

# 42Sp48 in Previtellogenic *Xenopus* Oocytes Is Structurally Homologous to EF-1 $\alpha$ and May Be a Stage-specific Elongation Factor

Nicholas J. Coppard, Knud Poulsen,\* Hans O. Madsen,† Jane Frydenberg, and Brian F. C. Clark

Division of Biostructural Chemistry, Aarhus University, Langelandsgade 140 DK-8000 Aarhus C, Denmark; \* Department of Oral Biology, Royal Dental College, Vennelyst Boulevard, DK-8000 Aarhus C, Denmark; † Tissue Typing Laboratory of the Department of Clinical Immunology, State University Hospital, DK-2000 Copenhagen N, Denmark

**Abstract.** We have isolated the cDNA for 42Sp48 and EF-1 $\alpha$  from mixed stage oocytes and tailbud (stage 22) *Xenopus laevis* cDNA libraries by use of the cDNA for human elongation factor-1 $\alpha$  (EF-1 $\alpha$ ) as probe. The nucleotide and deduced amino acid sequences of the entire coding region of 42Sp48 and EF-1 $\alpha$  cDNA were established. The proposed functional homology of the proteins is reflected in highly conserved amino acid sequences (91% identity), while the large number of silent mutations at the gene level may serve to prevent recombination at their loci.

42Sp48 is apparently encoded by two genes in *Xenopus*, while no sequences corresponding to 42Sp48 could be found in murine or human genomic DNA. 42Sp48 has been proposed to act as a stage-specific elongation factor in *Xenopus*. Comparison of the deduced amino acid sequences of 42Sp48 and EF-1 $\alpha$  with that of elongation factor Tu from *E. coli*, for which the three-dimensional structure including that of the GTP binding site have been determined, supports this hypothesis.

**O** GENESES in amphibians can be divided into two distinctive stages; previtellogenesis, in which the cytoplasm of the cell is devoid of yolk and is transparent, and vitellogenesis, in which yolk, lipids, glycogen, and ribosomes rapidly accumulate. An oocyte of the African clawed toad, *Xenopus laevis*, contains extremely large amounts of RNA, 95% of which is ribosomal in origin and 2% of which is tRNA, while the remainder is thought to be mRNA (Brown and Littna, 1964). The oocyte transcribes 5S rRNA and tRNA genes constantly during all stages of growth and during previtellogenesis, in which the amplified 28S and 18S rRNA genes are not transcribed, 80% of the total RNA content comprises 5S and tRNA species (Ford, 1971). These components are stored in two cytoplasmic ribonucleoprotein particles, the 7S and 42S ribonucleoprotein particles (Denis and Le Maire, 1983; Mattaj et al., 1983).

The 42S particle consists of tRNA, 5S rRNA, and two proteins, 42Sp48 and 42Sp43 of ~48 and 43 kD, respectively (Picard et al., 1980). 42Sp48, also referred to as thesaurin A (Picard et al., 1980; Kloetzel et al., 1981), binds to tRNA and since this tRNA is largely aminoacylated in the 42S particle (Denis and Le Maire, 1983), this binding probably protects the aminoacyl tRNA from hydrolysis. The protection of aminoacyl tRNA from nonenzymatic deacylation is achieved

in bacteria by the elongation factor Tu (EF-Tu)<sup>1</sup> (Pingoud and Urbanke, 1980) and probably in eukaryotes by the EF-Tu functional equivalent, EF-1 $\alpha$  (Clark, 1980). The aminoacyl tRNA-binding capacity of 42Sp48 and evidence that aminoacyl tRNA is transferred from 42S particles to ribosomes (Le Maire and Denis, 1987), suggested a similarity may exist with EF-1 $\alpha$  and indeed, immunological, functional and partial sequence similarities have since been reported (Mattaj et al., 1987; Viel et al., 1987).

In addition, the presence of large amounts of 42Sp48 in previtellogenic oocytes exclusively, in contrast to the expression of EF-1 $\alpha$  which increases greatly in later stages of development (Mattaj et al., 1987; Krieg et al., 1989), suggests that 42Sp48 is a developmental stage-specific elongation factor in *Xenopus laevis*.

The elongation factors are G-binding proteins. This class of proteins is characterized by their ability to bind GTP and they are involved in functions as various as signal transduction (Gilman, 1987), macromolecular biosynthesis (examples are PEPCK and GTP:AMP phosphotransferase), and cellular (for review of *ras* see Barbacid, 1987) and structural control (Mandelkow et al., 1985). Recently, studies on structural homologies of G-binding proteins have been published (for review see Woolley and Clark, 1989) which have re-

Nicholas J. Coppard's present address is Senetek PLC, Voldbjergvej 14A, DK-8240 Risskov, Denmark.

