Development 119, 433-446 (1993)
Printed in Great Britain © The Company of Biologists Limited 1993

δ-crystallin enhancer binding protein δEF1 is a zinc finger-homeodomain protein implicated in postgastrulation embryogenesis

Jun-ichi Funahashi[1], Ryohei Sekido[2], Kasumi Murai[2], Yusuke Kamachi[2] and Hisato Kondoh[2,]*

[1]Department of Molecular Biology, School of Science, Nagoya University, Nagoya 464-01, Japan
[2]Institute for Molecular and Cellular Biology, Osaka University, Yamadaoka 1-3, Suitashi, Osaka 565, Japan

*Author for correspondence

SUMMARY

We investigated nuclear factors that bind to δ1-crystallin enhancer core and regulate lens-specific transcription. A nuclear factor δEF1, which binds to the essential element of the δ1-crystallin enhancer core, was molecularly cloned from the chicken by a southwestern method. The protein organization of δEF1 deduced from the cDNA sequence indicated that it has heterogeneous domains for DNA-binding, two widely separated zinc fingers and a homeodomain, analogous to *Drosophila* ZFH-1 protein. The C-terminal zinc fingers were found to be responsible for binding to the δ1-crystallin enhancer core sequence. δEF1 had proline-rich and acidic domains common to various transcriptional activators. During embryogenesis, δEF1 expression was observed in the postgastrulation period in mesodermal tissues; initially, in the notochord, followed by somites, nephrotomes and other components. The expression level changed dynamically in a tissue, possibly reflecting the differentiation states of the constituent cells. Besides mesoderm, δEF1 was expressed in the nervous system and the lens, but other ectodermal tissues and endoderm remained very low in δEF1 expression. Cotransfection experiments indicated that this factor acts as a repressor of δ1-crystallin enhancer. Possession of heterogeneous DNA-binding domains and its dynamic change of expression in embryogenesis strongly suggest that δEF1 acts in multiple ways depending on the cell type and the gene under its regulation.

Key words: embryonic organogenesis, -crystallin enhancer, zinc finger-homeodomain protein, transrepression, lens-specific regulation

INTRODUCTION

Cell type-specific gene expression is mainly controlled in the process of transcriptional initiation by the differential activity of transcription factors. Cumulative evidence has suggested that cell-specific regulation is not necessarily elicited by the activity of a single kind of factor but generally by a combinatorial assortment of transcriptional regulators, each of which may have wider cell type distribution than the gene under its regulation. In addition, it has become increasingly clear that not only activators but also repressors play crucial roles in generating cell type-specific activation. Thus, to determine the mechanism of such regulation, it is essential to characterize each of the factors that participate and have a determinative role in eliciting the specificity.

We have been interested in the mechanism of lens-specific expression of the -crystallin gene, because this gene is the first to be turned on in lens cell differentiation during embryogenesis. In addition, since lens cells undergo terminal differentiation exceptionally early in embryogenesis through interaction of relatively simple cell lineages (Grainger, 1992), analysis of -crystallin gene regulation may also shed light on the mechanism of how transcription factors are regulated during embryonic development.

The lens specificity of the -crystallin expression is conferred by an enhancer located in the third intron of the gene (Hayashi et al., 1987; Goto et al., 1990), and a short stretch of the DNA sequence TTGCTCACCT in the enhancer core region was found to be essential for lens-specific activity of the enhancer (Funahashi et al., 1991). We have identified a nuclear factor, EF1, which binds specifically to the DNA sequence and is present in lens cells and many other non-lens cells. To assess the significance of this factor in lens-specific regulation of the -crystallin gene, EF1 cDNA was molecularly cloned and characterized.

It was found that EF1 has interesting features as a DNA-binding protein and as an embryonic gene regulator. It has multiple DNA-binding domains, two separated zinc finger clusters and a homeodomain, and represses -crystallin enhancer element. It is turned on in the early process of organogenesis in mesodermal tissues, the nervous system and the lens, and its activity changes dynamically along with tissue maturation. These characteristics place EF1 not only among the regulators of -crystallin enhancer but among the

factors important for the regulation of postgastrulation embryogenesis.

MATERIALS AND METHODS

cDNA libraries

Random-primed cDNAs of 13 day chicken embryonic lens and oligo(dT)-primed cDNAs of lens and brain of the same stage were prepared from poly$(A)^+$ RNAs using a cDNA synthesis kit (Invitrogen). cDNAs were ligated with *Eco*RI-*Not*I linkers (Invitrogen) or *Eco*RI-*Bam*HI-*Not*I linkers (Takara), inserted into the *Eco*RI sites of gt10 or gt11 and packaged using a packaging extract (Stratagene).

Southwestern screening (Vinson et al., 1988; Singh et al., 1988) of the cDNA libraries

The random-primed lens cDNA library made on gt11 was screened for binding of HNW sequence (Funahashi et al., 1991). 2×10^5 phages were plated with Y1090 on ten 120 cm^2 plates and incubated at 42°C for 3.5 hours. Then nylon filters (Nytran; Schleicher & Schuell) impregnated with 10 mM isopropylthio- -D-galactopyranoside (IPTG) were overlaid, incubated at 37°C overnight, lifted, air-dried and processed for the guanidine-HCl denaturation-renaturation procedure (Vinson et al., 1988) in binding buffer [10 mM Hepes, 50 mM NaCl, 1 mM Na_2HPO_4, 0.1 mM $ZnSO_4$, 1 mM dithiothreitol (DTT), 5% glycerol]. The filters were blocked with 5% skim milk (Difco) in binding buffer and washed twice in binding buffer containing 0.25% skim milk. A pair of filters were placed in a plastic bag containing binding buffer with 0.3 mg/ml poly(dA-dT)(Pharmacia), 0.25% skim milk and end-labeled, octamerized HNW probe (23 ng/ml, 3×10^5 disints/minute/ml), and incubated at room temperature for 2 hours. The filters were washed by several changes of binding buffer and subjected to autoradiography.

Expression and purification of JF12 fusion proteins

JF12 cDNA insert of the original gt11 clone was placed in pGEX-3X(*Not*I) for fusion with glutathione-S transferase (GST) or in pMAL-c2 vector (New England Biolabs) for fusion with a maltose-binding protein (MBP). To generate a GST fusion protein, the cDNA excised by *Not*I digestion was inserted into the *Not*I site of pGEX-3X(*Not*I) which was constructed by insertion of a *Not*I linker (New England Biolabs) at the *Sma*I site of pGEX-3X (Smith and Johnson, 1988) so that GST and JF12 coding sequences become in frame. To obtain MBP fusion protein, the cDNA was excised by *Eco*RI digestion and inserted in the *Eco*RI site of pMAL-c2. The fusion proteins were produced in *E. coli* and affinity-purified according to the published procedure (Smith and Johnson, 1988) or to the manufacturer's instruction. In a rough estimate, 1 l of bacterial culture yielded 1 mg of GST fusion protein and 2 mg of MBP fusion protein. GST fusion protein was further purified using FPLC mono Q column (Pharmacia).

Generation of antiserum and affinity-purification of antibodies

Anti-GST:JF12 antiserum was raised in a rabbit by subcutaneous injection of the Mono Q-purified GST:JF12 fusion protein using Freund's adjuvant. The antiserum was loaded on a column of CNBr-activated Sepharose 4B (Pharmacia) coupled with MBP:JF12 fusion protein. Specific antibodies were eluted with 10 mM glycine (pH 2.5) and 100 mM triethylamine (pH 11.5). Eluates were combined and dialyzed against phosphate-buffered saline (PBS).

Gel mobility shift assay and western blotting of nuclear extracts

Nuclear extracts were prepared and subjected to gel mobility shift assay according to Funahashi et al. (1991). Immunoprecipitation and western blotting were done according to Kato et al. (1990) except that an ECL system (Amersham) was used in place of ^{125}I-labeled protein A.

Transfection and luciferase assay

Construction of luciferase reporter gene and EF1 expression vector pCMVX- EF1 has been described by Kamachi and Kondoh (1993). The chicken embryonic cells in primary culture (Hayashi et al., 1987) were seeded at the density of 2×10^5 cells per 3.5 cm dish the day before, and transfected with 1 μg plasmid DNA according to Chen and Okayama (1987), washed after 6 hours, and harvested after 24 hours for luciferase assay (deWet et al., 1987). Luciferase activity was measured according to Kamachi and Kondoh (1993) at 25°C using a 1251 luminometer (Wallac).

Immunohistology

Embryos staged according to Hamburger and Hamilton (1951) were fixed with 3.5% paraformaldehyde in Hepes-buffered saline (HBS) at 4°C for 2-5 hours, washed, impregnated in a graded series of sucrose in HBS (10, 15 and 25%) at 4°C, embedded in OCT compound (Miles Scientific) and frozen on dry ice. Sections 4 μm thick were made in a cryostat (Bright), placed on glass slides coated with gelatin-chromealum and air-dried. Sections were incubated successively with purified anti- EF1 antibodies, biotinylated anti-rabbit Ig (Amersham), fluorescein-conjugated streptoavidine (Amersham) and 4 ,6 -diamidino-2-phenylindole dihydrochloride (DAPI; 20 μg/ml) in Tris-buffered saline (TBS) containing 0.1% Tween-20 and 10% skim milk (Difco), with washings with TBS between the steps. Finally, the sections were mounted in Gelvatol (PBS containing 20% polyvinyl alcohol, 20% glycerol, and 2.5% 1,4-diazabycyclo-[2,2,2]-octane) and examined under an Axioplan microscope (Zeiss).

