biol. Chem. 260, 3090-3096
14 Linz,J.E., Katayama,C., and Sypherd,P.S. (1986) Mol. Cell. Biol. 6, 593-600
15 Lenstra,J.A., Van Vliet,A., Arnberg,A.C., Van Hemert,F.J. and Möller,W. (1986) Eur. J. Biochem. 155, 475-483
16 Brands,J.H.G.M., Maassen,J.A., Van Hemert,F.J., Amons R., and Möller,W. (1986) Eur. J. Biochem. 155, 167-171
17 Rao,T.R., Slobin,L.I. (1986) Nucl. Acids Res. 14, 2409
18 Kohl,A., (1986) Diploma thesis, University of Heidelberg
19 Roth,W.W., Bragg,P.W. Corrias,M.V., Reddy,N.S., Dholakia,J.N., and Wahba,A.J. (1987) Mol. Cell. Biol. 7, 3929-3936
20 Maxam, A.M., and Gilbert,W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564
21 Sanger, F., Nicklen,S., and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467
22 Geliebter,J., Zeff, R.A., Melvold,R.W., and Nathenson,S.G. (1986) Proc. Natl. Acad. Sci. USA 83, 3371-3375
23 Hovemann,B., Walldorf,U., and Ryseck,R.-P. (1986) Mol. Gen. Genet. 204, 334-340
24 Maniatis,T., Fritsch,E.F.,and Sambrook,J. (1982) Molecular Cloning (A Laboratory Manual), Cold Spring Harbor
25 Roberts,D.B. (1986) In: "Drosophila: a practical approach", D.B. Roberts, ed. (Oxford, IRL Press), 1-38
26 Keller, E.B., and Noon, W.A. (1985) Nucl. Acids Res. 13 4971-4981
27 Ashburner,M., Bodmer,M., and Lemeunier,F. (1984) Develop. Genet. 4, 295-312
28 O'Connell,P., and Rosbash,M. (1984) Nucl. Acids Res. 12, 5495-5513
29 Sharp,P.M., Tuohy, T.M.F., and Mosurski,K.R. (1986) Nucl. Acids Res. 14, 5125-5143
30 Linz,J.E., Lira,L.M., and Sypherd,P.S. (1986) J. Biol. Chem. 261, 15022-15029
31 Dynan,W.S. (1986) Trends Genet. 2, 196-197

32 Huet,J., Cottrelle,P., Cool,M., Vignais,M.-L., Thiele,D., Marck,C., Buhler,J.-M., Sentenac,A., and Fromageot,P. (1985) EMBO J. 4, 3539-3547
33 Walldorf,U. (1985) Ph.D. thesis, University of Heidelberg
34 Ballinger,D.G., and Pardue,M.L. (1983) Cell 33, 103-114
35 Viel,A., Dje',M.K., Mazabraud,A., Denis,H., and le Maire,M. (1987) FEBS Letters 223, 232-236
36 Mattaj,I.W., Coppard,N.J., Brown,R.S., Clark,B.F.C., and De Robertis,E.M. (1987) EMBO J. 6, 2409-2413
37 Leberman,R., and Egner,U. (1984) EMBO J. 3, 339-341
38 Dever,T.E., Glynias,M.J., and Merrick,W.C. (1987) Proc. Natl. Acad. Sci. USA 84, 1814-1818