1. Abbreviations used in this paper: EF-Tu, elongation factor Tu; ORF, open reading frame.

P48-2 TCTGCTGTTGTCTAGTGCTCGCGGCCAAC  
P48-1  
xEF-1α TCTAATTAAT\*\*\*CA

50

ATGGGAAAGGAAAAGATTACATTAAACATCGTTGTGATCGGGCATGTGGACTCGGGCAAGTCCACCACAACCGGCCATCTCATCTACAAA P48-2

\*\*\*\*\*C\*\*\*\*\*C\*\*\*\*\*C\*\*T\*\*A\*\*C\*\*C\*\*T\*\*T\*\*A\*\*\*\*\*A\*\*\*\*\*T\*\*\*\*\*G\*\*\*\*\* xEF-1α

100

150

TGCGGGGCATCGACAAAAGAACCATCGAGAAGTTTGAAGAAGGCCAGAGATGGGTAAGGCTCCTTTAAATATGCTTGGGTCTCTG P48-2

\*\*T\*\*T\*\*A\*\*\*\*\*G\*\*\*\*\*T\*\*A\*\*\*\*\*C\*\*\*\*\*T\*\*C\*\*\*\*\*A\*\*\*\*\*C\*\*G\*\*\*\*\*C\*\*\*\*\*T\*\* xEF-1α

200

250

GATAAGTTGAAAGCAGAGCGGGAACTGGCATAACCATCGATATTTCTTGTGGAAGTTTGAAGTCTGAAAGTATTATATCACCATTATT P48-2

\*\*C\*\*A\*\*C\*\*G\*\*\*\*\*T\*\*\*\*\*T\*\*C\*\*\*\*\*T\*\*C\*\*C\*\*C\*\*A\*\*\*\*\*A\*\*\*\*\*CA\*\*C\*\*A\*\*C\*\*G\*\*T\*\*C\*\* xEF-1α

300

350

GATGCTCCTGGCCACAGAGACTTCATCAAAAACATGATTACTGGCACTTCTCAGGCTGACTGTGCTGTGCTCATTGTTGCTGGTGGTGT P48-2

\*\*\*\*\*A\*\*A\*\*\*\*\*T\*\*\*\*\*G\*\*\*\*\*C\*\*\*\*\*T\*\*\*\*\*C\*\*\*\*\*C\*\*G\*\*\*\*\*C\*\*\*\*\* xEF-1α

400

450

GGAGAATTTGAGGCTGGTATCTCTAAAAATGGACAGACCCGGGAACATGCTCTCTTGGCGTTCCACCTAGGTGTCAAGCAGCTTATAATT P48-2

\*\*\*\*\*A\*\*T\*\*\*\*\* P48-1

\*\*T\*\*\*\*\*A\*\*\*\*\*A\*\*G\*\*\*\*\*A\*\*\*\*\*T\*\*G\*\*\*\*\*C\*\*T\*\*C\*\*A\*\*T\*\*G\*\*\*\*\*T\*\*\*\*\*G\*\*TG\*\* xEF-1α

500

GGAGTTAACAAAATGGACTCCACTGAGCCTCCTTTTAGCCAGAAAAGGTTTGAAGAAATTACTAAGAAGTCAAGTGCCTACATTAAGAAG P48-2

\*\*\*\*\*G\*\*\*\*\*C\*\*\*\*\*A\*\*\*\*\* P48-1

\*\*TA\*\*\*\*\*T\*\*\*\*\*A\*\*A\*\*A\*\*AC\*\*\*\*\*G\*\*A\*\*A\*\*G\*\*\*\*\*GTA\*\*G\*\*\*\*\*CA\*\*A\*\*\*\*\*C\*\*\*\*\*A xEF-1α

550

600

ATTGGCTACAACCCAGCGACCGTTCCATTGTGCCAATATCTGGATGGCATGGAGACAACATGCTGGAGGCTAGCACCATATGCCCTGG P48-2

\*\*\*\*\*T\*\*\*\*\*T\*\*\*\*\*G\*\*\*\*\*T\*\*\*\*\*A\*\*C\*\*T\*\*\*\*\*T\*\*C\*\*C\*\*C\*\*\*\*\*T\*\* xEF-1α

650

700

TTTAAAGGCTGGAAGATTGAAAGGAAAGAAATGCCAGTGGTATAACTCTGCTAGAACACTTGACTGTATAATTCCTCCTCAGAGA P48-2

\*\*\*\*\*G\*\*\*\*\*G\*\*\*\*\*T\*\*\*\*\*G P48-1

\*\*\*\*\*G\*\*A\*\*\*\*\*A\*\*CAC\*\*C\*\*T\*\*\*\*\*G\*\*TC\*\*G\*\*C\*\*A\*\*CT\*\*C\*\*\*\*\*G\*\*\*\*\*T\*\*\*\*\*C\*\*TT\*\*G\*\*A\*\*ATCTC\*\*C xEF-1α

750

800

CCCACTAACAGCCTCTACGACTCCCTCTGCAGGATGTGTACAAAATGGTGGAAATGGCACTGTGCCTGTTGGCAGAGTGGAGACTGGT P48-2

\*\*\*\*\*G\*\*\*\*\*G\*\*\*\*\*T\*\*G\*\*\*\*\* P48-1

\*\*A\*\*CG\*\*T\*\*\*\*\*C\*\*T\*\*G\*\*\*\*\*C\*\*\*\*\*C\*\*T\*\*\*\*\*T\*\*\*\*\*A\*\*A\*\*\*\*\*TC\*\*T\*\*\*\*\* xEF-1α

850

900

GTCTTAAAGCCAGGCATGATTGTGACTTTTGTCTCCAAGTAATGTCACAGAAAGTAAAGTCTGTGGAATGCATCATGAGGCCTTGGTA P48-2

\*\*\*\*\*T\*\*T\*\*\*\*\*A\*\*\*\*\*T\*\*CA\* P48-1

\*\*A\*\*T\*\*A\*\*\*\*\*G\*\*G\*\*T\*\*C\*\*C\*\*T\*\*G\*\*T\*\*C\*\*A\*\*A\*\*T\*\*G\*\*A\*\*\*\*\*T\*\*\*\*\*C\*\*\*\*\*A\*\*C\*\*TACT xEF-1α