RESULTS

Southwestern cloning of putative δEF1 cDNA

For the purpose of cloning EF1 cDNA, the southwestern approach (Vinson et al., 1988; Singh et al., 1988) was employed. A cDNA library was constructed from lens poly$(A)^+$ RNA of 13 day chicken embryos in a gt11 expression vector, and screened for binding of the cDNA-encoded protein fused with an N-terminal portion of -galactosidase to the octameric HN fragment of the 1-crystallin enhancer. The HN fragment was 55 bp long, contained the binding site of EF1 near its 3 end, and exhibited stringently lens-specific enhancer activity in its multimeric form (Funahashi et al., 1991). In the previous study (Funahashi et al., 1991), we noted that EF1 binding was most clearly demonstrated in a gel mobility shift assay when poly(dA-dT) was used as competitor of non-specific DNA binding. Therefore, the same poly(dA-dT) was used in the screening.

From a total of 2×10^6 plaques screened, 5 positive clones were isolated. Analysis of binding specificity utilizing mutant probes (Funahashi et al., 1991) indicated that one of them encoded a protein that bound specifically to the EF1-binding site (blocks 5 through 6; Fig. 1A) (data not shown). This clone, JF12, was further characterized. Two of the remaining four, JF9 and JF10, coded for proteins with HMG boxes (JF9-encoded protein was a chicken HMG1) and bound preferentially to AT-rich blocks 10 to 3, and the other

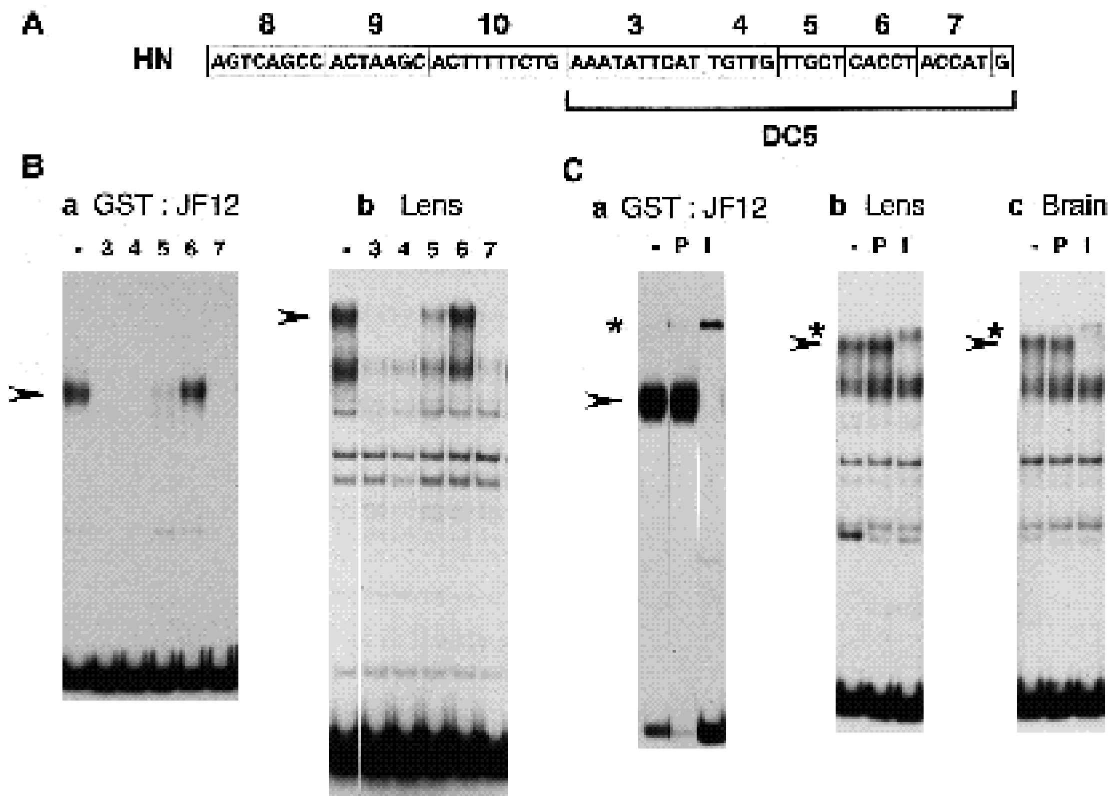

Fig. 1. Gel mobility shift assay of EF1 and GST:JF12 fusion protein, and the effect of anti-JF12 antibodies. (A) HN fragment sequence and its blocks. Mutant fragments had transversion type base alterations in the blocks indicated (Funahashi et al., 1991). DC5 is the 3 half of the HN fragment used for probe of gel shifts. (B) Effect of mutations of HN fragments as competitor of DC5 gel shift. (a) Affinity-purified GST:JF12 (5 ng per lane), and (b) lens nuclear extract (10 µg per lane). −, no sequence-specific competitor; 3-7, HN fragments with mutations in the designated blocks. Arrowheads indicate the position of the complex with GST:JF12 (a) and EF1 (b). Note that, for both GST:JF12 and EF1, the mutation in block 5 reduced and that in block 6 totally abolished binding of DC5 sequence. (C) Effect of anti-GST:JF12 fusion protein serum to factor-probe complexes. (a) GST:JF12 (100 ng per lane), (b) lens nuclear extract (10 µg per lane) and (c) brain nuclear extract (5 µg per lane). −, no serum; P, preimmune serum; I, anti-fusion protein serum. Arrowheads indicate the complexes of the fusion protein or EF1, and asterisk, the super-shifted complexes.

two, JF11 and JF13, without known DNA-binding motifs bound specifically to block 10 (Fig. 1A). These nucleotide sequences will appear in EMBL/GenBank/DDBJ databases under the accession numbers D14314 (JF9), D14315 (JF10), D14316 (JF11), D14317/8 (JF13).

The cDNA insert of JF12 was 826 base pairs long and an open reading frame (ORF) spanned the entire insert length. The ORF encoded three consecutive zinc fingers as putative DNA-binding motifs in its deduced amino acid sequence, as described below, and these fingers must have served as the DNA-binding domain of the -galactosidase:JF12 fusion protein.

Antibodies against JF12-encoded protein confirmed identity with δEF1

The cDNA sequence was excised from the phage vector and placed downstream of the GST sequence of a pGEX-based vector (Smith and Johnson, 1988). The GST:JF12 fusion protein was expressed in bacterial cells and purified using a glutathione column. This fusion protein had the same DNA-binding specificity as EF1 in gel mobility shift assay (Fig. 1B). In addition, this binding was dependent upon the presence of zinc ions and inhibited by EDTA, consistent with zinc fingers being the DNA-binding domain (data not shown).

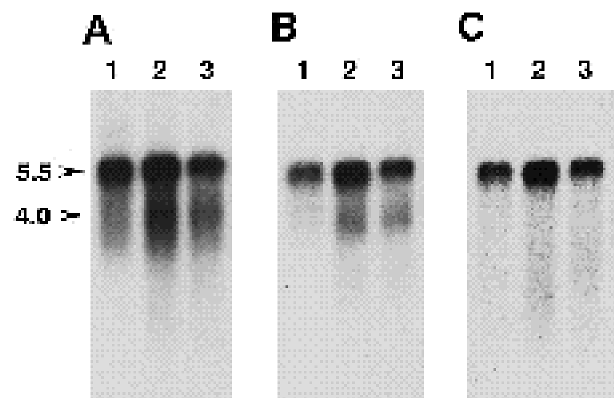

Fig. 2. Northern blot analysis of EF1 mRNA. 1.2 µg poly(A)$^{+}$ RNAs from 13 day embryonic lens (1), brain (2) and heart (3) were electrophoresed and hybridized with JF12 probe (A), RS12-113 (5) probe (B) and 3 probe derived from 3 half of RS12-32 (C) (see Fig. 3).

