CHARACTERIZATION OF GENES OF THE ELONGATION FACTOR 2 (EF-2) FAMILY OF CAENORHABDITIS ELEGANS

by

Esther Ngozi Ofulue

B.Sc. (honors), The University of Nigeria, 1982
M.Sc., The University of Ibadan, 1985

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

Department of Biochemistry

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

October 29, 1991

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of BIOCHEMISTRY

The University of British Columbia
Vancouver, Canada

Date Nov 5, 1991

ABSTRACT

A gene eft-1 encoding a protein synthesis elongation factor 2-like protein was isolated from chromosome III of C. elegans, and was mapped approximately 50 kb downstream from ubq-1. Five overlapping cosmids spanning 150 kb in this region were mapped by restriction endonuclease digestion, and hybridized to cDNA probes made from embryo polyA$^+$RNA. One positive fragment B3255 was analyzed by sequencing, and was also used as a specific probe to isolate a cDNA clone pEF1.35 encoding eft-1 mRNA. The entire eft-1 gene of 3.8 kb predicted a protein (EFT-1) of 849 amino acid residues which shared 38% overall identity with mammalian and Drosophila elongation factor 2 (EF-2) sequences. Sequence segments implicated in GTP-binding and GTPase activity in EF-2 were found in the N-terminal region, while segments characteristic of EF-2 and its prokaryotic counterpart EF-G were found at the corresponding C-terminal portion of EFT-1; the latter region shared 40 - 50% similarity with the hamster EF-2. However, the histidyl residue target for ADP-ribosylation and inactivation of EF-2 by diphtheria toxin, which is thought to be of functional importance in EF-2, was replaced by a tyrosyl residue in EFT-1.

By rapid amplification of C. elegans DNA sequences using primers specific for highly conserved regions of mammalian and Drosophila EF-2 using the polymerase chain reaction, cDNA and genomic clones containing a C. elegans gene (eft-2) were isolated and characterized. One cDNA clone, pCef6A, encoding the entire eft-2 mRNA predicted a polypeptide of 852 amino acid residues which shared greater than 80% identity with the hamster and Drosophila EF-2 sequences. The GTP-binding domains, ADP-modifiable histidyl residue, and high homology regions shared between EF-2 and EF-G were 80 - 100% conserved in the C. elegans protein. These results suggested that eft-2 and not eft-1 encoded the C. elegans homolog of EF-2. The conservation of functional domains of EF-2 in EFT-1 implied that the two genes were derived at least in part from a common ancestor. The

copy numbers of eft-1 and eft-2 were examined and their expression was monitored throughout nematode development. The results revealed that both genes are unique and each encodes a 3 kilobase mRNA species which does not appear to be under nutritional or developmental regulation.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS

aa	aminoacyl
ADP	adenosine 5'-diphosphate
ATP	adenosine 5'-triphosphate
bp	base pair(s)
BSA	bovine serum albumin
β-gal	β-galactosidase
cDNA	complementary DNA
C. elegans	Caenorhabditis elegans
CHO	Chinese hamster ovary
cpm	counts per minute
DEPC	diethylpyrocarbonate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dATP	deoxyadenosine 5'-triphosphate
dCTP	deoxycytidine 5'-triphosphate
dGTP	deoxyguanosine 5'-triphosphate
dTTP	deoxythymidine 5'-triphosphate
ddATP	dideoxyadenosine 5'-triphosphate
ddCTP	dideoxycytidine 5'-triphosphate
ddGTP	dideoxyguanosine 5'-triphosphate
ddTTP	dideoxythymidine 5'-triphosphate
DT	diphtheria toxin
DTT	dithiothreitol
E. coli	Escherichia coli
EDTA	ethylenediamine tetraacetic acid

EF	elongation factor
EMS	ethyl methane sulfonate
GTP	guanosine 5'-triphosphate
Gu-HCl	guanidinium hydrochloride
HEPES	N-2-hydroxyethylpiperazine-N-2-ethanesulfonic acid
IF	initiation factor
kb	kilobase pairs
kDa	kilodalton
Met	methionine
MOPS	3-(N-morpholino)propanesulfonic acid
mRNA	messenger RNA
MMLV	Moloney murine leukemia virus
NAD	nicotinamide adenine dinucleotide
PA	Pseudomonas aeruginosa exotoxin A
PCR	polymerase chain reaction
PMSF	phenylmethyl sulfonylfluoride
polyA$^+$	polyadenylated
RACE	rapid amplification of cDNA end
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal RNA
S. cerevisiae	Saccharomyces cerevisiae
SDS	sodium dodecyl sulfate
SSC	150 mM NaCl, 15 mM $Na_3Citrate.2H_2O$, pH 7.0
SSPE	180 mM NaCl, 1 mM EDTA, 10 mM NaH_2PO_4, pH 7.4

Taq	Thermus aquaticus
TBE	90 mM Tris-borate pH 8.3, 1 mM EDTA
TE	10 mM Tris-HCl pH 8.0, 1 mM EDTA
Tris	tris (hydroxymethyl) aminomethane
tRNA	transfer RNA
ubq-1	polyubiquitin gene

ACKNOWLEDGEMENTS

I wish to thank my supervisor, Dr. Peter Candido, for providing the materials and a good working environment for my studies, and for his readiness to guide and assist at all times. Special thanks to my colleagues Mike Hockertz, Eve Stringham, Dave Leggett, Rob Boissy, Mei Zhen and Don Jones for all the jokes and lively discussions that made it worthwhile going back to the lab to work on those discouraging experiments. Thanks also to Don Jones for advice and for the "ready-made" solutions.

To my other half, Anwuli, for all those lonely evenings and weekends, and for believing in me I say thanks and God bless. I am greatly indebted to my mum whose love, encouragement, and prayers helped see me through it all. I am grateful to Rev. Reuben and Mercy Unaegbu and all my friends and relatives for their love, moral, and prayerful support. I should acknowledge the Canadian Commonwealth Scholarship Plan and the Association of Universities and Colleges of Canada who provided the funds that made it possible for me to embark on this programme.

Finally, I give God all the glory for His unceasing care and faithfulness all through.

I. INTRODUCTION

Eukaryotic elongation factor 2 (EF-2) is a single polypeptide chain (Mr 93000 - 110000) which catalyzes the translocation of peptidyl-tRNA from the aminoacyl site to the peptidyl site of the ribosome, thus preparing the protein-synthesizing complex for the next step of polypeptide chain elongation (Skogerson and Moldave, 1968; Tanaka et al., 1977). This translocation step is coupled to the ribosome (Taira et al., 1972) and to GTP-hydrolysis (Chuang and Weissbach, 1972). The primary structure of EF-2 from many eukaryotic and archaebacterial sources including hamster, Drosophila, and Methanococcus vannielii has been determined by cloning and sequencing full-length cDNA and genomic clones (Kohno et al., 1986; Nakanishi et al., 1988; Grinblat et al., 1989; Lechner et al., 1988). The sites for EF-2 interaction with guanosine nucleotides and the ribosome have been located at the N-terminal and C-terminal portions of the protein respectively (Nilsson and Nygard, 1985; Kohno et al., 1986). EF-2 is post-translationally modified by phosphorylation (Ryazanov 1987) and ADP-ribosylation (Honjo et al., 1968), leading to inhibition of protein synthesis. EF-2 has been purified from rat liver (Galasinski and Moldave, 1969; Takamatsu et al., 1986) and a variety of other sources including pig and beef liver, wheat germ, yeast (Mizumoto et al., 1974; Brown and Bodley, 1979) and human placenta (Giovane et al., 1987). The N-terminal and C-terminal amino acids of the rat protein have been identified as valine and leucine, respectively (Comstock et al., 1977; Takamatsu et al., 1986). Furthermore, the sequence of nineteen N-terminal amino acids and of a fifteen amino acid peptide produced by tryptic digestion of ADP-ribosylated rat and bovine EF-2 have been determined (Takamatsu et al., 1986; Robinson et al., 1974; Brown and Bodley, 1979). The amino acid which is specifically ADP-ribosylated by diphtheria toxin has been identified as diphthamide, which is a post-translationally modified histidyl residue (Van Ness et al., 1980; Kohno et al., 1986).

A. BACKGROUND PERSPECTIVE

Most of the early research on protein biosynthesis focused on the cellular components required for the incorporation of amino acids into proteins in cell-free extracts in vitro. Keller and Zamecnik (1956) had shown that microsomes, a nondialyzable heat-labile fraction from the cytosol, an ATP-generating system, and GTP are required for in vitro translation. The discovery of the carboxyl activation of amino acids, by reaction with ATP, and of tRNA which accepted the aminoacyl (aa-) moiety from the enzyme (aa-tRNA synthetase) to form the aa-tRNA intermediate in the incorporation of amino acids into protein, soon followed (Davie et al., 1956; Hoagland et al., 1958). Purification of components such as aa-tRNA synthetases (Wong et al., 1960), aa-tRNAs (Grossi et al., 1959), and ribosomes (Fessenden and Moldave, 1961), and the development of strategies for assaying various intermediates allowed the examination of the role of soluble protein factors in eukaryotic protein biosynthesis.

1. Initiation factors

A ribosomal wash factor capable of forming a ternary complex with Met-tRNA$_1$ and GTP in the absence of ribosomes (Levin et al., 1972) was fractionated into several protein factors (initiation factors) including eIF-1, eIF-2, eIF-3, eIF-4, and eIF-5, which are required for the synthesis of globin or methionylpuromycin in cell-free extracts dependent on globin mRNA (Schreier et al., 1977). As in prokaryotes, the factors were shown to recycle between ribosomes and supernatant during protein synthesis (Freinstein and Blobel, 1975). eIF-3 contains at least nine major polypeptides ranging in molecular weight from 28,000 to 140,000 (Benne and Hershey, 1976), all of which bind to 40S ribosomal subunits. eIF-3 is responsible for releasing the 40S subunit from 80S ribosome, forming a stable intermediate complex with the 40S subunit. eIF-2, which is functionally equivalent to the prokaryotic IF-

2, forms a ternary complex containing equimolar amounts of factor, Met-tRNA$_1$ and GTP which interacts with the 40S•eIF-3 complex to form a stable 40S pre-initiation complex. Association of the pre-initiation complex with mRNA and the large (60S) ribosomal subunit is promoted by eIF-1, eIF-3, eIF-4, and eIF-5 (Hershey, 1980), and results in the formation of the 80S initiation complex which is stabilized by the bound mRNA. The process is accompanied by hydrolysis of GTP, causing the release of the factors. Factor eIF-2 has been purified to homogeneity from rabbit reticulocyte ribosomal washes (Safer et al., 1975). It has a molecular weight of 120,000 and is composed of three nonidentical subunits designated α, β and γ. The γ-subunit interacts with and promotes binding of Met-tRNA$_1$ to the small ribosomal subunit (Jagus et al., 1982). The α-subunit (Lloyd et al., 1980) binds GTP (Barrieux and Rosenfeld), which facilitates interaction of the γ-subunit with Met-tRNA$_1$. Binding of GDP inhibits interaction of eIF-2 and Met-tRNA$_1$. The function of the β-subunit is unclear, but it appears to be involved in the recycling of eIF-2 and may also react with guanine nucleotides (Jagus et al., 1982). The mechanisms of formation of eukaryotic and prokaryotic initiation complexes are virtually identical except for the involvement of more factors in the former. In addition, bacterial initiation factor 1 (IF-1) enhances the IF-2-mediated binding on fMet-tRNA$_1$ to 30S subunits whereas eIF-1 (together with eIF-3 and eIF-4) promotes mRNA binding to the 40S initiation complex.

2. Elongation factors

Studies with purified aa-tRNA synthetases, aa-tRNAs, and ribosomes provided evidence that a nondialyzable heat-labile cytosolic protein referred to as aminoacyltransferase I (Grossi et al., 1960), a dialyzable component which could only be replaced by GTP, as well as an activity (aminoacyltransferase II) present in microsomes are also essential for incorporation of amino acids from aa-tRNA into ribosomes (Fessenden and

Moldave, 1961). Fractionation of aminoacyltransferases I and II, now called elongation factors 1 and 2 (EF-1 and EF-2) , from the cytosol was achieved by ammonium sulfate precipitation (Fessenden and Moldave, 1963), and by gel filtration (Gasior and Moldave, 1965). Both factors have also been purified from rat liver and shown to correspond to the prokaryotic EF-Tu and EF-G (Schneir and Moldave, 1968; Galasinski and Moldave, 1969).

a. Elongation factor 1

The cytosolic EF-1 exists in the form of aggregates containing different polypeptide chains, EF-1α, EF-1β, and EF-1γ of molecular weight 26,000 - 53,000 (Slobin and Moller, 1976; Kaziro, 1978; van Damme et al., 1990). When EF-1 preparations from yeast (Richter and Lipmann, 1970) and wheat germ (Bollini et al., 1974) were incubated with GTP and aa-tRNA, as with the prokaryotic factor EF-Tu, a ternary complex was formed in which both EF-1α (or EFTu) and aa-tRNA were more stable than the corresponding free forms. In experiments involving multistep incubations in which some of the later additions included ribosomes, Ibuki and Moldave (1968) showed that the synthesis of the ternary complex [EF-1α•GTP•aa-tRNA] reflected the formation of an obligatory intermediate between aa-tRNA and ribosome-bound aa-tRNA. EF-1α catalyzed this intermediate step and the subsequent binding of the aa-tRNA to open A sites on ribosomes. This process is accompanied by GTP hydrolysis and release of GDP-bound EF-1α. In contrast to EF-Tu, EF-1α binds GTP somewhat more tightly than GDP. Analysis of the intermediate reactions of the elongation process revealed that after each round of elongation and GTP hydrolysis, EF-1α recycles via a process in which EF-1β catalyzes the exchange of bound GDP for GTP on EF-1α (Iwasaki et al., 1976; Janssen et al., 1988). EF-1β is therefore analogous in function to the prokaryotic EF-Ts. The function of the γ subunit is not clear, but it is believed to be the kinase moiety of the factor which reversibly phosphorylates EF-1β (Ejiri

and Honda, 1985). Peptide bond formation between the endogenous peptidyl moiety on the P site and the incoming amino acid at the A site is catalyzed by a ribosomal activity, peptidyltransferase, which does not require elongation factor or GTP. The tRNA-bound amino acid does not participate in peptide bond formation unless a peptidyl-tRNA is present at the ribosomal P site.

b. Elongation factor 2

Evidence for the role of EF-2 in translocation was obtained by Skogerson (1968) who isolated a ribosome•EF-2 complex containing GTP by ultracentrifugation of extracts containing these components and glutathione. EF-2 stimulates polypeptide chain elongation, translocation of peptidyl-tRNA (and its corresponding codon) from the puromycin-insensitive (A-) site to the puromycin-sensitive (P-) site on the ribosome, and ribosome-dependent GTP hydrolysis. EF-2 forms stoichiometric binary complexes with guanosine nucleotides, but unlike its prokaryotic counterpart EF-G, the complex with EF-2 is stable, GDP binding more tightly than GTP (Henriksen et al., 1975; Mizumoto et al., 1974). The association of the factor with ribosomes is strictly dependent on the nucleotide. The first complete amino acid sequence of a mammalian EF-2, that of hamster, was determined by cDNA cloning (Kohno et al., 1986). EF-2 (which has also been purified from a variety of eukaryotic sources; Brown and Bodley, 1979) is a large, acidic protein comprised of a single polypeptide with a molecular weight of about 100,000.

c. Elongation factor 3

A third elongation factor EF-3 is uniquely required for in vitro protein synthesis by yeast ribosomes, in addition to EF-1 and EF-2 (Skogerson and Engelhardt, 1977). This factor stimulates EF-1-dependent binding of aa-tRNA to yeast 40S ribosomal subunit ;

hydrolysis of purine nucleotide is necessary for the stimulation (Kamath and Chakraburtty, 1989). EF-3 which exhibits ribosome-dependent ATP and GTP hydrolysis, has been purified from several fungal species and found to consist of a single polypeptide chain with a molecular weight of 125,000 (Dasmahapatra and Chakraburtty, 1981; Uritani and Miyazaki, 1988).

3. Release factors

Mammalian cells appear to contain only a single release factor (RF) which recognizes all three termination codons (Tate et al., 1973). RF from rabbit reticulocytes is a dimer composed of identical subunits of molecular weight 56,000 (Caskey, 1977). The factor catalyzes the hydrolysis of GTP in a ribosome-dependent reaction stimulated by tetranucleotides containing the termination codon sequences.

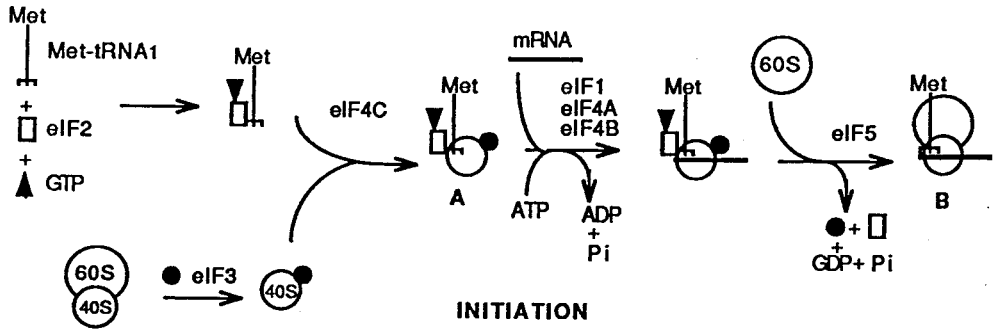
B. PATHWAY OF EUKARYOTIC PROTEIN BIOSYNTHESIS

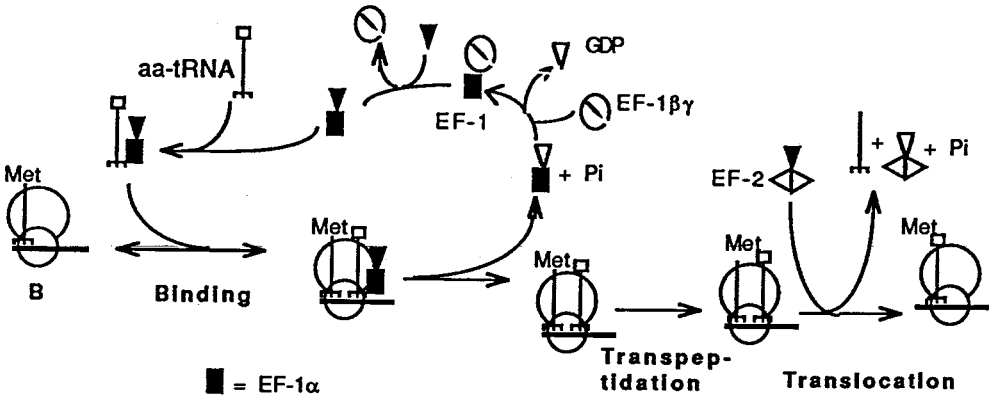
Eukaryotic protein biosynthesis is divided into three stages: initiation, elongation, and termination, each involving a distinct set of soluble protein factors called, respectively, initiation, elongation, and release factors. A schematic representation of the pathway of protein biosynthesis in eukaryotes as adapted from Hershey (1980) is shown in Figure 1.

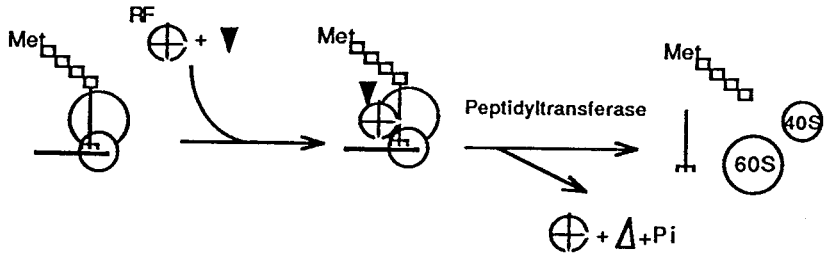
1. Initiation

The pathway of assembly of the 80S initiation complex has been studied most extensively with components purified from rabbit reticulocytes. Two major approaches were used: (i) intermediates were isolated and identified from crude cell lysates, often following the addition of an antibiotic that inhibits initiation, (ii) intermediate complexes were

Fig. 1. Schematic representation of the pathway for eukaryotic protein biosynthesis. The soluble protein factors involved at each stage are given as eIF (eukaryotic initiation factor) or EF (elongation factor). The reactions in the pathway are explained in the text (see section 1,B). Because of the enormous complexity of the initiation factors, their precise functional role in the pathway is not yet clear, and possibly more factors are involved (which are yet to be characterized) than are represented in this scheme (Hershey, 1980). The 40S pre-initiation complex and the 80S initiation complex are represented as A and B, respectively. This scheme is an adaptation from Hershey (1980). Some data indicate that GTP hydrolysis occurs concomitantly with or immediately following proper binding of the ternary complex to the 40S ribosomal subunits (Odom, et al., 1978; Kramer and Hardesty, 1980).


Met
Met-tRNA1
eIF2
GTP
Met
eIF4C
Met
A
mRNA
eIF1
eIF4A
eIF4B
ATP
ADP
+
Pi
Met
60S
eIF5
Met
B
+
GDP+ Pi
60S
40S
eIF3
40S
INITIATION


aa-tRNA
GDP
EF-1
EF-1βγ
+ Pi
Met
B
Binding
Met
Met
EF-2
+
+ Pi
Met
Met
= EF-1α
Transpep-
tidation
Translocation
ELONGATION


RF
+
Met
Met
Peptidyltransferase
Met
40S
60S
+ Δ+Pi
TERMINATION

assembled in vitro from purified components. Prior to the initiation stage, a free molecule of methionine is activated in the presence of ATP. The activated methionine is then attached to $tRNA_1^{Met}$ by the specific methionyl-tRNA synthetase, forming Met-$tRNA_1$ (which is modified by the addition of a formyl group to form N-formylMet-tRNA, in bacteria). The initiation process begins with the dissociation of the 80S ribosome into subunits by a factor activity that is poorly characterized in the mammalian system, but which is proposed to be the multicomponent factor eIF-3 (Thompson et al., 1977). In a parallel set of reactions, the cytosolic initiation factor eIF-2 forms a ternary complex with the Met-$tRNA_1$ and GTP which combines with the 40S ribosomal subunit (in the presence of eIF-3 and eIF-4C) to form the 40S (30S in prokaryotes) pre-initiation complex (A, see Figure 1). Following the binding of the mRNA which is promoted by eIF-4A and eIF-4B, and involves the hydrolysis of ATP, the large (60S, 50S in prokaryotes) ribosomal subunit associates with the 40S initiation complex in the presence of eIF-5 and an 80S (60S in prokaryotes) initiation complex (B, see Figure 1) is formed with concomitant hydrolysis of the GTP (Hershey, 1980) and release of the initiation factors. The Met-$tRNA_1$ bearing the first amino acid is now bound to the P-site on the ribosome.

2. Elongation

The elongation process is conveniently divided into three steps: the binding of aminoacyl-tRNA, peptide bond formation (transpeptidation), and translocation (see Figure 1). This process requires two factors, elongation factors 1 and 2 (EF-1 and EF-2), which are functionally equivalent to the prokaryotic factors EF-Tu/EF-Ts and EF-G, respectively. EF-1 is responsible for binding the aminoacyl-tRNA of the ternary complex EF-1•GTP•aminoacyl-tRNA to the empty A-site on the ribosome, a process accompanied by hydrolysis of the GTP. Once a peptide bond is formed by transfer of the amino acid or nascent

peptide from its tRNA in the P-site to the α-amino group of the aminoacyl-tRNA in the A-site, EF-2 promotes the translocation of the peptidyl-tRNA (and the corresponding codon) from the A-site to the P-site (Moldave 1985 and refs therein), i.e. the ribosome moves one codon down the mRNA chain thereby allowing a new aminoacyl-tRNA to enter the empty A-site. Association of EF-2 with the ribosome during translocation requires GTP and activates the GTPase centre of EF-2. As a result of GTP hydrolysis, the ribosome-EF-2 complex is destabilized and the factor leaves the ribosome. A detailed mechanism for EF2-mediated translocation is discussed elsewhere (Section 1, D) and the scheme for EF-2 and ribosome cycles is shown in Figure 2.

In vitro studies by Moazed and Noller (1989) on tRNA-ribosome complexes in the pre- and post-peptidyl transfer stages of the translational cycle revealed a third site, E-site, (in addition to the A- and P-sites) for tRNA binding to E. coli ribosome. In their hybrid site model for the movement of tRNA during translocation, after peptide bond formation, the deacylated tRNA and the peptidyl-tRNA shift from P to E and A to P sites, respectively. Both tRNAs first move with respect to the large subunit but maintain their locations on the small subunit. Finally, the EF-G-catalyzed step moves the anticodon ends of both tRNAs, together with their associated mRNA, relative to the small subunit. It is possible that this mechanism also exits in the eukaryotic system.

3. Termination

When the ribosome arrives at a termination codon (UAG, UAA, or UGA) which is recognized by a release factor RF, hydrolysis of the peptidyl-tRNA on the ribosome by peptidyltransferase releases the completed polypeptide and the last tRNA, and the two ribosomal subunits separate. In eukaryotic cells, RF binding to the ribosome requires GTP which is hydrolyzed when RF is released from the ribosome.

In addition to the cytosolic machinery described above, an independent translational machinery is also present in the mitochondria of eukaryotic cells. The mitochondrial translational apparatus is similar to that of prokaryotes and also uses a 70S class of ribosome. The protein components (including all ribosomal proteins, aa-tRNA synthetases, and soluble translational factors) are encoded by nuclear genes, synthesized in the cytoplasm, and transported into the mitochondrion. The similarities shared by bacterial and eukaryotic systems in protein synthesis are striking and indicate that many of the components and the overall mechanism have been highly conserved during evolution. However, the eukaryotic components differ from bacterial components primarily by being larger and more numerous, as observed especially for ribosomes and the soluble factors.

C. REGULATION OF EUKARYOTIC PROTEIN BIOSYNTHESIS

A change in the rate of protein synthesis can have effects on a wide range of processes including those on gene expression, cell physiology and cell proliferation (Moldave, 1985). Translational regulation in eukaryotes is more complex and less well characterized than in prokaryotes. Polypeptide initiation was thought to be the rate-limiting step, so most early research on translational regulation focused on the control of the initiation process.

1. Control of polypeptide initiation

A summary diagram including the mechanism of eukaryotic peptide initiation is shown in Figure 1. Control of polypeptide initiation is believed to occur at two levels: Met-tRNA$_1$ binding to 40S subunits, and mRNA binding to the 40S initiation complex. In addition, GDP has been reported as a potent inhibitor of the Met-tRNA$_1$•eIF-2•GTP complex formation, and may play a role in regulating initiation through the "energy charge" of the cell (Walton

and Gill, 1976). Kramer et al. (1977) and Grankowski et al. (1980) showed that reversible phosphorylation of eIF-2α by eIF-2 kinase systems and a counteracting phosphatase, may provide a physiologically important mechanism by which Met-tRNA$_1$ binding to 40S subunits can be reversibly inhibited thereby regulating the rate of protein synthesis. Since this modification does not inhibit the ability of eIF-2 to form ternary complex with Met-tRNA$_1$and GTP, it is believed not to be associated with activation/inactivation, but with perturbation of the recycling mechanism which promotes guanine nucleotide exchange on the factor (for review see Safer and Jagus, 1981). Also phosphorylation of protein S6 of the 40S subunit is thought to constitute a separate control system, distinct from the eIF-2 kinase system, that may also regulate peptide initiation at the level of Met-tRNA$_1$ binding to the 40S subunits (for review see Kramer and Hardesty, 1980).

The results of several studies indicate that a cap structure (7-methylguanosine) at the 5'-terminus of most mRNAs may play an important role in binding mRNA to 40S subunits during peptide initiation, by specifically interacting with one or more initiation factors (Shafritz et al., 1976 ; Kaempfer et al., 1978). Phosphorylation of eIF-3 and eIF-4 in intact rabbit reticulocytes (Benne et al., 1978), probably by a cyclic-AMP-independent protein kinase (Issinger et al., 1976), and glycosylation of eIF-3 have also been reported (Ilan and Ilan, 1976). The role of these modifications in the control of peptide initiation is still unclear.

2. Control of polypeptide elongation

Evidence that the elongation cycle also is subjected to regulation came from results that the cAMP-dependent activation of protein synthesis was due to dephosphorylation of EF-2 (Sitikov et al., 1988). Redpath and Proud (1989) also showed that phosphorylation of EF-2 inhibits translation of natural mRNA in a cell-free system where initiation, elongation, and termination take place, suggesting that EF-2 phosphorylation may represent a novel

mechanism of translational control. Furthermore, studies on transformed human amnion cells undergoing mitosis also showed that increased phosphorylation of EF-2 may partly explain the decline in the rate of protein synthesis observed during cell division (Celis et al., 1990). EF-2 kinase, which phosphorylates threonyl residues in EF-2 and effectively inhibits translational elongation in vitro (Ryazanov et al., 1988a,b), has been purified from mammalian cells (Palfrey 1983; Nairn et al., 1985) and its in vivo activity was shown to depend upon growth factors and other agents affecting the level of Ca^+ and cAMP (Nairn et al., 1987a). The kinase shows strong substrate specificity for EF-2 (Nairn et al., 1985), whereas other known protein kinases are unable to phosphorylate EF-2 to a significant extent. Phosphorylation of EF-2 interfered with ribosome•EF-2 complex formation by reducing the affinity of EF-2 for the ribosome (Carlberg et al., 1990). Thus, the inability of phosphorylated EF-2 to promote translocation of peptidyl-tRNA from the ribosomal A-site to the P-site may result from the reduced affinity for the ribosome.

EF-1α is also modified by methylation on lysyl residues and by phosphorylation (Fonzi et al., 1985). The degree of methylation alters the activity of the factor(Merrick et al., 1990). A direct role for a kinase activity in altering the interaction of EF-1α with the ribosome has been suggested by Davydova et al. (1984), and supplies another example of a regulatory role for elongation in translation. EF-1β is also modified by phosphorylation at a specific seryl residue. Modification of either the guanine nucleotide binding subunit, EF-1α , or the actual guanine nucleotide exchange subunit, EF-1β (by EF-1γ), could affect the rate of formation of ternary complex EF-1α•GTP•aminoacyl-tRNA (Janssen et al., 1988). Cavallius et al. (1986) have shown that the activity and amounts of EF-1α undergo cell cycle- and age-related changes in normal human fibroblasts and in SV40-transformed cells derived from them. The two genes encoding the Drosophila homologs of EF-1α are regulated, at least at the level of transcription, in a sex-specific and developmental stage-

specific manner (Hovemann et al., 1988; Walldorf et al., 1985).

3. Effect of insulin on protein synthesis

Insulin has been reported to increase protein synthesis exclusively by enhancing initiation (Monier and LeMarchand-Brustel, 1982) through a stimulation of the phosphorylation of ribosomal protein S6 (Smith et al., 1980; Thomas et al., 1982), and the dephosphorylation of eIF-2α (Towle et al., 1984). Phosphorylation of S6 may enhance recruitment of ribosomes containing the modified protein into translationally active polysomes (Thomas et al., 1982), whereas dephosphorylation of phospho-eIF-2α is necessary to allow it to recycle and form another initiation complex (Panniers and Henshaw, 1983). Insulin has also been shown to induce rapidly the synthesis of EF-2 predominantly or exclusively at the level of mRNA translation (Levenson et al., 1989).

4. Other factors affecting protein synthesis

A number of stimuli can increase the rate of protein synthesis apparently through increases in the elongation rate of nascent peptide chains. These stimuli include heat shock (Theodorakis et al., 1988; Ballinger and Pardue, 1983) or treatment with estrogen and progesterone (Palmiter, 1972; Gehrke et al., 1981), and serum (Nielsen and McConkey, 1980). For example, after withdrawal of serum from the medium of actively dividing vertebrate cells, the rate of elongation declines rapidly (Nielsen and McConkey, 1980) and alterations in elongation rate have been found after heat shock of Drosophila cells (Ballinger and Pardue, 1983). This decline in the rate of peptide chain elongation is associated with a sharp reduction of intrinsic activity of EF-1 (Fischer et al., 1980; Hassell and Engelhardt, 1976). Regulation of protein synthesis at the level of elongation has also been demonstrated during phorbol ester treatment (Gschwendt et al., 1988), growth factor stimulation (Thomas

and Thomas, 1986), transformation (Nielsen and McConkey, 1980), and ageing (Webster, 1985; Cavallius et al., 1986). However, the changes in the amounts and activities of various elongation factors which may be involved in this regulation are yet unknown. Recently, Riis et al. (1990) observed an irreversible decrease in the amount of EF-2 that could be ADP-ribosylated in aged and SV40-transformed human cell cultures, which could account for the slowing-down of protein synthesis during cell cycle arrest and during cellular ageing in culture.