950

GAAGCACTGCCTGGAGACAATGTGGGTTTCAATGTGAAGAATATATCCGTGAAGGACATTAGGAGAGGAAATGTAGCTGGGGATAGCAAA P48-2

\*\*G\*\*T\*\*\*\*\*T\*\*\*\*\*C\*\*\*\*\*T\*\*A\*\*\*\*\*C\*\*C\*\*\*\*\*T\*\*C\*\*\*\*\* P48-1

\*\*G\*\*CG\*\*C\*\*T\*\*\*\*\*C\*\*T\*\*\*\*\*T\*\*C\*\*C\*\*\*\*\*CG\*\*T\*\*\*\*\*G\*\*CC\*\*TC\*\*C\*\*C\*\*T\*\*\*\*\*T\*\*C\*\*\*\*\*G xEF-1α

Figure 1. Nucleotide sequence of *Xenopus laevis* 42Sp48 cDNA clones p48-1 and p48-2 and the cDNA of *Xenopus* EF-1α (xEF-1α). The suggested sites of initiation and termination of translation are boxed. Those base substitutions leading to amino acid substitutions in the deduced amino acid sequence of p48-2 are underlined. Numbering starts from the A of the ATG initiator codon. The introduction of spacing ( . . . ) in the 3' untranslated region of 48-1 yields maximal homology. Poly(A) tail lengths are indicated. These sequence data are available from EMBL/GenBank/DBJ under accession numbers X56698 and X56699 for 42Sp48-1 and 42Sp48-2, respectively.

1000 1050  
AATGACCCACCTATGCAAGCTGGCAGCTTTACAGCACAGGTGATTATACTCAATCACCAGGACAGATTAGTGCTGGCTATGCTCCAGTC P48-2  
\*\*\*\*\*T\*\*\*\*\*C\*\*\*\*\*T\*\*\*\*\* P48-1  
\*\*\*\*\*A\*\*G\*\*\*\*\*TTCA\*\*\*\*\*C\*\*\*\*\*T\*\*C\*\*C\*\*G\*\*C\*\*\*\*\*C\*\*\*\*\*G\*\*\*\*\*A\*\*\*\*\*C\*\*T\*\*G xEF-1 $\alpha$

1100 1150  
CTGGACTGTCATACAGCTCATATTGCTTGTAAAGTTGCTGAGCTGAAGCAAAGATTGACAGAAGAAGTGGCAAGAAGCTGGAAGACGAT P48-2  
\*\*\*\*\*C\*\*\*\*\*C\*\*\*\*\*T\*\*C P48-1  
T\*\*\*T\*\*\*C\*\*T\*\*\*C\*\*\*\*\*C\*\*\*\*\*C\*\*G\*\*G\*\*\*\*\*TC\*\*CC\*\*TTC\*\*T\*\*\*A\*\*\*\*\*A\*\* xEF-1 $\alpha$

1200 1250  
CCAAAATCTTGAAGTCTGGAGATGCAGCTATAGTGAAATGATCCCTGGGAAACCAATGTGTGTAGAAACCTTCTCAGACTATCCTCCA P48-2  
\*\*\*\*\*A\*\*\*\*\*T\*\*\*\*\*G\*\*C\*\*\*\*\*G\*\*T\*\*T\*\*\*\*\* P48-1  
\*\*C\*\*G\*\*C\*\*\*\*\*T\*\*\*\*\*T\*\*C\*\*T\*\*C\*\*\*\*\*A\*\*A\*\*G\*\*\*\*\*C\*\*G\*\*G\*\*G\*\*\*\*\*T\*\*\*\*\*C\*\*\*\*\*T xEF-1 $\alpha$

1300 1350  
CTTGGCCGATTTGCAGTCCGTGATATGAGGCAGACGGTGTGGCGTAATCAAAGGTGTGGACAAAAGTTGGCAAGCTCTGGGAAA P48-2  
\*\*\*\*\*T\*\*\*\*\*A\*\*\*\*\*GCA\*\*\*\*\*T\*\*\*\*\* P48-1  
\*\*\*\*\*T\*\*T\*\*T\*\*T\*\*C\*\*\*\*\*A\*\*\*\*\*C\*\*T\*\*\*\*\*A\*\*A\*\*\*\*\*G\*\*CA\*\*C\*\*G\*\*G\*\*GCA\*\*G\*\*A\*\*C\*\*C\*\* xEF-1 $\alpha$

1400  
GTGACAAAGTCTGCTGCAAAAGCTGGCAAGAAGTGAATCAATGTACAATGTAATATTGGTGTCTGTAATAACTTAATAAATGTTGATGC P48-2  
\*\*A\*\*\*\*\*TC\*\*\*\*\*A\*\*\*\*\*AT\*\*\*\*\*G\*\*T\*\*C\*\*\*\*\*A\*\*\*A P48-1  
\*\*C\*\*\*\*\*CAG\*\*\*\*\*C\*\*G\*\*A\*\*C\*\*A\*\*AG\*\*A\*\*TTCCA\*\*A\*\*TATC\*\*T\*\*CC\*\*AGTC\*\*T\*\*CTAGGGCTG\*\*A\*\*GA\*\*TGGT\* xEF-1 $\alpha$

1500  
TTTTF (A) 42 P48-2  
TG (A) 22 P48-1  
\*CAGAACTCTTTCTATCAATGGCCATCAGAGTTAATAGTCAAAGACTGGTAAATGATTACAATGCATCGCAAAGCTTCAGAAGGAAA xEF-1 $\alpha$

1600  
GAAATGCTTGGGACACATTTGTTTGTGGCAGTTTTAAGTTATTAGTGTTTTAAATCCAGTAATTTCTAAATGGGAGGCAACTTGACCAA xEF-1 $\alpha$

1700  
AATCTGTCACCAAATTTGAGACCATT (A) 35 xEF-1 $\alpha$

vealed the presence of a consensus in amino acid sequences involved in nucleotide and phosphoryl binding in the G-domain. The reported functional homology of 42Sp48 and EF-1 $\alpha$  suggests that structural homology may also exist between the two proteins and it was therefore of interest to determine the degree of homology between them and other G-binding proteins. In the present study we describe the isolation and characterization of the cDNA for 42Sp48 and EF-1 $\alpha$  from *Xenopus laevis*.