Antibodies were raised against the GST:JF12 fusion

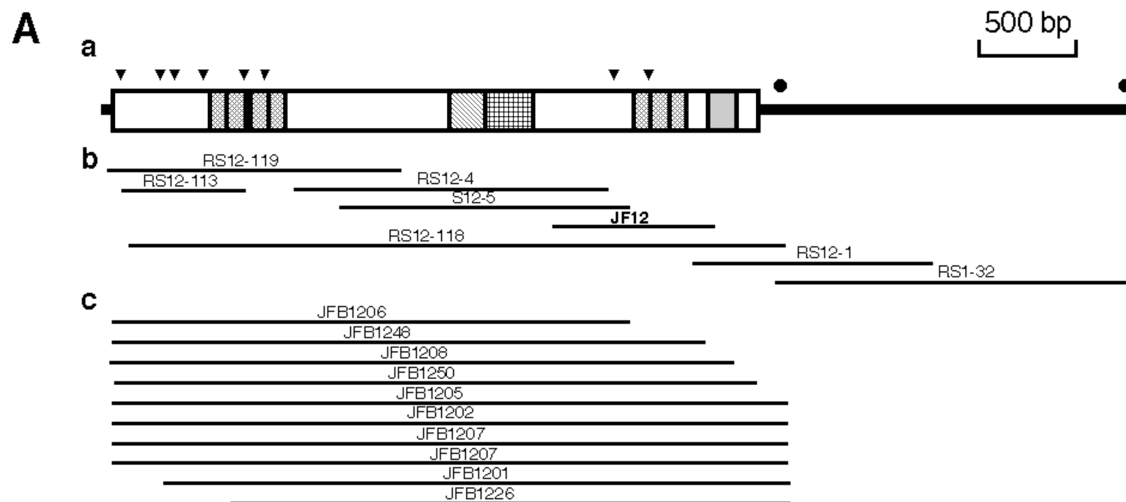

Fig. 3. cDNA clones of EF1. (A) A scheme of the cDNA and the clones. (a) Reconstructed full-length cDNA and assignment of the coding region. The coding region is boxed, and the characteristic domains, zinc fingers (▩), the homeodomain (▦), the proline-rich domain (▧) and the acidic domain (▢) are indicated. The wedges indicate locations of introns in the genomic sequence and the dots potential poly(A) addition sites. (b) Overlapping cDNA clones isolated from lens libraries. Clones RS12-32, RS12-118 and RS12-119 were from oligo(dT)-primed libraries and others from a random-primed library. (c) cDNA clones isolated from oligo(dT)-primed brain library screened with RS12-1 (5 half) and RS12-113 as probes. (B) Nucleotide sequence and encoded amino acid sequence. The nucleotide sequence was generated from RS12-119, RS12-118 and RS12-32. The last nucleotide shown in the figure (position 5271) is the point to which a complete match was found with the genomic sequence, and was followed by a poly(A) tract in clone RS12-32. Potential poly(A) addition signals are boxed. In the amino acid sequence, zinc fingers are numbered and underlined, with landmark cystine and histidine residues shown stippled; the homeodomain are underlined by a broken line; potential glycosylation sites are circled; potential phosphorylation sites are boxed and stop codon are indicated by an asterisk. This nucleotide sequence will appear in EMBL/GenBank/DDBJ databases under the accession number D14313.

protein and examined for their effect on DNA- EF1 complex. As shown in Fig. 1C, anti-GST:JF12 abolished the EF1-DC5 complex in the same way as the GST:JF12-DC5 complex, and produced the supershifted bands. The effects of the antibodies on EF1 were essentially the same regardless of its source (Fig. 1C for lens and brain).

Thus, because of having the same sequence specificity of DNA binding and the same antigenicity, we concluded that the cDNA sequence of the clone JF12 represented a portion of EF1 that included a DNA-binding domain.

Entire cDNA sequence and the encoded protein

The cDNA insert of JF12 was used to probe northern blots of poly(A)$^+$ RNAs prepared from various embryonic organs. A major RNA species of 5.5 kb was present in all organs examined (Fig. 2A), as expected from the apparently ubiquitous distribution of EF1-binding activity (Funahashi et al., 1991). In addition, a second mRNA species of 4 kb was detected in non-lens tissues.

Starting from JF12, overlapping cDNA clones were isolated from gt10 and gt11 libraries of embryonic lens cDNAs, as shown in Fig. 3A(b). The cDNA sequence reconstructed from the overlapping clones was 5271 bases long without including poly(A), consistent with the 5.5 kb mRNA detected on northern blots. The longest ORF spanned 3342 bases (Fig. 3A), and coded for a protein of 124×10^3 M_r. Western blotting of nuclear extracts of embryonic lens and brain with anti-JF12 (i.e. anti- EF1) serum detected a band corresponding to 170×10^3 M_r, apparently larger than the coding capacity of the cDNA. However, when the cDNA reconstructed from clones 12-118 and 12-119 was placed in an expression vector pCDM8 (Seed, 1987) and transfected to COS7 cells, the cDNA directed synthesis of the same 170×10^3 M_r band (Fig. 4, lane 1). This indicated successful cloning of the entire EF1 coding sequence.

In the reconstructed cDNA sequence, the ORF was preceded by an 11 base 5 untranslated sequence, and followed by a 1918 base 3 untranslated sequence. Examination of the 3 untranslated region indicated a potential alternative poly(A) addition site immediately downstream of the 3 end of the ORF. This alternative poly(A) addition site accounted for the 4 kb form of mRNA observed in non-lens tissues. In northern blots, the 5 proximal cDNA probe detected both RNA species (Fig. 2B), but the probe downstream of the first poly(A) site detected only the 5.5 kb species (Fig. 2C).

When an oligo(dT)-primed cDNA library of 13 day embryonic brain was screened with clones RS12-113 and the 5 half of RS12-1 [probes representing N-proximal and C-proximal regions, respectively, of the coding sequence; Fig. 3A(b)], the majority of the cloned cDNAs started near the initiator Met codon and terminated approximately 160 bases downstream of the termination codon [Fig. 3A(c)], consistent with the same mRNA start site between lens and brain, and also with the frequent use of the alternative poly(A) addition signal in the brain. cDNA sequences of the coding region were identical to those of the lens except for base changes to synonymous codons in a few places, which likely reflected genetic polymorphism in the chicken population. Thus, it was concluded that EF1 proteins are identical in amino acid sequence in the lens and brain, and perhaps in other non-lens tissues.