D. MECHANISM OF EF-2 MEDIATED TRANSLOCATION

EF-2 promotes the translocation of peptidyl-tRNA from the A site to the P site on the ribosome (Moldave, 1985). A schematic illustration (adapted from Carlberg et al., 1990) of the EF-2 and ribosome cycles during the translocation process is given in Figure 2. Association of the factor with the pre-translocation ribosome requires GTP (Nygard and Nilsson, 1984). After binding, the pre-translocation ribosome is converted to a post-translocation form (Moldave, 1985). This transition activates the GTPase centre of EF-2 and the factor-bound GTP is hydrolyzed to GDP and inorganic phosphate (Nygard and Nilsson, 1989), leading to a reduced affinity of the factor for the ribosome (Nygard and Nilsson, 1985). As a result, the factor leaves the ribosome, thereby allowing a new aa-tRNA to enter the empty A site (Moldave, 1985). The rate of GTP hydrolysis is dependent on the rates of the EF-2 and ribosome cycles. The two cycles have multiple reactions in common, i.e. association of the EF-2•GTP complex with the ribosome, conversion of the complex to a GTPase active form, hydrolysis of GTP and release of the EF-2•GDP complex from the ribosome (Nygard and Nilsson, 1989). The role of GTP can be described as that of an allosteric effector which alters protein tertiary structure (Kaziro, 1978) to expose the ribosome-binding site

E. POST-TRANSLATIONAL MODIFICATION OF EF-2

EF-2 is known to undergo two types of post-translational modification which result in inhibition of translational elongation.

1. Phosphorylation of EF-2

EF-2 is phosphorylated by a calcium $^{2+}$/calmodulin-dependent protein kinase III

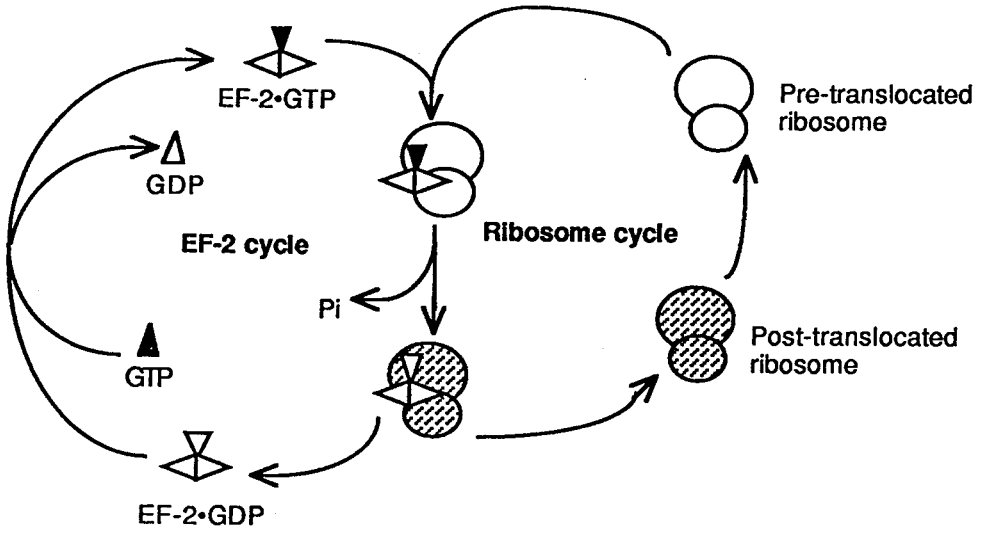


Fig.2. Schematic illustration of the EF-2 and ribosome cycles during the translocation process.

This model for the association of EF-2 with the ribosome and GTP during the translocation process is an adaptation from Carlberg et al. (1990). Protein synthesis could be inhibited due to a malfunction of the modified factor in any of the partial reactions of the EF-2 cycle, i.e. ribosomal binding of the factor, GTP hydrolysis, factor release from the ribosome and guanine nucleotide exchange.

(Ryazanov 1987) and has been reported as the major substrate for the kinase in vivo and in vitro (Nairn et al., 1987b; Kigoshi et al., 1989). After phosphorylation, EF-2 is completely inactive as a translocase due to reduced affinity for the ribosome; translation is inhibited, but EF-2 is still able to hydrolyze GTP (Carlberg et al., 1990). Dephosphorylation of EF-2 by phosphatase restores its activity (Ryazanov et al., 1988a).

2. ADP-ribosylation of EF-2

a. Effect of ADP-ribosylation

Diphtheria toxin causes the post-translational modification of EF-2 by ADP-ribosylation, resulting in the inactivation of its translocase activity. ADP- ribosylation was shown by Nygard and Nilsson (1985) to decrease the affinity of EF-2 for ribosomes in vitro leading to an inhibition of protein synthesis (Collier and Pappenheimer, 1964; Nilsson and Nygard, 1985). This modification also inhibits the GTPase (Raeburn et al., 1968) but not the GTP-binding function of EF-2 (Pappenheimer 1977; Sperti et al., 1971). Conversely, binding of GTP was shown to inhibit ADP-ribosylation of EF-2 (Sperti et al., 1971; Montanaro et al., 1971). The protective effect of GTP is related to a conformational change in EF-2 upon nucleotide binding (Nilsson and Nygard, 1985).

b. The diphthamide target for ADP-ribosylation

The EF-2 target for this modification is a unique amino acid, 2-[3-carboxylamido-3-(trimethylammonio) propyl]histidine, designated diphthamide, which itself is generated via a series of elaborate post-translational modifications of a histidyl residue (Van Ness et al., 1980; Dunlop and Bodley, 1983). Diphthamide has not been found in any other eukaryotic protein examined (Collier, 1975; Pappenheimer, 1977); on the other hand, EF-2 from a wide

variety of eukaryotes, ranging from mammals to yeast (Van Ness et al.,1978), and archaebacteria (Kessel and Klink 1980; Pappenheimer et al.,1983) contains a single residue of this unique amino acid and is irreversibly modified by diphtheria toxin in vitro. The physiological role of diphthamide is unknown, but the diphthamide forming enzymes appear to recognize a specific sequence and / or secondary structure in EF-2.

c. EF-2 ADP-ribosylating enzymes

ADP-ribosylation of EF-2 at the diphthamide residue is catalyzed by fragment A of diphtheria toxin (DT, Honjo et al., 1968), *P. aeruginosa* exotoxin A (PA, Iglewski and Kabat 1975) or by intracellular ADP-ribosyltransferase (Fendrick and Iglewski 1989; Marzouki et al., 1989). These enzymes transfer the ADP-ribose moiety of NAD^+ to the N-1 nitrogen of the imidazole ring of the diphthamide residue in EF-2. Diphthamide is essential for ADP-ribosylation of EF-2 by these toxins. DT and PA, which inhibit the translocase function of eukaryotic and archaebacterial EF-2's through this covalent modification, do not appear to modify any other eukaryotic or prokaryotic protein (Collier, 1975; Pappenheimer, 1977). The presence of a single molecule of fragment A of DT in the cytosol is sufficient to kill a cell (Yamaizumi et al., 1978). DT and PA have the same activity as NAD:EF-2-ADP-ribose transferase, but the sensitivities of different animal species to the two toxins differ greatly. Humans, monkeys, and hamsters are sensitive to DT, whereas mice and rats are not, but are very sensitive to PA. EF-2 prepared from all of these species can be ADP-ribosylated by PA or fragment A of DT in the presence of NAD^+ in a cell-free system (Kohno and Uchida, 1987). Several mono-ADP-ribosylating toxins, including DT and PA, carry a histidyl residue within the amino acid sequence that is conserved in spacing and location with respect to other critical residues. Histidine-426 of PA exotoxin A has been shown to be essential for the toxin's ADP-ribosyltransferase activity (Wozniak et al., 1988). This

residue is not associated with the proposed NAD^+ binding site.

F. STRUCTURE-FUNCTION RELATIONSHIPS

The approaches used to examine the relationships between structure and function of EF-2 include: (i) purification of EF-2 and analysis of its properties (Comstock et al., 1977; Mizumoto et al., 1974; Merrick et al., 1975; Robinson et al., 1974), (ii) comparison of deduced primary structures of cloned cDNAs, (iii) isolation and characterization of DT- and PA-resistant cells containing EF-2 that could not be modified by ADP-ribosylation (Moehring and Moehring, 1977; Gupta and Siminovitch, 1978). Comparisons of the amino acid sequence of EF-2 from several species reveal a high degree of conservation at the amino acid and nucleotide levels, suggesting important functions that need to be preserved. The primary structure of EF-2 from various species also reveals the structural and functional importance of certain evolutionarily conserved regions. EF-2s from phylogenetically distant organisms such as yeast, rat and wheat germ have common structural features, notably in the nonapeptide target for ADP-ribosylation by DT (Brown and Bodley, 1979) which accounts for the recognition of the proteins by the toxin. Studies have also confirmed the association of EF-2 with the ribosome and GTP during the elongation process (Nilsson and Nygard, 1985). The factor has a third functional domain, the catalytic centre responsible for GTP hydrolysis (Nilsson and Nygard, 1989). By affinity labeling with GTP analogues and/or radioactive NAD^+ in the presence of diphtheria toxin, followed by limited proteolysis of EF-2, Nilsson and Nygard (1985) showed that ADP-ribosylation decreases the affinity of EF-2 for ribosomes, and located the GTP-cleaving centre and the site of ADP-ribosylation in 48 kDa and 34 kDa tryptic fragments, respectively. The binding site for GTP is separated from the ADP-ribosylation site by a polypeptide sequence of 40 - 60 kDa. The three domains are interdependent since (a) GTP binding is strictly required for the association of the factor with

the ribosome, and (b) this interaction induces the GTP hydrolysis (Nilsson and Nygard 1984; Chuang and Weissbach 1972; Taira et al., 1972).

1. Conservation of EF-2 structure

The function of EF-2 includes the binding and hydrolysis of GTP, binding of peptidyl-tRNA, and recognition of and interaction with the 80S ribosome. The multifunctional nature of EF-2 presumably leaves little room for evolutionary divergence of the protein structure. The eukaryotic cytoplasmic EF-2s isolated to date (Comstock et al., 1977; Mizumoto et al., 1974; Merrick et al., 1975; Robinson et al., 1974) show a high degree of homology which is reflected in their physical properties such as pI (6.6 - 6.8, Takamatsu et al., 1986) and molecular weight. Their close relationship also becomes evident by comparing the deduced primary structure of various recently cloned and sequenced cDNAs. For example, hamster and Drosophila EF-2 share greater than 80% overall amino acid sequence identity (Grinblat et al., 1989), and the nucleotide sequences of hamster and rat share 89.7% homology with only two amino acid replacements (Oleinikov et al., 1989). The overall protein sequences for EF-2 of human and hamster, and human and rat differ in only 8 positions; only one amino acid difference is found in the GTP-binding region, and none in the 15 amino acid residue peptide which contains the histidyl residue that could be ADP-ribosylated. The sequence similarity of the DNA coding regions for hamster and human EF-2 is 87%, and for rat and human EF-2 is 88% (Rapp et al., 1989). Hamster and D. discoideum EF-2 have 61.3% identity and 87.4% similarity overall. The corresponding protein in archaebacteria (Methanococcus and Halobacteruim) displays 42% - 45% overall sequence similarity to eukaryotic EF-2 and 35% sequence similarity to the eubacterial EF-G ; sequence similarity between the two archaebacterial proteins is 62% (Lechner et al., 1988; Itoh, 1989). This striking degree of primary structural conservation of EF-2 in organisms which have been

evolving independently for ~ 1 billion years (Dayhoff, 1978), suggests an ancient origin and a vital function for the protein, consistent with its role in the translational apparatus. The low degree of similarity between the two archaebacterial EF-2s is thought to be possibly due to the high salt environment present in Halobacteria; however, the N-terminal region involved in GTP binding is highly conserved (68 - 82%) compared with regions in the C-terminal half (42 - 56%) associated with the ribosome (Itoh, 1989).

2. Homology with GTP-binding proteins

EF-2 was long known to be a guanine nucleotide-binding protein (Henriksen et al., 1975). It is suggested that a GTP-induced conformational change is a prerequisite for the ribosome-factor interaction and thereby for the hydrolysis of GTP. Thus the translocation of peptidyl-tRNA is coupled to a ribosome and to EF-2-dependent GTP-hydrolysis (Chuang and Weissbach, 1972; Taira et al., 1972). The polypeptide domains involved in this binding have only recently been identified. By in vitro and in vivo studies, bacterial elongation factors (Arai et al., 1980; Zengel et al., 1984) and mammalian ras proteins (Seeburg et al., 1984 ; McGrath et al., 1984) have been shown to share similarities of sequence and of GTP-binding and GTPase activities (Halliday, 1984). On the basis of primary sequence comparisons, Kohno et al. (1986) identified six highly similar regions corresponding to about 160 amino acids in the N-terminal third of the hamster EF-2, which are shared between these proteins and the bacterial initiation factor 2 alpha (Sacerdot et al., 1984), bovine transducin (Tanabe et al., 1985), and yeast elongation factor 1 alpha (Nagata et al., 1984). Five of these regions, termed G1 - G5, have since been shown to be conserved in all elongation factor sequences so far examined. Among all of the GTP-binding proteins compared, the most conserved sequence (Asn-Lys-Xaa-Asp) is in the G5 region. These amino acids are considered to play important roles in GTP and GDP binding. By

photoaffinity-labeling and X-ray crystallographic studies, Girshovich et al. (1979) and la Cour et al. (1986) have also shown that while all five domains in EF-Tu are involved in the binding of guanine nucleotides, the G1 domain may be more important for GTP-hydrolysis.

Guanosine nucleotide-binding proteins are often involved in regulatory reactions. They are believed to constitute a distinct class of proteins, because of (i) their ability to be ADP-ribosylated, (ii) the conformational alterations induced by the nucleotide binding and (iii) the similarities in subunit composition (cf. Hughes, 1983). EF-2 differs from other GTP-binding proteins in being composed of a single polypeptide chain. However, the proteolytic pattern of EF-2 and the stability of its trypsin- and chymotrypsin-derived polypeptides indicate the existence of a pseudo-subunit structure (Nilsson and Nygard 1985).

3. Homology with bacterial elongation factor EF-G

By comparison of the deduced amino acid sequence of hamster EF-2 and the E. coli functional homolog EF-G, Kohno et al. (1986) reported sequence segments with 34 - 75% similarity between these proteins. Grinblat et al. (1989) have shown that whereas segment E1 is common to several elongation factors, segments E2 - E4 are highly conserved between hamster and Drosophila EF-2, suggesting a general role for the region that includes segment E1, and a more specific role for the region containing E2 - E4. Kohno et al. (1986) suggested that E1, which is located in the GTP-binding region, might interact with peptidyl-tRNA and/or ribosomes rather than guanine nucleotides. In general, E1 - E4 may be involved in direct ribosome-binding as well as modulation of affinity of EF-2 for the ribosome (Nilsson and Nygard, 1985; Kohno et al., 1986).

4. The modifiable histidyl region

ADP- ribosylation by diphtheria toxin was shown to decrease the affinity of EF-2 for

ribosomes in vitro, (Nygard and Nilsson, 1985), suggesting that the site of ADP-ribosylation is located in the ribosome-binding domain. The specificity of DT and PA for EF-2 suggests that all EF-2s contain a functionally essential structural feature which is recognized by the toxins and is absent from all other proteins. Sequence analysis of the trypsin-derived 15 amino acid ADP-ribosyl peptide of rat liver (Robinson et al., 1974), yeast (Van Ness et al., 1978), beef liver and wheat germ (Brown and Bodley, 1979), identified the site of ADP-ribosylation as diphthamide (Van Ness et al., 1980). That the diphthamide residue results from modification of histidine has been confirmed from deduced amino acid sequences of EF-2 from a variety of eukaryotic and archaebacterial sources (Kohno et al., 1986; Grinblat et al., 1989; Gehrmann et al., 1985). Since only a single histidyl residue is modified within the EF-2 sequence, one can infer that the diphthamide forming enzymes must display a stringent sequence and / or secondary structure specificity. The amino acid sequences near this histidyl residue are conserved in all 4 peptides (Brown and Bodley, 1979), and in all eukaryotic and archaebacterial EF-2s so far examined, and thus seem important both for the function of EF-2 and for the recognition of EF-2 by the toxins. Since EF-G is not ADP-ribosylatable by DT, this region and the target histidyl residue (His-715 in the hamster sequence) are not conserved in eubacterial EF-G (Kohno et al., 1986).

a. Non-ADP-ribosylatable EF-2 mutants

Kohno et al. (1985) and Kaneda et al. (1984) determined that DT-resistant and co-dominant CHO, human and hamster mutant cell lines contain non-ADP-ribosylatable EF-2 which is cross-resistant to PA. Sequence analysis of cDNA clones encoding EF-2 from mutant hamster cells revealed a single point mutation which replaced Gly-717 in the wild-type EF-2 cDNA by arginine, and which conferred resistance to DT and PA in transient and long-term expression assays (Kohno and Uchida, 1987). The mutation was found to be co-

dominant. Four independently isolated cell lines, generated by a one-step EMS mutagenesis contained the same mutation which was found in one allele of the EF-2 gene. One cell line (KE1) maintained about 50% of the normal level of cellular protein synthesis in the presence of DT or PA and contained equivalent amounts of ADP-ribosylatable and non-ribosylatable EF-2 molecules (Kohno et al., 1985). It is not clear how this mutation confers resistance to ADP-ribosylation by DT and PA. It is possible that the substitution blocks the interaction of EF-2 with the toxins and/or prevents modification of the histidyl residue to diphthamide, thus indirectly blocking ADP-ribosylation.

By two-step mutation of hamster cells with EMS Kohno and Uchida (1987) also isolated a mutant cell line KEE1 which showed full resistance to DT and PA and produced only non-ribosylatable EF-2. The first EMS treatment introduced a single point mutation in codon 717 in one allele and the second treatment produced a point mutation in the other allele of KEE1 cells resulting in the substitution of leucine for Pro-207. Transient expression assays in mouse L cells using the Pro-207 mutant showed loss of the native translocase activity of EF-2. This result indicated that EF-2 synthesized by only one allele may be sufficient for normal protein synthesis without causing retardation of cell growth, since KEE1 cells grew normally with or without toxin. No mutation of the histidyl residue was isolated (Kohno and Uchida, 1987), suggesting that this histidine in EF-2 is essential for recognition of the appropriate ribosomal site for translocation during the elongation step. EF-2 with a substitution at this residue might be non-functional even if not recognized by toxins.

G. EF-2 GENE ORGANIZATION

To date, only two eukaryotic EF-2 genes have been characterized based on genomic clones: the hamster and Drosophila EF-2 genes. The entire hamster EF-2 gene is 5.7 kb in length and has thirteen exons separated by introns of 90 - 200 bases except the first and

third introns which are about 1 kb and 400 bases long, respectively (Nakanishi et al., 1988). The Drosophila EF-2 gene is about 5 kb and includes four exons which are separated by introns of about 100 - 900 bases (Grinblat et al., 1989). The organization of the EF-2 gene is thought to reflect EF-2 function and its contribution to the efficiency of the translational process (Nakanishi et al., 1988). In both genes, the first exon is short (84 bases in the hamster and 75 bases in Drosophila) and largely untranslated, containing only the initiator methionine codon. The second exon of the Drosophila gene encodes all of the GTP-binding homologies as well as the conserved E1 region, while exons 3 and 4 encode the EF-2/EF-G-specific homologies E2 and E3-E4, respectively. The second intron is located within the poorly conserved segment between amino acids 240 and 274. In the hamster gene, the proposed GTP-binding domains are not interrupted by any of the 12 introns. Each exon appears to correspond to a functional domain of EF-2.

1. Copy number of EF-2 genes

Although multiple copies of the gene encoding EF-1α have been reported for a number of eukaryotes (Linz et al., 1986; Hovemann et al., 1988), the EF-2 gene from eukaryotic sources is present in a single copy per haploid genome as determined by genomic Southern blot analysis, or in the case of Drosophila, by in situ hybridization to polytene chromosomes and analysis of multiple genomic clones (Kohno and Uchida, 1987; Grinblat et al., 1989; Toda et al., 1989). However, Koide et al. (1990) reported that mouse cells contain about 70 copies of amplified EF-2-related sequences or processed pseudogenes, called MERs, per haploid genome, in addition to a single copy of the EF-2 gene.

2. Chromosomal localization of EF-2 genes

As in eubacteria (Nomura et al., 1984), the gene encoding the Methanococcus

vannielii EF-2 is closely linked to (i.e., precedes) that coding for EF-1α (Lechner et al., 1988). Also, the Halobacterium halobium EF-2 gene lies downstream from those of ribosomal proteins H-S12 and H-S7 (Itoh, 1989). This is analogous to the str operon of E. coli, in which genes are linked in the order S12-S7-EFG-EFTu (Lindahl and Zengel, 1986). The Drosophila EF-2 gene is found at position 39E-F at the base of the 2L chromosomal arm (Grinblat et al., 1989). Deletions in this location produced a Minute phenotype (which is characterized by slower growth rate and a smaller adult body size) as a result of a dominant action of many recessive lethal mutations at about 50 loci throughout the genome. This phenotype is consistent with the predicted consequences of loss of an essential component of the translational machinery. One Minute locus has been identified as a gene encoding a ribosomal protein (Kongsuwan et al.,1985), and a number of others are believed to encode ribosomal components (Kay and Jacobs-Lorena, 1987).

3. Promoter activity

EF-2 belongs to the family of housekeeping genes whose function is present across the entire evolutionary spectrum. By deletion mapping, Nakanishi et al. (1988) showed that the region from positions -31 to -92 may be the most effective element of the hamster EF-2 gene promoter. This region with very high (79%) G+C content, contains many GC repeats fitting the consensus for the GC box (Melton et al., 1986), and a CACCC-like sequence which is repeated several times in this region and at position -170. CACCC is known to be a regulatory sequence of housekeeping genes (Lawn et al., 1980), and is also observed in other genes (Nonaka et al., 1986; McDonald et al., 1986). The promoter of the EF-2 gene may be highly active because of the large number of such repeats. A functional TATA box and a cAMP-responsive element, TGACGTCA (Montiminy et al., 1986), whose function

here is not known, are found at positions -30 and -10, respectively.

The EF-1α gene, also a housekeeping gene, expresses one of the most abundant proteins in eukaryotic cells (Slobin, 1980). Hovemann et al. (1988) studied the Drosophila genes F1 and F2 encoding EF-1α and found sequences with strong homology to the general promoter enhancer motif, the HOMOL box (Huet et al. 1985). This motif has been identified in front of the EF-1α gene and several ribosomal protein genes in the yeast Saccharomyces cerevisiae, but is not found in the hamster EF-2 gene promoter.

The content of EF-2 mRNA is high (about 0.1% of the total mRNA) in growing mammalian cultured cells, and it is used at each step of addition of an amino acid to the growing polypeptide chain (Nakanishi et al., 1988).

H. TRANSLATION FACTOR-LIKE PROTEINS

1. IF_{EMC}

An eIF-3-like factor referred to as IF_{EMC} is required specifically in addition to eIF-3 for translation of RNA of encephalomyocarditis (EMC) virus. This factor was purified to homogeneity from ascites cell supernatant and has a molecular weight of ~53,000 (Wigle and Smith, 1973). A similar factor from a reticulocyte ribosomal wash reportedly stimulates the translation of α-globin but not β-globin mRNA (Nudel et al., 1970).

2. LC/p43

Partial sequence analysis of tryptic peptides (including one containing a GTP-binding domain) of a tumor-related 43 kDa cytoplasmic protein (antigen) LC/p43 isolated from human hepatoma cell lines revealed that this protein shares 50 - 70% homology to different domains of EF-1α and EF-Tu (Koch et al., 1990). The function of LC/p43 is not known, but this

antigen is almost always found in tumor cells (but not normal adult tissues), indicative of an important physiological role. Although LC/p43 differs in size from EF-1α (53 kDa), the shared homology of these proteins may imply similar biochemical properties.

3. Cyclophilin

Cyclophilin is a specific, high-affinity binding protein for the immunosuppressant cyclosporin A (Handschumacher et al.,1984). A ten amino acid domain at the N-terminus of bovine cyclophilin and hamster EF-2 share 50% similarity. Another domain, from amino acid 115-157 in the former, which contains the putative binding site for cyclosporin A (Dalgarno et al., 1986) shares 50% similarity with residues 312-356 of hamster EF-2. Both proteins contain the sequence Glx-Xaa-Gly-Xaa-Xaa-Gly, which is characteristic of nucleotide binding proteins (Wierenga and Hol, 1983). These sequence similarities indicate that cyclophilin may be a nucleotide binding protein and that it may play a role in protein biosynthesis (Gschwend et al., 1988). In this context it is intriguing that cyclosporin A is an effective inhibitor of phorbol ester-induced protein synthesis (Gschwend et al., 1988).

4. Beta integrin chain

The Drosophila EF-2 gene was cloned by virtue of its weak cross-hybridization with the myospheroid gene which encodes a Drosophila beta integrin chain (Grinblat et al., 1989). Drosophila EF-2 is recognized by monoclonal antibodies directed against beta integrin, suggesting that both proteins (which are functionally unrelated) share an epitope. Drosophila integrin is a cell surface protein implicated in cell-extracellular matrix interactions. The identity of the shared epitope is not obvious from sequence comparisons (Grinblat et al., 1989); the longest contiguous match between Drosophila EF-2 and beta integrin, which is also conserved in hamster EF-2, is a pentapeptide, FDAIM, located at position 301 in the

hamster sequence (Figure 11). Oligopeptides of comparable sizes have been shown to confer antigenic specificity (Geysen et al., 1988). Thus, the shared epitope may be assembled from non-contiguous sequences upon protein folding.

I. CAENORHABDITIS ELEGANS AS A MODEL SYSTEM

Caenorhabditis elegans, a member of the family Rhabditidae, is a small free-living nematode which feeds primarily on bacteria. Adult nematodes are found as males or hermaphrodites of 1 mm in length with a life cycle of about 3.5 days when grown on E. coli at 20°C. After fertilization, the eggs are laid at about the 30-cell stage (gastrulation). Each hermaphrodite produces 200 - 300 progeny. Following embryogenesis, a juvenile containing about 550 cells hatches from the egg and develops through four larval stages, L1 - L4, before reaching the adult stage. Many cell divisions occur during the larval period, the somatic cell number increasing to 959 in the adult hermaphrodite and 1031 in the adult male. The first larval stage contains only two germ line cells while the adult contains 1000 - 2000 germ line nuclei. L2 stage larvae can develop into a resistant stage known as dauer larvae under adverse environmental conditions such as starvation (Cassasa and Russell, 1975). In this state, the nematode is resistant to starvation, desiccation, and harmful chemicals such as detergents, but resumes normal development when nutrients become available.

C. elegans is a simple organism both anatomically and genetically and is easily maintained in the laboratory. Individual animals are easily manipulated and large numbers can be grown routinely in mass culture. The animals are transparent throughout the life cycle, which has allowed the complete cell lineage to be elucidated (Sulston et al., 1983). Its small size has allowed a complete anatomical description of the animal and of the wiring of its nervous system at the electron microscope level (White et al., 1986). Homologous transformation by microinjection has recently been developed, allowing the study of C.

elegans gene expression in an appropriate context (Fire, 1986).

J. CAENORHABDITIS ELEGANS GENOME MAP

C. elegans, with only six haploid chromosomes (five autosomal and one sex chromosome), is ideally suited to genetic and molecular biological analyses. The haploid genome size is 10^8 bp, approximately eight times that of the yeast S. cerevisiae or half that of Drosophila (Wood, 1988) and about 80% of the C. elegans genome is composed of single copy sequences.

1. Genetic map

C. elegans has a well characterized transposon, Tc1, which is useful for transposon tagging mutagenesis and cloning (Moerman et al., 1986). In addition, mutants are readily obtained following chemical mutagenesis or exposure to ionizing radiation, and the sexual system of self-fertilization makes it easy to isolate recessive mutations on all chromosomes. To date, thousands of mutations have been mapped to about 500 genes. Many include genes which affect behavior and morphology. Brenner (1974) estimated the total number of essential genes in C. elegans to be approximately 2000 based on the frequency of induced visible mutations and X chromosome lethals. Clark et al. (1988) have revised this number to approximately 3500. By complementation, deletion, and recombination techniques, many of these genes have been genetically mapped to the six linkage groups. A strong tendency for the clustering of genes has been reported (Brenner, 1974).

2. Physical map

The simplicity of the nematode genome has sparked a cooperative effort to physically map the entire C. elegans genome. The map initially consisted of a collection of overlapping

cosmid clones which were ordered using a fingerprinting technique (Coulson et al., 1986). A group of overlapping cosmid clones (known as a contig) can be mapped to a chromosome by in situ hybridization (Albertson, 1985). Recent improvements in physical mapping make use of large segments of DNA cloned in yeast artificial chromosome (YAC) vectors (Coulson et al., 1988). By reciprocal hybridization of YACs and cosmids, cosmid contigs can be linked together and gaps in the map filled. At the time of writing, nearly 85% of the genome had been categorized into about 102 segments (contigs) with an average size of 550 kb, and some greater than 850 kb. It is now a relatively straightforward procedure for a researcher to locate most DNA sequences on the genome map. The physical map now provides an accurate representation of distances between genes, which was only approximated by classical genetic mapping. By assigning genetically mapped fragments to the physical map, 85 x 10^6 bp of DNA is now aligned with the genetic map. The genome map has been essential in the recently initiated project to sequence the entire genome (Roberts, 1990).

3. Localization of ubq-1

Graham et al. (1989) cloned and characterized the major polyubiquitin gene ubiA of C. elegans. Employing a ubiA gene clone λUb1, Sulston and coworkers isolated overlapping cosmids (i.e. a contig) containing ubiA (subsequently named ubq-1) as part of their C. elegans genome mapping project (J. Sulston, personal communication). One of these cosmids (C16A7) was used as a probe for in situ hybridization to C. elegans chromosomes by D. Albertson (Medical Research Council Laboratory of Molecular Biology, Cambridge, England). Using these procedures, ubq-1 was mapped to the middle of chromosome III, between the genetic markers ceh-10 and col-8, both of which had also been physically mapped (Edgley and Riddle, personal communication). A representation of the physical map of the ubq-1 region is shown in Figure 3, p. 56.

K. THE PRESENT STUDY

A great deal of progress has been made since 1986 when the initial experiments to physically map the C. elegans genome were reported, i.e. the contig number and average size have gone from 899 contigs and 58 kb to 102 contigs and 550 kb, and about 85 million bp of DNA has been aligned with the genetic map. However, gaps still exist between the 102 clusters (contigs) mapped to date. Within contigs, many regions, including those around the ubq-1 locus (see Figure 3) remain genetically undefined. For example, given that the six haploid chromosomes are roughly of equal size (Wood, 1988), one chromosome (arbitrarily divided into 100 units) contains ~17 x 10^6 bp of DNA. Thus each unit on the scale of Figure 3 will contain ~170 kb of DNA. This means that between the genetic markers ceh-10 and ubq-1, and ubq-1 and col-8 (see Figure 3) are regions containing ~170 kb and ~9 00 kb of DNA, respectively, which are yet to be defined genetically. The analysis of such regions by classical methods, e.g. the mapping of visible and lethal mutations, would require a great deal of effort, and would not yield the identities of the genes thus mapped. Furthermore, members of multi-gene families with similar or identical functions, or those yielding e.g. dominant lethal phenotypes, would be missed by such an analysis. The most straightforward and informative approach, given the availability of the cloned DNA, is to physically map and sequence such regions, and identify corresponding transcripts.

The present investigation was undertaken in order to (i) create a finer physical (restriction) map of the region around ubq-1, (ii) define genes and transcripts using molecular techniques, and (iii) further characterize selected genes in the region.

The discovery (in the region around ubq-1) of a gene eft-1 predicting a polypeptide with features identical to eukaryotic elongation factor 2 (EF-2) prompted a search for the authentic gene eft-2 encoding C. elegans EF-2.

II. EXPERIMENTAL PROCEDURES

A. MAINTENANCE OF COSMID CLONES

The overlapping cosmid clones ZK48 (which contains ubq-1), C16A7, ZK331, C41E11 and T01D10 were obtained in E. coli HB101 cells (see Appendix B) suspended in 2 x YT (YT= 0.8% tryptone, 0.5% yeast extract, 0.5% NaCl pH 7) medium containing kanamycin (30 μg/ml, for the lorist-based cosmids ZK48, ZK331 and T01D10) or ampicillin (75 μg/ml, for the pJB8 cosmids C16A7 and C41E11), from J. Sulston (MRC Cambridge) and stored frozen at -70°C in 20% glycerol.

1. Cosmid DNA preparation

A single cosmid colony was used to grow a 3 ml culture in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl pH 7.5) medium containing the appropriate antibiotic. Cosmid DNA was prepared essentially by the alkaline lysis protocol of Birnboim and Doly (1979) as modified by Ish-Horowicz and Burke (1981), starting from 1.5 ml portions of the culture. After two ethanol precipitations, the product (1-3 μg) was dissolved in 50 μl of TE (TE: 10 mM Tris-HCl pH 8.0, 1 mM EDTA) containing 40 μg/ml RNase A. Large (250 ml) cultures were grown to saturation from 1.0 ml portions of the minicultures. Preparation of large quantities of cosmid DNA was as described by Radloff et al. (1967) and involved a differential precipitation step and centrifugation to equilibrum in cesium chloride gradients containing saturating amounts of ethidium bromide. The product (200-250 μg) was dissolved in 400 μl TE and stored at -20°C in 50 μl aliquots.