## Materials and Methods

### Screening of cDNA Libraries

The *Xenopus*  $\lambda$ gt10 cDNA libraries were a gift of Professor Eddy De Robertis (University of California, Department of Biological Chemistry, Los Angeles, CA). The phages were plated on *E. coli* Y1090 and positive plaques were identified by in situ hybridization (Benton and Davis, 1977; Sambrook et al., 1989), purified by replating, and DNA was isolated from 20-ml phage cultures. As a label in the hybridizations ( $\alpha$ -<sup>32</sup>P) dATP was used, which was incorporated into the DNA probes by nick translation. In the high stringency screenings the final posthybridization wash was in 0.1 $\times$  SSC, 0.1% SDS, 0.1% sodium pyrophosphate at 67°C. For low stringency washing we used 1 $\times$  SSC, 0.1% SDS, 0.1% sodium pyrophosphate at 37°C.

### Sequence Analysis

The Eco RI insert of positive  $\lambda$ gt10 recombinant clones to be sequenced was subcloned into M13 mpi8 (Yanisch-Perron and Sambrook, 1989). Sequencing strategy of xEF-1 $\alpha$ , p48-1, and p48-2 utilized the internal sites for the restriction enzymes Pst I and Bam HI, Pst I and Hin dIII, and Pst I and Sma I, respectively, to create clones containing deletions of these Eco RI subclonings. Thereby, we obtained overlapping sequences for each of the three clones. Partial sequences of the remaining lambda clones were obtained by subcloning and sequencing the ends of the Eco RI inserts. A very large poly (A) tail blocked the sequencing of the 3' end of some clones. The sequence of individual clones was determined by the dideoxynucleotide chain termination method (Sanger et al., 1977; Biggin et al., 1983) using the universal primer, the Sequenase DNA sequencing kit (United States Biochemical Corp., Cleveland, OH), and ( $\alpha$ -<sup>35</sup>S) dATP.

### Southern Blot Analysis

Approximately 10  $\mu$ g of genomic DNA was restricted with the appropriate enzyme, size fractionated by agarose gel electrophoresis, and transferred to nitrocellulose filters (Schleicher & Schuell, Inc., Keene, NH) (Sambrook et al., 1989). The Southern blot hybridizations were performed using stringent conditions (see above). In the rehybridization experiment the filter was washed at 95°C in water for 5 min in between hybridizations.

## Results

Functional (Le Maire and Denis, 1987; Mattaj et al., 1987),

		1		
P48-2	MGKEK....IHINIVVIGHVDSGKSTTTGHLIYKCGGIDKRTIEKFEKEAAEMGKGSFK			54
xEF-1a	*****.T*****			54
hEF-1a	*****.T*****			54
EF-Tu	*S***FERTKP*V*VGT***H**T*L*AAITTVLAKTYGGAAR.....			44
		2		
P48-2	YAVVLDKLAERERGITIDISLWKFETGKYIITIIDAPGHRDFIKNMITGTSQADCAVLI			114
xEF-1a	*****S***V*****			114
hEF-1a	*****S***V*****			114
EF-Tu	AFDQI*NAPE*KA*****NT*HVEYD*PTRHYAHV**A*YV*****AA*M*G*I*V			104
		3		
P48-2	VAGGVGEFEAGISKNGQTRHALLAFTLGVKQLIIGVKNKMDSTEPPFSQKRFEEITKEVS			174
P48-1	*****			22
xEF-1a	**A*****Y*****V**I*****Y***Y***V****			174
hEF-1a	**A*****Y*****V**I*****Y***Y***V****			174
EF-Tu	**ATD*PM.....P*****I**GRQV**PYI*VFL**K*.MVDDEELLELVEMEVR			155
		4		
P48-2	AYIKKIGYNPATVPFVPISGWHGDNMLEASTNMPWFKGWKIERKEGNASGITLLEALDC.			233
P48-1	*****V*****			81
xEF-1a	T*****D**A*****N*****P*****T***SG**T*****			233
hEF-1a	T*****D**A*****N*****P*****VT**D*****T*****			233
EF-Tu	L.LSQYDFPGDD.....T.PIVRG*ALKALE.....GDA*WE*KILE*AGF**SY			198
P48-2	IIPPQRPTNKPLRLPLQDVYKIGGIGTVPVGRVETGVLKPGMIVTFAPSNV...TTEVKS			290
P48-1	*****A*****			138
xEF-1a	*L**S**D*****I***V*****V**...*****			290
hEF-1a	*L**T**D*****V*****V**...*****			290
EF-Tu	*PE*E*AID**FL**IE**FS*S*R***VT***R**II*V*EE*EIVGIKETQKS*CTG.			257
P48-2	VEMHHEALVEALPGDNVGFNVKNISVKDIRRGVAGDSKNDPPMQAGSFTAQVIILNHGP			350
P48-1	*****Q*****			198
xEF-1a	*****T**V*****V***V*****E*****			350
hEF-1a	*****S*****V***V*****E*AG*****			350
EF-Tu	***FRKL*D**GRA*E***VLLRG*KREE*E**Q*LAKPGTIK*HTK..*ESE*Y**SKDE			315
P48-2	QISA....GYAPVLDCHTAHIACKFAELKQK.IDRRSGKKLEDDPKFLKSGDAAIVEMI			404
P48-1	****.*****			252
xEF-1a	**G*....*****E*.*****N*****D**			404
hEF-1a	****.*****E*.*****G*****D*V			404
EF-Tu	GGRHTPFFK**R*QFYFR*TDVTGTI.**PEGVE.....MVMP**.....			355
P48-2	PGKPMCETFSDYPPL.G.RFAVRDMRQTVAVGVIKGVDKKLASSGKVTKSAKAGKK			460
P48-1	*****S*****.*.*****A*****V*****			309
xEF-1a	*****S*****.*.*****A**E**A*G*****Q**A*TK			461
hEF-1a	*****S*****.*.*****A**A*GA*****Q**Q*AK			461
EF-Tu	.NIK*V*TLIHPIAMDD*L***I*EGGR**GA**VAK*LS			393