B

```
1    GCGAGAGGATCATGGCGGATGGCCCCAGG TGTAAGCGCAGGAAGCAGGCGAACCCGAGG CGCAACAACGTCACCAATTA CAATAACGTGATAGAAGCAA ACTCAGATTCAGATGACGAA GATAAATTGCATATAGTGGA AGAAGAAAGT
                               M  A  D  G  P  R  C  K  R  R  K  Q  A  N  P  R  R  N  N  V  T  N  Y  N  N  V  I  E  A  N  S  D  S  D  D  E  D  K  L  H  I  V  E  E  E  S

150  ATAACTGATGCAGCAGACTGCGATGCAAGT GTGCCAGAAGATGACTTGCC AACAGACCACACCGTGTTAC CAGAAAACAGTGAGAGGGAA GGGAGCACAAACAGCTGCTG GGAAGATGAAGGAAAAGAAA CAAAGGAAATCCTGGGGCCT
     I  T  D  A  A  D  C  D  A  S  V  P  E  D  D  L  P  T  D  H  T  V  L  P  E  N  S  E  R  E  G  S  T  N  S  C  W  E  D  E  G  K  E  T  K  E  I  L  G  P

300  GAAGCTCAGTCAGATGAAGTTGGATGTACA GTAAAAGAAGATGAATGTGA TTCTGATGCAGAAAATGAAC AGAACCATGACCCTAATGTT GAAGAATTCCTTCAACAAGA AGATACAGCTGTTATTTACC CTGAAGCACCTGAGGAGGAC
     E  A  Q  S  D  E  V  G  C  T  V  K  E  D  E  C  D  S  D  A  E  N  E  Q  N  H  D  P  N  V  E  E  F  L  Q  Q  E  D  T  A  V  I  Y  P  E  A  P  E  E  D

450  CAGAGACAAGGCACACCAGAAGCTAGTGGT CAGGATGAAAATGGAACGCC CGATGCATTTTCCCAGCTGC TCACCTGCCCGTACTGTGAC AGAGGGTACAAACGCTTCAC CTCTCTGAAGGAACACATTA AATACCGCCATGAAAAGAAC
     Q  R  Q  G  T  P  E  A  S  G  Q  D  E  N  G  T  P  D  A  F  S  Q  L  L  T  C  P  Y  C  D  R  G  Y  K  R  F  T  S  L  K  E  H  I  K  Y  R  H  E  K  N
                                                                       1

600  GAGGATAACTTCAGTTGCTCCCTCTGCAGT TACACGTTTGCGTATAGAAC ACAGCTGGACCGCCACATGA CGTCGCACAAGTCAGGACGA GATCAGAGACATGTGACGCA GTCCAGTGGTAATCGAAAAT TCAAGTGCACTGAATGTGGA
     E  D  N  F  S  C  S  L  C  S  Y  T  F  A  Y  R  T  Q  L  D  R  H  M  T  S  H  K  S  G  R  D  Q  R  H  V  T  Q  S  S  G  N  R  K  F  K  C  T  E  C  G
           2                                                                                                                 3

750  AAAGCTTTTAAATATAAACATCATCTAAAG GAACACCTACGAATTCACAG TGGAGAGAAGCCATATGAGT GCCCAAACTGCAAGAAACGT TTTTCCCATTCTGGTTCATA CAGCTCACACATAAGCAGTA AGAAGTGTATTGGTTTGATG
     K  A  F  K  Y  K  H  H  L  K  E  H  L  R  I  H  S  G  E  K  P  Y  E  C  P  N  C  K  K  R  F  S  H  S  G  S  Y  S  S  H  I  S  S  K  K  C  I  G  L  M
                                                                4

900  CCCGTAAAGGGTCGAGCTCGGTCAGGGCTC AAGACGTCTCAGTGCTCCTC CCCTTCCCTTTCTGCATCGC CCGGTAGCCCAGCAAGACCA CAGATACGACAAAAGATAGA AAATAAACCCTTGCAAGAGC AACTTCCTGTTAACCAAATT
     P  V  K  G  R  A  R  S  G  L  K  T  S  Q  C  S  S  P  S  L  S  A  S  P  G  S  P  A  R  P  Q  I  R  Q  K  I  E  N  K  P  L  Q  E  Q  L  P  V  N  Q  I

1050 AAAACTGAACCTGTGGATTATGAATTCAAG CCCATAGTGGTTGCTTCAGG AATTAATTGTTCGACCCCTT TGCAGAATGGGGTTTTTAGT GGTGGTAGCCCGTTGCAGGC AACCAGTTCTCCTCAGGGTG TGGTGCAAGCTGTTGTTCTG
     K  T  E  P  V  D  Y  E  F  K  P  I  V  V  A  S  G  I  N  C  S  T  P  L  Q  N  G  V  F  S  G  G  S  P  L  Q  A  T  S  S  P  Q  G  V  V  Q  A  V  V  L

1200 CCAACAGTGGGTCTGGTGTCTCCCATAAGC ATCAACTTAAGTGACATTCA AAATGTACTTAAAGTGGCAG TGGATGGTAATGTAATAAGG CAAGTATTGGAAAACAATCA CGCTAATCTTGCATCCAAAG AACAAGAAACAATCAGCAAT
     P  T  V  G  L  V  S  P  I  S  I  N  L  S  D  I  Q  N  V  L  K  V  A  V  D  G  N  V  I  R  Q  V  L  E  N  N  H  A  N  L  A  S  K  E  Q  E  T  I  S  N

1350 GCATCTATACAACAAGCTGGCCATTCCCTC ATTTCAGCTATCAGTCTTCC TTTGGTTGACCAAGATGGGA CAACCAAAATTATCATCAAC TACAGCTTGGAGCAGCCAAG TCAACTTCAGGTTGTTCCCC AAAATCTAAAAAAAGAACAT
     A  S  I  Q  Q  A  G  H  S  L  I  S  A  I  S  L  P  L  V  D  Q  D  G  T  T  K  I  I  I  N  Y  S  L  E  Q  P  S  Q  L  Q  V  V  P  Q  N  L  K  K  E  H

1500 TCTGTTCCTACAAACAGTTGCAAAAATGAA AAGTTACCAGAAGATCTTAC GGTGAAGTCTGAGAAAGATA AAAACTTTGAAGGAGAGACC AACGATAGCACTTGTCTTCT TTGTGATGACTGTCCAGGAG ATCTTAATGCACTTCAAGAA
     S  V  P  T  N  S  C  K  N  E  K  L  P  E  D  L  T  V  K  S  E  K  D  K  N  F  E  G  E  T  N  D  S  T  C  L  L  C  D  D  C  P  G  D  L  N  A  L  Q  E

1650 TTAAAGCACTATGAAACAAAAAATCCTCCT CAGCTTCCCCAGTCCAGTGG AACAGAAGCTGAGAAGCCCA GCTCCCCTGCCCCATCAGAA ACTGGGGAGAACAACTTATC TCCTGGTCAGCCACCCTTAA AGAACCTTTTATCGCTCCTA
     L  K  H  Y  E  T  K  N  P  P  Q  L  P  Q  S  S  G  T  E  A  E  K  P  S  S  P  A  P  S  E  T  G  E  N  N  L  S  P  G  Q  P  P  L  K  N  L  L  S  L  L

1800 AAAGCATATTATGCATTAAATGCACAACCA AGCGCAGAAGAGCTTTCAAA AATCGCCGATTCCGTGAACC TACCACTGGATGTGGTAAAA AAGTGGTTTGAAAAAATGCA AGCTGGACAAATTTCTGTGC AGTCTTCTGGACCATCTTCT
     K  A  Y  Y  A  L  N  A  Q  P  S  A  E  E  L  S  K  I  A  D  S  V  N  L  P  L  D  V  V  K  K  W  F  E  K  M  Q  A  G  Q  I  S  V  Q  S  S  G  P  S  S

1950 CCTGAACAAGTTAAAATAAGCAGTCCCACA GATAACGATGATCAAGCAGC AACTACAAATGAGAGTGAAC CCCAGAACAGCACAAACAAC TCACAAAATCCGGCCAATAC AAGTAAATCTCAGACTTCAT CAGGGGGATCAACTCAGAAT
     P  E  Q  V  K  I  S  S  P  T  D  N  D  D  Q  A  A  T  T  N  E  S  E  P  Q  N  S  T  N  N  S  Q  N  P  A  N  T  S  K  S  Q  T  S  S  G  G  S  T  Q  N

2100 GGTTCGCGCAGTAGCACGCCATCCCCATCA CCACTAAACCTTTCTTCATC AAGAAATTCACAGGGTTATA CGTACACAGCAGAGGGTGTA CAAGAAGAGCCACAAATGGA ACCTCTTGACCTTTCACTAC CAAAGCAACATGGAGAACTG
     G  S  R  S  S  T  P  S  P  S  P  L  N  L  S  S  S  R  N  S  Q  G  Y  T  Y  T  A  E  G  V  Q  E  E  P  Q  M  E  P  L  D  L  S  L  P  K  Q  H  G  E  L

2250 TTGGAAAGATCTACCATAACTAGTGTTTAC CAGAACAGTGTTTATTCTGT CCAGGAAGAACCTTTGAACT TAACTTGTGCAAAAAAAGAA CCACAAAAGGACAACAGTAT TACAGACTCTGATCCTATTG TAAATGTAATCCCACCAAGT
     L  E  R  S  T  I  T  S  V  Y  Q  N  S  V  Y  S  V  Q  E  E  P  L  N  L  T  C  A  K  K  E  P  Q  K  D  N  S  I  T  D  S  D  P  I  V  N  V  I  P  P  S

2400 GCCAATCCCATAAATATTGCTATACCTACA GTCACTGCCCAGTTACCTAC AATTGTTGCCATTGCTGACC AGAACAGTGTTCCATGCTTG AGAGCTCTTGCTGCCAATAA GCAAACCATTTTGATTCCCC AGGTGGCTTATACGTACTCT
     A  N  P  I  N  I  A  I  P  T  V  T  A  Q  L  P  T  I  V  A  I  A  D  Q  N  S  V  P  C  L  R  A  L  A  A  N  K  Q  T  I  L  I  P  Q  V  A  Y  T  Y  S

2550 ACTACAGTTAGTCCTGCAGTTCAGGAAACA CCACCAAAACAGACCCAAGC CAACGGAAGTCAGGACGAAA GGCAGGACACTAGCTCAGAA GGAGTATCGAATGTAGAAGA TCAAAATGATTCTGATTCAA CACCCCCAAAGAAAAAGATG
     T  T  V  S  P  A  V  Q  E  T  P  P  K  Q  T  Q  A  N  G  S  Q  D  E  R  Q  D  T  S  S  E  G  V  S  N  V  E  D  Q  N  D  S  D  S  T  P  P  K  K  K  M

2700 AGAAAGACAGAAAATGGGATGTATGCATGT GATTTGTGTGACAAAATATT CCAGAAGAGCAGCTCATTAT TGAGACATAAGTATGAACAC ACAGGTAAAAGACCTCATGA GTGTGGAATCTGTAAAAAAG CCTTCAAACACAAACACCAT
     R  K  T  E  N  G  M  Y  A  C  D  L  C  D  K  I  F  Q  K  S  S  S  L  L  R  H  K  Y  E  H  T  G  K  R  P  H  E  C  G  I  C  K  K  A  F  K  H  K  H  H
                        5                                                                         6

2850 TTGATCGAACACATGCGACTGCATTCTGGG GAGAAACCCTACCAATGTGA CAAATGCGGGAAGCGCTTCT CACACTCGGGGTCGTACTCT CAGCACATGAATCATCGCTA CTCCTATTGCAAAAGAGAGG CAGAGGAGCGTGACAGCACA
     L  I  E  H  M  R  L  H  S  G  E  K  P  Y  Q  C  D  K  C  G  K  R  F  S  H  S  G  S  Y  S  Q  H  M  N  H  R  Y  S  Y  C  K  R  E  A  E  E  R  D  S  T
                                      7

3000 GAGCAGGAGGAGGTGGGACAGGAGGTGCTC AGCAGCGAGCATGCTGGTGC CAGAGCATCGCCATCGCAGA TCGACTCCGATGAGAGAGAG AGTCTAACCAGGGAGAAGAA GGAAGATAGTGAAAAAGAGG AAGAGGAGGAGGAAGAAAAG
     E  Q  E  E  V  G  Q  E  V  L  S  S  E  H  A  G  A  R  A  S  P  S  Q  I  D  S  D  E  R  E  S  L  T  R  E  E  E  E  D  S  E  K  E  E  E  E  E  E  E  K

3150 GACGTAGAAGGACTTCAGGAAGAAAAAGAA TGTAGAAAACTACAAGATGT AGAGGAGGAAGAAGAAGTAG AAGAAGAAGAAGAAGAAGAG GAAGGGAAAACTGAAGGTAA CAAGAACGATGACGTTGTAA ATCGAGCAAGCAATGCAGAA
     D  V  E  G  L  Q  E  E  K  E  C  R  K  L  Q  D  V  E  E  E  E  V  E  E  E  E  E  E  E  G  K  T  E  G  N  K  N  D  D  V  V  N  R  A  S  N  A  E

3300 CCAGAAGTTATACAGAGCAATGGGCAGGTG TCAGAAGAAAAAACAAATAA AGCTTAAATGATAGTTTTTT CTATATGGAAAAAAATAAAT CAGTAATGAAGATCGTTCTA TATTATACATGCTTTTCATG GAAACACGGTAGCCTGCATA
     P  E  V  I  Q  S  N  G  Q  V  S  E  E  K  T  N  K  A  *

3450 CTGTGATTTCTGTTCACTACTGTGTAAAAA AATATATATATAAAAAAAAA GAAAAAAGAAAAAAAAAAGA AATCCAGGTGTGCCTGAACC TCAAACTTAGTAATTTTTCACACAGTTTTC AAAGTTAGGAACAAGTTTGT AACATGAAGC
3600 AGATTAGAAAAATACAGTGACTCTGGAAGC AAAGGTTTAACAGGTTGTAG GAAACCTGATTTATTAGACA AGCATCTGGCATTGTTTGTT TTATCAGTATTAATTCTCATTCTTACGTTC ATTAATTTTTAAGCTGTACA ATTGGGAGAG
3750 ATTTATATGATTTTTTTTTTTTTTTATGGCA AAACATTTCCTAAATCCGCT TCAGTATTTTATTATGTTTT TAAAATGTGAGAACTTCTGC ACTACAGAATTCCCTTCTGCAGAAACCTCT AACAGTGCTTGGCTACCTTT GTAAATTCAA
3900 CCTAGAAGGTAGCGGTTATAACATTAGATG TTGTAGTAGAGCATATTATC ATTTAAAGTGTATTGTTAGC CTTAAGAAAGCAGATGATAG AAGAACCGAAGTTTCTTACTCATGTGGTTA AAAATGTAGTTGAAAAGATT GTTGTTGAGT
4050 TCTGATTGCAGGGACTAACAGTGTTAATAC CGGTAAGGACAGCAAAGTCG TCAAAATCAGTGTTTGTAAT TGTGTTTTGAGTATGTAGTA ACAATATGAAGGATATGATATGAAGCTTTG TATCTCCTTTGGCCTTATGC AAGACCTGTG
4200 TGCTGTAAGTGCCATTTATCAGTATTACAA AGGCTCTCACCGGCCTTTCT CCAATGCGTGGCCTACAATA ACTAGCATTTGTTGATTTGT ATCAAACTTCTAATCTAATCATGGGGGAAA AAACACTCAGCCATCAGAGA AGTAAGAACA
4350 CTGGTATCTTTCATTATTTAATTATTTGGT GGTATTTTTGGTTAATACCT GACTTGCAGAATTTCTCTAC GATAATTACAAGGGAAATTT TCTAATCTGCTTTATTGCTTTATCAGTTTA CTTTAATTTCAGACACATAC ATGACATGTT
4500 ATCTGGCTCATAAGTATTTTCAAATGTTAG TTATTATGGACCCCATTTAT TAACAATGTTAGCTGATTTT TACCTATCAGTATTATTTTA TTTCTTTTAGTTTATAGATCTGTGCAACAT TTTGTACTGTATGTCTCCAA ACCTGGCAGT
4650 ATTAATACCCTTCTTACTGACATATGTACC TTTTTAGTTTTAGAAAACTT TTATATTTATGTGTCTTATT TTTATATTTCTTTATTTATT ACACAGTGTAGTGTATAATACTGTAGTTTG TATTAATACAATAATATATT TTAGTATGAA
4800 AATTTGGAAAGTTGATAAGATTTAAAGTAG AGATGCAATCGGTTCTCTTG CATTGAGATTTGATTTAACA GTGTTATGTTAACATTTATA CTTGCCTTGGACTGTAGAACAGATGAATTA AAATGGGAATGTATTAATTT TACAATTGCA
4950 ATCAATTCATTTTACCTTTACCCAGTTTTT AATATAAAGCTCTTAAATTT TGAAATTCACTGTGTGACTA ATAGCATGATGCTCTGCAGT TTTATTAAGAGATTAGCCTAACCATAAAAC TCTCATTTCCTTAGTAAGCC AAATTAGAGT
5100 AATCTTATAAACAGTGTTGGGAACAACGTT TAACTTTTTTGTGCCAATTT GTTCCTGTATTCATGTATGT ATGTTACAAATCTGAAACTT CATTTTTAAGTTCCTTGTTACATCATGGTC ATTTTCTAGTTTTTTACCAG ACTCCCCATC
5250 TCACAATAAAATACGTCAACAA
```