2. Cosmid transcript mapping

Cosmid DNA (0.5-1 μg) was digested with restriction endonuclease for 2 or more

hours. Fragments were separated by electrophoresis on a 0.7% agarose gel and blotted onto nitrocellulose or nylon membranes (Amersham) in 20 x SSPE (1 x SSPE: 180 mM NaCl, 1 mM EDTA, 10 mM NaH_2PO_4, pH 7.4). After transfer, the membrane was hybridized with ^{32}P-labeled first strand cDNA probes made from embryo polyA+RNA or pooled polyA+RNA selected from embryo, larvae, and young adult nematode populations. Hybridization conditions were: 42°C for 18 hours in a buffer containing 50% deionized formamide (Sambrook et al., 1989).

3. Cosmid restriction mapping

Restriction mapping of cosmids was carried out by the cosmid mapping procedure of Rackwitz et al. (1985) with modifications. Since the cohesive site in linearized cosmid DNA appeared to be cleaved more efficiently than that in circular DNA, cosmid DNA was first partially digested with an appropriate restriction enzyme and the cohesive site in the linearized cosmid DNA was then cleaved with a crude λ DNA terminase preparation (see below). The λ DNA terminase reaction was carried out at 22°C for 30 minutes in a 26 µl-reaction mixture that contained 4 µl (1 - 2 µg) of linearized cosmid DNA, 2µl of DPB-ATP (85 µl of DPB: 6 mM Tris-HCl, pH 7.4, 18 mM $MgCl_2$, 30 mM spermidine, 60 mM putrescine, mixed with 100 mM ATP), 15 µl of DPA (20 mM Tris-HCl, pH 8, 3 mM $MgCl_2$, 1 mM EDTA, 7 mM β-mercaptoethanol), and 5 µl of λ terminase fraction. The DNA was extracted twice with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1), and once with 5 volumes of ether. After removal of the ether by heating the open tubes to 75°C for 5 minutes, the DNA was precipitated with 2.5 volumes of 95% ethanol and resuspended in TE. Alternatively, 10 µg of cosmid DNA (sufficient for five partial digestions) was first cleaved with the terminase in a 50 µl-reaction mixture. After extractions with phenol-

chloroform-isoamyl alcohol and ether, aliquots of the DNA (without precipitation) were partially digested with the appropriate restriction enzyme. DNA was precitipated with 95% ethanol and resuspended in TE . The alternative method provided bulk terminase reactions from which aliquots could be obtained for the different enzyme digestions.

Aliquots of the partial restriction digest were divided into two equal parts, and each was hybridized with ^{32}P-5'-end labeled dodecamers L or R complementary to the left or right cohesive λ DNA terminus, respectively (see Appendix A1), in a 10 µl reaction. Five µl of gel-loading buffer (36 mM Tris-HCl, pH 7.7, 30 mM NaH_2PO_4, 60 mM EDTA, 50% glycerol, and 0.1% bromophenol blue) was added and the sample was loaded onto a 0.4% agarose gel. After electrophoresis at 1.5 volts per centimeter for 24 hours in TBE (90 mM Tris-borate, 8.3, 1 mM EDTA), the gel was dried onto Whatman DE cellulose paper at 60°C for 45 minutes and autoradiographed for 2 - 3 days with Kodak X-Omat AR film and intensifying screen.

B. PREPARATION OF LAMBDA DNA TERMINASE

The E. coli strain AZ1935 containing plasmid pCM101, which produces large amounts of λ DNA terminase upon temperature induction, was kindly provided by H. Murialdo (University of Toronto). Induction of gene expression was carried out as described by Chow et al. (1987) with some modifications. 400 ml of LB broth (10 g tryptone, 10 g NaCl, 5 g yeast extract, per liter, pH 7.5) containing 20 µg/ml of ampicillin in a 2 liter flask, was inoculated with 3 ml of an overnight culture of AZ1935 and incubated in a gyratory shaker at 31°C until the culture attained a density of about 10^8 cells/ml. For thermoinduction, the temperature of the culture was quickly increased to 44 - 45°C by partial submersion of the flask in a 70°C water bath while shaking rapidly. The flask was then transferred to a 45°C shaking water bath for 15 minutes. The temperature of the culture was then lowered to 40-

41°C by partial submersion in an ice-water mixture, the culture was incubated in an air shaker at 40°C for 15 minutes, and finally chilled in ice-water with swirling.

After centrifugation at 6,000 rpm in a GSA rotor for 10 minutes at -4°C, the pellet was resuspended in and made up to 4 ml with DPA buffer, and sonicated in a 50 ml Falcon tube submerged in an ice-NaCl mixture, using 20-second bursts until a translucent suspension was achieved. After centrifugation at 6,000 rpm (SS34 rotor) for 6 minutes, the supernatant (containing the terminase) was mixed with an equal volume of 100% glycerol; PMSF (made up in DMSO, Sigma) was added to 0.1 mM and Aprotinin (Sigma) to 50 μg/ml final concentration. This terminase preparation was stored at -20°C without appreciable loss of activity for 4 months.

1. Terminase cleavage activity assay

The measurement of λ DNA terminase activity was according to Murialdo et al. (1981). The substrates used for the cleavage activity assay were the cosmids C16A7.3 and C41E5.5 (Fig. 5) linearized with PstI. These cosmids were obtained by digestion of cosmids C16A7 and C41E11 with HindIII and religation of the resultant 7.3 and 5.5 kb fragments, containing the pJB8 vector sequences and 1.5 kb or 200 bp C. elegans genomic DNA insert, respectively. Cosmid DNA (1 μg) was digested with PstI in a 12 μl reaction volume. Aliquots of 4 μl each were mixed with 5 μl of 3 x diluted (with DPA) or undiluted terminase preparation, 2 μl of DPB-ATP and 15 μl of DPA , and incubated at room temperature for 30 minutes. Five μl of the gel-loading buffer was added and the samples were resolved by electrophoresis on a 0.5% agarose gel. Four μl of each linearized cosmid was run along with the terminase-cleaved DNA to serve as standards. The DNA was visualized by ethidium bromide staining.

C. GROWTH AND MAINTENANCE OF NEMATODES

C. elegans (Bristol N2 strain) and C. briggsae were maintained on NGM (0.3% NaCl, 0.25% bactotryptone, 5 µg/ml cholesterol, 1 mM $MgSO_4$, 1 mM $CaCl_2$, and 25 mM KH_2PO_4 pH 6.0) plates containing E. coli OP50 as described by Brenner (1974), at 22°C. To obtain large quantities, nematodes were grown in liquid culture as described by Sulston and Brenner (1974) with frozen *E. coli* K12 (Grain Processing, Muscatene, Iowa) as food source. Synchronous populations were obtained by inoculating S medium (Basal S: 0.1 M NaCl, 50 mM KH_2PO_4 pH 6.0; supplemented with 0.01 mg/ml cholesterol, 2 mM potassium citrate pH 6.0, 0.3 mM $MgSO_4$, 0.3 mM $CaCl_2$, 1.3 µM $FeSO_4$, 2.5 µM EDTA, 0.5 µM $ZnSO_4$, 0.5 µM $MgCl_2$, and 0.05 µM $CuSO_4$) with embryos (2.5 grams/liter) prepared by dissolving gravid adults in alkaline sodium hypochlorite (Emmons et al., 1979). The 4 liter culture was allowed to hatch and arrest as L1 larvae prior to feeding (with 80 grams of frozen E. coli) to improve the degree of synchrony (Cruzen and Johnson, 1987). The culture was heavily aerated to ensure nematode viability. Frothing due to lysed bacterial protein and larval cuticle accumulation was suppressed by addition of sterile antifoam A emulsion (Sigma) as required. For harvesting successive stages, 1 - 2 liters were removed at the desired stage and the culture was made up to 4 liters with S medium, fed, and allowed to continue to the next stage. Dauer larvae were obtained by starvation of cultures at the L2 stage; dauer stage were identified by their altered morphology and activity level and by testing for their survival longer than 30 minutes in 0.1% SDS (Cassada and Russell, 1975). Typically, 2 grams of embryos inoculated in a 4 liter culture yielded 40-45 grams of gravid adults and about 10 grams of viable embryos.

1. Collection and freezing of larvae and adult nematodes

Cultures (1 - 2 liter) were placed on ice for 1 hour, then kept at 4°C until the

nematodes had settled. This technique removed much of the residual bacteria (which remained in suspension) and reduced the volume of culture to be processed. The supernatant was aspirated and nematodes were harvested by centrifugation (400 x g for 5 minutes at 4°C). Nematodes were purified from debris by two successive flotations on a 30% sucrose cushion (2000 x g, 30 sec, room temperature, Sulston and Brenner, 1974). Floated animals were washed twice with 4 volumes of sterile 0.14 M NaCl, suspended in an equal volume of sterile Basal S containing 30% glycerol, and frozen as pellets by dropping 100 µl at a time into liquid nitrogen. Worm pellets were stored at -70°C.

D. PREPARATION AND ANALYSIS OF NEMATODE RNA

Total cellular RNA was isolated by the method of Antonucci (1985). Frozen nematode pellets (3 - 5 grams) were powdered in a chilled mortar and pestle and dissolved in 3 ml of homogenization buffer (7.5 M guanidinium hydrochloride; 25 mM sodium citrate, pH 7.0; 0.1 M β-mercaptoethanol). The homogenate was passed ten times through a 21 g needle to reduce sample viscosity by shearing genomic DNA, and gently layered onto a 1 ml cushion of cesium chloride solution (5.7 M CsCl, 25 mM sodium citrate, pH 5.0) which had been treated with 0.05% diethylpyrocarbonate (DEPC) and sterilized by passing through a 0.22 µm Millipore filter. After centrifugation (220,000 x g) for 16 hours at 22°C, the RNA pellet was dissolved by suspension in 300 ml of sterile, DEPC-treated dH_2O and incubated at 50°C for 15 minutes. The RNA was then precipitated overnight at -20°C with 0.1 volumes of DEPC-treated 3 M sodium acetate pH 5.2 and 2.5 volumes of 95% ethanol, and finally redissolved in DEPC-treated dH_2O. Typical yields were 3 mg of total cellular RNA per gram of starting material.

1. Selection of polyA+RNA

PolyA+RNA selection was achieved using the PolyATract™ Magnetic mRNA isolation system (Promega). Total cellular RNA (up to 5 mg) was heated to 65°C for 10 minutes and incubated with a biotinylated oligo(dT) probe, in 0.5 x SSC at room temperature for 30 minutes (i.e. until completely cooled) to hybridize the probe to the 3'-polyA tail region of the mRNA species. The hybrids were captured using streptavidin covalently coupled to paramagnetic particles. After the removal of non-specific hybrids by high stringency washing (four times in 0.1 x SSC), the polyA+RNA was eluted into ribonuclease-free deionized water. The concentration of total and polyA+RNA was determined by absorbance at A260. This technique proved fast and efficient, yielding an essentially pure fraction of polyA+RNA after only a single round of purification. Typically, 50 µg polyA+RNA was recovered per mg of total RNA.

2. Electrophoresis of RNA and Northern transfers

RNA (in 25 mM EDTA containing 0.1% SDS) was denatured with formaldehyde as described by Sambrook et al. (1989). Prior to loading, RNA samples were treated with 0.3 µg/ml ethidium bromide. After electrophoresis in a 1% agarose gel in 5 x formaldehyde gel buffer (100 mM MOPS, pH 7; 40 mM sodium acetate; 5 mM EDTA, pH 8), the RNA was transferred to a nylon membrane (Amersham) as described (Sambrook et al., 1989). Treatment of RNA with ethidium bromide was found not to interfere with either the transfer of RNA to nylon or the hybridization of the probe to the filter-blotted RNA.

E. ISOLATION OF NEMATODE GENOMIC DNA

High molecular weight genomic DNA was isolated from nematodes by the method of Emmons et al., (1979) as modified by Jones et al., (1986). Frozen nematodes were

suspended in 20 ml protease K buffer (100 mM Tris-HCl, pH 8.5; 50 mM EDTA; 200 mM NaCl; 1% SDS); protease K was added to a final concentration of 200 µg/ml and the solution was incubated at 65°C for one hour. The sample was extracted three times with phenol in a separatory funnel and once with chloroform in polypropylene tubes. Phase separation was achieved by centrifugation (2000 x g for 3 minutes). The aqueous DNA supernatant was chilled to 4°C and two volumes of ice-cold 95% ethanol was gently layered over it. The precipitated DNA was recovered by gently winding it upon a glass rod as the tube was rotated at an angle to disturb the interface. The ethanol was changed three times during winding. This procedure rid the DNA of co-purifying RNA which remained in the ethanol washes (any RNA left with the DNA after the ethanol washes was routinely digested with RNaseA in subsequent restriction enzyme reactions). The DNA was rinsed with 70% ethanol and redissolved in TE. Typically, 3 mg of DNA was obtained per gram of frozen nematodes. For C. elegans, 1 µg represents approximately 10 million genome equivalents.

F. GENERAL DNA TECHNIQUES

1. Restriction endonuclease digestion of DNA

a. Partial digestion

Partial digestion was carried out on DNA samples (1 - 2 µg) in a 105 µl reaction mixture containing the appropriate 1 x restriction enzyme buffer as described by Sambrook et al. (1989). The reaction mixture was aliquoted into a series of tubes such that tube 1 contained 30 µl, while the other tubes each contained 15 µl of reaction mixture. Restriction enzyme (2 - 4 units) was added to tube 1, mixed rapidly and placed on ice. Using a different pipet tip, 15 µl was removed from tube 1 into tube 2 which was also mixed rapidly

and placed on ice. The serial dilution process was continued by successively pipeting 15 µl from tube 2 to 3, 3 to 4, until the enzyme was added to the last tube, making it up to 30 µl. The tubes were incubated at 37°C for 30 minutes and the reaction was terminated by the addition of 1 µl of 200 mM EDTA per tube. The reactions were pooled and the DNA was precipitated with ethanol and resuspended in TE.

b. Complete digestion

Complete digestion of DNA (0.5 - 1 µg) was carried out in 1 or 2 x One-Phor-All Buffer *PLUS* (Pharmacia) or in the restriction enzyme buffer system described by Sambrook et al. (1989), in a total volume of 15 µl. For each reaction, 1 - 2 units of restriction enzyme (Bethesda Research Laboratories, New England Biolabs, Boehringer or Pharmacia) was used and digestion was carried out for 2 - 3 hours. For overnight digestions, bovine serum albumin (ultrapure grade) was added to a final concentration of 100 µg/ml. For high molecular weight DNA such as phage DNA and C. elegans genomic DNA, 5 µg of RNaseA was also included in the reaction mixture.

2. Electrophoresis of DNA and Southern blot analysis

PCR products and restriction endonuclease fragments of DNA were analyzed on 0.5 - 1% agarose gels by electrophoresis in 1 x TBE (90 mM Tris-borate pH 8.3, 1mMEDTA) containing 0.5 µg/ml ethidium bromide. DNA bands were visualized and photographed under ultraviolet light using a Polaroid camera and type 57 film. For Southern transfers, DNA was denatured and blotted onto nylon or nitrocellulose membrane in 20 x SSPE as described (Southern, 1975), except that the depurination step was omitted.

3. Recovery of specific DNA fragments

DNA fragments were recovered from agarose gels by electroelution onto DEAE membranes (Schleicher and Schuell, NA45). The membrane was cut to the desired size, pre-moistened with NET buffer (150 mM NaCl, 0.1 mM EDTA, 20 mM Tris-HCl, pH 8), and inserted into the gel such that after electrophoresis for 10 - 20 minutes, the DNA had completely migrated onto the membrane. The membrane was rinsed with NET and the DNA was eluted into 150 µl of high salt NET (1 M NaCl, 0.1 mM EDTA, 20 mM Tris-HCl, pH 8) by incubation at 55°C for 15 minutes. After ethidium bromide extraction with 3 volumes of butanol, the DNA was precipitated with 2.5 volumes of 95% ethanol and 3 µl of 0.25% polyacrylamide as carrier (Gaillard and Strauss, 1990).

4. Purification of synthetic oligonucleotides

All of the oligonucleotides (see Appendix A) used in this study were synthesized by Tom Atkinson (The University of British Columbia, Vancouver) using an Applied Biosystems 380A DNA synthesizer. The λ cohesive end complementary oligonucleotides were obtained in purified form while the rest of the oligonucleotides were purified using C_{18} SEP-PAK (Millipore) chromatography in 20% acetonitrile/80% water as described by Atkinson and Smith (1984), except that the purification step by gel electrophoresis was omitted.

G. PREPARATION OF RADIOACTIVE DNA PROBES

1. First strand cDNA probes

First strand cDNA probes were synthesized using the cDNA Synthesis System (Amersham) in a 30 µl standard reaction mix containing 3 µg of embryo polyA+RNA, 60 units of AMV reverse transcriptase (Amersham), 30 µCi each of [α-^{32}P] dGTP and [α-

^{32}P] dCTP, and 5 μg of oligo dT(12/18) primer. The labeled cDNA product was separated from unincorporated nucleotides by spinning the sample through a 1 ml mini-column packed with Sephadex G-50 superfine (Pharmacia). Alternatively, cDNA probes were synthesized from pooled polyA+RNA (5 μg) selected from embryo, larvae and young adult stages of nematode development using the procedure described below (section I-1) for first strand cDNA synthesis, except that 20 μCi each of [α-^{32}P] dATP and [α-^{32}P] dCTP, 0.1 mM each of dATP and dCTP, 10 mM each of dGTP and dTTP were used.

2. Nick translation

Nick-translated radiolabeled DNA probes were prepared by standard procedures (Sambrook et al., 1989) with [α-^{32}P] dATP and [α-^{32}P] dGTP.

3. Primer extension M13 probes

An M13 single-stranded template Mp18-810 which contains 864 bp XhoI-HindIII fragment in the 5' region of the eft-1 sequence. The oligonucleotide EO4 was used as the extension primer such that when the product was digested with HindIII-EcoRI a 407 bp of fragment including the putative initiation ATG and about 400 nucleotides used to generate single-stranded probes for nuclease S1 analysis. The annealing reaction containing 5 μl of template (0.5 μg), 2 μl (1 pmol) EO4 primer (5' TCATGATATGAGGGCAGTCC 3'), 2 μl of 10 x annealing buffer (100 mM Tris-HCl, pH 7.5, 600 mM NaCl, 70 mM $MgCl_2$) was incubated at 65°C for 15 minutes. The mixture was allowed to gradually cool to room temperature, and 1 μl each of 1 mM dCTP, dTTP, dH_2O, 1.5 μl (15 μCi) each of [α-^{32}P] dATP and [α-^{32}P] dGTP, and 0.5 U of E. coli DNA polymerase I (Klenow fragment) were added. After incubation at 37°C for 10 minutes, the reaction was chased with 1 μl each of 1 mM dATP, dGTP and 2 μl dH_2O by incubation for another 10 minutes, before termination by

heating at 70°C for 10 minutes. The product was then digested with HindIII-EcoRI at 37°C for 30 minutes.

4. End-labeling of oligonucleotides

The 5'-ends of oligonucleotides were labeled with [γ-^{32}P] ATP as described (Sambrook et al., 1989) in a 10 µl reaction volume containing 1 µl 10 x kinase buffer (0.5 M Tris-HCl, pH 7.5, 70 mM $MgCl_2$, 0.1 M DTT), 10 µCi [γ-^{32}P] ATP, 20 pmol oligonucleotide, and 10 U T4 polynucleotide kinase (Promega). Labeled oligonucleotide was separated from unincorporated nucleotides in a 1 ml mini spun-column packed with Sephadex G-25 superfine (Pharmacia) that was equilibriated in 5 mM EDTA.

H. NUCLEIC ACID HYBRIDIZATION

Pre-hybridization and hybridization during Southern and Northern blot analyses were performed according to standard procedures (Sambrook et al., 1989) in Seal-a-Meal® bags. To reduce background due to non-specific hybridization, 5 µg/ml of heparin (sodium salt, Sigma) or 100 µg/ml of tRNA (or calf thymus DNA) was included in the hybridization mixture.

1. Oligonucleotide hybridization

After terminase cleavage reactions, aliquots of cosmid DNA partially digested with restriction enzymes were divided into two equal parts, each containing about 0.5 -1 µg (0.015 - 0.030 pmol) of DNA. Aqueous hybridization with 0.2 pmol of ^{32}P-5'-end labeled dodecamers L (5'-dAGGTCGCCGCCC-3') or R (5'-dGGGCGGCGACCT-3') complementary to the left or right cohesive λ DNA terminus, respectively, was carried out in a 10 µl mixture containing 100 mM NaCl (Rackwitz et al., 1984). The mixture was incubated

for 2 minutes at 75°C (to denature cos ends) and immediately transferred to a 45°C waterbath for 30 minutes.

I. PCR ANALYSIS

1. First strand cDNA synthesis

First strand cDNA was synthesized by a combination of the methods of Sambrook et al. (1989) and Frohman et al. (1988). Ten µg of total cellular RNA or 2 to 5 µg of polyA+RNA in 7.5 µl of DEPC-treated dH_2O was denatured by incubation (room temperature, 15 min) with an equal volume of freshly prepared 40 mM methylmercuric hydroxide. The mixture was then snap-frozen in a dry ice/ethanol bath. A 30 µl cocktail containing 33 µM DTT, 10 µl of 5 x Moloney murine leukemia virus (MMLV) reverse transcriptase buffer (1 x = 50 mM Tris-HCl, pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT), 1 U RNasin (Promega), 833 µM dNTP mix (i.e. dATP, dCTP, dGTP, dTTP), 200 pmol anchor oligo-dT RACE oligonucleotide (see Appendix A) was prepared and centrifuged into the frozen denatured RNA. Upon thawing, the mixture was flicked to mix and immediately 4.5 µl (900 units) of MMLV reverse transcriptase (BRL) was added. The reaction was incubated at 37 °C for 45 minutes and was terminated by freezing quickly in a dry-ice ethanol bath.

2. Rapid amplification of cDNA ends (RACE)

Rapid amplification of the cDNA was as described by Frohman et al. (1988) with some modifications. A 2 µl aliquot (4%) of the first strand cDNA reaction containing RNA:cDNA hybrids was used for the amplification reaction in which 50 pmol each of the anchor oligo-dT RACE adaptor, and OPC3 or OPC4 (see Appendix A-2) were added. OPC3

and OPC4 are 24-fold degenerate, 28-mer primers derived from amino acid residues 634 to 640 and 666 to 672, respectively, of the Drosophila EF-2 sequence (Grinblat et al., 1989). These primers contain BamHI sites at their 5' ends to facilitate cloning, and were optimized for C. elegans codon usage (Wada et al., 1990)]. The amplification mix, also containing 45 µl of PCR cocktail consisting of 50 µM of each dNTP, 1 x PCR buffer (10 mM Tris-HCl, pH 8.4, 0.05% Tween 20, 0.05% Nonidet P-40, 0.5 mM $MgCl_2$), and 1 unit of Taq DNA polymerase (Cetus Corp. or Promega), was overlaid with 60 µl of light mineral oil (LifeBrand®) and inserted into a pre-heated (94°C) Ericomp Twinblock thermocycler. Amplification was carried out for 35 to 40 cycles (denaturation for 30 seconds at 94°C, annealing for 1 minute at 55°C, polymerization for 1 minute at 72°C) with a final incubation for 10 min at 72°C. The annealing temperature varied for other oligonucleotides, depending on length and sequence.

3. Analysis and cloning of PCR products

PCR products were analyzed by agarose gel electrophoresis, and the desired fragments were recovered using a DEAE membrane (Schleicher and Schuell, NA45) as previously described. For ease of cloning, all PCR oligonucleotides were synthesized with restriction enzyme recognition sites at the 5' or 3' end. The purified fragments from the RACE-OPC3 (665 bp) or RACE-OPC4 (761 bp) amplifications were digested and cloned into pUC18 or pUC19 using the BamHI site and the SalI site in the OPC and RACE primers, respectively.

J. SCREENING OF RECOMBINANT DNA LIBRARIES

1. Screening of Bacteriophage λZAP cDNA library

A C. elegans cDNA library prepared in λZAP was kindly provided by R.J. Barstead

and R.H. Waterston (Washington University School of Medicine, St. Louis). The cDNA had been size selected to eliminate products less than 500 bp, methylated with EcoRI methylase, ligated to EcoRI linkers, and cleaved with EcoRI restriction enzyme. The cDNA was then ligated to EcoRI-digested bacteriophage λZAP vector (Barstead and Waterston, 1989). The phage was propagated in an E. coli BB4 host strain (see Appendix B) grown in TB (1.2% tryptone, 2.4% yeast extract, 0.4% glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4 pH 7.35) broth supplemented with 0.2% maltose and 10 mM $MgSO_4$. The library was screened with the nick translated, ^{32}P-labeled 761 bp PCR fragment (see Figs. 12 and 14B) by the method of Benton and Davis (1977). Six positive clones, including one (pCef6A) containing the entire coding region, were isolated from 1,200 plaques screened. An 870 bp EcoRI- ScaI probe from the 5' end of pCef6A was used to screen 4,000 phage of the library and 11 more clones were isolated. The library was also screened with nick-translated B3255 (a 3.3 kb HindIII fragment containing part of eft-1) and only one positive clone was isolated out of 4700 transformants screened. All positive clones were purified by a second screen.

a. Excision of λZAP phage clones

The positive phage cDNA clones were converted into pBluescript plasmids by superinfection with the helper phage R408 (Stratagene). For this purpose, 200 µl of A_{600} =1 BB4 cells, 200 µl of phage stock (containing > 10^5 phage particles), and 10 µl (10^8 pfu) of the helper phage, were mixed and incubated at 37°C for 15 minutes. Five ml of TB broth was added and the incubation was continued for 4 to 8 hours. The culture was heated at 70°C for 20 minutes and centrifuged (1000 x g, 5 minutes) in the cold room. The supernatant (which contained the pBluescript plasmid packaged in the M13 or f1 phage particle) was decanted and could be stored at 4°C for 1 to 2 months. To plate the rescued plasmid, 200 µl

of A_{600} =1 BB4 cells and 200 µl of the plasmid stock were combined and incubated at 37°C for 15 minutes. One to 100 µl of this sample were plated on LB plates containing 200 µg/ml ampicillin, and incubated overnight at 37°C.

2. Screening of Bacteriophage λEMBL4 genomic library

A λEMBL4 library derived from a partial MboI digest of C. elegans genomic DNA and cloned into the BamHI site of the vector, was constructed and kindly provided by Chris Link (University of Colorado, Boulder). The phage library was propagated in an E. coli Q358 host strain (see Appendix B) grown in NZYC (1% NZ-amine, 0.1% yeast extract, 0.4% glycerol, 0.1% casamino acids, 0.5% NaCl, 10 mM $MgCl_2$ pH 7) medium. About 40,000 phage, containing approximately four C. elegans genome equivalents (based on the conservative estimate that the DNA insert in each recombinant is about 10 kb), was screened by the method of Benton and Davis (1977). The positive recombinant clones were purified by rescreening.

3. Isolation of Bacteriophage DNA

Bacteriophage DNA preparation from 20 ml cultures was as outlined by Sambrook et al. (1989) with some modifications. Following the pre-incubation (15 minutes at 37°C) of 200 µl of an overnight culture of Q358 cells with 200 µl phage stock (10^6 phage) in lambda dilution buffer, 20 ml of NZYC was added. The culture was incubated at 37°C with vigorous shaking for 5 - 8 hours (until lysis occurred). Chloroform (3 ml) was added and incubation was continued for an additional 10 minutes after which the culture was stored at 4°C overnight. The culture was decanted away from the chloroform, and centrifuged twice (12,000 x g for 10 minutes) to remove all the bacterial debris. After the addition of 3 ml of 5 M NaCl and 3 g of polyethyleneglycol powder (PEG 8000) to the supernatant, the solution

was thoroughly mixed, left at 4°C for two or more hours, and then centrifuged (12,000 x g for 10 minutes) to precipitate the phage particles. The phage pellet was resuspended in 500 µl of DNase buffer (50 mM HEPES pH 7.5; 5 mM $MgCl_2$; 0.5 mM $CaCl_2$) to which 100 µg RNase A and 5 µg DNase I were added . Following incubation at 37°C for 60 minutes, 150 µg of proteinase K and 50 µl of 10 x SET (1 x SET: 10 mM Tris-HCl, pH 7.5; 20 mM EDTA; 0.5% SDS) were added and the sample was incubated for 60 min at 65°C. After extraction with an equal volume of phenol:chloroform (1:1) and then with chloroform, the phage DNA in the aqueous phase was precipitated with two volumes of 95% ethanol. Following centrifugation at 12,000 x g for 5 minutes, the phage DNA pellet was quickly resuspended in 50 µl TE buffer containing 40 µg/ml RNase A. Typically, 3 to 5 µg of phage DNA was obtained.

K. TRANSFORMATIONS

The E. coli strain DH5α was used to propagate (in TB medium) the pUC- and pGEM-derived recombinants, while strain JM109 was used for transformation with M13-derived clones and was grown in YT medium. JM109 cells were made competent for transformation using 50 mM $CaCl_2$ as described by Messing (1983), and DH5α cells were made competent using the methods of Hanahan (1983) which allowed for bulk preparation of competent cells that could be stored frozen at -70°C in aliquots for 2 to 3 months. Typically, competent bacteria had a transforming efficiency of 10^7 transformants per microgram of input plasmid. For some transformations, commercially obtained competent DH5α cells (BRL, > 10^6 transformants per microgram of input plasmid) were used.

Ligations were carried out with 20 ng of vector DNA and 50 - 100 ng of insert DNA in 10 µl reactions containing 50 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 10 mM DTT, 1 mM spermidine, 1 mM ATP, 100 µg/ml bovine serum albumin, and 3 to 10 units of T4 DNA ligase

(Pharmacia), at 15°C for at least 2 hours. Up to 5 µl of the ligation mixture could be used for transformation. Plasmid and M13 derivatives were screened for inserts by their inability to cleave 5-bromo-4-chloro-3-indolylgalactose (X-gal) and produce a blue color, as described by Messing (1983). E. coli cells were screened for efficient transformation with plasmids by their acquired resistance to ampicillin (Sambrook et al., 1989).

L. PURIFICATION OF PLASMID DNA

Plasmid DNA and M13 replicative form (RF) DNA were prepared by the alkaline lysis method of Birnboim and Doly (1979) starting from 1.5 ml of overnight bacterial cultures. Typically, 10 µg of plasmid DNA and 2 µg of M13 RF DNA were obtained.

M. PREPARATION OF M13 SINGLE-STRANDED DNA

An overnight phage culture (1.5 ml) was centrifuged at 12,000 x g for 5 minutes and 1.3 ml of the supernatant was transferred to a tube containing 0.3 ml of 2.5 M NaCl-20% PEG 6000. The sample was mixed by inverting the tube 10 times, kept for 15 minutes at room temperature, and centrifuged as before. The resultant pellet was resuspended in 200 µl low tris buffer (20 mM Tris-HCl pH 7.5, 20 mM NaCl, 1 mM EDTA), extracted once with an equal volume of phenol and twice with phenol:chloroform (1:1), then precipitated from 0.3 M sodium acetate (pH 7) and 2.5 volumes of 95% ethanol (12,000 x g for 5 minutes at 4°C). The pellet was air-dried and resuspended in 25 µl of the low Tris buffer. Typical yield was 3 to 5 µg of M13 single-stranded DNA.

N. PREPARATION OF NESTED DELETION CLONES

Unidirectional digestion of plasmid DNA with exonuclease III (Pharmacia) was carried out as described by Henikoff (1984), in order to generate a series of nested deletion

clones for sequencing. Essentially, 5 μg of purified plasmid DNA (sufficient for about 10 individual aliquots) was used for the reaction.

O. DNA SEQUENCING

1. Single-stranded DNA sequencing

M13mp18 and M13mp19 clones containing single-stranded DNA were sequenced by the dideoxy chain termination method of Sanger et al. (1977) using T7 DNA polymerase (Sequenase®, US Biochemical). DNA (1-2 μg) was heated with 1 pmol of oligonucleotide sequencing primer in 10 μl of Sequenase buffer (40 mM Tris-HCl, 20 mM $MgCl_2$, 50 mM NaCl) at 65°C for 2 minutes. Following a 15 minute annealing period at room temperature, DTT was added to 7 mM, [α-^{32}P] dATP to 10 μCi, and labeling nucleotide mix (dGTP, dCTP, dTTP) to 0.1 mM of each nucleotide. The labeling reaction was allowed to proceed for 5 minutes at room temperature after addition of 3 units of Sequenase®. Aliquots were then removed and placed into tubes marked G, A, T, or C pre-warmed to 42°C and each containing the termination mix (80 μM dGTP, dCTP, dTTP, dATP, and 50 mM NaCl) in addition to 8 μM ddGTP, 8 μM ddATP, 8 μM ddTTP, 8 μM ddCTP, respectively. After 5 minutes of primer extension at 42°C, formamide dye mix (which contained 0.05% bromophenol blue and xylene cyanol dyes) was added to final concentration of 20% (v/v). Prior to loading onto 6% or 8% polyacrylamide slab gels, samples were boiled for 3 minutes and immediately placed on ice.

2. Double-stranded DNA sequencing

All double-stranded plasmids sequenced were isolated from bacterial hosts deficient in recombination (recA1) and nuclease (endA) activities (e.g. DH5α -see Appendix B).