Figure 2. The amino acid sequence of 42Sp48 as deduced from cDNA clones p48-1 and p48-2. The amino acid sequences of *Xenopus* EF-1 $\alpha$  (xEF-1a), deduced from the cDNA clone xEF-1a, that of the human cDNA clone CEF4 (Madsen et al., 1990) (hEF-1a), and that published for EF-Tu from *E. coli* (Arai et al., 1980) are also shown. The initiator methionine has been designated amino acid position 0. The sequences are aligned around the conserved phosphoryl and guanine binding structures (boxed) (Nyborg and la Cour, 1989) and

immunological (Mattaj et al., 1987), and partial protein sequence homology (Viel et al., 1987) between eukaryotic EF-1 $\alpha$  and 42Sp48 suggest that the two proteins are closely related and that their corresponding mRNA sequences are homologous. We have previously reported that 42Sp48 disappears soon after the onset of vitellogenesis, while the EF-1 $\alpha$  concentration remains essentially unchanged during this transition (Mattaj et al., 1987). Therefore, in an attempt to isolate cDNAs for both 42Sp48 and EF-1 $\alpha$ , we screened  $\lambda$ gt10 cDNA libraries of polyA mRNA prepared from mixed-stage oocytes (stages 1–12) and tailbud (stage 22) toads. EF-1 $\alpha$  is very well conserved throughout eukaryotic evolution. Therefore, the 1.7-kb human EF-1 $\alpha$  cDNA clone, CEF4 (Madsen et al., 1990), which contains the entire coding region of 1,386 bp, was used as a probe for screening the  $\lambda$ gt10 cDNA libraries.

### Isolation of EF-1 $\alpha$ cDNA Sequences

The two cDNA libraries were screened with the CEF4 probe at high stringency. Partial sequencing of three positive clones isolated from the tailbud and two clones from the mixed stage oocyte cDNA library, revealed that they contained sequences almost identical in overlapping regions and very similar to that published previously for *Xenopus* EF-1 $\alpha$  (Krieg et al., 1989). The complete sequence of the largest of these five clones (xEF-1 $\alpha$ ) was determined (Fig. 1). The sequence of 1,661 bp revealed a large open reading frame (ORF) capable of encoding the entire EF-1 $\alpha$  protein of 462 amino acids (Fig. 2). This ORF is preceded 15 bp by 5' leader sequence and followed by 257 bp with a poly A tail at its extreme 3' end. Comparison of the xEF-1 $\alpha$  sequence with that previously reported for *Xenopus* EF-1 $\alpha$  cDNA reveals several silent mutations in the coding region and one mutation that leads to the incorporation of asparagine at position eight (Fig. 2) in the deduced amino acid sequence replacing lysine. These distinct cDNA sequences of almost identical coding potential are indicative of allelic variation within the *Xenopus laevis* EF-1 $\alpha$  gene(s).

### Isolation of 42Sp48 cDNA Sequences

Reduction in the stringency of the screening procedure revealed an additional set of clones in the mixed stage oocyte library that were low stringency positive and high stringency negative using CEF4 as probe. Three such clones were selected and analyzed for insert size. The largest of the clones, p48-1, was sequenced (Fig. 1). Comparison of this cDNA sequence with that obtained for xEF-1 $\alpha$  demonstrated limited homology, as predicted by their positivity in low stringency and negativity at high stringency hybridization conditions. Since p48-1 was not long enough to encode the entire 42Sp48 sequence, it was subsequently used to re-screen the mixed stage oocyte library at high stringency, yielding 10 clones which were assessed for insert size. The clone containing the largest insert, p48-2, was sequenced.

The insert of p48-2 was found to be 1,473 bp in length and contained 29 bp of the untranslated 5' leader sequence followed by an ORF of 1,386 bp including the TGA stop codon and a 3' untranslated region of 58 bp. The 3' region contained the polyadenylation signal AATAAA located 15 bp upstream of a poly (A) stretch at the ultimate 3' end. The ORF starting at the first ATG initiation codon has the potential of encoding a protein of 50,130 D. This value is close to the 48 kD determined for the 42Sp48 protein by SDS gel electrophoresis. Though no in-frame stop codons were found in the preceding 29-bp sequence, the assignment of this ATG triplet in Fig. 1 as the initiation codon agrees with the consensus sequence for translation initiation (Kozak, 1984).