Southern blot of chicken DNA using the EF1 cDNA fragments as probe indicated that the EF1 gene is likely present as a single copy per haploid genome (data not shown). Comparison of the nucleotide sequence of EF1 cDNA with that of the cloned genomic gene (R. Sekido, M, Okanami, unpublished result) indicated that the cDNA sequence was divided into nine exons, as shown in Fig. 3A(a), and that the initial Met codon of the cDNA ORF was located in the proximity of the promoter region.

Multiple DNA-binding domains of δEF1: zinc finger clusters and a homeodomain

Amino acid sequence of EF1 deduced from the cDNA sequence revealed an interesting feature of this DNA-binding protein (Fig. 3B). In addition to three zinc fingers near the C terminus (C fingers) initially identified in clone JF12, there were four additional zinc fingers of Cys_2-His_2 type near the N terminus (N fingers). The fourth N finger (finger 4) had a cystine in the place of a histidine in ordinary

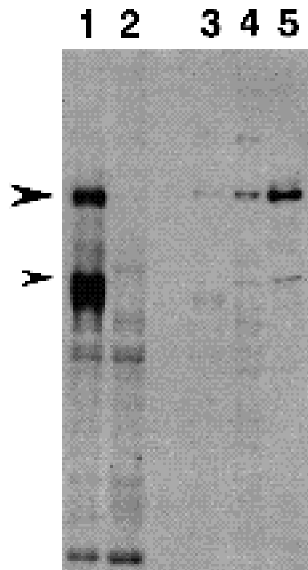

Fig. 4. Western blotting of EF1 in embryonic tissues and expressed from cDNA using an anti-GST:JF12 antiserum. Lane 1, A nuclear extract of COS7 cells (6 μg) transfected with EF1 cDNA (a composite of RS12-119 and RS12-118) on pCDM8 vector (Seed, 1987); lane 2, nuclear extract untransfected COS7 cells (6 μg); lane 3, chicken embryonic lens extract (6.5 μg); lane 4, chicken embryonic brain extract (5 μg); lane 5, immunoprecipitate of the same brain extract of a five times larger amount, showing that only the bands indicated by the arrowheads are the authentic antigens. The upper band indicated by the large arrowhead represents EF1, and corresponds to the size of 170×10^3 M_r. The lower band indicated by the small arrowhead (120×10^3 M_r) and observed in the transfected COS7 cells and in brain extract represents proteolytic products of EF1 generated during preparation, since inclusion of a protease inhibitor cocktail in the extraction medium significantly reduced the intensity of the bands.

fingers, but other characteristics (location of phenylalanine, clustering of basic amino acid residues adjacent to a histidine, and spacings between typical amino acid residues) classified this finger as the Cys_2-His_2 type.

Very interestingly, a homeodomain was found in between the zinc finger clusters. The amino acid sequence of the homeodomain had the conserved landmark amino acid residues, and was similar to POU homeodomains [especially *Drosophila* I-POU (Treacy et al., 1992) and rat Brn-3 (He et al., 1989)] in the regions assigned to helices. An intriguing characteristic of the EF1 homeodomain was that it lacks a basic amino acid cluster, which is usually found on its N-terminal side.

In the middle of the amino acid sequence, there was a domain rich in proline, and to the C-terminal side of the C fingers there was a domain highly rich in glutamic acid (acidic domain; Fig. 3A(a) and B), features common to transcriptional activators (Mitchell and Tjian, 1989; Seipel et al., 1992). There are several potential phosphorylation sites and a number of potential glycosylation sites (Fig. 3B).

A search for sequence similarity in DNA sequence data base revealed that the C-terminal half of the EF1 was almost identical to human Nil-2-a (Williams et al., 1991). This will be discussed later.