DNA sequencing was by the dideoxy method of Sanger et al. (1977) as modified for double-stranded plasmids by Hattori and Sakaki (1986). Plasmid DNA (3 µg) was treated with 10 µg/ml RNaseA at 37°C for 30 minutes, mixed with 30 µl of 2.5 M NaCl-20% PEG 6000, and kept on ice for at least one hour. The sample was centrifuged (12,000 x g for 5 minutes at 4°C) and the resultant pellet (consisting of covalently closed plasmid DNA) was rinsed with 70% ethanol, air-dried, and redissolved in 18 µl TE. This procedure removed contaminating RNA and nicked plasmid DNA. To denature the DNA for sequencing, freshly prepared NaOH was added to a final concentration of 0.2 M, and the mixture was kept at room temperature for 5 minutes. Ammonium acetate (pH 5.2) was added to final concentration of 0.5 M, and the DNA was precipitated with 95% ethanol and redissolved in 10 µl of Sequenase buffer. Annealing to oligonucleotide sequencing primer, and primer extension reactions were as previously described for the single-stranded templates.

P. NUCLEASE S1 ANALYSIS

Nuclease S1 protection analysis was performed as described by Berk and Sharp (1977). Single-stranded DNA probes prepared by primer extension of M13 templates (as described previously) were purified on a 4% denaturing acrylamide-urea gels, and recovered from the gels in 0.5 x TBE buffer containing 0.05% SDS. The purified probe (approximately 6 x 10^6 cpm) was mixed with 6 µg of nematode polyA+RNA in 30 µl of hybridization buffer (60% formamide, 0.4 M NaCl, 40 mM PIPES pH 6.4, 1 mM EDTA, 3 µl of 0.25% polyacrylamide) containing 1 U RNasin (Promega). The solution was heated to 70°C for 15 minutes and immediately transferred to 50°C and incubated overnight. Hybridization was terminated by the addition of 170 µl of ice-cold nuclease S1 buffer (0.28 M NaCl, 50 mM sodium acetate pH 4.6, 4.5 mM $ZnSO_4$) containing 150 units of nuclease S1 (Pharmacia). The mixture was incubated at 37°C for one hour and 40 µl of 4 M ammonium acetate-0.1 M

EDTA solution was added to terminate the nuclease digestion. Nuclease S1 protected DNA fragments were precipitated with one volume of 2-propanol using 100 µg/ml of E. coli tRNA as carrier. Samples were dissolved directly in 5 µl formamide dye mix and analyzed on a thin 6% polyacrylamide gel containing 8 M urea.

III. RESULTS

A. ISOLATION OF EFT-1 *GENE*

Figure 3 shows the physical map of the region around ubq-1 which maps to the centre of chromosome III of C. elegans. The five overlapping cosmids (ZK48, C16A7, ZK331, C41E11, and T01D10) which cover approximately 150 kb in this region are shown. In order to define transcript coding genes contained in this region, HindIII digests of the cosmids (except C16A7 which was not yet available at the time of this study) were hybridized with labeled first strand cDNA probes synthesized from embryo polyA$^+$RNA. Three fragments of 1.2, 3.3 and 6.6 kb in size from cosmid ZK331 which hybridized with the probes (Fig. 4A, lanes 2 and 3) were isolated. The restriction fragments of 0.9, 1.0 and 2.3 kb in size from cosmid ZK48 (lane 1), which contain the ubq-1 gene (ubiA, Graham et al., 1989) were also detected whereas none of the other cosmids appeared to contain genes expressed at detectable levels in embryos (Fig. 4A, lanes 4-7). Restriction mapping of ZK331 (Fig. 7) and DNA sequence analyses revealed that the 3.3 kb fragment (B3255) and the 6.6 kb fragment from ZK331 are separated by a 204 bp HindIII fragment, and that these three fragments are contained within the same gene, hereafter referred to as eft-1. The 204 bp band was not detected with the cDNA probe possibly because of its relatively small size. The four fragments from ZK331 and their subfragments (separately inserted into pGEM3Z), as well as a series of nested deletion clones derived from B3255 were further characterized by restriction mapping and sequence analysis. The 1.2 kb fragment proved to contain repetitive sequences when analyzed by genomic Southern blotting (Fig. 19B) and its identity is unknown.

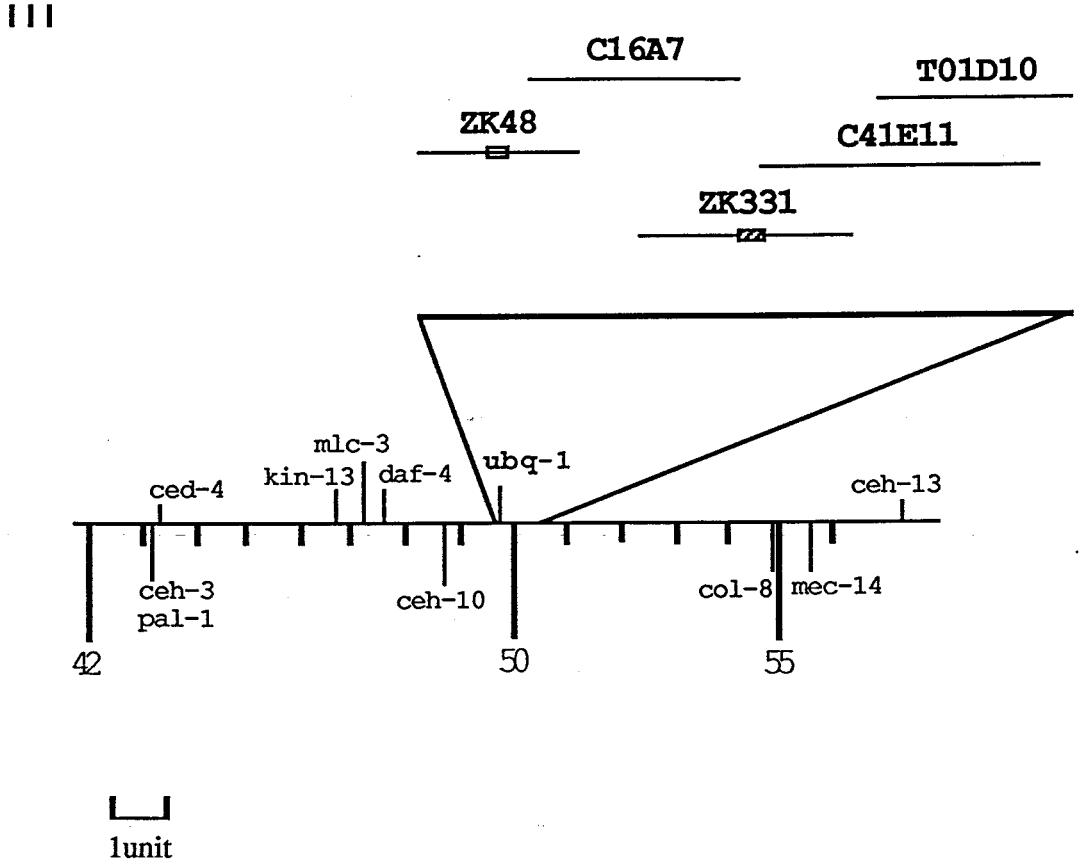


Fig. 3. Physical map of the region around ubq-1.

Partial map of chromosome III of C.elegans showing ubq-1 and flanking markers. The five overlapping cosmid clones (ZK48, C16A7, ZK331, C41E11, and T01D10) covering about 150 kb (~1 unit) of the region around ubq-1 are shown at the top. The open box in ZK48 represents the location of ubq-1 and the hatched box in ZK331 shows the position of eft-1.

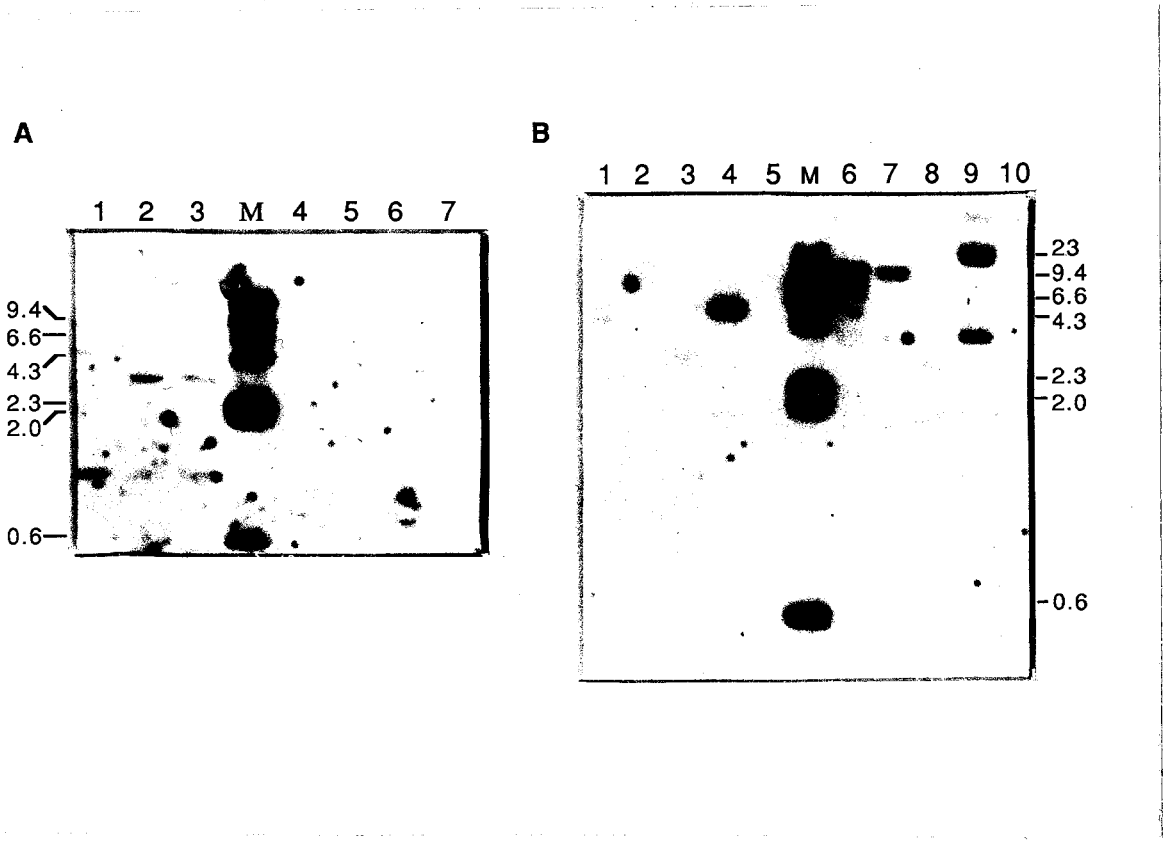


Fig.4. Cosmid genomic Southern blot analysis.

A. DNA (500 ng) from cosmid ZK48 (lane 1), ZK331 (lanes 2 and 3), C41E11 (lanes 4 to 6), and T01D10 (lane 7) was digested with HindIII for 2 hours, resolved by electrophoresis on a 0.7% agarose gel, blotted onto nitrocellulose membrane, and hybridized with a cDNA probe made from C. elegans embryo polyA+RNA. **B.** DNA (1 μg) from cosmid ZK48, C16A7, ZK331, C41E11, and T01D10 was digested with HindIII (lanes 1 to 5, respectively) or PstI (lanes 6 to 10, respectively) overnight, separated on 0.8% agarose gel, blotted onto nylon membrane, and hybridized with labeled cDNA probes synthesized from pooled polyA+RNA selected from embryo, larvae, and young adult nematode populations. Southern transfers and hybridization conditions were as stated in "Experimental Procedures". M, size markers in kilobase pairs.

B. DETECTION OF OTHER MESSENGER RNA CODING FRAGMENTS

Although no restriction fragments of the cosmids, C41E11, and T01D10 were detected when analyzed with cDNA probes made from embryo RNA, it was possible that these cosmids could contain genes which are expressed at detectable levels at other stages of nematode development. To investigate this possibility, HindIII and PstI-digested fragments of all 5 cosmids were hybridized with cDNA probes prepared from pooled polyA$^+$RNA obtained from all the major (embryo, L1-L4, dauer and young adult) stages of C. elegans development. As shown in Figure 4B, no additional fragments of ZK331 (lane 3) were detected, whereas a 5.3 kb HindIII (lane 4) and 3.8 kb PstI (lane 9) fragments of C41E11, and a 15 kb PstI fragment of C16A7 (lane 7) hybridized strongly with the probes. Also, two HindIII fragments of 5 kb and 7 kb (lane 1) and a 9.5 kb PstI fragment (lane 6) from ZK48 were detected in addition to the fragments including a 15 kb PstI fragment contained in ubq-1. The positions of these fragments are represented with shaded boxes on the contig restriction map (Fig. 7).

C. RESTRICTION MAPPING OF CONTIG

1. Terminase activity assay

The plasmid pCM101 contained in E. coli strain AZ1935 (kindly provided by H. Murialdo of University of Toronto) which expresses large amounts of λDNA terminase enzyme upon temperature induction (Murialdo et al., 1987), was used to obtain a crude preparation of this enzyme that cleaves the cos site of bacteriophage λ DNA during packaging. To assay the cos-cleavage activity of the enzyme preparation, cosmids C16A7.3 and C41E5.5 (Fig. 5A) were linearized with PstI and incubated with diluted and undiluted

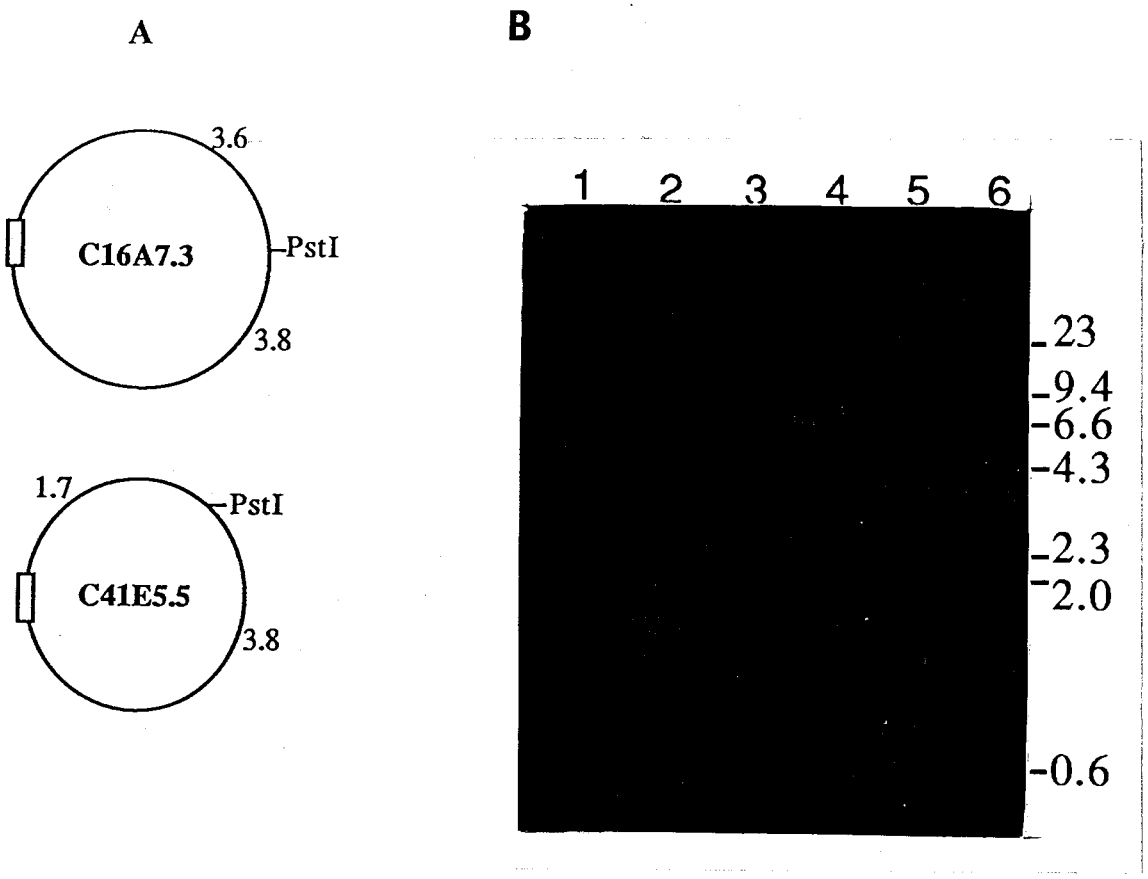


Fig. 5. Terminase activity assay.

A. Schematic representation of cosmids C16A7.3 and C41E5.5 (both derived from pJB8; see section B1, p. 37) which served as the substrates for the terminase assay. The box represents the cos site and the unique PstI site is shown. **B.** Ethidium bromide-stained gel of fragments derived from C41E5.5 (lanes 1-3) and C16A7.3 (lanes 4-6) after PstI digestion and terminase cleavage of cos site. DNA (1 μg) was digested with PstI and equal aliquots were cleaved with either 3 x diluted (lanes 3 and 6) or undiluted terminase preparation (lanes 2 and 5). The samples were resolved on a 0.5% agarose gel. Sizes of fragments generated by terminase cleavage after PstI digestion are given in kilobase pairs on the diagrams. Aliquots of DNA (lane 1, C41E5.5 and lane 4, C16A7.3) digested with PstI alone were included to serve as standards.

crude extracts of the terminase. Because cos was not equidistant from the unique PstI site, cleavage of the cosmids by the terminase generated fragments of different sizes. The resultant fragments were separated by agarose gel elctrophoresis and visualized by staining with ethidium bromide. The amount of DNA in each band was estimated by comparison with known concentrations of HindIII-digested λDNA. Figure 5B shows that roughly 60% of the cosmid DNA was cleaved by the undiluted terminase preparation (lanes 2 and 5). Little, if any, decrease in cleavage efficiency was observed with a three-fold dilution of the enzyme-containing extract (lanes 3 and 6). Chow et al. (1987) reported an 80% efficiency in conditions in which the crude enzyme:substrate molar ratio was about 1000:1, and suggested that at least 20% of the terminase molecules in the extracts are inactive for cleavage but still bind tightly (practically irreversibly) to the cos site.

2. Restriction map of the overlapping cosmids

Because of their relatively large sizes (40 to 50 kb), an attempt to map individual cosmid clones by multiple endonuclease digestions proved unsuccessful, since a large number of fragments was generated. The alternative approach used was to cleave each cosmid at the unique cos site using a crude preparation of λDNA terminase, before or after partial restriction enzyme digestion. The right (R) or left (L) cohesive ends thus generated were labeled, in separate reactions, by annealing with radioactive dodecameric oligonucleotides complementary to the sequence of the right or left end. As an example, the autoradiograms of electrophoretic ladders of restriction fragments, all with a common end, generated by partial digestion of cosmid C41E11 with HindIII, EcoRI, PstI, and XhoI, and cosmid ZK331 with PstI restriction enzymes are shown in Figure 6. Similarly, autoradiograms for all five cosmids were obtained following partial digestion with HindIII, EcoRI, PstI, or XhoI, and annealing of terminase-generated ends to labeled R or L

Fig. 6. Autoradiograms of gels showing partial digestion patterns of cosmids C41E11 and ZK331.

Cosmids (**A**, ZK331; **B**, C41E11) were partially digested with appropriate restriction enzyme. After cleavage of the cos site with a crude extract of λ DNA terminase, the generated cohesive ends were annealed to γ^{32}P-end-labeled oligonucleotides with complementary sequences. DNA fragments were resolved on a 0.4% agarose gel at 1.5 V/cm for 24 hours. Gels were transferred onto Whatman cellulose paper and dried at 60°C for 45 minutes. Autoradiography was carried out with Kodak X-Omat AR film and intensifying screen. LE, LH, LX, LP (or RE, RH, RX, RP), Left (or Right) cos end-labeled fragments after partial digestion with EcoRI, HindIII, XhoI, and PstI, respectively. Size markers are given in kilobase pairs.

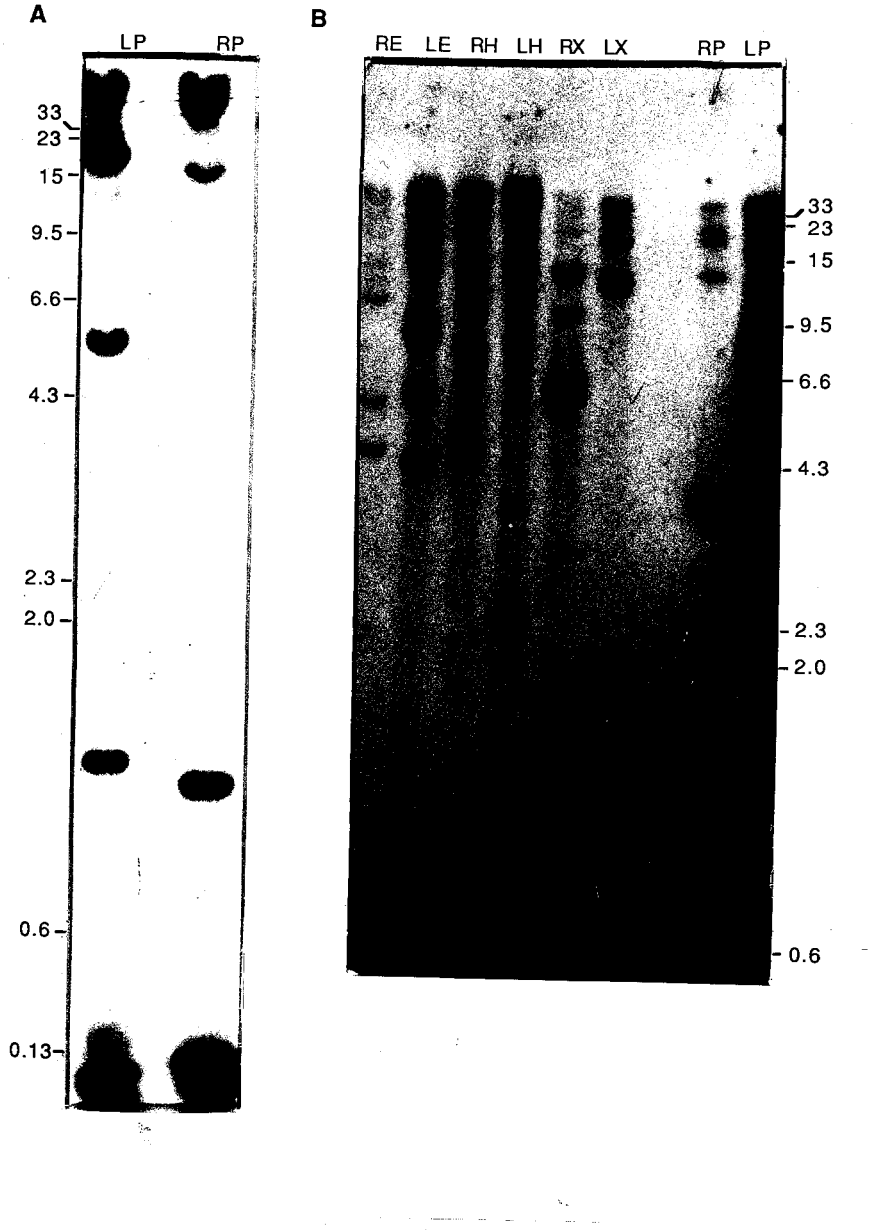
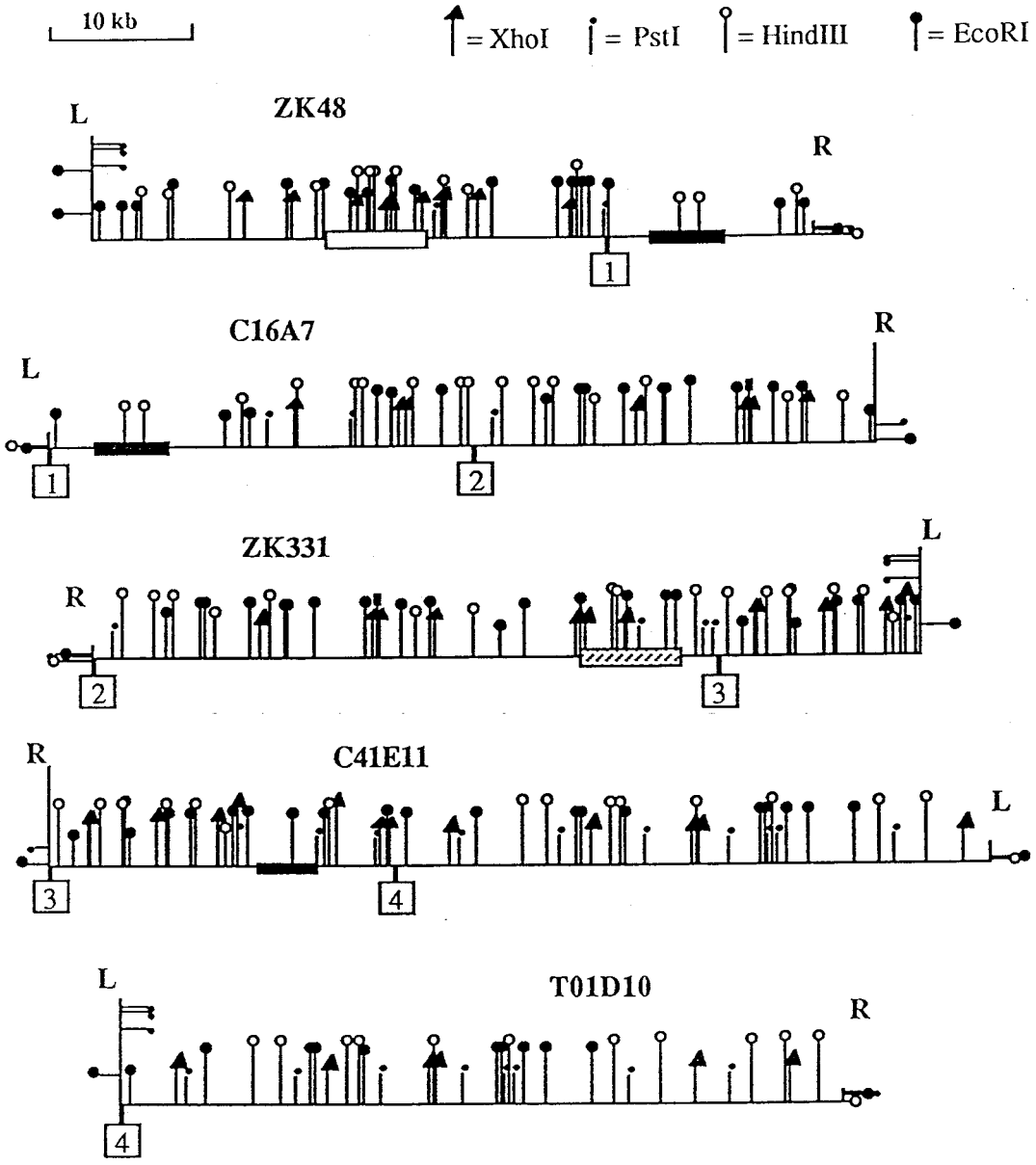

A
LP
RP
33
23
15
9.5
6.6
4.3
2.3
2.0
0.6
0.13
B
RE
LE
RH
LH
RX
LX
RP
LP
33
23
15
9.5
6.6
4.3
2.3
2.0
0.6

Fig. 7. Restriction map of the region around ubq-1 on chromosome III. Restriction map of five overlapping cosmids ZK48, C16A7, ZK331, C41E11, and T01D10 which span about 150 kb around the ubq-1 locus (see also Figure 3). The map was derived from autoradiograms of gels of partial restriction enzyme patterns of cosmids after cos cleavage with terminase and hybridization of generated cohesive ends with end-labeled complementary oligonucleotides. The beginning of overlap between cosmids is represented with successively numbered boxes. The vector DNA is shown at the ends of the linearized cosmids as a vertical line marked "L" or "R" for left or right cohesive ends, respectively. Open and stippled boxes on ZK48 and ZK331 represent the ubq-1 and eft-1 loci, respectively, and bold horizontal lines represent other putative coding regions. The coding regions were determined by hybridization of restriction fragments with polyA+RNA-derived first strand cDNA probes. L, left cohesive end; R, right cohesive end.


10 kb
= XhoI
= PstI
= HindIII
= EcoRI
ZK48
L
R
1
C16A7
L
R
1
2
ZK331
R
L
2
3
C41E11
R
L
3
4
T01D10
L
R
4

complement oligonucleotide. The positions of restriction sites as deduced from the autoradiograms are represented in Figure 7. The cosmids altogether cover about 150 kb in this region around ubq-1. Also, results of the restriction mapping of the cosmids revealed that eft-1 mapped approximately 50 kb from ubq-1, in a region in ZK331 that did not overlap with any other cosmid.

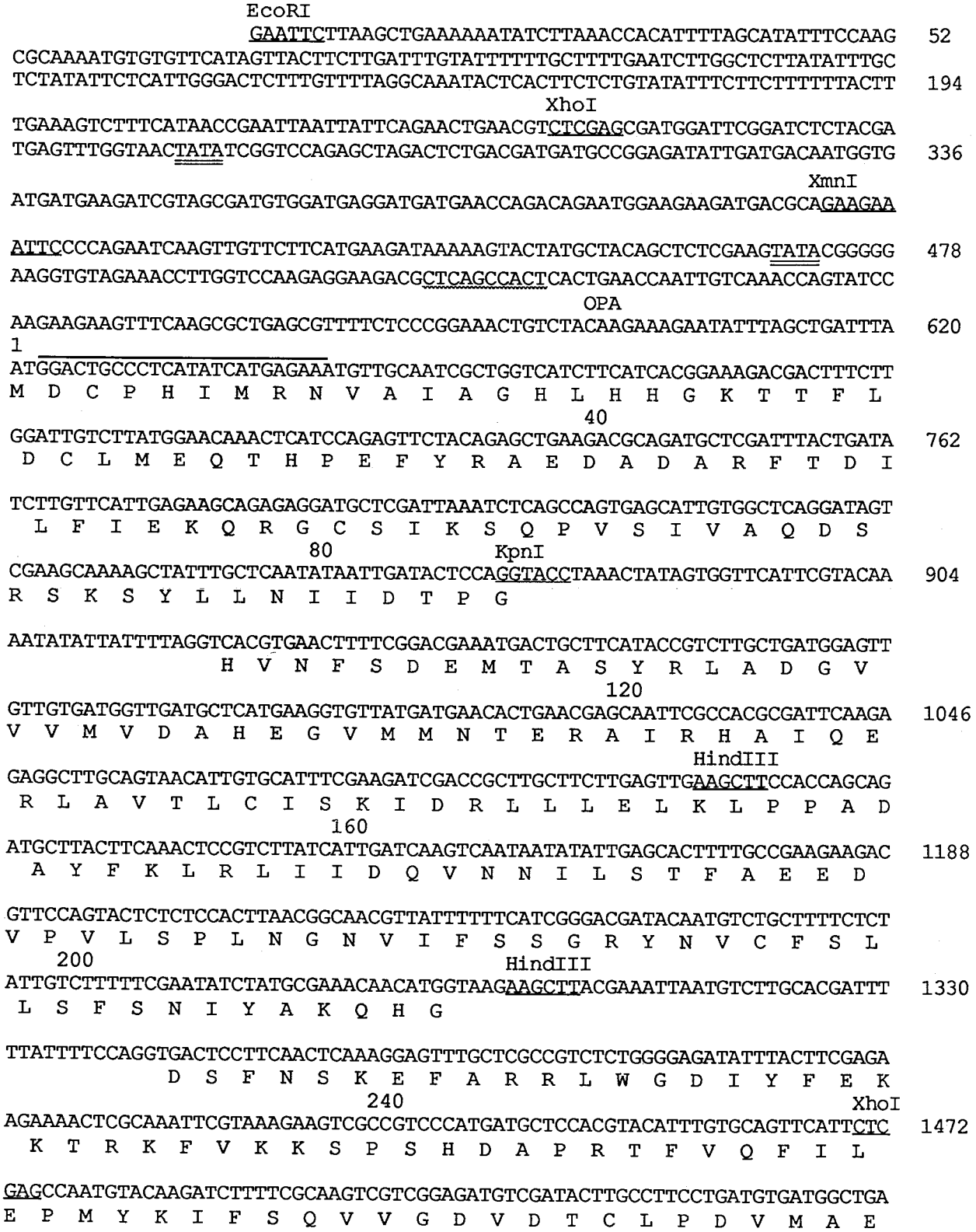
D. ANALYSIS OF EFT-1 GENE SEQUENCE

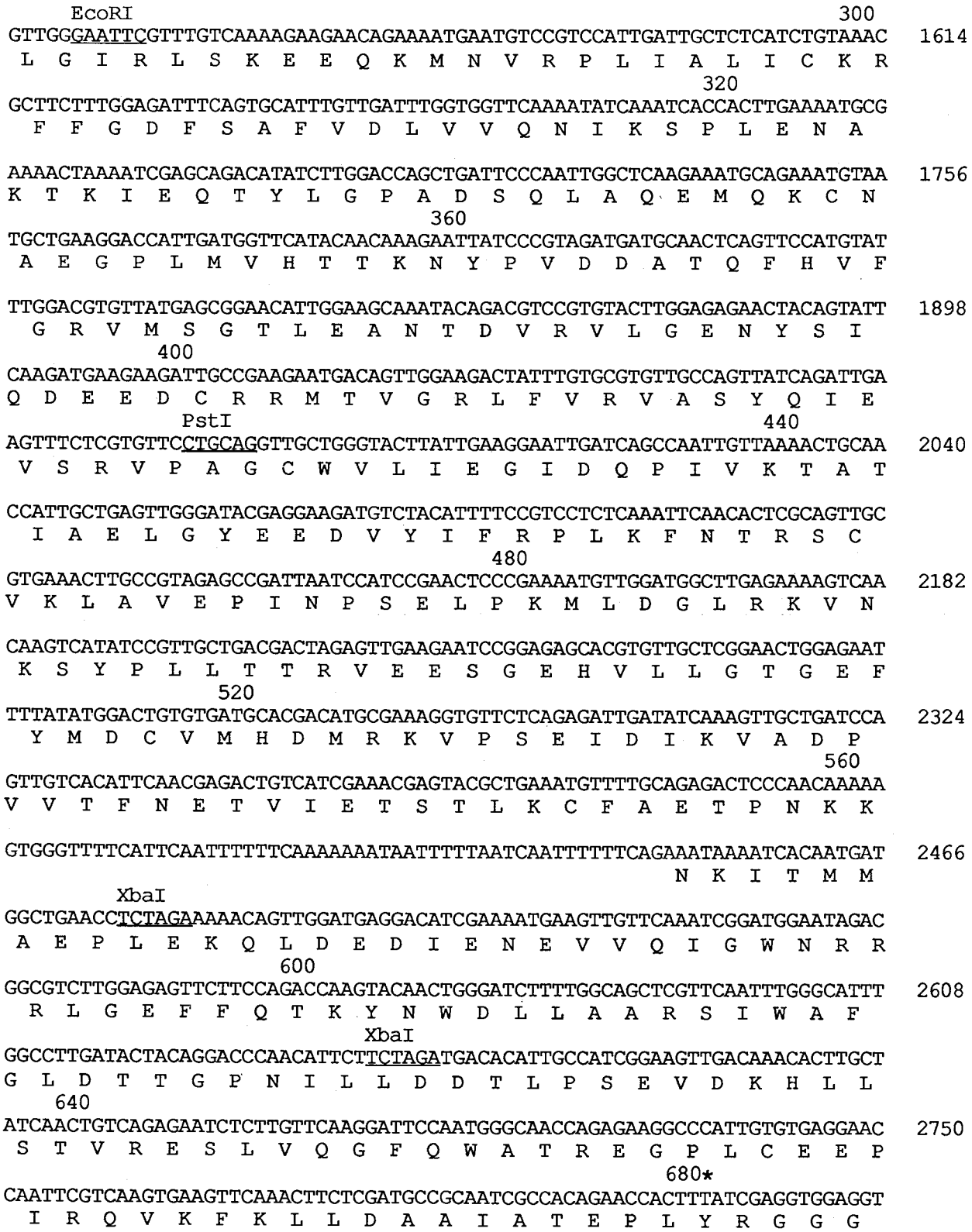
Figure 8 shows the combined nucleotide and derived amino acid sequences of eft-1 (elongation factor 2 gene 1). The 3.8 kb DNA sequence revealed a 620 bp 5'-flanking region followed by a putative translation initiation codon ATG (amino acid position 1), a 2547 bp open reading frame (ORF) interrupted by four short introns of 46-75 bases (based on sequence comparison with the mammalian and Drosophila EF-2 sequences), and a 420 bp 3'-flanking region. The initiation codon was taken to be the first ATG codon (nucleotides 622 - 624) appearing downstream of the in-frame nonsense codon TGA (nucleotides 525-527). An analysis of the sequence upstream of the putative initiation codon revealed the sequence CTCAGCCACT 100 bp away, which resembles the CAP site consensus sequence of Corden et al. (1980). A potential TATA box is seen 40 nucleotides upstream of this sequence. No other promoter-like sequences were found in the 5' flanking region. The four introns of eft-1 have exon-intron boundary sequences which are completely consistent with the consensus sequences for 5' and 3' splice sites (Mount, 1982; Blumenthal and Thomas, 1988; Frendewey and Keller, 1985). Intron 4 separates codons Ser-751 and Ala-752, while introns 1, 2 and 3 interrupt the codons of Gly-85, Gly-210 and Lys-561, respectively.