The insert of p48-1 contains 1,026 bp of the 3' end of the 42Sp48 cDNA including a poly A signal 12 bp upstream of a poly (A) tail. It differs from p48-2 in 56 of 927 nucleotides in the coding region sequenced, in contrast to the two deduced 42Sp48 proteins which differ only in 6 of 309 amino acid positions in this region. Comparison of the sequences shown for 42Sp48 in Fig. 1 reveals that the great majority of differences in the DNA sequences of p48-1 and p48-2 are silent mutations (45 of 56) while 11 lead to amino acid substitutions in the 6 positions in the deduced amino acid sequence.

The high degree of homology in the 3' untranslated regions between p48-1 and p48-2 (see Fig. 1) suggests that these two cDNA clones represent allelic forms of the 42Sp48 gene(s). Interestingly, the nucleotide sequence homology between the coding regions of 42Sp48 and xEF-1 $\alpha$  cDNA clones is 79% in contrast to the homology of 91% between the two proteins. The failure of the two cDNAs to cross-hybridize is thought to be due to the lack of long stretches (>16 bp) of identical sequence in the cDNAs.

### Estimation of the Expression of 42S p48 and EF-1 $\alpha$ by Screening of cDNA Libraries

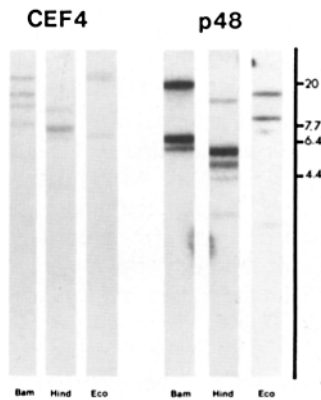
In an attempt to estimate the abundance of messenger RNA for 42Sp48 and EF-1 $\alpha$  in the RNA preparations used in the construction of cDNA libraries from *Xenopus* at various stages of development, libraries from mixed stage oocytes (whole ovary) and tailbud were screened with <sup>32</sup>P-labeled p48-2 and *Xenopus* EF-1 $\alpha$  cDNA at high stringency. Of 20,000 plaques analyzed from the whole ovary library 20 hybridized with the EF-1 $\alpha$  probe while 50 hybridized with the p48 probe. In the tailbud library, of 50,000 plaques analyzed, 50 hybridized with the EF-1 $\alpha$  probe while no strong binding was detected using the p48 probe.

### Estimation of Gene Copy Number by Southern Blot Hybridization

Genomic DNA from *Xenopus* was cleaved with Bam HI, Hind III, Eco RI, or Sau 3A and analyzed by Southern blot hybridization using p48-2 and xEF-1 $\alpha$  cDNA probes (Fig.

---

spacings (...) have been introduced to give maximal homology between all sequences. Identities with p48-2 are marked with \* and substitutions indicated with the substituted amino acid. The G-domain of *E. coli* EF-Tu runs from the NH<sub>2</sub>-terminal to around position 196 and its COOH-terminal end is underlined. Key to boxed sequences: (1 and 2) phosphoryl binding sequences; (3) guanine-specific sequence; (4) base interactive loop sequence.



**Figure 3.** Southern blot analysis of *Xenopus* genomic DNA using the 42Sp48 cDNA, p48-2, as hybridization probe. The filter was subsequently rehybridized with the human EF-1 $\alpha$  cDNA, CEF4, as probe. The restriction enzymes used for digestion of the genomic DNA are indicated below each lane: (*Bam*) Bam HI; (*Hind*) Hind III; and (*Eco*) Eco RI. Size markers in kilobases are indicated to the right of the figure. The observed difference in band intensity between CEF4 and p48 is due to the rehybridization procedure.

3). Few major bands were resolved in these digestions with both probes indicating that in *Xenopus* the genes for p48 and EF-1 $\alpha$  are low in copy number. Corresponding Southern blots of Eco RI restricted human and mouse genomic DNA were also hybridized with each of the two probes (results not shown). Multiple bands were resolved with the xEF-1 $\alpha$  probe in both the human and mouse genomic DNA preparations. In contrast, no hybridization was found in human or mouse genomic DNA preparations with the p48-2 cDNA probe.

## Discussion

Partial protein sequences have been published previously for 42Sp48 purified from 42S particles (Viel et al., 1987). Comparison of the published 42Sp48 sequence with those deduced here for p48-1 and p48-2 reveal homologies that are not shared with EF-1 $\alpha$ . In addition to general sequence homology, the published sequence shows identity with the deduced p48 sequence at the hypervariable position 299 and at positions 282 and 314 in the deduced sequence, positions in which the two p48 sequences are identical but which differ from EF-1 $\alpha$ . Significantly, these three positions include the only two nonconservative substitutions in the deduced 42Sp48 sequence compared to EF-1 $\alpha$ .

The accumulation of 42Sp48 in stage 1 oocytes has been reported previously (Dixon and Ford, 1982; Mattaj et al., 1983), and the appearance of EF-1 $\alpha$  immunoreactivity in stage three oocytes has been found to coincide with the disappearance of 42Sp48 at this stage (Mattaj et al., 1987). The results obtained by screening stage-specific cDNA libraries with p48 and EF-1 $\alpha$  probes suggest that the levels of message for 42Sp48 were high in the mRNA used to prepare the cDNA library from mixed stage oocytes and low or absent in the tailbud (stage 22) preparations. This result supports the proposal that 42Sp48 is transcribed only in the early stages of oocyte development, whereas EF-1 $\alpha$  transcription continues throughout this process.