A combination of multiply clustered zinc fingers and interposed homeodomain(s) has recently been reported for *Drosophila* ZFH proteins (ZFH-1 and 2) (Fortini et al., 1991; Lai et al., 1991) and for a rat -fetoprotein enhancer binding protein (Morinaga et al., 1991). It is interesting to note that EF1 and ZFH-1 not only have an organization of DNA-binding domains analogous to each other (Fig. 5A), but also have similar amino acid sequences in these DNA-binding domains (Fig. 5B). Especially fingers 3 to 4 of EF1 were very similar to fingers 4 to 5 of ZFH-1 and also fingers 5 to 7 of EF1 were similar to fingers 7 to 9 of ZFH-1 (Fig. 5B). It was also noted that fingers 3 to 4 and 6 to 7 of EF1, i.e., the last two N fingers and C fingers are similar to each other (Fig. 5C). In regions other than the DNA-binding domains, no significant similarity of amino acid sequence was observed between EF1 and ZFH-1.

Transrepression by δEF1

To address the question of what activity EF1 has as a transcriptional regulator, we set up an experiment in which EF1 expression vector was cotransfected with reporter luciferase gene carrying the HN fragment octamer (Fig. 6A). The HN fragment has strictly lens-specific enhancer activity (Funahashi et al., 1991). The luciferase gene was driven by the minimum -crystallin promoter that has no tissue preference (Hayashi et al., 1987; Kamachi and Kondoh, 1993). Octamerized HN enhancer fragment carrying the wild-type sequence (HNW) was placed upstream of the promoter to activate expression. To control the extent of activation, we used octamerized HN6 fragment carrying a mutation that totally inactivated the enhancer and EF1 binding (Funahashi et al., 1991; Fig. 6B). In the absence of exogenous EF1 expression, octameric HNW fragment elicited thirteen-fold activation over HN6 mutant in lens cells. By cotransfection of an increasing amount of EF1 expression vector, there was a progressive decrease of the enhancer activity of the HN fragment (Fig. 6C). Under the condition that the molar input of EF1 expression vector was nearly the same as that of the reporter luciferase gene, enhancer activity was decreased by fivefold. By contrast, the luciferase gene carrying the HN6 mutant fragments was not affected by exogenous EF1 (Fig. 6E,F). The repressing effect of co-transfected EF1 expression vector was observed only in lens cells (Fig. 6C,D). These results indicated that EF1 functions as a repressor of the enhancer by binding to the HN fragment.

Expression of δEF1 during embryogenesis

Utilizing highly purified anti- EF1 antibodies, tissue distribution of EF1 was histologically examined focusing on early developmental stages of the chicken embryos. The results are summarized in Table 1. In brief, the major sites of EF1 expression in the embryo were the mesodermal tissues, neuroectoderms, neural crest derivatives and the lens. In most of these tissues, EF1 expression became detectable after outline of the tissue is established. Amnion expressed EF1 from the beginning.

At stage 8 when three germ layers had just been established in the rostral half, no EF1 expression was detectable (Fig. 7A,B). As the development proceeded past stage 10, EF1 expression was initiated in the mesoderm, initially in

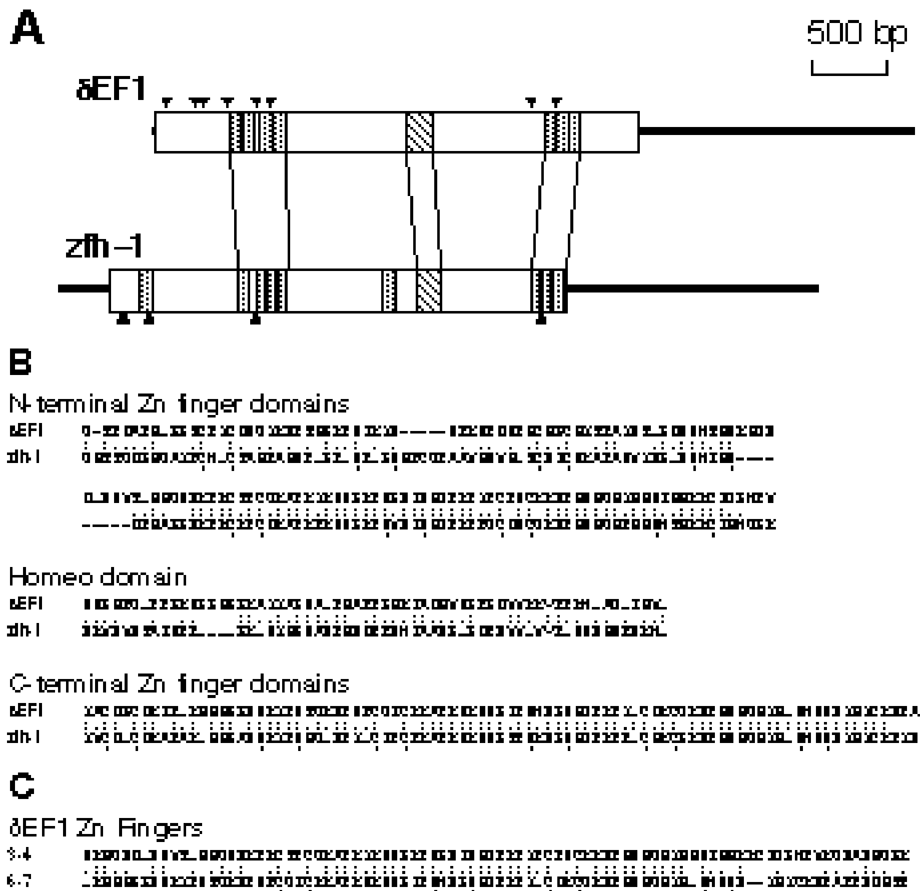

Fig. 5. Comparison of zinc fingers and homeodomains between EF1 and ZFH-1. (A) Schematic alignment of zinc finger domains (cross-hatched) and homeodomains (hatched) on cDNA sequences where coding regions are boxed. (B) Comparison of amino acid sequences between DNA-binding domains. The cystine and histidine residues involved in zinc ion binding are indicated by asterisks. Identical amino acid residues are indicated by **:**, and the residues of similar hydrophobicity by **.** (C) Comparison of N fingers 3-4 and C fingers 6-7 of EF1.

Table 1. Summary of immunohistological analysis of EF1 expression at various developmental stages

	Stage (number of somites)											
Tissue	8 (4)	8+ (5)	10 (10)	10+ (11)	11− (12)	11 (13)	12− (15)	12+ (17)	13− (18)	13+ (20)	14 (22)	18 (30-36)
Ectoderm												
Neural tube	−	−	−	−	−	−	−	+*	+*	++*	++	++
Retina	/	/	/	/	/	/	/	/	+*	++*	++	++
Lens	/	/	/	/	/	/	/	/	/	/	−	+*
Rathke's pouch	/	/	/	/	/	/	/	/	/	/	−	−
Otic vesicle	/	/	/	−	−	−	−	−	−	−	−	−
Neural crest derivatives												
Cephalic mesenchyme	/	/	−	+	+	+	++	++	++	++	++	++
Mesenchyme of visceral arches	/	/	/	/	/	/	/	/	/	/	+	+
Mesoderm												
Notochord	−	−	+	+	+	++	++	++	+++	+++	+++	+++
Somites	−	−	−	−	−	+*	+*	+*	++*	dm +++ sc ++	d ++ m +++ sc ++	++ ++++ +++
Lateral plate	−	−	−	−	−	−	−	−	+*	++*	++*	+++*
Nephrotome	/	/	−	−	−	−	−	−	+*	++*	++* W +	+++* −
Heart	/	/	+	+	+	++	++	+++	+++	+++	+++	+++
Endoderm												
Pharynx/ Digestive tract	/	/	−	−	−	−	−	−	−	−	−	−
Amnion	/	+	++	++	++	++	+++	+++	+++	+++	+++	+++

The stage of development was determined according to Hamburger and Hamilton (1951).
/, the tissue has not developed yet; −, no significant expression; + to ++++, weak to very strong expression.
*, scattered or non-homogeneous distribution of the EF1-expressing cells.
dm, dermamyotome; sc, sclerotome; d, dermatome; m, myotome; W, Wolffian duct.

Fig. 6. Transrepression of HN fragment-bearing luciferase gene by cotransfection of EF1 expression vector. (A) Schemes of the EF1 expression vector driven by cytomegalovirus promoter (top) and the luciferase-encoding reporter gene driven by the minimum -crystallin promoter and carrying octamerized HN fragment in the upstream (bottom). (B) The nucleotide sequences of wild-type HN fragment monomer (HNW) and of HN6 mutant. Linker sequences are in lower case letters and base sequence alteration is indicated by the stippled box. (C-F) The change of relative luciferase activity with increasing amount of EF1 expression vector, taking the luciferase activity without exogenous EF1 as 100%. Total copy number of expression vector plasmids and total amount of DNA (1 μg) were kept constant by inclusion of insert-less pCMVX vector and inert pUC plasmid in the transfection mixture, which minimized the effect of titration of transcription factors by exogenous DNA (Goto et al., 1990). (C) With wild-type HN fragments (HNW) in lens cells, (D) with HNW fragments in fibroblasts, (E) with HN6 mutant fragments in lens cells and (F) with HN6 mutant fragments in fibroblasts. The relative luciferase activity without exogenous EF1 compared to HN fragment-less luciferase gene was 31, 0.5, 2.4 and 1.6 for C, D, E and F, respectively.

the notochord and then in the somites, the nephrotomes and the lateral plates (Fig. 7C-F). This was followed by expression in the neural tube which began at stage 12. The beginning of EF1 expression in these tissues was characterized by the scattered appearance of EF1-positive nuclei in a tissue without remarkable polarity of distribution (Fig. 7C,E; Table 1, asterisks). With development, the fraction of EF1-positive nuclei increased to nearly 100% in a tissue.