E. ISOLATION AND ANALYSIS OF A CDNA CLONE ENCODING EFT-1

In order to confirm the positions of the splice sites and determine the poly(A)addition

Fig. 8. Complete nucleotide and deduced amino acid sequences of eft-1. Numbers at right and within the sequence indicate the positions of nucleotide and amino acid residues from the beginning of the sequence and the initiator methionine, respectively. The presumptive CAP site (Corden et al., 1980) (wavy underline), TATA sequences (double underline), and restriction enzyme sites (single underline) are shown. The sequence (positions 624 to 643) complementary to the oligonucleotide E04 (see Appendix A) which was used for the S1 nuclease protection analysis is shown with an overline. The asterisk at amino acid 681 denotes the tyrosyl residue at a position corresponding to the ADP-ribosylatable histidyl residue in EF-2. The beginning of the poly(A) tail of clone pEF1.35 is marked by an up-arrow (after nucleotide 3502) and the dots identify the corresponding potential poly(A) addition signal(s).


EcoRI
GAATTCTTAAGCTGAAAAAATATCTTAAACCACATTTTAGCATATTTCCAAG 52
CGCAAAATGTGTGTTCATAGTTACTTCTTGATTTGTATTTTTTGCTTTTGAATCTTGGCTCTTATATTTGC
TCTATATTCTCATTGGGACTCTTTGTTTTAGGCAAATACTCACTTCTCTGTATATTTCTTCTTTTTTACTT 194
XhoI
TGAAAGTCTTTCATAACCGAATTAATTATTCAGAACTGAACGTCTCGAGCGATGGATTCGGATCTCTACGA
TGAGTTTGGTAACTATATCGGTCCAGAGCTAGACTCTGACGATGATGCCGGAGATATTGATGACAATGGTG 336
XmnI
ATGATGAAGATCGTAGCGATGTGGATGAGGATGATGAACCAGACAGAATGGAAGAAGATGACGCAGAAGAA
ATTCCCCAGAATCAAGTTGTTCTTCATGAAGATAAAAAGTACTATGCTACAGCTCTCGAAGTATACGGGGG 478
AAGGTGTAGAAACCTTGGTCCAAGAGGAAGACGCTCAGCCACTCACTGAACCAATTGTCAAACCAGTATCC
OPA
AAGAAGAAGTTTCAAGCGCTGAGCGTTTTCTCCCGGAAACTGTCTACAAGAAAGAATATTTAGCTGATTTA 620
1
ATGGACTGCCCTCATATCATGAGAAATGTTGCAATCGCTGGTCATCTTCATCACGGAAAGACGACTTTCTT
M D C P H I M R N V A I A G H L H H G K T T F L
40
GGATTGTCTTATGGAACAAACTCATCCAGAGTTCTACAGAGCTGAAGACGCAGATGCTCGATTTACTGATA 762
D C L M E Q T H P E F Y R A E D A D A R F T D I
TCTTGTTCATTGAGAAGCAGAGAGGATGCTCGATTAAATCTCAGCCAGTGAGCATTGTGGCTCAGGATAGT
L F I E K Q R G C S I K S Q P V S I V A Q D S
80
KpnI
CGAAGCAAAAGCTATTTGCTCAATATAATTGATACTCCAGGTACCTAAACTATAGTGGTTCATTCGTACAA 904
R S K S Y L L N I I D T P G
AATATATTATTTTAGGTCACGTGAACTTTTCGGACGAAATGACTGCTTCATACCGTCTTGCTGATGGAGTT
H V N F S D E M T A S Y R L A D G V
120
GTTGTGATGGTTGATGCTCATGAAGGTGTTATGATGAACACTGAACGAGCAATTCGCCACGCGATTCAAGA 1046
V V M V D A H E G V M M N T E R A I R H A I Q E
HindIII
GAGGCTTGCAGTAACATTGTGCATTTCGAAGATCGACCGCTTGCTTCTTGAGTTGAAGCTTCCACCAGCAG
R L A V T L C I S K I D R L L L E L K L P P A D
160
ATGCTTACTTCAAACTCCGTCTTATCATTGATCAAGTCAATAATATATTGAGCACTTTTGCCGAAGAAGAC 1188
A Y F K L R L I I D Q V N N I L S T F A E E D
GTTCCAGTACTCTCTCCACTTAACGGCAACGTTATTTTTTCATCGGGACGATACAATGTCTGCTTTTCTCT
V P V L S P L N G N V I F S S G R Y N V C F S L
200
HindIII
ATTGTCTTTTTCGAATATCTATGCGAAACAACATGGTAAGAAGCTTACGAAATTAATGTCTTGCACGATTT 1330
L S F S N I Y A K Q H G
TTATTTTCCAGGTGACTCCTTCAACTCAAAGGAGTTTGCTCGCCGTCTCTGGGGAGATATTTACTTCGAGA
D S F N S K E F A R R L W G D I Y F E K
240
XhoI
AGAAAACTCGCAAATTCGTAAAGAAGTCGCCGTCCCATGATGCTCCACGTACATTTGTGCAGTTCATTCTC 1472
K T R K F V K K S P S H D A P R T F V Q F I L
GAGCCAATGTACAAGATCTTTTCGCAAGTCGTCGGAGATGTCGATACTTGCCTTCCTGATGTGATGGCTGA
E P M Y K I F S Q V V G D V D T C L P D V M A E


EcoRI
300
GTTGGGAATTCGTTTGTCAAAAGAAGAACAGAAAATGAATGTCCGTCCATTGATTGCTCTCATCTGTAAAC 1614
L G I R L S K E E Q K M N V R P L I A L I C K R
320
GCTTCTTTGGAGATTTCAGTGCATTTGTTGATTTGGTGGTTCAAAATATCAAATCACCACTTGAAAATGCG
F F G D F S A F V D L V V Q N I K S P L E N A
AAAACTAAAATCGAGCAGACATATCTTGGACCAGCTGATTCCCAATTGGCTCAAGAAATGCAGAAATGTAA 1756
K T K I E Q T Y L G P A D S Q L A Q E M Q K C N
360
TGCTGAAGGACCATTGATGGTTCATACAACAAAGAATTATCCCGTAGATGATGCAACTCAGTTCCATGTAT
A E G P L M V H T T K N Y P V D D A T Q F H V F
TTGGACGTGTTATGAGCGGAACATTGGAAGCAAATACAGACGTCCGTGTACTTGGAGAGAACTACAGTATT 1898
G R V M S G T L E A N T D V R V L G E N Y S I
400
CAAGATGAAGAAGATTGCCGAAGAATGACAGTTGGAAGACTATTTGTGCGTGTTGCCAGTTATCAGATTGA
Q D E E D C R R M T V G R L F V R V A S Y Q I E
PstI
440
AGTTTCTCGTGTTCCTGCAGGTTGCTGGGTACTTATTGAAGGAATTGATCAGCCAATTGTTAAAACTGCAA 2040
V S R V P A G C W V L I E G I D Q P I V K T A T
CCATTGCTGAGTTGGGATACGAGGAAGATGTCTACATTTTCCGTCCTCTCAAATTCAACACTCGCAGTTGC
I A E L G Y E E D V Y I F R P L K F N T R S C
480
GTGAAACTTGCCGTAGAGCCGATTAATCCATCCGAACTCCCGAAAATGTTGGATGGCTTGAGAAAAGTCAA 2182
V K L A V E P I N P S E L P K M L D G L R K V N
CAAGTCATATCCGTTGCTGACGACTAGAGTTGAAGAATCCGGAGAGCACGTGTTGCTCGGAACTGGAGAAT
K S Y P L L T T R V E E S G E H V L L G T G E F
520
TTTATATGGACTGTGTGATGCACGACATGCGAAAGGTGTTCTCAGAGATTGATATCAAAGTTGCTGATCCA 2324
Y M D C V M H D M R K V P S E I D I K V A D P
560
GTTGTCACATTCAACGAGACTGTCATCGAAACGAGTACGCTGAAATGTTTTGCAGAGACTCCCAACAAAAA
V V T F N E T V I E T S T L K C F A E T P N K K
GTGGGTTTTCATTCAATTTTTTCAAAAAAATAATTTTTAATCAATTTTTTCAGAAATAAAATCACAATGAT 2466
N K I T M M
XbaI
GGCTGAACCTCTAGAAAAACAGTTGGATGAGGACATCGAAAATGAAGTTGTTCAAATCGGATGGAATAGAC
A E P L E K Q L D E D I E N E V V Q I G W N R R
600
GGCGTCTTGGAGAGTTCTTCCAGACCAAGTACAACTGGGATCTTTTGGCAGCTCGTTCAATTTGGGCATTT 2608
R L G E F F Q T K Y N W D L L A A R S I W A F
XbaI
GGCCTTGATACTACAGGACCCAACATTCTTCTAGATGACACATTGCCATCGGAAGTTGACAAACACTTGCT
G L D T T G P N I L L D D T L P S E V D K H L L
640
ATCAACTGTCAGAGAATCTCTTGTTCAAGGATTCCAATGGGCAACCAGAGAAGGCCCATTGTGTGAGGAAC 2750
S T V R E S L V Q G F Q W A T R E G P L C E E P
680*
CAATTCGTCAAGTGAAGTTCAAACTTCTCGATGCCGCAATCGCCACAGAACCACTTTATCGAGGTGGAGGT
I R Q V K F K L L D A A I A T E P L Y R G G G

```
CAGATGATCCCAACTGCACGCCGATGTGCCTATTCTGCATTCCTTATGGCCACGCCAAGATTAATGGAGCC   2892
Q  M  I  P  T  A  R  R  C  A  Y  S  A  F  L  M  A  T  P  R  L  M  E  P
                                720
ATACTATACAGTGGAAGTTGTTGCACCAGCTGATTGTGTAGCTGCTGTATATACAGTGTTAGCTAAACGAC
 Y  Y  T  V  E  V  V  A  P  A  D  C  V  A  A  V  Y  T  V  L  A  K  R  R

GTGGTCACGTCACCACTGATGCACCAATGCCAGGATCACCTATGTACACTATCAGCGTGAGTTTAAATATA   3034
  G  H  V  T  T  D  A  P  M  P  G  S  P  M  Y  T  I  S

ATGTTAAAAGAGAACACACATTAAAATAGTTTTATGATATTCTAAAATTATGATTTTCAGGCGTACATTCC
                                                              A  Y  I  P
              760                      EcoRI
AGTAATGGACTCGTTCGGATTCGAAACTGATCTTCGAATTCACACACAGGGACAAGCATTCTGTATGTCTG   3176
 V  M  D  S  F  G  F  E  T  D  L  R  I  H  T  Q  G  Q  A  F  C  M  S  A
                                                           800
CTTTCCATCATTGGCAACTCGTACCAGGAGATCCACTCGATAAATCCATTGTTATCAAGACTCTCGACGTC
  F  H  H  W  Q  L  V  P  G  D  P  L  D  K  S  I  V  I  K  T  L  D  V

CAGCCAACTCCACATCTCGCCAGAGAGTTCATGATCAAGACCAGAAGACGCAAGGGTTTGTCTGAAGATGT   3318
Q  P  T  P  H  L  A  R  E  F  M  I  K  T  R  R  R  K  G  L  S  E  D  V
                                840                                849
CTCCGTTAACAAGTTCTTCGATGATCCGATGCTTCTCGAACTTGCAAAGCAGCAAGATTATACTGGATTC
 S  V  N  K  F  F  D  D  P  M  L  L  E  L  A  K  Q  Q  D  Y  T  G  F

TAAATTGCGATTTCCTAACTTATTTTCTCATTGTTTCACACTTTTTATCTTTCAGTGTTCTTTCCCGTTTC   3459
OCH

ATTGAGTTGTTTGATCACAATAGTAAAATGAAAATTCTTGTTTATCTATCTATCAAATGATATTCAGAGCA
                          ••••••••••••
            EcoRI
GGGACCAAAATTGAATTCTCTTAGTTTCGTCAAACAAGCTGAATGTCCGTTGAAGGTATTACTCTTGGACA   3601
ACTTTTTTGAAAAAAAAAAAACTGAAGATCAACATAAGCCTACGCTTAATTTTAAGTCTAAGCATAAGCTA
TAACCGACTTGGAGGCCCCACGAAAAGGGGAGCAGAACGAAAAGGGGATCTGCAAAAAGGGGATCTGCGAA   3742
AAGGGGAGCAACGAAAAGGGGAGCTGGCACTGTGCAAACGGACAAAACGCATTTTCTCACGCAGCGCACCG
TTTG                                                                       3818
```

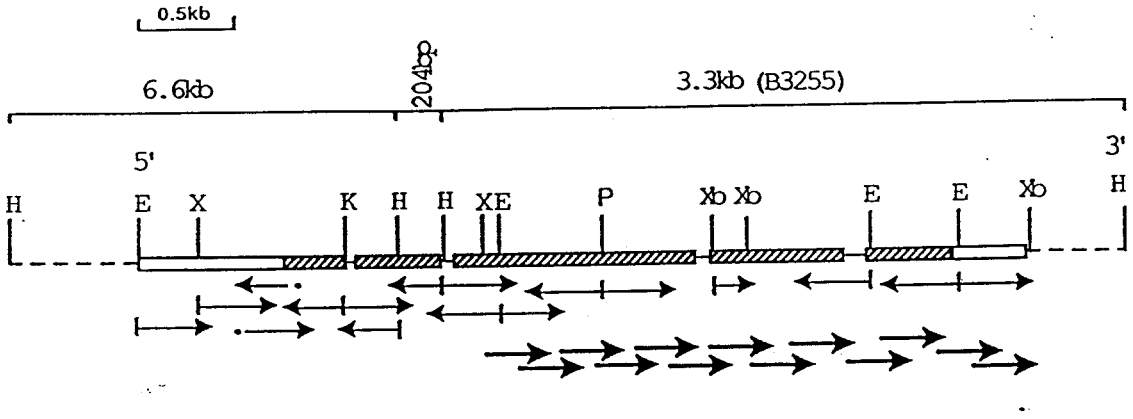


Fig. 9. Restriction map of the eft-1 locus.

Restriction map of eft-1 and the strategy used for nucleotide sequencing are represented. The three Hind III fragments of cosmid ZK331 containing the eft-1 locus are shown at the top. Stippled boxes represent exons and adjoining lines indicate introns. The open boxes represent 5'- (left) and 3'- (right) untranslated regions. The horizontal arrows indicate the direction and extent of the sequence obtained from the restriction sites, or from synthesized primers (•). Bold arrows show sequences obtained from nested deletion clones of B3255 (see "Experimental Procedures"). The cDNA clone pEF1.35, with a 1.35 kb insert encoding the 3' half of eft-1, beginning at nucleotide 2115 (Fig. 8), was also sequenced on both strands. Restriction sites are: H, HindIII; E, EcoRI; X, XhoI; Xb, XbaI; P, PstI; K, KpnI. kb denotes kilobase pairs.

and 5' cap sites for eft-1 mRNA, it was necessary to isolate cDNA clones encoding this transcript. When a total of 4700 transformants of a C. elegans cDNA library in bacteriophage λZAP vector was screened with nick-translated B3255, only one positive clone was isolated. Nucleotide sequencing confirmed that the clone pEF1.35, with a 1.3 kb insert, encoded the 3' half of eft-1, beginning at nucleotide 2115 (Fig. 8). In an attempt to isolate more and possibly full-length cDNA clones, the cDNA insert of pEF1.35 was nick translated and used as a probe to screen approximately 38,000 phage of the λZAP library, but no positive clones were isolated. Sequence analysis of the cDNA clone pEF1.35 confirmed the position of introns 3 and 4, and revealed the poly(A) addition site between nucleotides 3502 and 3503 (111 bp after the TAA termination codon). The sequences AGTAAA and ATGAAA were located 17 and 11 nucleotides upstream of the poly(A)tail, respectively, and may serve as the poly(A) addition signal(s) for eft-1 (Park et al., 1986). The high degree of similarity of the eft-1-encoded polypeptide to eukaryotic EF-2 allowed identification of the 5'-most exons which were not covered by the cDNA sequence. That the codon usage in these putative exons is similar to that from other unequivocal exons provides corroborative evidence that these are in fact, correctly identified.

A detailed restriction map of the eft-1 locus and the restriction fragments and deletion clones that were sequenced are shown in Figure 9.

F. NUCLEASE S1 PROTECTION ANALYSIS

In an attempt to locate the 5' end of eft-1 mRNA within the genomic sequence, nuclease S1 protection analysis was performed using a single stranded probe synthesized by primer extension of clone Mp18-810 which contained the 864 bp XhoI-HindIII fragment in the 5' region of the eft-1 sequence (Fig. 8). The oligonucleotide EO4 was used as the extension primer such that when the product was digested with HindIII and EcoRI a 407 bp

fragment including the putative initiation ATG and about 400 nucleotides of 5'-flanking region was generated. The ^{32}P-labeled 407 bp fragment was hybridized to nematode embryo polyA+RNA and the unprotected region of the fragment was digested with S1 nuclease. Analysis on a 6% polyacrylamide-8 M urea gel yielded an ambigous result in which multiple bands that were resistant to S1 nuclease digestion were detected. An attempt to investigate further the possibility of eft-1 mRNA initiation at multiple sites by reverse transcription primer extension mapping proved unsuccessful due to premature termination of polymerization by reverse transcriptase as a result of secondary structures in the 5' region of the mRNA.

G. PRIMARY STRUCTURE OF EFT-1

The 2547 bp ORF encodes a protein (EFT-1) of 849 amino acid residues (Figs. 8 and 17) with a calculated molecular weight of 96,151. A comparison between EFT-1 and related sequences using the FASTA computer program of Pearson and Lipman (1988) revealed an overall sequence identity of 38% (56% at the nucleotide level) between EFT-1 and the hamster or Drosophila EF-2 sequences. As shown in Figure 11, the regions of sequence similarity (i.e. including conservative changes) were dispersed throughout the length of the primary structure. However, the highest degree of similarity was found in the regions G1-G5 (which are implicated in GTP binding and GTPase activity in EF-2), and E1-E4 which are shared among elongation factors and which may be involved in direct ribosome binding

Fig. 10. Amino acid sequence comparison of EFT-1 and elongation factor 2 from hamster (HamEF2), Drosophila melanogaster (DmEF2), and C. elegans (CeEF2). The numbering is according to the hamster sequence. Amino acid identity with HamEF2 is denoted with a dash (-), the up-arrows indicate the positions of the four splice sites in EFT-1 as deduced from the nucleotide sequence and down-arrows denote the positions of two introns so far identified in CeEF2. The GTP-binding regions G1 - G5 and the regions highly conserved among elongation factors (E1 - E4, Kohno et al., 1986, Grinblat et al., 1989) are indicated. The positions corresponding to the degenerate primers OPC3 and OPC4 used for the PCR analysis are shown. The histidyl residue target for ADP-ribosylation by diphtheria toxin is marked with an asterisk (position 715). Gaps were introduced to maximize alignment.

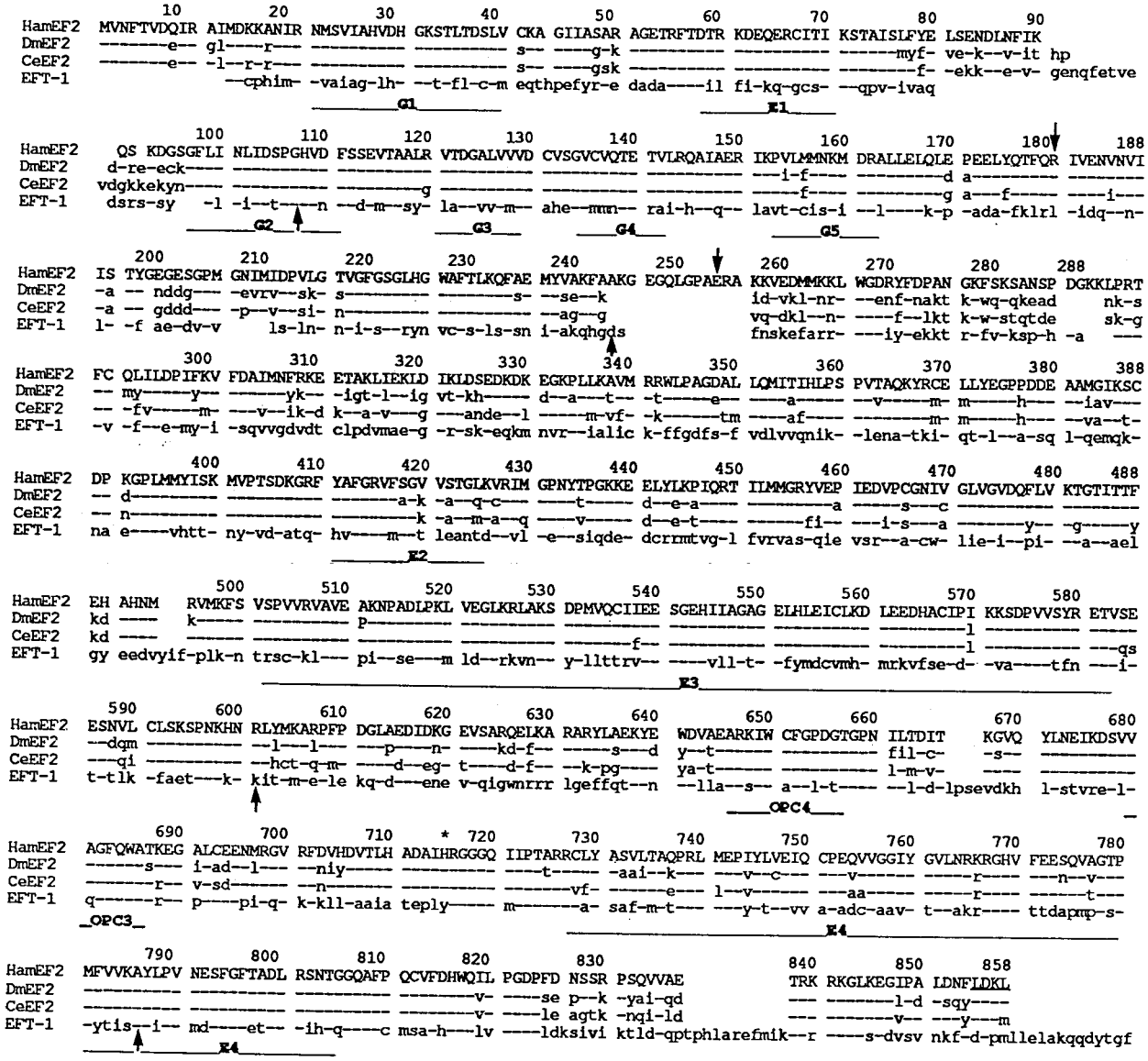
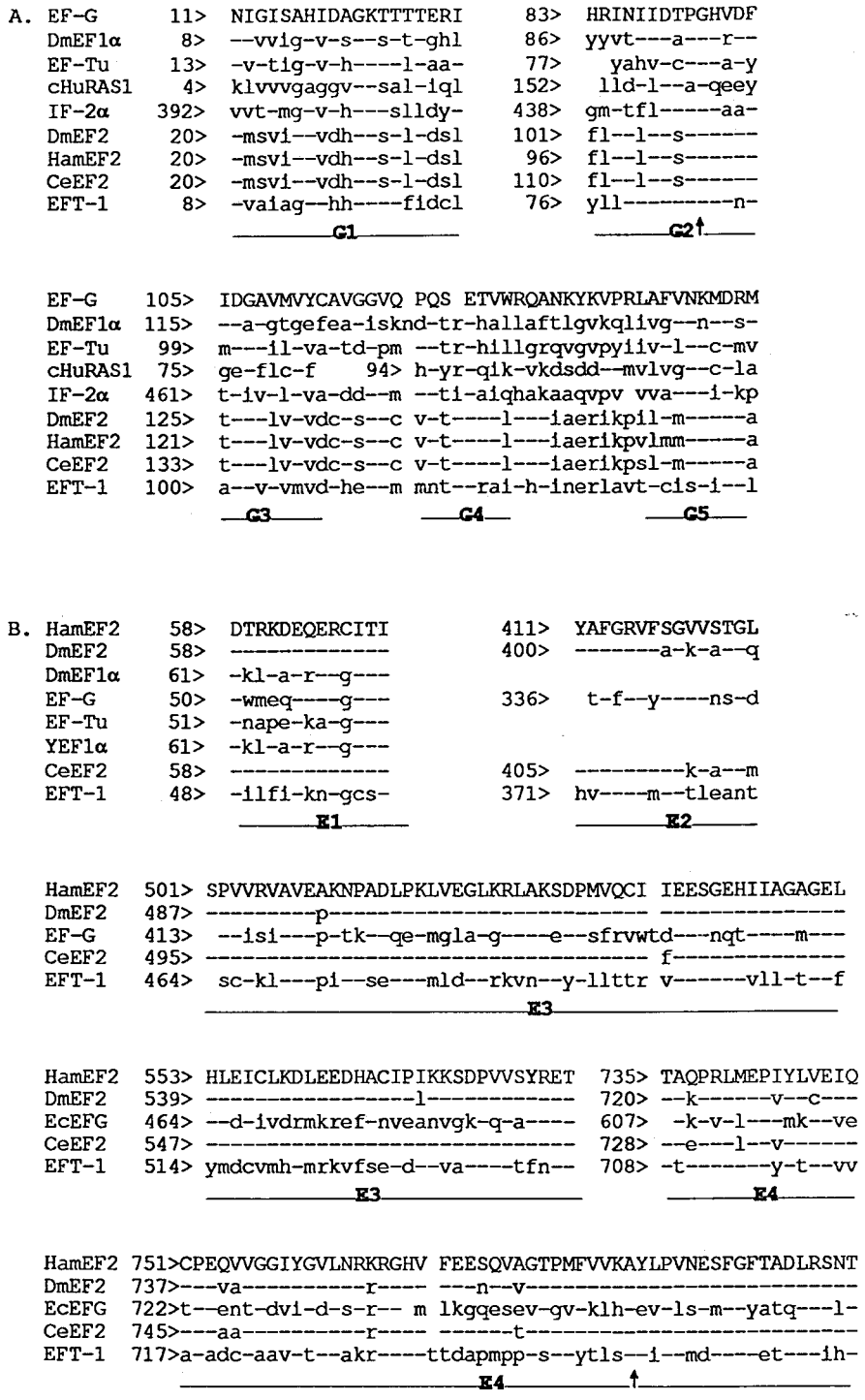

HamEF2
DmEF2
CeEF2
EFT-1
G1
E1
G2
G3
G4
G5
E2
E3
OPC4
OPC3
E4

Fig. 11. Alignment of the amino acid sequence of C. elegans EFT-1 and EF-2 with related proteins from other species.

A. Alignment of the highly conserved regions G1 - G5. Sequences compared are: C. elegans EFT-1 and EF-2 (CeEF2), E. coli elongation factors G and Tu (EF-G and EF-Tu, Montandon and Stutz, 1983; Zengel et al., 1984), D. melanogaster elongation factor 1-alpha (DmEF1α, Hovermann et al., 1988), human RAS1 protein (cHuRAS1, Capon et al., 1983), bacterial initiation factor 2-alpha (IF-2α, Sacerot et al., 1984), HamEF2 and DmEF2. **B.** Comparison of EFT-1 and CeEF2 in the regions E1 - E4 (highly conserved regions in elongation factors), with HamEF2, DmEF1α, DmEF2, E-FG, EF-Tu, and yeast elongation factor 1-alpha (YEF1α, Nagata et al., 1984). Amino acid identity with the top sequence is marked with a dash (-). The numbers at the left indicate the residue number starting from the initial methionine of the respective protein. Gaps are introduced to maximize alignment.