This observation and the homology of the partial protein sequence of 42Sp48 published previously with that deduced from the cDNA of p48, confirm that p48-1 and p48-2 are cDNAs for 42Sp48.

The elongation factors function as carrier molecules trans-

ferring aminoacyl-tRNA to the ribosome. The binding of a GTP molecule is necessary for the formation of the elongation factor-aminoacyl-tRNA complex and is hydrolyzed on the binding of the aminoacyl-tRNA to the ribosome. Secondary functions including protection of the aminoacyl group against hydrolysis have also been attributed to them.

The diversity of binding activities displayed by the elongation factors have also been reported for 42Sp48 (Le Maire and Denis, 1987; Mattaj et al., 1987; Viel et al., 1987), and analysis of the deduced protein sequence of 42Sp48 presented here indicates a high degree of homology with EF-1 $\alpha$ . The 461 amino acid sequence of 42Sp48 differs from that of EF-1 $\alpha$  in 43 positions, 15 of which are conservative substitutions.

Those structural elements of EF-1 $\alpha$  predicted from tertiary structure studies on EF-Tu to be involved in G-binding (Woolley and Clark, 1989) are completely conserved. The elucidation of the three-dimensional structure of the G-binding protein EF-Tu (Nyborg and la Cour, 1989) has allowed prediction of the structure involving the conserved sequences in other G-binding proteins. The G-domain of EF-Tu contains two phosphoryl binding sequences, GHVDGK and DCPG at positions 18–24 and 80–84, respectively, and the G-specific sequence NKCD at positions 135–138. These sequences are highly conserved among G-binding proteins including EF-1 $\alpha$  and the *ras* protein p21 (Dever and Merrick, 1989; Jornak, 1988). They are also found in the protein sequence deduced from p48-2 cDNA at positions 13–19, 90–93, and 152–155, respectively. These consensus sequences are shown in Fig. 2. The alignment of the conserved elements is based on the three-dimensional structure of EF-Tu and allows the greatest degree of structural homology. 42Sp48 aligns with EF-1 $\alpha$  in this model and appears therefore to possess the structural characteristics of this factor. The structure indicates it is guanine specific and therefore the stimulation of 42Sp48 activity by ATP reported previously (Le Maire and Denis, 1987) is probably due to stimulation of the aminoacylation of tRNA.

The sequence ERxRGITI, (at arg-68 in EF-1 $\alpha$ ) thought to be typical for GTP-dependent proteins which bind noninitiator tRNA to the ribosome (Möller et al., 1987) is also conserved in 42Sp48, reinforcing the proposed role of 42Sp48 as an elongation factor in previtellogenic oocytes (Mattaj et al., 1987). A major function of 42Sp48, in association with the 42S particle, may be the protection from hydrolysis of the high levels of aminoacyl-tRNA present at this stage of development.

Considerable variation is observed in the degree of homology between 42Sp48 and EF-1 $\alpha$  at the DNA and protein level. Comparison of the cDNA sequences of p48 and EF-1 $\alpha$  reveals a level of homology of >91% at the protein level and <80% at the DNA level. This would suggest that although 42Sp48 and EF-1 $\alpha$  may be functionally very closely related, diversion of the coding sequences has occurred to prevent genetic recombination while retaining functional homology. The prevention of conversion would also be important for the preservation of the functional difference and stage-specific expression of 42Sp48 over that of EF-1 $\alpha$ , which is expressed at all stages of development. The results of the Southern blot analysis of *Xenopus* genomic DNA with EF-1 $\alpha$  and 42Sp48 cDNA probes are shown in Fig. 3, and indicate that these genes are present in low copy numbers. We estimate, based

on the number, size, and intensity of the hybridizing fragments, that two copies of each gene are present.

The *Xenopus* EF-1 $\alpha$  cDNA probe hybridizes to a multiplicity of bands in mouse and human genomic DNA preparations as expected from reports of multiple copy retro-pseudogenes for EF-1 $\alpha$  in mammals (Madsen et al., 1990). The failure of the 42Sp48 cDNA probe to hybridize to sequences in human and mouse genomic DNA may indicate that mammals do not possess a functional equivalent of 42Sp48.

Received for publication 27 June 1990 and in revised form 13 September 1990.

*Note Added in Proof.* Very recently, evidence for another stage-specific form of EF-1 $\alpha$  has been found in *Xenopus* oocytes (Djé, M. K., A. Mazabraud, A. Viel, M. le Maire, H. Dennis, E. Crawford and D. Brown 1990. Three genes under different developmental control encode elongation factor 1 $\alpha$  in *Xenopus laevis*. *Nucleic Acids Res.* 18:3489–3493). This new form, termed EF-1 $\alpha$ O, appears to be highly homologous with the 42Sp48 sequences reported here and may represent an allelic variation.