At stage 14 when somitic tissues differentiated, the level of EF1 expression increased in the myotome and this became the most prominent site of expression (Fig. 8A,B for stage 18 embryo). In the same section, it is seen that expression of EF1 in the neural tube was lower than in the somites. Migrating neural crest cells of the dorsal trunk did not express EF1 (Fig. 8A, arrowheads), but those that colonized in the spinal ganglia (Fig. 8A,B) and in the mesenchymal compartment of the visceral arches (Fig. 8C,D) expressed EF1. In the mesonephros, Wolffian duct initially expressed a modest level of EF1 but soon ceased its expression (data not shown). Condensed mesenchymes and primitive mesonephric tubules strongly expressed EF1, but once tubular structures were complete and connected to Wolffian duct, they stopped EF1 expression (Fig. 8E,F).

When the lens placode was formed in the ectoderm and then began to invaginate (stage 14), -crystallin expression was clearly demonstrated but EF1 was not detectable (data not shown). EF1 became barely detectable only at the stage when the lens vesicle was formed (stage 18; Fig. 9A,B). This indicates that EF1 itself is probably not required for initiation of -crystallin expression. However, as lens devel-

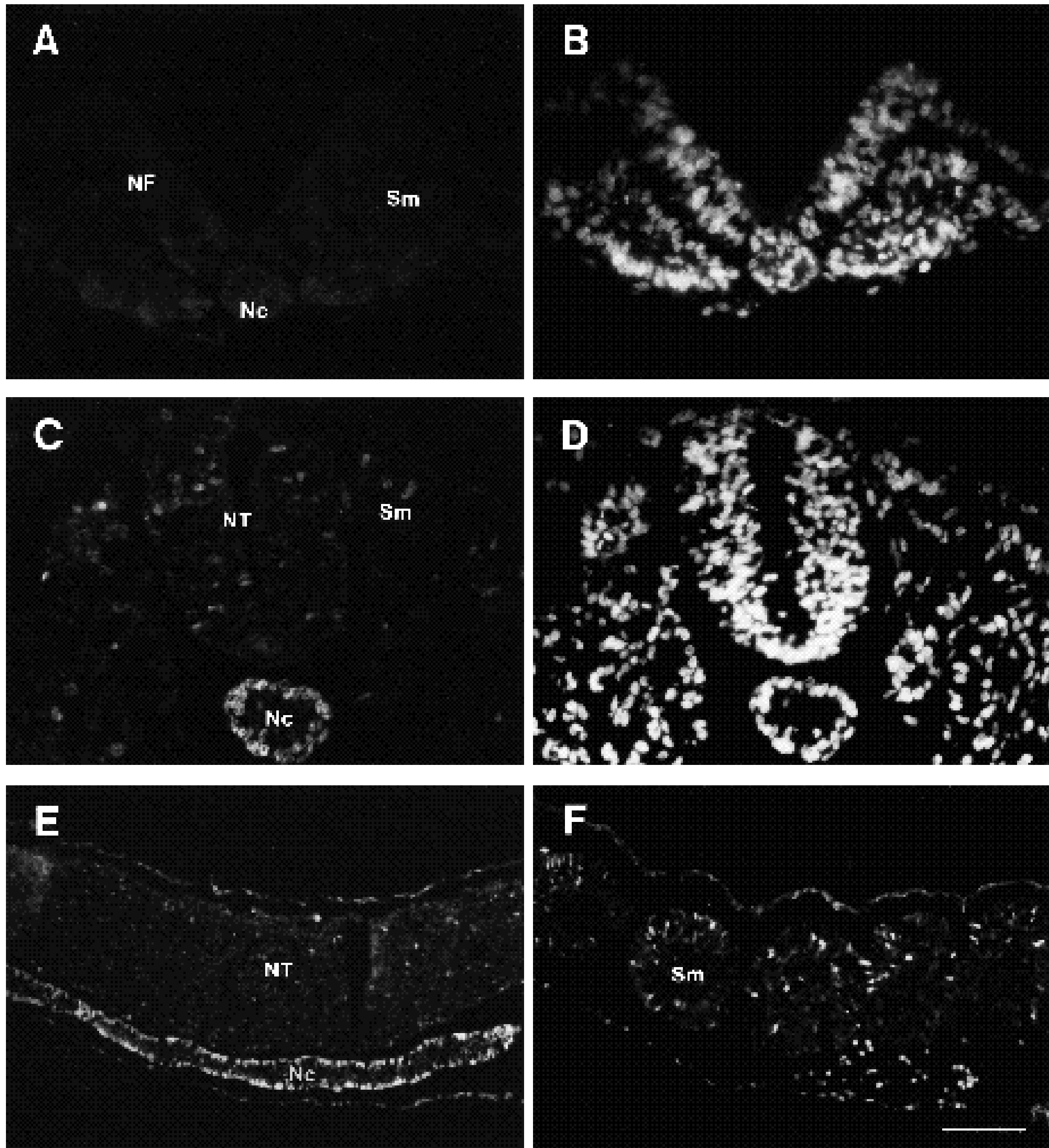

Fig. 7. Immunohistology of EF1 expression in the trunk of postgastrulation embryos. (A,B) Stage 8 cross section through neural fold (NF), somite (Sm) and notochord (Nc) stained with anti- EF1 for FITC-fluorescence (A) and with DAPI (B). (C,D) Stage 14 cross section stained with anti- EF1 (C) and DAPI (D). Neural tube (NT), somite (Sm) and notochord (Nc) are indicated. Note that most of the nuclei in the notochord are EF1 positive, whereas the positive nuclei in the neural tube and the somite are scattered. (E,F) Sagittal (E) and parasagittal (F) sections of a stage 14 embryo through notochord and somites, respectively. Right is rostral. Note high EF1 expression level throughout the notochord (E) and that, in the somites, EF1-positive nuclei are present almost exclusively in the external epithelial layer (F). The bar indicates 50 μm for A to D and 100 μm for E and F.

opment proceeded and reached the stage of lens fiber formation, EF1 became clearly detectable in all lens cell nuclei (Fig. 9C,D). Expression of EF1 in the placodal components of the ectoderm was specific to the lens. Otic vesicles and nasal pit epithelia, for instance, remained EF1 negative (data not shown). Later in lens development, e.g., at stage 33 (day 8), EF1 became expressed more strongly in the lens fiber cells than in the epithelial cells (Fig. 9E,F). Thus, it appeared that EF1 was abundant in the nuclei of the cells that highly expressed the -crystallin gene.

Endodermal tissues and most other ectodermal tissues remained negative or very low in EF1 expression.