A.
EF-G 11> NIGISAHIDAGKTTTTERI 83> HRINIIDTPGHVDF
DmEF1α 8> --vvig-v-s--s-t-ghl 86> yyvt---a---r--
EF-Tu 13> -v-tig-v-h----l-aa- 77> yahv-c---a-y
cHuRAS1 4> klvvvgaggv--sal-iql 152> lld-l--a-qeey
IF-2α 392> vvt-mg-v-h---slldy- 438> gm-tfl-----aa-
DmEF2 20> -msvi--vdh--s-l-dsl 101> fl--l--s------
HamEF2 20> -msvi--vdh--s-l-dsl 96> fl--l--s------
CeEF2 20> -msvi--vdh--s-l-dsl 110> fl--l--s------
EFT-1 8> -vaiag--hh----fidcl 76> yll---------n-
G1
G2
EF-G 105> IDGAVMVYCAVGGVQ PQS ETVWRQANKYKVPRLAFVNKMDRM
DmEF1α 115> --a-gtgefea-isknd-tr-hallaftlgvkqlivg--n--s-
EF-Tu 99> m---il-va-td-pm --tr-hillgrqvgvpyiiv-l--c-mv
cHuRAS1 75> ge-flc-f 94> h-yr-qik-vkdsdd--mvlvg--c-la
IF-2α 461> t-iv-l-va-dd--m --ti-aiqhakaaqvpv vva---i-kp
DmEF2 125> t---lv-vdc-s--c v-t----l---iaerikpil-m-----a
HamEF2 121> t---lv-vdc-s--c v-t----l---iaerikpvlmm-----a
CeEF2 133> t---lv-vdc-s--c v-t----l---iaerikpsl-m-----a
EFT-1 100> a--v-vmvd-he--m mnt--rai-h-inerlavt-cis-i--l
G3
G4
G5
B.
HamEF2 58> DTRKDEQERCITI 411> YAFGRVFSGVVSTGL
DmEF2 58> ---------- 400> -------a-k-a--q
DmEF1α 61> -kl-a-r--g---
EF-G 50> -wmeq----g--- 336> t-f--y----ns-d
EF-Tu 51> -nape-ka-g---
YEF1α 61> -kl-a-r--g---
CeEF2 58> ---------- 405> ---------k-a--m
EFT-1 48> -ilfi-kn-gcs- 371> hv----m--tleant
E1
E2
HamEF2 501> SPVVRVAVEAKNPADLPKLVEGLKRLAKSDPMVQCI IEESGEHIIAGAGEL
DmEF2 487> ---------p---------- ----------
EF-G 413> --isi---p-tk--qe-mgla-g----e--sfrvwtd---nqt----m---
CeEF2 495> ---------- f----------
EFT-1 464> sc-kl---pi--se---mld--rkvn--y-llttr v------vll-t--f
E3
HamEF2 553> HLEICLKDLEEDHACIPIKKSDPVVSYRET 735> TAQPRLMEPIYLVEIQ
DmEF2 539> ----------l---------- 720> --k------v--c---
EcEFG 464> --d-ivdrmkref-nveanvgk-q-a---- 607> -k-v-l---mk--ve
CeEF2 547> ---------- 728> --e---l--v------
EFT-1 514> ymdcvmh-mrkvfse-d--va----tfn-- 708> -t-------y-t--vv
E3
E4
HamEF2 751>CPEQVVGGIYGVLNRKRGHV FEESQVAGTPMFVVKAYLPVNESFGFTADLRSNT
DmEF2 737>---va----------r---- ---n--v----------
EcEFG 722>t--ent-dvi-d-s-r-- m lkgqesev-gv-klh-ev-ls-m--yatq---l-
CeEF2 745>---aa----------r---- ------t----------
EFT-1 717>a-adc-aav-t--akr----ttdapmpp-s--ytls--i--md----et---ih-
E4

(Nilsson and Nygard, 1985; Kohno et al., 1986; Grinblat et al., 1989), as well as the 15 amino acid region immediately following the modifiable histidyl residue. A more detailed comparison of EFT-1 with GTP-binding proteins and elongation factors is presented in Figure 11A and B , respectively. Relative to other EF-2's, the extent of similarity ranged between 44% in G4 and 93% in G2; between E regions, similarities ranged from 40% in E2 to 57% in E4. EFT-1 contains a tyrosyl residue in a position corresponding to the modifiable histidyl residue in EF-2 (Nilsson and Nygard, 1985; Honjo et al., 1968). In the 15 amino acid region immediately following this tyrosyl residue, EFT-1 shared 80% and 86.7% identity with the Drosophila and hamster EF-2 sequences, respectively.

1. Analysis of the modifiable histidyl region

The lack of a modifiable histidyl residue in EFT-1 and the relatively low degree (38%) of amino acid sequence identity with EF-2 compared to the greater than 80% identity between the hamster and Drosophila sequences (Grinblat et al., 1989) suggested that EFT-1 was not likely to be the C. elegans EF-2. In order to isolate the C. elegans EF-2, first strand cDNA obtained by reverse transcription of embryo RNA was amplified as described in "Experimental Procedures". Sequencing of the 761 bp fragment resulting from the RACE-OPC4 amplification revealed a region including the modifiable histidyl residue of EF-2. Sequence comparison (Fig. 12) of this region with those of EFT-1, hamster and Drosophila EF-2 revealed 87% identity with hamster or Drosophila EF-2, and 57% identity with C. elegans EFT-1.

H. DEVELOPMENTAL EXPRESSION OF EFT-1 MRNA

Analysis of C. elegans mRNA using labeled B3255 as a probe revealed that eft-1 encodes a single mRNA species of 3 kilobases (Fig. 13A) which is expressed throughout

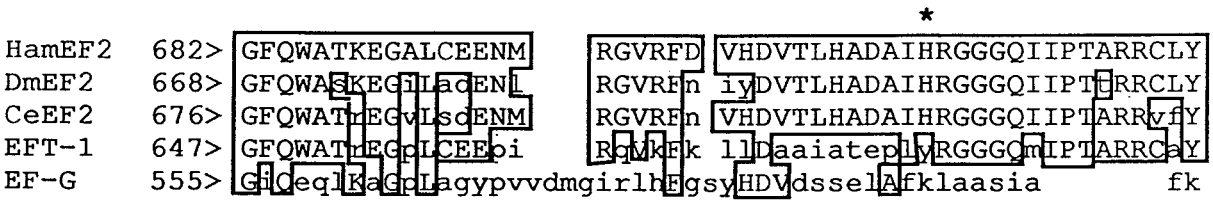


Fig. 12. Alignment of the deduced amino acid sequences of EFT-1 and EF-G with DmEF2, HamEF2, and C. elegans elongation factor 2 (CeEF2) in the ADP-modifiable histidyl region.

Comparison of the deduced amino acid sequence (CeEF2) in the modifiable histidyl region of the 761 bp fragment with those of EFT-1, EF-G, and the hamster (HamEF-2) and Drosophila EF-2 (DmEF-2). Gaps have been introduced to provide maximum alignment with EF-G. The histidyl residue modifiable by ADP-ribosylation is marked with an asterisk at position 715 of HamEF2. Residues that are identical with those of HamEF2 are shown in upper case and enclosed in boxes.

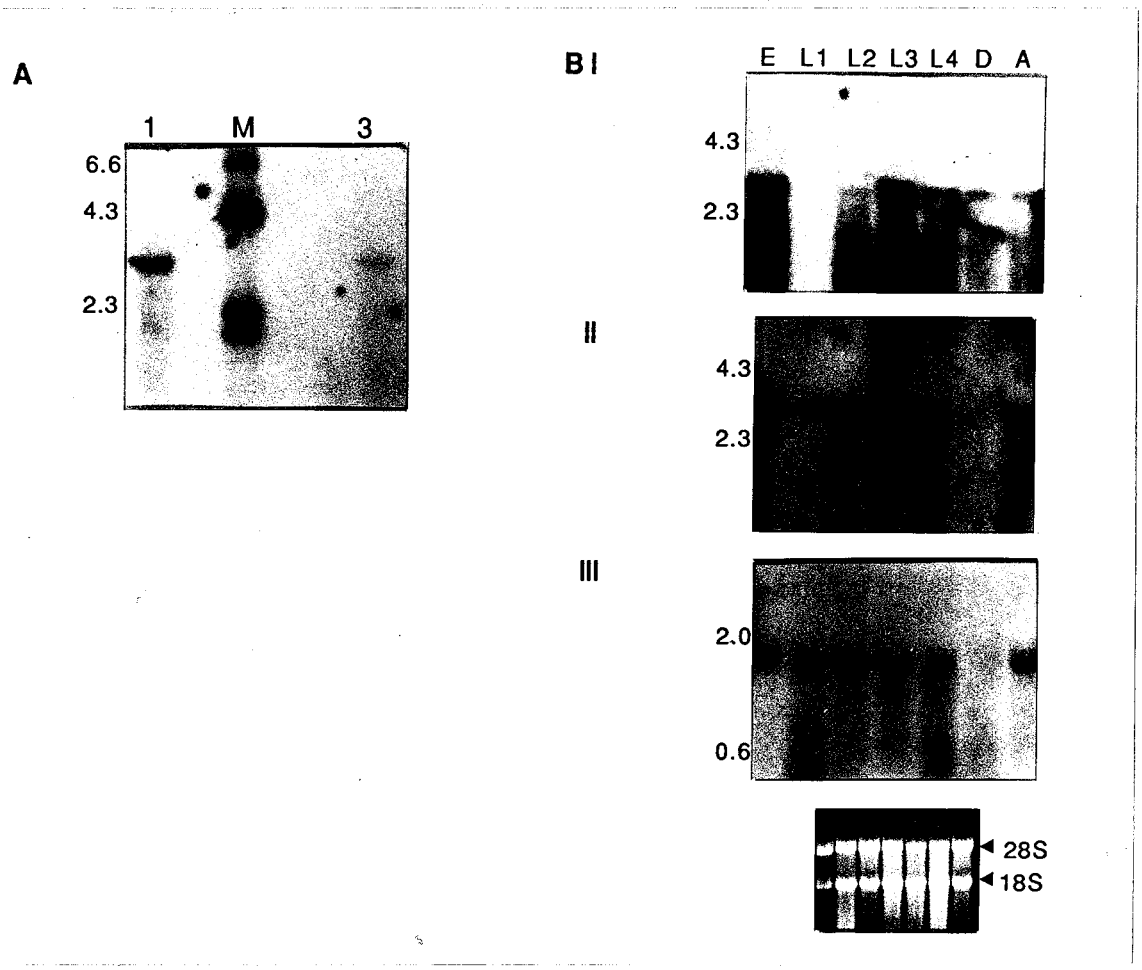


Fig. 13. Northern blot analysis of eft-1 mRNA.

A. 5 µg polyA+RNA (lane 1) and 20 µg total cellular RNA (lane 3) from embryos, and **B**, 5 µg polyA+RNA (row I) and 25 µg total RNA (rows II and III) from all stages of development were hybridized with nick-translated ^{32}P-labeled B3255 (panels A, BI and BII) or with labeled pCeA7 (C. elegans actin gene 1, panel BIII) as described in "Experimental Procedures". Filters were washed at high stringency (Sambrook et al., 1989) and exposed to X-ray film for 2-3 days at -70°C using Dupont Cronex intensifying screens. E, embryo; L1 to L4, larvae; D, dauer larva; A, young adult; M, size markers. Lower panel is the ethidium bromide-stained gel before blotting onto membrane. Sizes of marker bands are given in kilobases, and positions of ribosomal RNA bands are indicated by arrows.

nematode development, from embryo to adulthood (Fig. 13BI and BII), but appeared to be relatively low in abunbance in the first larval (L1) stage. The level of expression of this transcript and of the actin gene 1 mRNA (Fig. 13BIII) also appeared to be somewhat reduced in the dauer larval stage, although staining of the ribosomal RNA bands indicated comparable loadings of RNA from all stages (Fig. 13B, lower panel).

I. ISOLATION OF CDNA CLONES ENCODING EFT-2

Sequence analysis of the 665 bp and 761 bp (Fig. 12) PCR products derived from rapid amplifications with the RACE-OPC3 and RACE-OPC4 oligonucleotide combinations, respectively, revealed that they both encoded the 3' region of the C. elegans homolog of EF-2. When a C. elegans cDNA library prepared in λZAP was screened using the 761 bp fragment as a nick translated, ^{32}P-labeled probe, six positive clones including one (pCef6A) containing the entire coding region were isolated from 1,200 plaques screened. An 870 bp EcoRI- ScaI probe from the 5' end of pCef6A was used to isolate another 11 clones from 4,000 phage of the library. The positive lambda ZAP clones were converted into pBluescript plasmids by superinfection with the helper phage R408 (Stratagene). The plasmids and a series of nested deletions (Henikoff, 1984) generated from pCef6A were sequenced as previously described (Sanger et al., 1977; Hattori and Sakaki, 1986).

J. ISOLATION OF GENOMIC CLONES ENCODING EFT-2

To isolate the C. elegans gene (eft-2) encoding EF-2, recombinant clones (40,000) of a C. elegans λEMBL4 genomic library were screened with the 761 bp PCR fragment as a probe. Twelve positive clones were obtained from which phage DNAs were prepared. By restriction digestion of the phage DNA and hybridization with the entire cDNA insert of pCef6A, a 7 kb EcoRI-PstI fragment (7E/P) containing the entire coding region was isolated

from one clone (gEMBg2) and inserted into pBluescript(+) as pSK7E/P. Further analysis by restriction mapping and hybridization with probes corresponding to different portions of the insert in pCef6A allowed orientation of the coding region within the clone. Figure 14 shows the restriction map and orientation of the eft-2 locus derived from gEMBg2. Subfragments of 7E/P were inserted (Messing, 1983) into pGEM3Z or pGEM4Z vectors and partially sequenced. The results revealed two introns of 48 and 44 base pairs. The locations, sequences and splice junctions of these introns are shown in Figure 14B.

K. LOCALIZATION OF THE EFT-2 GENE

The genomic clone gEMBg2 was digested with restriction enzymes and the restriction fragment pattern obtained was compared by computer with the entire data base of C. elegans genomic clones, as part of the C. elegans genome mapping project (Coulson et al., 1986). Cosmids containing the clone with the most likely matches were then identified. Using this fingerprinting procedure, Alan Coulson (Medical Research Council Laboratory, Cambridge, England) identified a series of overlapping cosmids containing the eft-2 gene and which map to the right half of chromosome I between unc-29 and ceh-5 (Figure 15).

L. ANALYSIS OF EFT-2 GENE SEQUENCE

Figure 16 shows the complete nucleotide and deduced amino acid sequences of the cDNA insert in pCef6A. The initiation codon was assigned to the first ATG (nucleotides positions 46 - 48) occurring downstream of the in-frame terminator TGA (positions 19 - 21), and which specified the longest open reading frame (ORF). The 2556 bp ORF predicted a polypeptide of 852 amino acid residues (Fig. 17, calculated molecular weight of 94,564) with six N-terminal amino acid residues (underlined in Fig. 16) which are identical to the

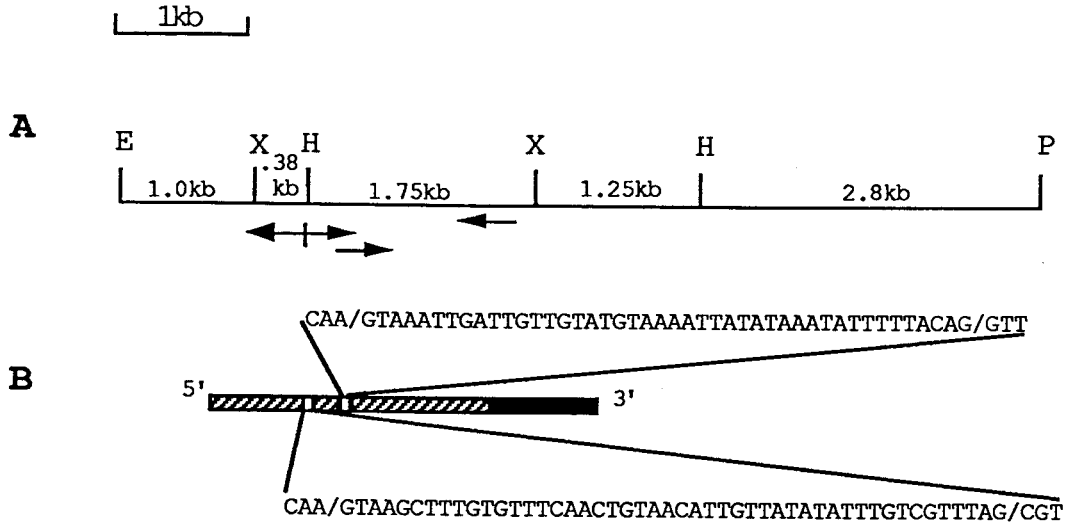


Fig. 14. Structure of the eft-2 locus.

A. Restriction map of the 7 kb Eco RI-PstI fragment isolated from genomic clone gEMBg2, containing eft-2. The restriction sites are: E, EcoRI, H, HindIII, X, XhoI, and P, PstI. Kb, kilobase pairs. **B.** Regions represented in the cDNA clone pCef6A are indicated by stippled boxes aligned to corresponding genomic sequences. Alignment was based on detailed restriction and Southern blot analyses using probes corresponding to portions of the coding region in the cDNA clone. The locations (open boxes), sizes and sequences of two introns identified by partial sequencing of the genomic clone are shown. The direction and extent of sequencing are indicated by arrows. The position of the 761 bp PCR-amplified fragment used as a probe to isolate gEMBg2 and pCef6A is indicated by a solid box.

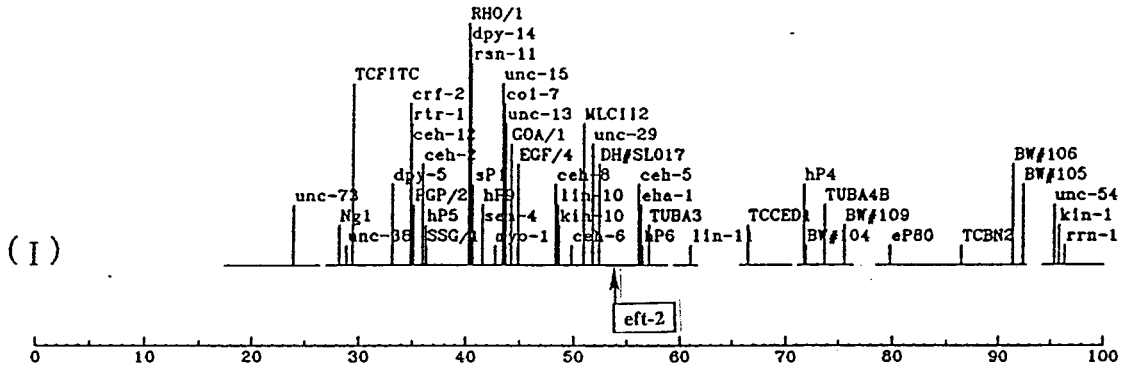


Fig. 15. Chromosomal location of the eft-2 locus of C. elegans.

Physical map of chromosome I of C. elegans showing the location of eft-2. The genomic clone gEMBg2 isolated from a λEMBL4 genomic library and which encodes the entire eft-2 mRNA, was sent to A. Coulson (MRC Laboratory of Molecular Biology, Cambridge, England). It was digested with restriction enzymes and the restriction fragment pattern was compared by computer with the entire data base of C. elegans genomic clones. The cosmid contig containing the clone with the most likely matches was identified and physically mapped to about 53 units from the left end of chromosome I (A. Coulson, personal communication). The position of eft-2 (boxed) is shown by an arrow between unc-29 and ceh-5.

sequences obtained from a partial tryptic fragment of purified rat EF-2 (Takamatsu et al., 1986). The insert of pCef6A also contained 45 bp of 5'-untranslated region and 115 bp of 3'-untranslated region preceding the poly(A) addition site; these regions have a relatively high A-T content (78% in the 5' and 75% in the 3' regions) compared to that of the coding region (49%).

In another group of cDNA clones (including pCef1d-g) encoding CeEF-2, the poly(A) tract was found 55 nucleotides upstream of the presumptive poly(A) addition site used in pCef6A. These putative poly(A) addition signals (AATAAA) are located 14 and 12 nucleotides, respectively, upstream of the poly(A) tail, and are indicated by dots (Fig. 16). A sequence TGTGCTAA resembling the consensus sequence (YGTGTTYY) implicated in efficient RNA 3' end formation (McLauchlan et al., 1985) is located 12 nucleotides upstream of the distal polyadenylation signal.

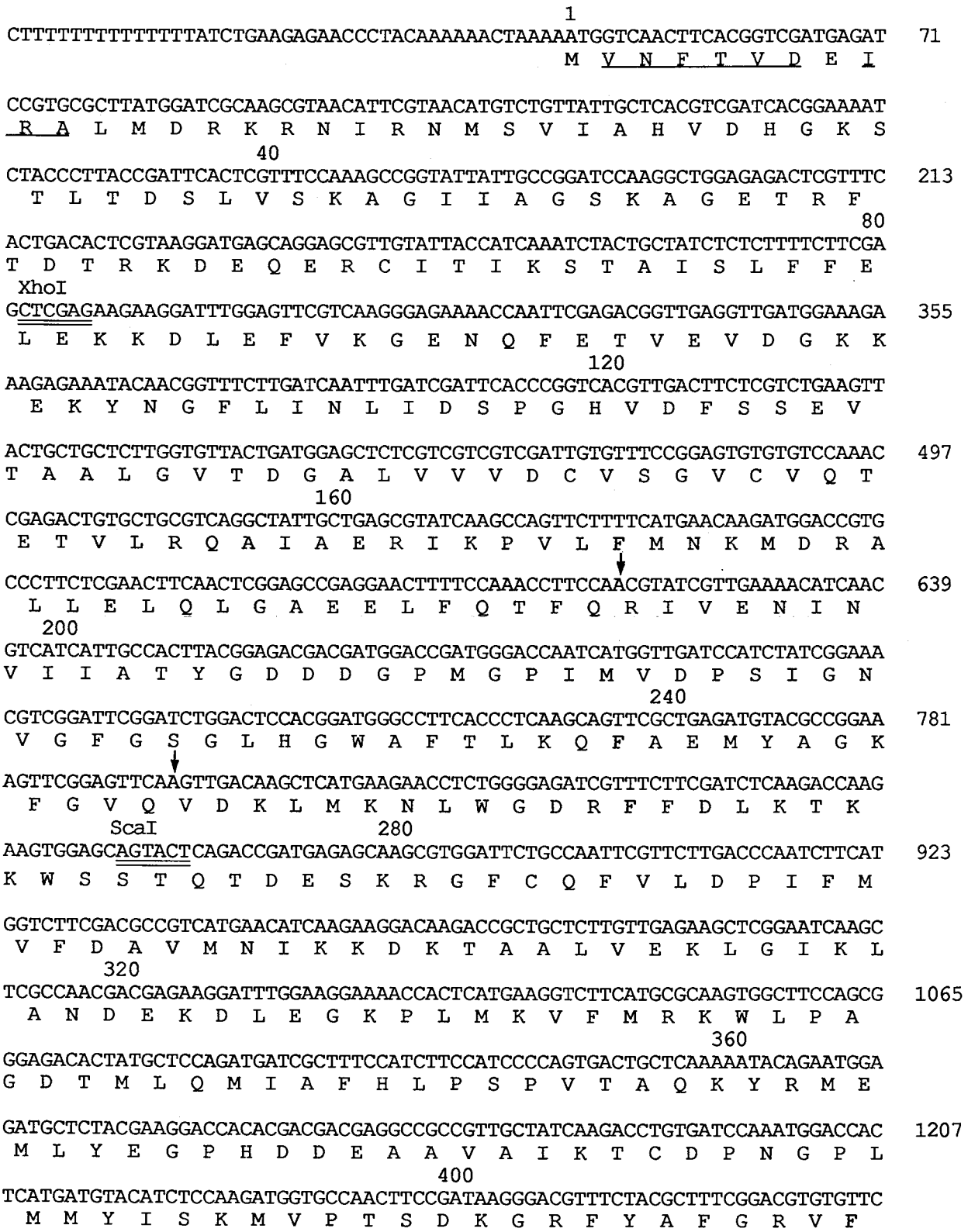
Partial sequencing of the genomic clone gEMBg2 encoding CeEF-2 has thus far revealed 2 short introns (Fig. 14B), one of 48 bases interrupting the ORF after nucleotide position 618 and the other of 44 bases after position 795 (Fig. 16). At the nucleotide level, the overall identities are 73% and 76% between eft-2 and the Drosophila and hamster EF-2 sequences, respectively, and 56% between eft-2 and eft-1.

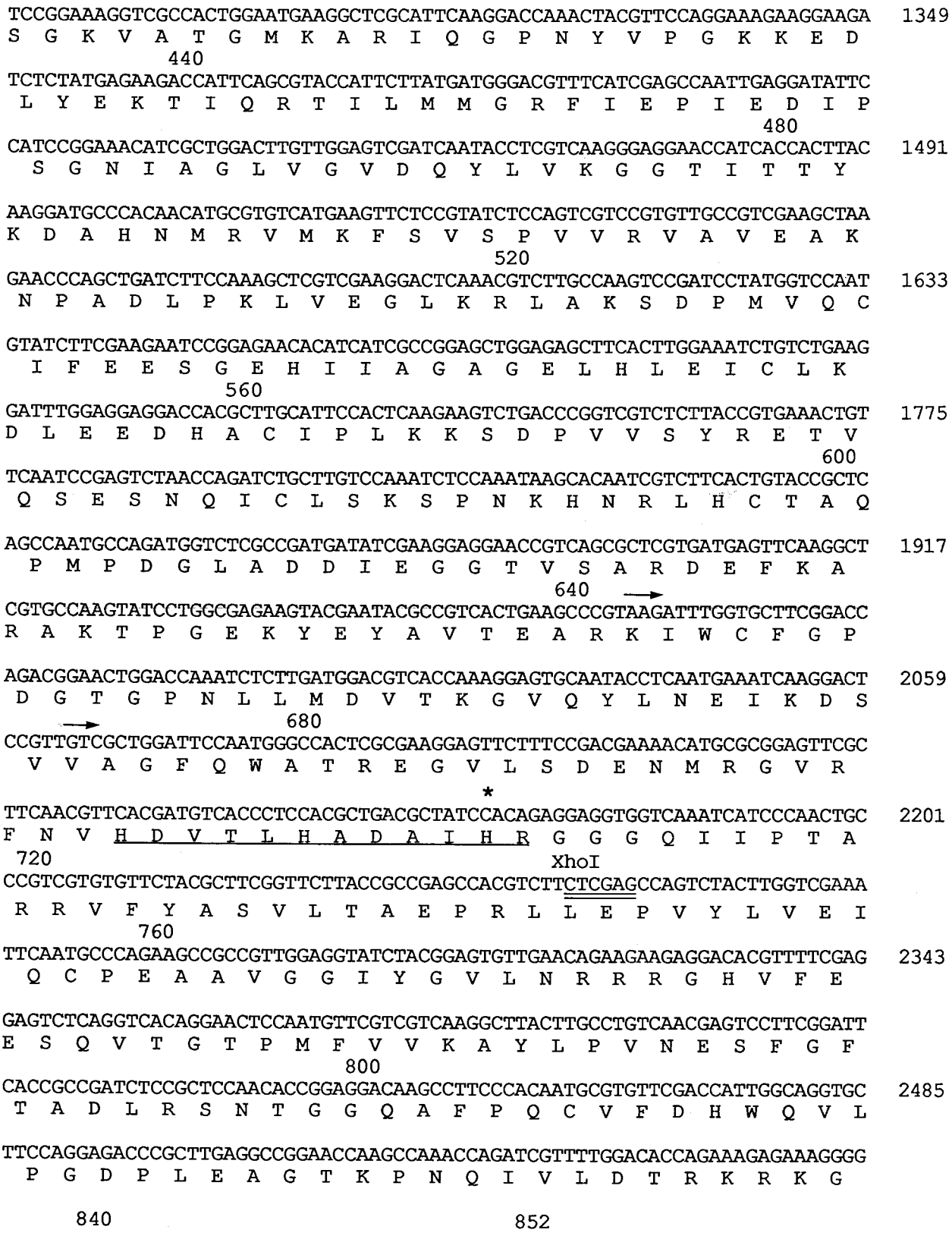
M. PRIMARY STRUCTURE OF CEEF-2

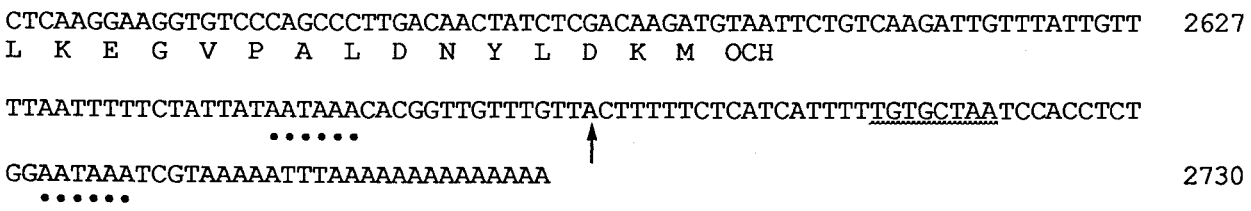
Computer analysis (Pearson and Lipman, 1988) showed that the deduced amino acid sequence of CeEF-2 shares an overall identity of 80% and 87% with Drosophila and hamster EF-2 sequences, respectively. The lower degree of identity at the nucleotide level compared to the amino acid sequence level was due mainly to base changes at the third positions of codons, reflecting a strong bias of codon usage in C. elegans (Wada et al., 1990). Alignment of the amino acid sequences of these proteins with that of EFT-1 is shown in Figure 10.

Fig. 16. Complete nucleotide and deduced amino acid sequence of the cDNA encoding C. elegans EF-2.

The nucleotide and amino acid residues are numbered from the beginning of the sequence and the initial methionine, respectively. Amino acid residues (positions 2-7, 9-11 and 698-710) which are identical in the sequence of purified rat EF-2 (Robinson et al., 1974; Takamatsu et al., 1986) are underlined. The down-arrows indicate the positions of introns identified by partial sequencing of the genomic clone. Horizontal arrows mark the beginning of the degenerate primers OPC4 and OPC3 used for the PCR analysis. The *Xho* I sites are doubly underlined and his-709 (*) is the target for ADP-ribosylation by diphtheria toxin. The poly(A) addition site in the prototypical cDNA clone pCef1d-g is indicated by an up-arrow (between nucleotides 2661 and 2662) while dots identify potential polyadenylation signals. A sequence potentially involved in efficient mRNA 3' end formation (McLauchlan et al., 1985) is indicated by a wavy line.