### References

- Arai, K., B. F. C. Clark, L. Duffy, M. Jones, Y. Kaziro, R. Laursen, J. L'Italien, D. Miller, S. Nagarkatti, S. Nakamura, K. Nielsen, T. Petersen, K. Takahashi, and M. Wade. 1980. Primary structure of elongation factor-Tu from *E. coli*. *Proc. Natl. Acad. Sci. USA.* 77:1326–1330.
- Barbacid, M. 1987. *ras* genes. *Annu. Rev. Biochem.* 56:779–827.
- Benton, W. D., and R. W. Davis. 1977. Screening  $\lambda$ gt11 recombinant clones by hybridization to single plaques *in situ*. *Science (Wash. DC).* 196:180–182.
- Biggin, M. D., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and S<sup>35</sup> label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA.* 80:3963–3965.
- Brown, D. D., and E. Littna. 1964. RNA synthesis during the development of *Xenopus laevis*, the South African clawed toad. *J. Mol. Biol.* 8:669–687.
- Clark, B. F. C. 1980. The elongation step of protein biosynthesis. *Trends Biochem. Sci.* 5:207–210.
- Denis, H., and M. Le Maire. 1983. Thesaurisomes, a novel kind of nucleoprotein particle. In *Subcellular Biochemistry*. Vol. 9. D. B. Roodyn, editor. Plenum Publishing Corp., New York. 263–298.
- Dever, T. E., and W. C. Merrick. 1989. The GTP-binding domain revisited. In *The Guanine-Nucleotide Binding Proteins. Common Structural and Functional Properties*. EMBO/NATO/CEC Adv. Res. Workshop, Renesse, The Netherlands, Aug. 1988. L. Bosch, B. Kraal, and A. Parmeggiani, editors. Plenum Press, NY. 35–48.
- Dixon, L. K., and P. J. Ford. 1982. Regulation of protein synthesis and accumulation during oogenesis in *Xenopus laevis*. *Dev. Biol.* 93:478–497.
- Ford, P. J. 1971. Non-coordinated accumulation and synthesis of 5S ribonucleic acid by ovaries of *Xenopus laevis*. *Nature (Lond.)*. 233:561–564.
- Gilman, A. G. 1987. G proteins: transducers of receptor-generated signals. *Annu. Rev. Biochem.* 56:615–649.
- Jurnak, F. A. 1988. The three-dimensional structure of c-H-ras p21: Implications for oncogene and G protein studies. *Trends Biochem. Sci.* 13:195–198.
- Kloetzel, P., W. Whitefield, and J. Summerville. 1981. Analysis and reconstruction of an RNP particle which stores 5S RNA and tRNA in amphibian oocytes. *Nucleic Acids Res.* 9:605–621.
- Kozak, M. 1984. Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. *Nucleic Acids Res.* 12:857–872.
- Krieg, P., S. Varnum, M. Wormington, and D. Melton. 1989. The mRNA encoding Elongation Factor 1- $\alpha$  (EF-1 $\alpha$ ) is a major transcript at the midblastula transition in *Xenopus*. *Dev. Biol.* 133:93–100.
- le Maire, M., and H. Denis. 1987. Biochemical research on oogenesis. *J. Biol. Chem.* 262:654–659.
- Madsen, H. O., K. Poulsen, O. Dahl, B. F. C. Clark, and J. P. Hjorth. 1990. Retro-pseudogenes constitute the major part of the human elongation factor 1 $\alpha$  gene family. *Nucleic Acids Res.* 18:1513–1516.
- Mandelkow, E.-M., M. Hermann, and U. Rühl. 1985. Tubulin domains probed by limited proteolysis and subunit-specific antibodies. *J. Mol. Biol.* 185:311–327.
- Mattaj, I. W., S. Leinhard, R. Zeller, and E. M. De Robertis. 1983. Nuclear exclusion of transcription factor IIIA and the 42S particle transfer RNA-binding protein in *Xenopus* oocytes: a possible mechanism for gene control? *J. Cell Biol.* 97:1261–1265.
- Mattaj, I. W., N. J. Coppard, R. S. Brown, B. F. C. Clark, and E. M. De Robertis. 1987. 42Sp48—the most abundant protein in previtellogenic *Xenopus* oocytes—resembles elongation factor 1 $\alpha$  structurally and functionally. *EMBO (Eur. Mol. Biol. Organ.) J.* 6:2409–2413.
- Möller, W., A. Schipper, and R. Amons. 1987. A conserved amino acid sequence around Arg-68 of *Artemia* elongation factor 1 $\alpha$  is involved in the binding of guanine nucleotides and aminoacyl transfer RNAs. *Biochimie.* 69:983–990.
- Nyborg, J., and T. la Cour. 1989. New structural data on elongation factor Tu: GDP based on x-ray crystallography. In *The Guanine-Nucleotide Binding Proteins. Common Structural and Functional Properties*. EMBO/NATO/CEC Adv. Res. Workshop, Renesse, The Netherlands, Aug. 1988. L. Bosch, B. Kraal, and A. Parmeggiani, editors. Plenum Press, NY. 3–14.
- Picard, B., M. Le Maire, M. Wegnez, and H. Denis. 1980. Biochemical research on oogenesis: composition of the 42S storage particles of *Xenopus laevis* oocytes. *Eur. J. Biochem.* 109:359–368.
- Pingoud, A., and C. Urbanke. 1980. Aminoacyl transfer ribonucleic acid binding site of the bacterial Elongation Factor Tu. *Biochemistry.* 19:2180–2112.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual*. 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.
- Sanger, F., S. Nicklson, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.
- Viel, A., M. Djé, A. Mazabraud, H. Denis, and M. Le Maire. 1987. Thesaurin a, the major protein of *Xenopus laevis* previtellogenic oocytes, present in the 42 S particles, is homologous to elongation factor EF-1 $\alpha$ . *FEBS (Fed. Eur. Biochem. Soc.) Letts.* 223:232–236.
- Woolley, P., and B. F. C. Clark. 1989. Homologies in the structures of G-binding proteins: an analysis based on elongation factor EF-Tu. *Biotechnology.* 7:913–920.
- Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains; nucleotide sequences of the M13 mp18 and pUC19 vectors. *Gene (Amst.)*. 33:103–119.