1
CTTTTTTTTTTTTTTATCTGAAGAGAACCCTACAAAAAACTAAAAATGGTCAACTTCACGGTCGATGAGAT 71
M V N F T V D E I
CCGTGCGCTTATGGATCGCAAGCGTAACATTCGTAACATGTCTGTTATTGCTCACGTCGATCACGGAAAAT
R A L M D R K R N I R N M S V I A H V D H G K S
40
CTACCCTTACCGATTCACTCGTTTCCAAAGCCGGTATTATTGCCGGATCCAAGGCTGGAGAGACTCGTTTC 213
T L T D S L V S K A G I I A G S K A G E T R F
80
ACTGACACTCGTAAGGATGAGCAGGAGCGTTGTATTACCATCAAATCTACTGCTATCTCTCTTTTCTTCGA
T D T R K D E Q E R C I T I K S T A I S L F F E
XhoI
GCTCGAGAAGAAGGATTTGGAGTTCGTCAAGGGAGAAAACCAATTCGAGACGGTTGAGGTTGATGGAAAGA 355
L E K K D L E F V K G E N Q F E T V E V D G K K
120
AAGAGAAATACAACGGTTTCTTGATCAATTTGATCGATTCACCCGGTCACGTTGACTTCTCGTCTGAAGTT
E K Y N G F L I N L I D S P G H V D F S S E V
ACTGCTGCTCTTGGTGTTACTGATGGAGCTCTCGTCGTCGTCGATTGTGTTTCCGGAGTGTGTGTCCAAAC 497
T A A L G V T D G A L V V V D C V S G V C V Q T
160
CGAGACTGTGCTGCGTCAGGCTATTGCTGAGCGTATCAAGCCAGTTCTTTTCATGAACAAGATGGACCGTG
E T V L R Q A I A E R I K P V L F M N K M D R A
CCCTTCTCGAACTTCAACTCGGAGCCGAGGAACTTTTCCAAACCTTCCAACGTATCGTTGAAAACATCAAC 639
L L E L Q L G A E E L F Q T F Q R I V E N I N
200
GTCATCATTGCCACTTACGGAGACGACGATGGACCGATGGGACCAATCATGGTTGATCCATCTATCGGAAA
V I I A T Y G D D D G P M G P I M V D P S I G N
240
CGTCGGATTCGGATCTGGACTCCACGGATGGGCCTTCACCCTCAAGCAGTTCGCTGAGATGTACGCCGGAA 781
V G F G S G L H G W A F T L K Q F A E M Y A G K
AGTTCGGAGTTCAAGTTGACAAGCTCATGAAGAACCTCTGGGGAGATCGTTTCTTCGATCTCAAGACCAAG
F G V Q V D K L M K N L W G D R F F D L K T K
ScaI
280
AAGTGGAGCAGTACTCAGACCGATGAGAGCAAGCGTGGATTCTGCCAATTCGTTCTTGACCCAATCTTCAT 923
K W S S T Q T D E S K R G F C Q F V L D P I F M
GGTCTTCGACGCCGTCATGAACATCAAGAAGGACAAGACCGCTGCTCTTGTTGAGAAGCTCGGAATCAAGC
V F D A V M N I K K D K T A A L V E K L G I K L
320
TCGCCAACGACGAGAAGGATTTGGAAGGAAAACCACTCATGAAGGTCTTCATGCGCAAGTGGCTTCCAGCG 1065
A N D E K D L E G K P L M K V F M R K W L P A
360
GGAGACACTATGCTCCAGATGATCGCTTTCCATCTTCCATCCCCAGTGACTGCTCAAAAATACAGAATGGA
G D T M L Q M I A F H L P S P V T A Q K Y R M E
GATGCTCTACGAAGGACCACACGACGACGAGGCCGCCGTTGCTATCAAGACCTGTGATCCAAATGGACCAC 1207
M L Y E G P H D D E A A V A I K T C D P N G P L
400
TCATGATGTACATCTCCAAGATGGTGCCAACTTCCGATAAGGGACGTTTCTACGCTTTCGGACGTGTGTTC
M M Y I S K M V P T S D K G R F Y A F G R V F


TCCGGAAAGGTCGCCACTGGAATGAAGGCTCGCATTCAAGGACCAAACTACGTTCCAGGAAAGAAGGAAGA 1349
S G K V A T G M K A R I Q G P N Y V P G K K E D
440
TCTCTATGAGAAGACCATTCAGCGTACCATTCTTATGATGGGACGTTTCATCGAGCCAATTGAGGATATTC
L Y E K T I Q R T I L M M G R F I E P I E D I P
480
CATCCGGAAACATCGCTGGACTTGTTGGAGTCGATCAATACCTCGTCAAGGGAGGAACCATCACCACTTAC 1491
S G N I A G L V G V D Q Y L V K G G T I T T Y
AAGGATGCCCACAACATGCGTGTCATGAAGTTCTCCGTATCTCCAGTCGTCCGTGTTGCCGTCGAAGCTAA
K D A H N M R V M K F S V S P V V R V A V E A K
520
GAACCCAGCTGATCTTCCAAAGCTCGTCGAAGGACTCAAACGTCTTGCCAAGTCCGATCCTATGGTCCAAT 1633
N P A D L P K L V E G L K R L A K S D P M V Q C
GTATCTTCGAAGAATCCGGAGAACACATCATCGCCGGAGCTGGAGAGCTTCACTTGGAAATCTGTCTGAAG
I F E E S G E H I I A G A G E L H L E I C L K
560
GATTTGGAGGAGGACCACGCTTGCATTCCACTCAAGAAGTCTGACCCGGTCGTCTCTTACCGTGAAACTGT 1775
D L E E D H A C I P L K K S D P V V S Y R E T V
600
TCAATCCGAGTCTAACCAGATCTGCTTGTCCAAATCTCCAAATAAGCACAATCGTCTTCACTGTACCGCTC
Q S E S N Q I C L S K S P N K H N R L H C T A Q
AGCCAATGCCAGATGGTCTCGCCGATGATATCGAAGGAGGAACCGTCAGCGCTCGTGATGAGTTCAAGGCT 1917
P M P D G L A D D I E G G T V S A R D E F K A
640
CGTGCCAAGTATCCTGGCGAGAAGTACGAATACGCCGTCACTGAAGCCCGTAAGATTTGGTGCTTCGGACC
R A K T P G E K Y E Y A V T E A R K I W C F G P
AGACGGAACTGGACCAAATCTCTTGATGGACGTCACCAAAGGAGTGCAATACCTCAATGAAATCAAGGACT 2059
D G T G P N L L M D V T K G V Q Y L N E I K D S
680
CCGTTGTCGCTGGATTCCAATGGGCCACTCGCGAAGGAGTTCTTTCCGACGAAAACATGCGCGGAGTTCGC
V V A G F Q W A T R E G V L S D E N M R G V R
*
TTCAACGTTCACGATGTCACCCTCCACGCTGACGCTATCCACAGAGGAGGTGGTCAAATCATCCCAACTGC 2201
F N V H D V T L H A D A I H R G G G Q I I P T A
720
XhoI
CCGTCGTGTGTTCTACGCTTCGGTTCTTACCGCCGAGCCACGTCTTCTCGAGCCAGTCTACTTGGTCGAAA
R R V F Y A S V L T A E P R L L E P V Y L V E I
760
TTCAATGCCCAGAAGCCGCCGTTGGAGGTATCTACGGAGTGTTGAACAGAAGAAGAGGACACGTTTTCGAG 2343
Q C P E A A V G G I Y G V L N R R R G H V F E
GAGTCTCAGGTCACAGGAACTCCAATGTTCGTCGTCAAGGCTTACTTGCCTGTCAACGAGTCCTTCGGATT
E S Q V T G T P M F V V K A Y L P V N E S F G F
800
CACCGCCGATCTCCGCTCCAACACCGGAGGACAAGCCTTCCCACAATGCGTGTTCGACCATTGGCAGGTGC 2485
T A D L R S N T G G Q A F P Q C V F D H W Q V L
TTCCAGGAGACCCGCTTGAGGCCGGAACCAAGCCAAACCAGATCGTTTTGGACACCAGAAAGAGAAAGGGG
P G D P L E A G T K P N Q I V L D T R K R K G
840
852


CTCAAGGAAGGTGTCCCAGCCCTTGACAACTATCTCGACAAGATGTAATTCTGTCAAGATTGTTTATTGTT 2627
L K E G V P A L D N Y L D K M OCH
TTAATTTTTCTATTATAATAAACACGGTTGTTTGTTACTTTTTCTCATCATTTTTGTGCTAATCCACCTCT
GGAATAAATCGTAAAAATTTAAAAAAAAAAAAAA 2730

CeEF-2 shares 38% amino acid sequence identity with EFT-1. A single insertion of 12 amino acids is present in CeEF-2 relative to hamster EF-2 at position 90, and CeEF-2 lacks 13 and 4 amino acid residues respectively, present at positions 238-250 and 281-284 of the hamster protein. The regions G1 - G5, which are implicated in GTP binding and GTPase activity in EF-2 (Nilsson and Nygard, 1985; Kohno et al, 1986), and E2 - E4 which are highly conserved in EF-2 and EF-G, were found in CeEF-2. In the E1 region (highly conserved among elongation factors), CeEF-2 was 100% identical to the hamster and Drosophila sequences (Fig. 11B). The extent of identity in the G1-G5 and E2-E4 domains ranged from 80% in E2 to 100% in G1-G4 between CeEF-2 and the hamster or Drosophila sequences (Fig. 11). The histidyl residue which is ADP-ribosylated by diphtheria toxin in EF-2 was found (at amino acid position 709, Fig. 16) in CeEF-2 in a sequence context similar to that seen in other eukaryotic EF2s (Fig. 12, and Robinson et al., 1974; Brown and Bodley, 1979).

1. Comparison of EFT-1 and CeEF-2 structures

Sequence identities between EFT-1 and CeEF-2 in the G and E-regions are similar to those between EFT-1 and hamster or Drosophila EF-2 sequences. In addition, CeEF-2 is 12 and 24 amino acids longer than EFT-1 at the C-terminus and between positions 90-100 (Fig. 10), respectively. The latter region was shown to have diverged greatly in length and sequence between the hamster and Drosophila EF-2 sequences (Grinblat et al., 1989), as well as in EFT-1; both eft-1 and eft-2 mRNAs contain an intron in this region. Other divergent regions include positions 494-500, 666-670, and 837-840 (in the hamster sequence, Fig. 10), where EFT-1 is 3, 5, and 11 amino acids, respectively, longer than CeEF-2. Also EFT-1 is 13 amino acid longer than CeEF-2 at the C-terminus. A comparison of the amino acid compositions of CeEF-2 and EFT-1 (Figure 17) reveals

CeEF2			
Amino acid	Count	Percent (%)	% by Weight
V (Val)	75	8.80	7.86
G (Gly)	71	8.33	4.28
L (Leu)	67	7.86	8.01
K (Lys)	64	7.51	8.67
A (Ala)	60	7.04	4.51
E (Glu)	58	6.81	7.92
D (Asp)	56	6.57	6.81
I (Ile)	47	5.52	5.62
T (Thr)	46	5.40	4.92
R (Arg)	41	4.81	6.77
P (Pro)	41	4.81	4.21
F (Phe)	40	4.69	6.22
S (Ser)	38	4.46	3.50
M (Met)	31	3.64	4.30
N (Asn)	30	3.52	3.62
Q (Gln)	29	3.40	3.93
Y (Tyr)	21	2.46	3.62
H (His)	17	2.00	2.46
C (Cys)	13	1.53	1.42
W (Trp)	7	0.82	1.38

EFT-1			
Amino acid	Count	Percent (%)	% by Weight
L (Leu)	76	8.95	8.94
V (Val)	69	8.13	7.11
E (Glu)	58	6.83	7.79
A (Ala)	57	6.71	4.21
D (Asp)	54	6.36	6.46
T (Thr)	52	6.12	5.46
I (Ile)	50	5.89	5.88
R (Arg)	49	5.77	7.95
K (Lys)	47	5.54	6.26
S (Ser)	47	5.54	4.25
P (Pro)	45	5.30	4.54
F (Phe)	44	5.18	6.73
G (Gly)	42	4.95	2.49
Q (Gln)	32	3.77	4.26
M (Met)	29	3.42	3.95
N (Asn)	29	3.42	3.44
Y (Tyr)	25	2.94	4.24
H (His)	20	2.36	2.85
C (Cys)	17	2.00	1.82
W (Trp)	7	0.82	1.35

Fig. 17. Comparison of the amino acid compositions of C. elegans EFT-1 and EF-2. The amino acid compositions of CeEF2 and EFT-1 as deduced from the nucleotide sequences of eft-1 and eft-2, respectively, were determined using the GeneWorks program of M.J. Glynias (Intelligenetics).

comparable contents of acidic (13.38% and 13.19%) and basic amino acids (14.32% and 13.67%, respectively), as well as high contents of valine, leucine and proline and low contents of cysteine and tryptophan in both proteins.

N. DEVELOPMENTAL EXPRESSION OF EFT-2 MRNA

Analysis of embryo RNA revealed a single 3.0 kilobase mRNA species which hybridized strongly with the 761 bp fragment (Fig. 18A). To determine the developmental expression profile of this mRNA, RNA from each major stage of C. elegans development was analyzed using the same probe. As shown in Fig. 18-BI, eft-2 mRNA was expressed at all stages examined, from embryo to adulthood. However, the level of expression of the transcript appeared to be somewhat reduced in the first larval (L1) stage. The level of actin mRNA was also low in this sample and in the dauer larval stage (Fig. 18-BII), while the loadings of ribosomal RNA were comparable for all stages (lower panel).

O. GENOMIC SOUTHERN ANALYSIS OF EFT-1 AND EFT-2

To estimate the copy numbers of eft-1 and eft-2, genomic DNA from C. elegans and from a closely related species, C. briggsae, was digested with various restriction endonucleases and examined by blot hybridization with labeled B3255 (the 3.2 kb HindIII fragment of eft-1) or the 2.8kb insert of pCef6A as probes. As shown in Figure 19, a single major band was detected in each digest of C. elegans DNA (lanes 3-8), whereas with C. briggsae DNA, EcoRI yielded a major and a minor band, and HindIII digest showed a major and two minor bands (lanes 1 and 2) when hybridized to pCef6A. The sizes of the EcoRI-PstI, XhoI and HindIII bands obtained with pCef6A probe (lanes 3, 7, 8) were consistent with the respective sizes of fragments obtained by restriction mapping of the genomic clone (Fig. 14A). When hybridized with the eft-1 probe, C. elegans DNA digested with PstI, XhoI,

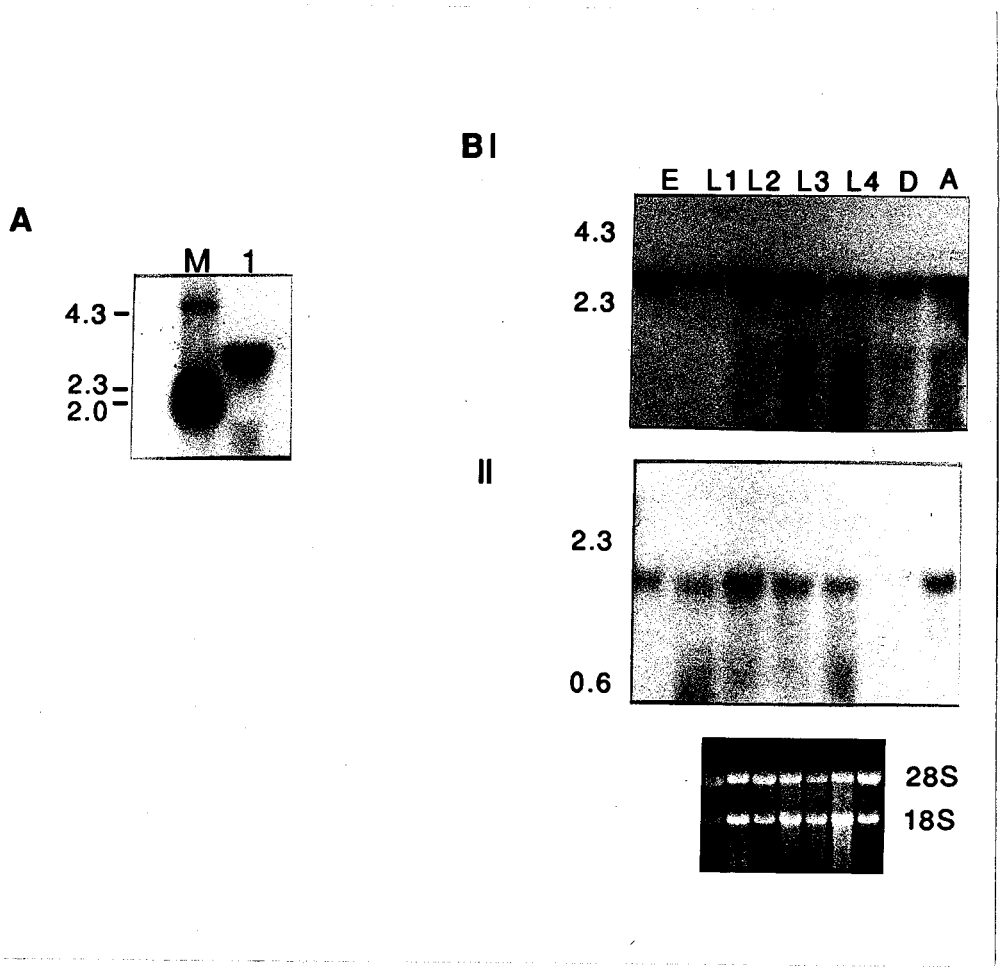


Fig. 18. Northern blot analysis of eft-2 mRNA.

Samples of total cellular RNA (20-25 μg) isolated from **A**, embryo and **B**, various stages of nematode development, were denatured, separated on a 1% agarose-formaldehyde gel and blotted onto a nylon membrane. (Sambrook et al., 1989). After hybridization with the 761 bp PCR fragment (panel A), the 2.8 kb insert of the cDNA clone pCef6A (BI), or the 3 kb actin gene 1 insert of pCeA7 (BII), filters were washed as described in "Experimental Procedures" and exposed to X-ray film for 5h (panel BI) or overnight at -70°C using Dupont Cronex intensifying screens. Lower panel, ethidium bromide-stained gel prior to blotting. E, embryo; L1 to L4, larvae; D, dauer larva; A, young adult; M, size markers. Sizes of marker bands are given in kilobases.

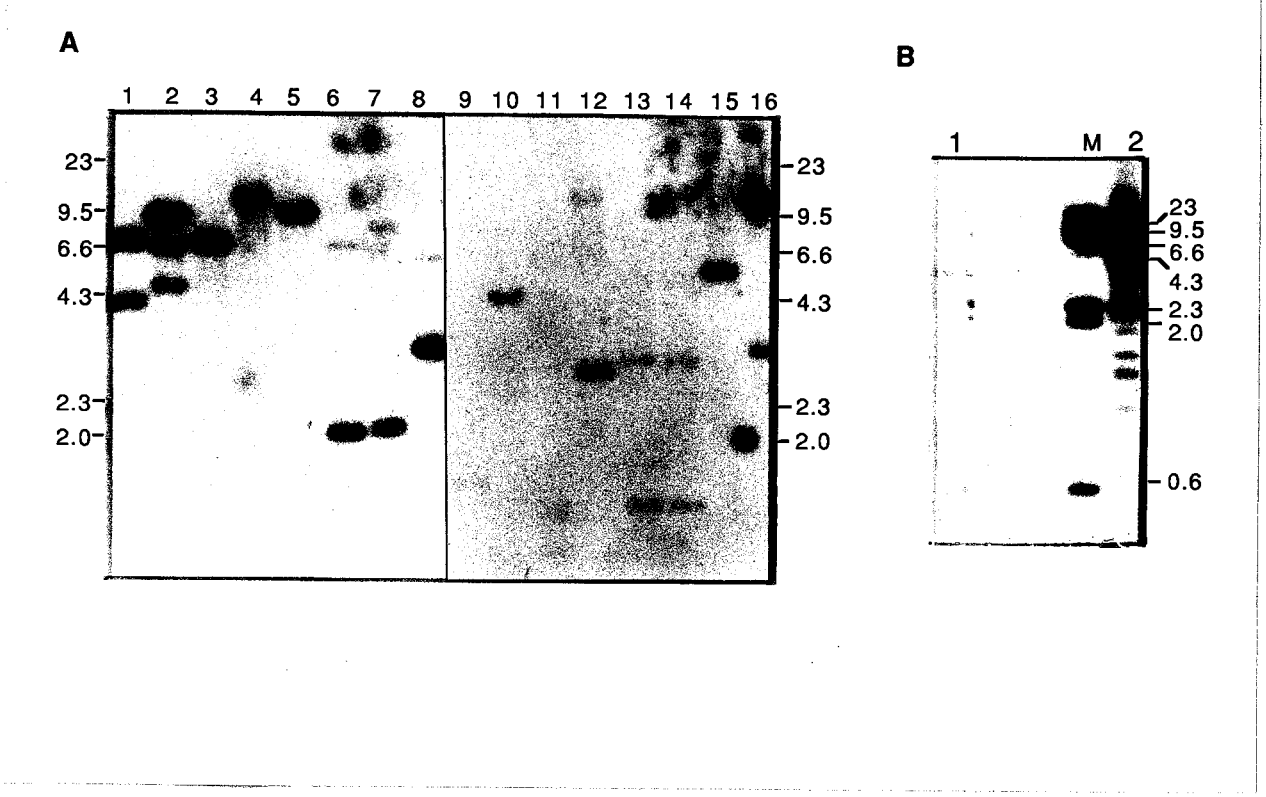


Fig. 19. Genomic Southern blot analysis of eft-1 and eft-2.

A. Genomic DNA (1-2 µg) from C. elegans (lanes 3 to 8 and 11 to 16) and C. briggsae (lanes 1, 2, 9 and 10) was digested to completion with various restriction enzymes, separated on a 0.7% agarose gel, blotted onto a nylon membrane, and hybridized with the 2.8 kb insert of pCef6A (encoding eft-2, lanes 1 to 8) or fragment B3255 (contained in eft-1, lanes 9 to 16) labeled by nick-translation as described in "Experimental Procedures". Lanes 1,5,9 and 13, EcoRI; 2,8,10 and 16, HindIII; 3 and 11, EcoRI-PstI; 4 and 12, PstI; 6 and 14, EcoRI-XhoI; 7 and 15, XhoI. **B.** C. elegans DNA was hybridized with B3255 (lane 1) or the 1.2 kb HindIII fragment of cosmid ZK331 (lane 2) which previously hybridized with cDNA probes made from embryo RNA (see Figure 4A, lanes 2 and 3). Positions of size markers are shown in kilobases.

and HindIII showed a single major band each (lanes 12, 15 and 16), while EcoRI-PstI, EcoRI, and EcoRI-XhoI digests gave two bands each (lanes 11, 13 and 14). With C. briggsae DNA, EcoRI and HindIII digests yielded only a single major band each (lanes 9 and 10).

IV. DISCUSSION

A. PHYSICAL MAP OF THE UBQ-1 REGION OF CHROMOSOME III

The availability of overlapping cosmid clones of C. elegans DNA facilitated the restriction mapping of the 150 kb region around the ubq-1 locus on chromosome III using four different restriction enzymes (Fig. 7). Transcript mapping of this region using cDNA probes synthesized from nematode polyA$^+$RNA revealed a number of putative mRNA coding regions including one gene, subsequently named eft-1, which mapped approximately 50 kb away from ubq-1, in a region on the chromosome (between ubq-1 and col-8) which was genetically undefined.

B. EFT-1 GENE STRUCTURE

The gene organization and complete nucleotide sequence of eft-1, which defines the primary structure of a 96 kdalton protein, EFT-1, have been determined. The cDNA clone pEF1.35 encoding the 3' half of the eft-1 sequence was also characterized. The entire gene consists of 4 short introns and 5 exons in a total length of 3.8 kb. Putative TATA and CAP site sequences were found 140 and 100 bp, respectively, upstream of the presumptive initiation codon. However, since the TATA box influences the accuracy of transcript initiation in eukaryotes, and is almost always found 20 to 35 nucleotides upstream of the cap site of eukaryotic genes transcribed by RNA polymerase II (Nevins, 1983), this TATA is unlikely to be functional. The lack of a functional TATA sequence may explain the heterogeneity of the 5' end of eft-1 mRNA transcripts observed in the nuclease S1 analysis. The promoter regions of most housekeeping genes have been shown to lack the consensus TATA (Dynan, 1986) or CAAT sequence, resulting in multiple transcription initiation sites in agreement with the assigned function for the TATA sequence (Park et al., 1986). No other

strong consensus promoter-like sequences were found in this 5' flanking region of the eft-1 sequence. No canonical poly(A) addition signal (AATAAA) was found in the 114 bp of 3'-untranslated region preceeding the poly(A)tail. However, the sequences AGTAAA and ATGAAA were located 17 and 11 nucleotides upstream of the poly(A)tail, respectively, and may serve as the poly(A) addition signal(s) for eft-1, since in addition to the consensus hexamer AAUAAA, other related sequences including AGUAAA have been seen in different messages and appear to be fully functional in the appropriate context (Park et al., 1986).

The short introns of eft-1 (46-75 bases) are characteristic of 64% of the introns found in C. elegans genes (Fields, 1990), and all of the exons are bounded by good consensus splice junctions which when brought together generate a single continuous open reading frame. Together with 3'- and 5'-untranslated regions, and with due allowance for a poly(A)tail, the theoretical size of the eft-1 mRNA is consistent with the result from the Northern blot analysis.

C. EFT-1 PRIMARY STRUCTURE

Evaluation of the deduced amino acid sequence revealed striking similarities between EFT-1 and EF-2, namely: 38% identity with the hamster EF-2 (Kohno et al., 1986), Drosophila EF-2 (Grinblat et al., 1989) or CeEF-2 and an mRNA species of similar size (3 kb) encoding a polypeptide of 849 amino acid residues (Mr 96,151). The regions of highest similarity between EFT-1 and eukaryotic EF-2 are in the G1-G5 regions (implicated in GTP binding and GTPase activity in EF-2, Nilsson and Nygard, 1985; Kohno et al., 1986) and in the E1-E4 regions which are conserved among elongation factors and may be involved in direct ribosome binding (Nilsson and Nygard, 1985; Kohno et al., 1986; Grinblat et al., 1989), as well as the region immediately following the modifiable histidine (Figs. 10 and 12). The regions G1, E1, G2, and G5 designated G1, G2, G3, and G4, respectively, in $p21^{ras}$ are

critical in GDP/GTP exchange, GTP-induced conformational change, and GTP hydrolysis (Bourne et al., 1991). It is tempting to speculate that these regions may serve similar functions in EFT-1. Interestingly, the regions of least similarity with EF-2 (amino acid positions 90-100 and 230-260 of HamEF2, Fig. 10) were shown to have diverged considerably in sequence and size between hamster and Drosophila EF-2 (Grinblat et al., 1989). The second splice site for eft-1 (Figs. 8 and 10), Drosophila EF-2, and eft-2 mRNAs (Grinblat et al., 1989) occurs in the latter region. EFT-1, however, lacks the modifiable histidyl residue which appears to be important for the function of EF-2, especially in the recognition of the appropriate ribosomal site for the translocation step (Kohno and Uchida, 1987). This result and the relatively low degree (38%) of amino acid sequence identity with the hamster EF-2 (which is greater than 80% identical with Drosophila EF-2; Grinblat et al., 1989) suggests that EFT-1 likely does not function as the C. elegans EF-2 but may represent a closely related protein with a distinct function. Indeed by PCR analysis of C. elegans cDNA (using primers constructed from highly conserved regions in the hamster and Drosophila sequences), a fragment encoding part of the C. elegans EF-2 including the ADP-ribosylatable histidine was isolated. In this region (Fig. 12), EFT-1 and the C. elegans EF-2 fragment are 57% and 87% identical, respectively, with the hamster EF-2 sequence.

Conservation of the functional regions of elongation factors, in particular EF-2, in EFT-1 suggests that the genes encoding these proteins are at least in part derived from a common ancestor. In light of the highly divergent region preceding the modifiable histidyl residue, the highly (>80%) conserved 15 amino acid sequence immediately following this residue is very unlikely to have evolved by accident in EFT-1. Of particular interest is the report that a single point mutation two codons away from the ADP-ribosylatable histidyl residue which results in the substitution of arginine for glycine in EF-2, is sufficient to confer

resistance to ADP-ribosylation by DT or PA. Glycine is found in the corresponding position in EFT-1 just as in the ADP-ribosylatable EF-2 (Fig. 12). It is interesting also that the amino acid corresponding to Gly12 in ras proteins is replaced by isoleucine in EFT-1 (Fig. 11A). This substitution in ras proteins strongly enhances their transforming ability (Seeburg et al., 1984). Replacement of this amino acid by valine or isoleucine in elongation and initiation factors has been reported (Kohno et al., 1986).

Taken as a whole, the data suggest that EFT-1 may be a GTP-binding protein with other conserved domains whose functions may be similar to the corresponding domains in EF-2.

D. EFT-2 GENE STRUCTURE

By PCR analysis employing rapid amplification of cDNA ends (RACE, Frohman et al., 1988), we have amplified a region of a C. elegans EF-2 gene which is highly conserved in Drosophila and hamster EF-2. Using this fragment as a probe, a number of cDNA and genomic clones encoding C. elegans EF-2 (CeEF-2) were isolated. The identity of the clones was established by DNA sequence analysis and comparison with the known hamster EF-2 (Kohno et al., 1986) and Drosophila EF-2 (Grinblat et al., 1989) sequences. The sequence AGAGAACCC was found 16 nucleotides upstream from the initiator ATG codon in the 5'-untranslated region (Fig. 16). A similar sequence with the consensus A/TGAGAAT/CCC is found 16 to 20 nucleotides upstream of the initiator ATG codon of many eukaryotic EF-2 genes including hamster and rat, and in the archaebacterium H. halobium (Kohno et al., 1986; Grinblat et al., 1989; Itoh , 1989). Its function is unknown. The intron boundary sequences agree with the consensus sequences for 5' and 3' splice sites (Mount, 1982, Frendewey and Keller, 1985, Blumenthal and Thomas, 1988). The small introns (48 and 44 base pairs) are in the size range of 64% of the introns characterized in C. elegans

genes (Fields, 1990). Two potential polyadenylation signals (AATAAA; Proudfoot and Brownlee, 1976) were found in the 3'-untranslated region of pCef6A. The second sequence was located 12 nucleotides upstream from the poly(A)tail of this clone, and the first one was found 52 nucleotides further upstream (Fig. 16). The isolation of cDNA clones which were polyadenylated 14 nucleotides from the first signal suggests that both signals are functional in eft-2. Although multiple AATAAA sequences have also been reported for Drosophila (Grinblat et al., 1989), rat and hamster EF-2 genes (Nakanishi et al., 1986), only one signal has so far been shown to be functional in each case. It is not known whether the 2.8 kb insert of pCef6A with 45 bp of 5'-untranslated region encodes the entire eft-1 mRNA. In an attempt to isolate cDNA clones with inserts extending to or beyond the 5' residue of pCef6A, the cDNA library was screened with a probe from the 5' end of pCef6A. The clone (out of 11 positive clones isolated) with the longest insert, pCef2h, encoded CeEF-2 lacking the first 54 amino acid residues.

Not much is known about the promoters of C. elegans housekeeping genes. Nakanishi et al., (1988) have characterized the hamster EF-2 gene (which is a housekeeping gene) and shown that its promoter is highly efficient, is not tissue-specific, and retains high activity in all mammalian cells. Presumably, the promoter of the C. elegans EF-2 gene is also efficient as judged from the abundance of eft-2 mRNA (Fig. 18).

E. EFT-2 PRIMARY STRUCTURE

The deduced amino acid sequence of CeEF-2 shows greater than 80% identity with the hamster and Drosophila sequences (Fig. 10). A high degree of conservation was found in domains G1 - G5 (implicated in GTP-binding and GTPase activity), E1 - E4 (shared among elongation factors and possibly involved in ribosome binding), and in the modifiable histidine region (Figs. 11 and 12), all of which appear to be important for the function of EF-2 (Kohno

et al., 1986; Grinblat et al., 1989; Nilsson and Nygard, 1985; Kohno and Uchida, 1987). These facts strongly suggest that CeEF-2 represents a functional EF-2 in C. elegans. Amino acid insertions and deletions in CeEF-2 occur at regions (residues 90-100 and 237-250) which have diverged considerably between the hamster and Drosophila proteins (Grinblat et al., 1989) and in EFT-1. A third divergent region was found between residues 280 and 290 of the hamster sequence (Fig. 10), where the C. elegans EF-2 is 4 amino acids shorter and EFT-1 is 3 amino acids shorter relative to hamster EF-2.

Neither of the two introns so far identified in the C. elegans EF-2 gene (eft-2) correspond in position to those of the hamster gene (Nakanishi et al., 1988). The first intron of eft-2 (following Gln-191 of CeEF-2) is located in a highly conserved region of EF-2 proteins, while the second (following Gln-250) occurs in a highly variable region (residues 237-260 of hamster EF-2, Fig. 10), where the second splice site of eft-1 mRNA also occurs.

F. EVOLUTIONARY AND FUNCTIONAL RELATIONSHIPS

To determine evolutionary relationships by comparative analysis of molecular sequences, the sequences must change but slowly with time, so that correct relationships among very distant species can be determined. There appears to be extensive identity of amino acid sequence in most translational factors characterized to date in that within mammalian species, usually greater than 99% identity is observed, in contrast with the conservation of amino acid sequences in the α or β haemoglobin chains which range from 75 to 90% identity between other mammals and humans. Extreme examples are rabbit EF-1α which is 100% identical to human EF-1α and rabbit eIF-4AI and eIf-4AII which are 100% identical to the corresponding mouse factors in the amino acids sequenced whereas a comparison of the primary structure of EF-1α to EF-Tu indicates an overall sequence

identity of 33%; however, within the amino-terminal 180 amino acids (the GTP-binding domain), there are found regions of much greater (59%) identity (Merrick et al., 1990). Also among all eukaryotes (including rat, hamster, Drosophila, and C. elegans) so far studied, identity in the amino acid sequence of EF-2 is greater than 80%. The high degree of conservation of the translational factor sequences among different organisms is presumed to reflect not only conservation of active sites (nucleotide, tRNA, or mRNA binding) but also exterior surfaces which allow for the complex factor/factor and factor/ribosome interactions which occur during the various steps of protein synthesis. In a similar sense, one anticipates that proteins involved in other macromolecular processes such as DNA replication, RNA transcription or mRNA processing are likely to also have quite highly conserved sequences (Merrick et al., 1990). A protein of similar function to EFT-1 from any eukaryotic source would be expected to share amino acid identity with the C. elegans protein. Interestingly, the amino acid insertions and deletions in CeEF-2 occur at highly divergent regions between the hamster and Drosophila proteins and where there is low homology between these proteins and EFT-1; the second intron of eft-1 and eft-2 is located in one of these regions. In another divergent region in all eukaryotic EF-2 the C. elegans EF-2 and Drosophila EF-2 were each 4 amino acids shorter and the former showed more similarity to the Drosophila than to the hamster or EFT-1 proteins. EFT-1 was also found to be 3 amino acids shorter relative to hamster EF-2 in this region.

G. EFT-1 AND EFT-2 GENE EXPRESSION DURING DEVELOPMENT

Northern blot analyses showed that eft-1 and eft-2 each expresses a single mRNA species 3 kilobases in length (Figs. 13 and 18). This is similar to the size of the EF-2 mRNA from hamster (Kohno and Uchida, 1987), Drosophila (Grinblat et al., 1989), human (Rapp et al., 1989), and the slime mold D. discoideum (Toda et al., 1989). eft-1 and eft-2

transcripts were present at all stages of nematode development as is the case for Drosophila (Grinblat et al., 1989), but at somewhat reduced levels in L1 and dauer larvae. The low level detected in the dauer larval stage seem to have been due to selective degradation of mRNA, since the level of actin mRNA was also low in this sample, while the loadings of ribosomal RNA were comparable for all stages. There was no appreciable decrease in eft-2 mRNA levels at the adult stage. Toda et al. (1989) have reported that the expression of EF-2 mRNA in D. discoideum is high in vegetative cells, becomes maximal at the aggregation stage, and decreases thereafter through development. In Drosophila EF-2 mRNA first becomes detectable by 4 hours of embryogenesis and persists throughout development and into adulthood. It increases somewhat in abundance during the late embryonic, late larval, and early pupal stages of active organogenesis (Grinblat et al., 1989).

In general, the eft-2 mRNA was relatively more abundant than the eft-1 mRNA at all stages, as judged by the intensity of the signal on Northern blots, i.e. larger amounts of RNA and longer exposure times are required to detect the eft-1 transcript. This conclusion is supported by the higher representation of clones encoding the eft-2 transcript in the cDNA library studied. Nakanishi et al. (1988) have reported that the content of the hamster EF-2 transcript is high (approximately 0.1% of total mRNA) in growing mammalian cultured cells. The developmental profile of expression of eft-1 mRNA indicates that the function of EFT-1 may be required throughout C. elegans development. Whether the relatively low level of expression of EFT-1 relative to EF-2 is due to transcriptional or post-transcriptional processes remains to be determined.

H. EFT-1 AND EFT-2 GENE COPY NUMBER

Genomic Southern blot analyses showed that the band obtained with HindIII digested C. elegans DNA was the same size as the probe B3255 (Fig. 19, lane 16), thus confirming

the results with the cDNA probe (Fig. 4A); these data and the genomic Southern blot analysis of the closely related species (C. briggsae) DNA suggest that eft-1 is a single copy gene in both species. A similar analysis using pCef6A as probe showed that eft-2 is also unique in C. elegans and C. briggsae genomes. The restriction fragment patterns obtained with the two gene probes were different, thus confirming that the two genes are distinct and did not cross-react under the hybridization conditions used for the analysis. Also, whereas eft-1 was isolated from chromosome III, eft-2 was localized to chromosome I of C. elegans. The EF-2 gene has been shown to be unique in other eukaryotic organisms examined, including hamster (Kohno and Uchida, 1987), Drosophila (Grinblat et al., 1989), and mouse (Koide et al., 1990). In addition, mouse cells contain about 70 copies of amplified EF-2-related sequences or processed pseudogenes, called MERs, per haploid genome. The genes are thought to have been generated by the integration of one copy of MER, derived from the mature, polyadenylated mRNA for EF-2, into the mouse genome and subsequent amplification of the MER and flanking sequences (Koide et al., 1990).

I. POTENTIAL AREAS OF FUTURE STUDY

The genomic Southern analysis showed that eft-1 exists in a closely related species C. briggsae. In order to elucidate the function of EFT-1, it would be of interest to determine whether similar proteins exist in other eukaryotic systems. By analogy with the situation in C. elegans, such a protein might be expected to resemble nematode EFT-1 more closely than EF-2 from the same or other species. Immunolocalization experiments, as well as in vitro and in vivo functional tests, such as an attempt to complement the function of EF-2 in protein synthesis, should also be informative. Furthermore, direct structure-function correlations by a systematic characterization of mutants will offer further possibilities for analysis of the functional and structural organization of the protein. In this way, the

significance of the substitution of a tyrosine in EFT-1 for the histidyl residue which is supposed to be of importance in EF-2 function could be investigated.

It would be interesting to know if eft-1 is also subject to post-translational modification and as such, if this modification is a regulatory step in any biochemical process.

Not much is known about the regulation of transcription from constitutively expressed genes. The C. elegans EF-2 gene promoter would be a good model system from which to characterize regulatory factors which interact with constitutively expressed genes. In addition, the function of the sequence A/TGAGAAT/CCC found in the 5'-untranslated region (20 to 16 nucleotides upstream of the initiator ATG codon) of many eukaryotic EF-2 genes needs to be investigated.

V. REFERENCES

1. Albertson, D.G. 1985. EMBO J. 4, 2493-2498.

2. Antonucci, T.K. 1985. Recombinant DNA Techniques 6, 22-24, University of Michigan.

3. Arai, K., B.F.C. Clark, L. Duffy, M.D. Jones, D. Kaziro, R.A. Laursen, J. L'Italien, D.L. Miller, S. Nagarkatti, S. Nakamura, K.M. Nielsen, T.E. Petersen, K. Takahashi, and M. Wade. 1980. Proc. Natl. Acad. Sci. USA 77, 1326-1330.

4. Atkinson, T., and M. Smith. 1984. In Oligonucleotide Synthesis, a Practical Approach (M.J. Gait, ed.) pp. 35-81, IRL Press Ltd., Oxford.

5. Ballinger, D.G., and M.L. Pardue. 1983. Cell 33, 103-114.

6. Barrieux, A., and M.G. Rosenfeld. 1977. Biochemistry 16, 514-518.

7. Barstead, R.J. and R.H. Waterston. 1989. J. Biol. Chem. 264, 10177-10185.

8. Benne, R., and J.W.B. Hershey. 1976. Proc. Natl. Acad. Sci. USA 73, 3005-3009.

9. Benne, R., J. Edman, R.R. Traut and J.W.B. Hershey. 1978. Proc. Natl. Acad. Sci. USA 75, 108-112.

10. Benton, W.D. and R.W. Davis. 1977. Science 196, 180-182.

11. Berk, A.J., and P.A. Sharp. 1977. Cell 12, 721-732.

12. Birnboim, H.C., and J. Doly. 1979. Nucleic Acid Res. 7, 1513-1523.

13. Blumenthal, T., and J. Thomas. 1988. Trends Genet. 4, 305-308.

14. Bollini, R., A.N. Soffientini, A. Bertani, and G.A. Lanzani. 1974. Biochemistry 13, 5431-5445.

15. Bourne, H.R., D.A. Sanders, and F. McCormick. 1991. Nature 349, 117-126.

16. Brenner, S. 1974. Genetics 77, 71-94.

17. Brown, B.A., and J.W. Bodley. 1979. FEBS Lett. 103, 253-255.

18. Capon, D.J., E.Y. Chen, A.D. Levinson, P.H. Seeburg, and D.V. Goeddel. 1983. Nature 302, 33-37.

19. Carlberg, U., A. Nilsson, and O. Nygard. 1990. Eur. J. Biochem. 191, 639-645.

20. Caskey, C.T. 1977. In The Mechanisms of Protein Biosynthesis, (H. Weissbach, and S. Pestka, eds.) pp. 443-465, Academic Press, New York.

21. Cassada, R.C., and R.L. Russell. 1975. Dev. Biol. 46, 326-342.

22. Cavallius, J., S.I.S. Rattan, and B.F.C. Clark. 1986. Exp. Gerontol. 21, 149-157.

23. Celis, J.E., P. Madsen, and A.G. Ryazanov. 1990. Proc. Natl. Acad. Sci. USA 87, 4231-4235.

24. Chow, S., E. Daub, and H. Murialdo. 1987. Gene 60, 277-289.

25. Chuang, D. M. and H. Weissbach. 1972. Arch. Biochem. Biophys. 152, 114-124.

26. Clark, D.V., T.M. Rogalski, L.M. Donati, and D.L. Baillie. 1988. Genetics 119, 345-353.

27. Collier, R.J. 1975. Bacteriol. Rev. 39, 54-85.

28. Collier, R.J., and A.M. Pappenheimer Jr. 1964. J. Exp. Med. 120, 1019-1039.

29. Collins, J.F., S. Raeburn, and E.S. Maxwell. 1971. J. Biol. Chem. 246, 1049-1055.

30. Comstock, J.P., and N.T. Van. 1977. Biochem. Biophys. Acta 477, 199-220.

31. Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon. 1980. Science 209, 1406-1414.

32. Coulson, A., J. Sulston, S. Brenner, and J. Karn. 1986. Proc. Natl. Acad. Sci. USA 83, 7821-7825.

33. Coulson, A., R. Waterston, J. Kiff, J. Sulston, and Y. Kohara. 1988. Nature 335, 184-186.

34. Cruzen, M., and T. Johnson. 1987. Worm Breeders Gazette 10, 141.

35. Dalgarno, D.C., M.W. Harding, A. Lazarides, R.E. Handschumacher, and I.M. Armitage. 1986. Biochemistry 25, 6778-6784.

36. Dasmahapatra, B., and K. Chakraburtty. 1981. J. Biol. Chem. 256, 9999-10004.

37. Davie, E.W., V.V. Koningsberger, and F. Lipmann. 1956. Arch. Biochem. Biophys. 65, 21-28.

38. Davydova, E.K., A.S. Sitikov, and L.P. Ovchinnikov. 1984. FEBS Lett. 176, 401-405.

39. Dayhoff, M.O. ed. 1978. National Biomedical Research Foundation, Washington, D.C.

40. Dunlop, P.C., and J.W. Bodley. 1983. J. Biol. Chem. 258, 4754-4758.

41. Dynan, W.S. 1986. Trends Genet. 2, 196-197.

42. Ejiri, S., and H. Honda. 1985. Biochem. Biophys. Res. Commun. 128, 53-60.

43. Emmons, S.W., M.R. Klass, and D. Hirsh. 1979. Proc. Natl. Acad. Sci. USA 76, 1333-1337.

44. Fendrick, J.L., and W.J. Iglewski. 1989. Proc. Natl. Acad. Sci. USA 86, 554-557.

45. Fessenden, J.M., and K. Moldave. 1961. Biochem. Biophys. Res. Commun. 6, 232-235.

46. Fessenden, J.M., and K. Moldave. 1963. J. Biol. Chem. 238, 1479-1484.

47. Fields, C. 1990. Nucleic Acids Res. 18, 1509-1512.

48. Filipowicz, W., J.M. Sierra, and S. Ochoa. 1975. Proc. Natl. Acad. Sci. USA 72, 3947-3951.

49. Fire, A. 1986. EMBO J. 5, 2673-2680.

50. Fischer, I., S.M. Arfin, and K. Moldave. 1980. Biochemistry 19, 1417-1425.

51. Fonzi, W.A., C. Katayama, T. Leathers, and P.S. Syphered. 1985. Mol. Cell Biol. 5, 1100-1103.

52. Freinstein, C., and G. Blobel. 1975. Proc. Natl. Acad. Sci. USA 72, 3392-3396.

53. Frendewey, D., and W. Keller. 1985. Cell 42, 355-367.

54. Frohman, M.A., M.K. Dush, and G.R. Martin. 1988. Proc. Natl. Acad. Sci. USA 85, 8998-9002.

55. Gaillard, C., and F. Strauss. 1990. Nucleic Acids Res. 18, 378.

56. Galasinski, W., and K. Moldave. 1969. J. Biol. Chem. 244, 6527-6532.

57. Gasior, E., and K. Moldave. 1965. J. Biol. Chem. 240, 3346- 3352.

58. Gehrke,L., R.E. Bast, and J. Ilan. 1981. J. Biol. Chem. 256, 2522-2530.

59. Gehrmann, R., A. Henschen, and F. Klink. 1985. FEBS Lett. 185, 37-42.

60. Geysen, H.M., T.J. Mason, and S.J. Rodda. 1988. J. Mol. Recognition 1, 32-41.

61. Gill, D.M., A.M. Pappenheimer Jr., R. Brown, and J.J. Kurnick. 1969. J. Exp. Med. 129, 1-21.

62. Giovane, A., L. Servillo, L. Quagliuolo, and C. Balestrieri. 1987. Biochem. J. 244, 337-344.

63. Girshovich, A.S., E.S. Bochkareva, T.V. Kurtskhalia, V.A. Pazdnyakov, and Y.A. Ovchinnikov. 1979. Methods Enzymol. 60, 726-745.

64. Graham, R.W., D. Jones, and E.P.M. Candido. 1989. Mol. Cell. Biol. 9, 268-277.

65. Granknowski, N., D. Lehmusvirta, G. Kramer, and B. Hardesty. 1980. J. Biol. Chem. 255, 310-317.

66. Grinblat,Y., N.H. Brown, and F. Kafatos. 1989. Nucleic Acids Res. 17, 7303-7313.

67. Grossi, L.G., and K. Moldave. 1959. Biochim. Biophys. Acta 35, 275-277.

68. Grossi, L.G., and K. Moldave. 1960. J. Biol. Chem. 235, 2370-2374.

69. Gschwendt, M., W. Kittstein, and F. Marks. 1988. Biochem. Biophys. Res. Commun. 150, 545-551.

70. Gschwendt, M., W. Kittstein, and F. Marks. 1988. Biochem. J. Lett. 256, 1061.

71. Gupta, R.S., and L. Siminovitch. 1978. Somat. Cell Genet. 4, 553-571.

72. Halliday, K.R. 1984. J. Cyclic Nucleotide Protein Phosphorylation Res. 9, 435-448.

73. Hanahan, D. 1983. J. Mol. Biol. 166, 557-580.

74. Handschumacher, R.E., M.W. Harding, J. Rice, and R.J. Drugge. 1984. Science 226, 544-546.

75. Hassell, J.A., and D.L. Engelhardt. 1976. Biochemistry 15, 1375-1381.

76. Hattori, M. and Y. Sakaki. 1986. Anal. Biochem. 152, 232-238.

77. Henikoff, S. 1984. Gene 28, 351-359.

78. Henriksen, O., E.A. Robinson, and E.S. Maxwell. 1975. J. Biol. Chem. 250, 720-724.

79. Hershey, J.W.B. 1980. In Cell Biology: A Comprehensive Treatise, (D.M. Prescott and L. Goldstein, eds.) Vol. 4, pp. 1-68, Academic Press, New York.

80. Hoagland, M.B., M. L. Stephenson, J.F. Scott, L.I. Hecht, and P.C. Zamecknik. 1958. J. Biol. Chem. 231, 241-257

81. Honjo, T., Y. Nishizuka, O. Hayashi, and I. Kato. 1968. J. Biol. Chem. 243, 3553-3555.

82. Hovemann, B., S. Richter, U. Walldorf, and C. Cziepluch. 1988. Nucleic Acids Res. 16, 3175-3194.

83. Huet, J., P. Cottrelle, M. Cool, M.-L. Vignais, D. Thiele, C. Marck, J.-M. Buhler, A. Sentenac, and P. Fromageot. 1985. EMBO J. 4, 3539-3547.

84. Hughes, S.M. 1983. FEBS Lett. 164, 1-8.

85. Ibuki, F., and K. Moldave. 1968. J. Biol. Chem. 243, 791-798.

86. Iglewski, B.H. and D. Kabat. 1975. Proc. Natl. Acad. Sci. USA 72, 2284-2288.

87. Ilan, J., and J. Ilan. 1976. J. Biol. Chem. 251, 5718-5725.

88. Ish-Horowicz, D., and J.F. Burke. 1981. Nucleic Acid Res. 9, 2989-2998.

89. Issinger, O.G., R. Benne, J.W.B. Hershey, and R.R. Traut. 1976. J. Biol. Chem. 251, 6471-6474.

90. Itoh, T. 1989. Eur. J. Biochem. 186, 213-219.

91. Iwasaki, K., K. Motoyoshi, S. Nagata, and Y. Kaziro. 1976. J. Biol. Chem. 251, 1843-1845.

92. Jagus, R., D. Crouch, A. Konieczny, and B. Safer. 1982. Current Topics in Cellular Regulation 21, 35-63.

93. Janssen, G.M.C., G.D.F. Maessen, R. Amons, and W. Moller. 1988. J. Biol. Chem. 263, 11063-11066.

94. Jones, D., R.H. Russnak, R.J. Kay, and E.P.M. Candido. 1986. J. Biol. Chem. 261, 12006-12015.

95. Kaempfer, R., R. Hollender, W.R. Abrams, and R. Israeli. 1978. Proc. Natl. Acad. Sci. USA 75, 209-213.

96. Kamath, A. and K. Chakraburtty. 1989. J. Biol. Chem. 264. 15423-15428.

97. Kaneda, Y., M.C. Yoshida, K. Kohno, T. Uchida, and Y. Okada. 1984. Proc. Natl. Acad. Sci. USA 81, 3158-3162.

98. Kay, M.A., and M. Jacobs-Lorena. 1987. Trends Genet. 3, 347-351.

99. Kaziro, Y. 1978. Biochim. Biophys. Acta 505, 95-127.

100. Keller, E.B., and P.C. Zamecnik. 1956. J. Biol. Chem. 221. 45-59.

101. Kessel, M., and F. Klink. 1980. Nature 287, 250-251.

102. Kigoshi, T., K. Uchida, and S. Morimoto. 1989. J. Steroid Biochem. 32, 381-385.

103. Koch, I., P.H. Hofschneider, F. Lottspeich, C. Eckerskorn, and R. Koshy. 1990. Oncogene 5, 839-843.

104. Kohno, K., and T. Uchida. 1987. J. Biol. Chem. 262, 12298-12305.

105. Kohno, K., T. Uchida, E. Mekada, and Y. Okada. 1985. Somatic Cell Mol. Genet. 11, 421-431.

106. Kohno, K., T. Uchida, H. Ohkubo, S. Nakanishi, T. Nakanshi, T. Fukui, E. Ohtsuka, M. Ikehara, and Y. Okada. 1986. Proc. Natl. Acad. Sci. USA 83, 4978-4982.

107. Koide, T., M. Ishiura, N. Hazumi, T. Shiroishi, Y. Okada, and T. Uchida. 1990. Genomics 6, 80-88.

108. Kongsuwan, K., Q. Yu, A. Vincent, M.C. Frisardi, M. Roshbash, J.A. Lengyel, J. Merriam. 1985. Nature 317, 555-558.

109. Kramer, G., A.B. Henderson, P. Pinphanichakarn, M.H. Wallis, and B. Hardesty. 1977. Proc. Natl. Acad. Sci. 74, 1445-1449.

110. Kramer, G., and B. Hardesty. 1980. In Cell Biology: A Comprehensive Treatise, (D.M. Prescott and L. Goldstein, eds.) Vol. 4, pp. 69-105, Academic Press, New York.

111. la Cour, Y.F.M., J. Nyborg, S. Thirup, and B.F.C. Clark. 1985. EMBO J. 4, 2385-2388.

112. Lawn, R.M., A. Efstratiadis, C. O'Coonnell, and T. Maniatis. 1980. Cell 21, 647-651.

113. Lechner, K., G. Heller, and A. Bock. 1988. Nucleic Acids Res. 16, 7817-7826.

114. Levenson, R.M., E.V. Maytin, and D.A. Young. 1989. Anal. Biochem. 158, 294-301.

115. Levin, D.H., D. Kyner, and G. Acs. 1973. Proc. Natl. Acad. Sci. USA 70, 41-45.

116. Lindahl, L., and J.M. Zengel. 1986. Annu.Rev. Genet. 20, 297-326.

117. Linz, J.E., C. Katayama, and P.S. Sypherd. 1986. Mol. Cell. Biol. 6, 593-600.

118. Lloyd, M.A., J.R. Osborne, B. Safer, G.M. Powell, and W.C. Merrick. 1980. J. Biol. Chem. 255, 1189-1193.

119. Marzouki, A., J.P. Lavergne, J.P. Reboud, and A.M. Reboud. 1989. FEBS Lett. 255, 72-76.

120. McDonald, J.D., F.K. Lin, and E. Goldwasser. 1986. Mol. Cell. Biol. 6, 842-848.

121. McGrath, J.P., D.J. Capon, D.V. Goeddel, and A.D. Levinson. 1984. Nature (London) 310, 644-649.

122. McLauchlan, J., D. Gaffney, J.L. Whitton, and J.B. Clements. 1985. Nucleic Acids. Res. 13, 1347-1368.

123. Melton, D.W., C. McEwan, A.B. McKie, A. M. Reid. 1986. Cell 44, 319-328.

124. Merrick, W.C., T.E. Dever, T.G. Kinzy, S.C. Conroy, J. Cavallius, and C.L. Owens. 1990. Biochim. Biophys. Acta 1050, 235-240.

125. Merrick, W.C., W.M. Kemper, and W.F. Anderson. 1975. J. Biol. Chem. 250, 2620-2625.

126. Messing, J. 1983. Meth. Enzymol. 101, 20-78.

127. Mizumoto, K., K. Iwasaki, M. Tanaka, Y. Kaziro. 1974. J. Biochem. 75, 1047-1056.

128. Moazed, D., and H.F. Noller. 1989. Nature 342, 142-148.

129. Moehring, T.J., and J.M. Moehring. 1977. Cell 11, 447-454.

130. Moerman, D.G., G.M. Benian, and R.H. Waterston. 1986. Proc. Natl. Acad. Sci. USA 83, 2579-2583.

131. Moldave, K. 1985. Annu. Rev. Biochem. 54, 1109-1149.

132. Monier, S., and Y. LeMarchand-Brustel. 1982. FEBS Lett. 147, 211-214.

133. Montanaro, L., S. Sperti, and A. Mattioli. 1971. Biochim. Biophys. Acta 238, 493-497.

134. Montandon, P.E., and E. Stutz. 1983. Nucleic Acids Res. 11, 5877-5892.

135. Montiminy, M.R., K.A. Sevarino, J.A. Wagner, G. Mandel, and R.H. Goodman. 1986. Proc. Natl. Acad. Sci. 83, 6682-6686.

136. Mount, S. 1982. Nucleic Acids Res. 10, 459-472.

137. Murialdo, H., A. Davidson, S. Chow, and M. Gold. 1987. Nucleic Acids Res. 15, 119-140.

138. Murialdo, H., W.L. Fife, A. Becker, M. Feiss, and J. Yochem. 1981. J. Mol. Biol. 145, 375-404.

139. Nagata, S., K. Iwasaki, and Y. Kaziro. 1977. J. Biochem. 82, 1633-1646.

140. Nagata, S., K. Nagashima, Y. Tsunetsugu-Yokota, K. Fujimura, M. Miyazaki, and Y. Kaziro. 1984. EMBO J. 3, 1825-1830.

141. Nairn, A.C., and H.C. Palfrey. 1987b. J. Biol. Chem. 262, 17299-17303.

142. Nairn, A.C., B. Bhagat, and H.C. Palfrey. 1985. Proc. Natl. Acad. Sci. USA 82, 7939-7943.

143. Nairn, A.C., R.A. Nichols, M.J. Brady, and H.C. Palfrey. 1987a. J. Biol. Chem. 262, 14265-14272.

144. Nakanishi, T., K. Kohno, M. Ishiura, H. Ohashi, and T. Uchida. 1988. J. Biol. Chem. 263, 6384-6391.

145. Nevins, J.R. 1983. Annu. Rev. Biochem. 52, 441-466.

146. Nielsen, P.J., and E.H. McConkey. 1980. J. Cell. Physiol. 104, 269-281.

147. Nilsson, L. and O. Nygard. 1985. Eur. J. Biochem. 148, 299-304.

148. Nilsson, L., and O. Nygard . 1984. Biochim. Biophys. Acta 782, 49-54.

149. Nomura, M., R. Gourse, and G. Baughman. 1984. Annu. Rev. Biochem. 53, 75-117.

150. Nonaka, M., H. Kimura, Y.D. Yeul, S. Yokoyama, K. Nakayama, and M. Takahashi. 1986. Proc. Natl. Acad. Sci. USA 83, 7883-7887.

151. Nudel, U., B. Lebleu, and M. Revel. 1973. Proc. Natl. Acad. Sci. USA 70, 2139-2144.

152. Nygard, O., and L. Nilsson. 1985. Biochim. Biophys. Acta 824, 152-162.

153. Nygard, O., and L. Nilsson. 1989. Eur. J. Biochem. 179, 603-608.

154. Odom, O.W., G. Kramer, A.B. Henderson, P. Pinphanichakarn, and B. Hardesty. 1978. J. Biol. Chem. 253, 1807-1816.

155. Oleinikov, A.V., G.G. Jokhadze, and Y.B. Alakhov. 1989. FEBS Lett. 248, 131-136.

156. Pain, V.M. 1986. Biochem. J. 235, 625-637.

157. Palfrey, H.C. 1983. FEBS Lett. 157, 183-190.

158. Palmiter, R.D. 1972. J. Biol. Chem. 247, 6770-6780.

159. Panniers, R., and E.C. Henshaw. 1983. J. Biol. Chem. 258, 7928-7934.

160. Pappenheimer, A.M. 1977. Annu. Rev. Biochem. 46, 69-94.

161. Pappenheimer, A.M., P.C. Dunlop, K.W. Adolph, and J.W. Bodley. 1983. J. Bact. 153, 1342-1347.

162. Park, J., H.V. Hershey, and M.W. Taylor. 1986. In: Molecular Genetics of Mammalian Cells: a Primer in Developmental Biology. (G.M. Malacinski, C.C. Simonsen, and M. Shepard. eds.) pp. 79-98. Macmillan, New York.

163. Pearson, W.R., and D.J. Lipman. 1988. Proc. Natl. Acad. Sci. USA 85, 2444-2448.

164. Rackwitz, H.R., G. Zehetner, A. Frischauf, and H. Lehrach. 1984. Gene 30, 195-200.

165. Rackwitz, H.R., G. Zehetner, G. Murialdo, H. Delius, J.H. Chai, A. Poustka, A. Frischauf, and H. Lehrach. 1985. Gene 40, 259-266.

166. Radloff, R., W. Bauer, J. Vinograd. 1967. Proc. Natl. Acad. Sci. USA 57, 1514.

167. Raeburn, S., R.S. Goor, J.A. Schneider, and E.S. Maxwell. 1968. Proc. Natl. Acad. Sci. USA 61, 1428-1434.

168. Rapp, G., J. Klaudiny, G. Hagendorff, and M.R. Luck. 1989. Biol. Chem. Hoppe-Seyler 370, 1071-1075.

169. Redpath, N.T., and C.G. Proud. 1989. Biochem. J. 262, 69-75.

170. Richter, D., and F. Lipmann. 1970. Nature 227, 1212-1214.

171. Riis, B., S.I.S. Rattan, A. Derventzi, and B.F.C. Clark. 1990. FEBS Lett. 266, 45-47.

172. Roberts, L. 1990. Science 248, 1310-1313.

173. Robinson, E.A., O. Henriksen, and E.S. Maxwell. 1974. J. Biol. Chem. 249, 5088-5093.

174. Ryazanov, A.G. 1987. FEBS Lett. 214, 331-334.

175. Ryazanov, A.G., E.A. Shestakova, and P.G. Natapov. 1988a. Nature (London) 334, 170-173.

176. Ryazanov, A.G., P.G. Natapov, E.A. Shestakova, F.F. Severin, and A.S. Spirin. 1988b. Biochimie 70, 619-626.

177. Sacerdot, C., P. Dessen, J.W.B. Hershey, J.A. Plumbridge, and M. Grunberg-manago. 1984. Proc. Natl. Acad. Sci. USA 81, 7787-7789.

178. Safer, B., and R. Jagus. 1981. Biochimie 63, 709-717.

179. Safer, B., S.L. Adams, W.F. Anderson, and W.C. Merrick. 1975. J. Biol. Chem. 250, 9076-9075.

180. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

181. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. Proc. Natl. Acad Sci. USA 74, 5463-5467.

182. Schneir, M., and K. Moldave. 1968. Biochim. Biophys. Acta 166, 58-67.

183. Schreier, M.H., B. Erni, and T. Staehelin. 1977. J. Mol. Biol. 116, 727-753.

184. Seeburg, P.H., W.W. Colby, D.J. Capon, D.V. Goeddel, and A.D. Levinson. 1984. Nature 312, 71-75.

185. Shafritz, D.A., J.A. Weinstein, B. Safer, W.C. Merrick, L.A. Weber, E.D. Hickey, and C. Baglioni. 1976. Nature 261, 291-294.

186. Sitikov, A.S., P.N. Simonenko, E.A. Shestakova, A.G. Ryazanov, and L.P. Ovchinnikov. 1988. FEBS Lett. 228, 327-331.

187. Skogerson, L. and K.Moldave. 1968. J. Biol. Chem. 243, 5354-5360.

188. Skogerson, L., and D. Engelhardt. 1977. J. Biol. Chem. 252, 1471-1475.

189. Skogerson, L., and E. Wakatama. 1976. Proc. Natl. Acad. Sci. USA 73, 73-76.

190. Slobin, L.I. 1980. Eur. J. Biochem. 110, 555-563.

191. Slobin, L.I., and W. Moller. 1976. Eur. J. Biochem. 69, 351-375.

192. Smith, C.J., C.S. Rubin, and O.M. Rosen. 1980. Proc. Natl. Acad. Sci. USA 77, 2641-2645.

193. Southern, E.M. 1975. J. Mol. Biol. 98, 503-518.

194. Sperti, S., L. Montanaro, and A. Mattioli. 1971. Chem. Biol. Interact. 3, 141-148.

195. Sulston, J. E., E. Schierenberg, and J.G. White. 1983. Dev. Biol. 100, 64-119.

196. Sulston, J., and S. Brenner. 1974. Genetics 77, 95-104.

197. Taira, H., S. Ejiri, and K. Shimura. 1972. J. Biochem. (Tokyo) 72, 1527-1535.

198. Takamatsu, K., T. Uchida, Y. Okada. 1986. Biochem. Biophys. Res. Commun. 134, 1015-1021.

199. Tanabe, T., T. Nukada, Y. Nishikawa, K. Sugimoto, H. Suzuki, H. Takahashi, M. Noda, T. Haga, A. Ichiyama, K. Kangawa, N. Minamino, H. Matsuo, and S. Numa. 1985. Nature (London) 315, 242-245.

200. Tanaka, M., K. Iwasaki, and Y. Kaziro. 1977. J. Biochem. (Tokyo) 82, 1035-1043.

201. Theodorakis, N.G., S.G. Banerji, and R.I. Morimoto. 1988. J. Biol. Chem. 263, 14579-14585.

202. Thomas, G. and G. Thomas. 1986. J. Cell. Biol. 103, 2137-2144.

203. Thomas, G., J. Martin-Perez, M. Siegmann, and A.M. Otto. 1982. Cell 30, 235-242.

204. Thompson, H.A., I. Sadnik, J. Scheinbuks, and K. Moldave. 1977. Biochemistry 16, 2221-2230.

205. Toda, K., M. Tasaka, K. Mashima, K. Kohno, T. Uchida, and I. Takeuchi. 1989. J. Biol. Chem. 264, 15489-15493.

206. Towle, C.A., H.J. Mankin, J. Avruch, and B.V. Treadwell. 1984. Biochem. Biophys. Res. Commun. 121, 134-140.

207. Uritani, M., and M. Miyazaki. 1988. J. Biochem. 103, 522-530.

208. Van Ness, B.G., J.B. Howard, and J.W. Bodley. 1978. J. Biol. Chem. 253, 8687-8690.

209. Van Ness, B.G., J.B. Howard, and J.W. Bodley. 1980. J. Biol. Chem. 255, 10710-10716.

210. van Damme, H.T.F., R. Amons, R. Karssies, C.J. Timmers, G.M.C. Janssen, and W. Moller. 1990. Biochim. Biophys. Acta 1050, 241-247.

211. Wada, K., S. Aota, R. Tsuchiye, F. Ishibashi, T. Gojobori, and T. Ikemura. 1990. Nucleic Acids Res. 18, 2367-2411.

212. Walldorf, U., B. Hovemann, and E.K.F. Bautz. 1985. Proc. Natl. Acad. Sci. USA 82, 5795-5799.

213. Walton, G.M., and G.N. Gill. 1976. Biochim. Biophys. Acta 418, 195-203.

214. Webster, G.C. 1985. In: Molecular Biology of Ageing: Gene Stability and Gene Expression. (Sohal, R.S., L. Birnbaum, and R.G. Cutler. eds.) pp. 263-289. Raven, New York.

215. White, J.G., E. Southgate, J.N. Thomson, and S. Brenner. 1986. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314, 1-340.

216. Wierenga, R.K., and W.G.J. Hol. 1983. Nature 302, 842-844.

217. Wigle, D.T., and A.E. Smith. 1973. Nature New Biol. 242, 136-140.

218. Wong, K.K., A. Meister, and K. Moldave. 1960. Biochim. Biophys. Acta 36, 531-533.

219. Wood, W.B. 1988. The Nematode Caenorhabditis elegans, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

220. Wozniak, D.J., L. Hsu, and D.R. Galloway. 1988. Proc. Natl. Acad. Sci. USA 85, 8880-8884.

221. Yamaizumi, M., E. Mekada, T. Uchida, and Y. Okada. 1978. Cell 15, 245-250.

222. Zengel, J.M., R.H. Archer, and L. Lindahl. 1984. Nucleic Acids Res. 12, 2181-2192.

APPENDIX

A. LIST OF OLIGONUCLEOTIDES AND THEIR SEQUENCES

1. Cos complement oligonucleotides

Left cos complement 'L' 5' dAGGTCGCCGCCC 3'

Right cos complement 'R' 5' dGGGCGGCGACCT 3'

2. PCR oligonucleotides

OPC3 5' dGAGGATCCGT(T,C)GC(T,C,A)GGATT(T,C)CA(A,G)TGGGC 3'

OPC4 5' dGAGGATCCAA(A,G)AT(T,C,A)TGGTG(T,C)TT(T,C)GGACC 3'

EO3 5' dGGTCTTCGACGCCGTCATGA 3'

EO5 5' dAGGGATCCGGATAGCGTCAGCGTGGAGG 3'

RACE 5' dCGAGCATGCGTCGACAGGCATTTTTTTTTTTTTTTT 3'

Oligonucleotides were synthesized with 5' restriction enzyme (BamHI) recognition sequences (underlined) in order to facilitate cloning of PCR products.

3. Oligonucleotides for sequencing

EO1 5' dGACTGCAGAAGAAATTCCCCAGAAT 3'

EO2 5' dTTTAGGTACCTGGAGTATCA 3'

EO4* 5' dTCATGATATGAGGGCAGTCC 3'

* Also used to synthesize the M13 single-stranded probe for the S1 nuclease experiment.

*B. SUMMARY OF E. COLI STRAINS AND THEIR GENOTYPES**

BB4	hsdR514 supE44 supF58 galK2 galT22 trpR55 metB1 tonA ΔlacU169 F'[proAB $lacI^q$ lacZΔM15 Tn10(tet^r)]
DH5α	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 ΔlacU169 (ϕ80 lacZΔM15)
HB101	recA13 hsdS20($r_B^-m_B^-$) ara-14 proA2 lacY1 galK2 rpsL20 xyl-5 mtl-1
JM109	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 Δ(lac- proAB) Nal^r F'(traD36 $lacI^q$ lacZΔM15)
Q358	hsdR supE $\phi 80^r$

BB4 and DH5α were used to propagate plasmid vectors, HB101 for cosmid vectors, JM109 for M13 vectors, and Q358 was used for λEMBL4 bacteriophage.

* Taken from Sambrook et al. (1989).