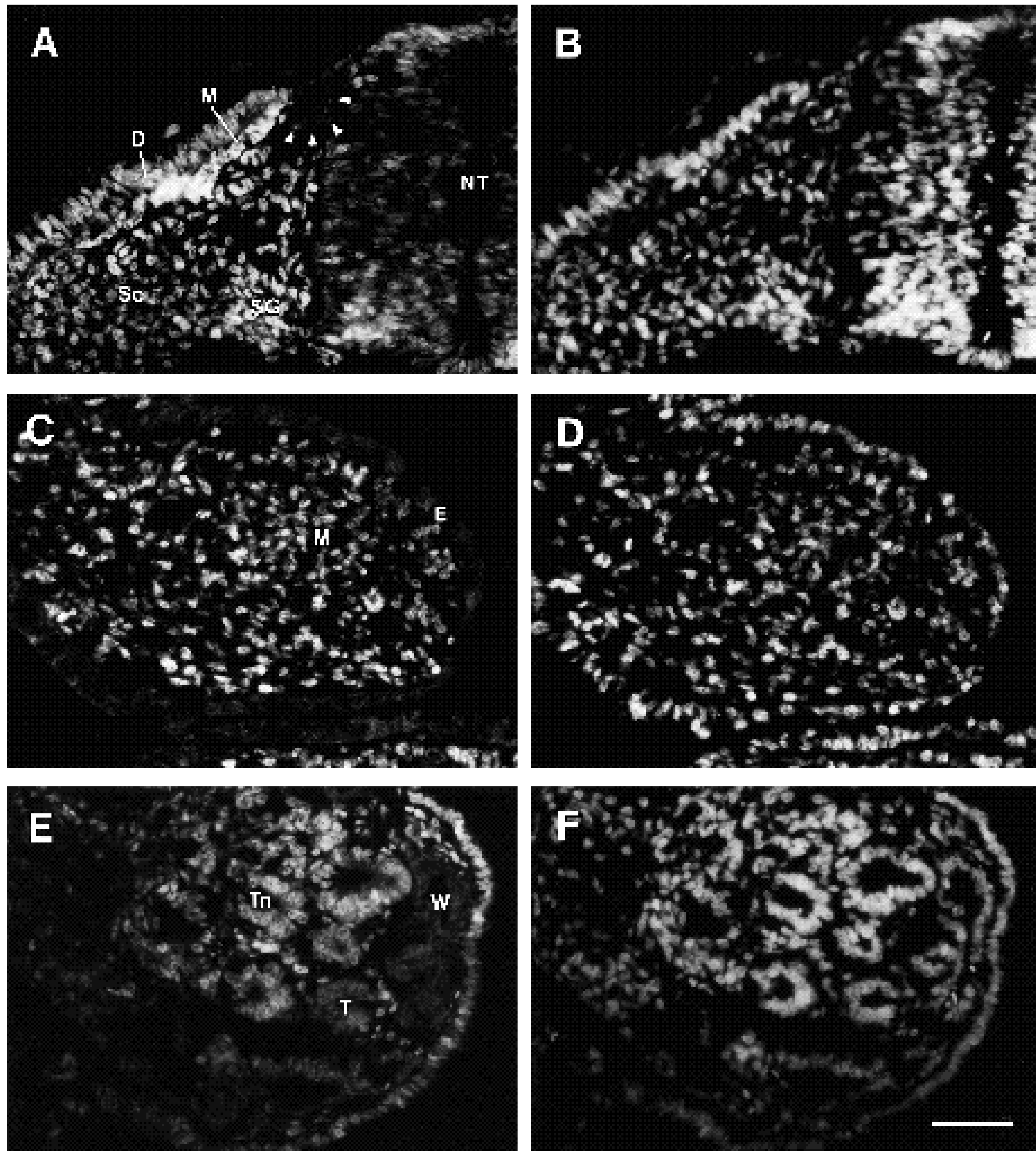

Fig. 8. Expression of EF1 in mesoderm and neural crest derivatives. (A,B) A part of a cross section through a stage 18 embryo stained for EF1 immunofluorescence (A) and for DAPI fluorescence (B). D, dermatome; M, myotome; NT, neural tube; Sc, sclerotome; SG, presumptive spinal ganglion. A cluster of migrating neural crest cells is indicated by the arrowheads. (C,D) A cross section through a visceral arch (hyoid arch) of a stage 18 embryo stained for EF1 fluorescence (C) and for DAPI fluorescence (D). E, epithelial component derived from ecto- and endoderms; M, mesenchymal component mostly derived from the neural crest. (E,F). Mesonephros of a stage 23 embryo stained for EF1 immunofluorescence (E) and for DAPI fluorescence (F). Note that nascent tubules (Tn) express high EF1, while Wolffian duct (W) expresses little EF1 at this stage and the tubules connected to the duct (T) express a very reduced level of EF1. The bar indicates 50 μm.

DISCUSSION

The relationship with previously described Nil-2-a

The homology search for nucleotide/amino acid sequences in data bases identified human Nil-2-a (Williams et al., 1991) as having sequences very similar to EF1. Nil-2-a was defined by a cDNA clone coding for a protein that binds to the −100 to −105 region of IL2 promoter, which has significant similarity to the block 6 of the HN fragment

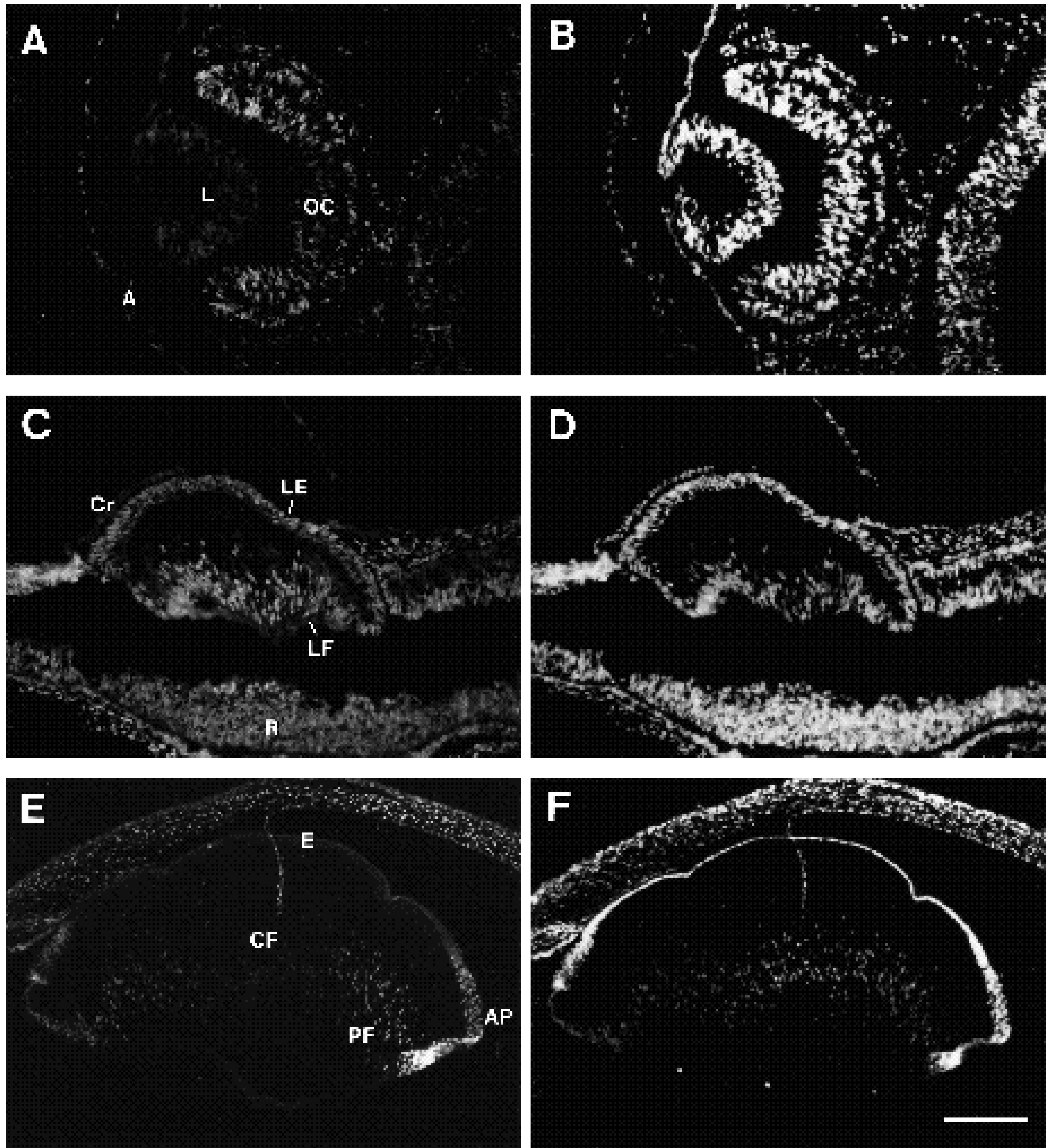

Fig. 9. EF1 expression in the developing lens. (A,C,E) EF1 immunofluorescence; (B,D,F) DAPI fluorescence. (A,B) Stage 18 embryo when lens vesicle is formed. A, amnion, L, lens vesicle and OC, optic cup. Note that EF1 level is very low in the lens vesicle and in the continuing ectoderm, compared to the optic cup. (C,D) Stage 23 embryonic lens when lens fibers had elongated. Both external lens epithelium (LE) and internal lens fiber (LF) contained a high level of EF1 in their nuclei, while the ectoderm remained very low in the EF1 expression. R, retina; Cr, cornea. (E,F) Stage 33 (8 day) lens. EF1 was positive in the annular pad (AP) and the peripheral fibers (PF) but negative in most of the epithelia area (E) and the cortical fiber cells (CF). The bar indicates 100 μm for A to D and 200 μm for E and F.

of the 1-crystallin enhancer. Alignment of EF1 and Nil-2-a amino acid sequence (Fig. 10) indicated that the latter matches very well to the C-terminal two-thirds of the EF1 sequence. The sequence similarity extends to the region upstream of the first methionine residue of the Nil-2-a sequence, which is valine in the EF1 sequence. This suggested that Nil-2-a-encoding cDNA is a partial cDNA of human EF1 rather than an alternative form of EF1. This has been supported by our comparative analysis of chicken and mouse EF1 genes (R. Sekido, T. Takagi and Y. Higashi, unpublished result). Organization of the EF1 gene is very well conserved between the chicken and the

```
DEF1LS    MADGPRCKRR KQANPRRNNV TNYNNVIEAN SDSDDEDKLH IVEEESITDA ADCDASVPED DLPTDHTVLP ENSEREGSTN SCWEDEGKET KEILGPEAQS
HUMNIL2A  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

DEF1LS    DEVGCTVKED ECDSDAENEQ NHDPNVEEFL QQEDTAVIYP EAPEEDQRQG TPEASGQDEN GTPDAFSQLL TCPYCDRGYK RFTSLKEHIK YRHEKNEDNF
HUMNIL2A  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

DEF1LS    SCSLCSYTFA YRTQLDRHMT SHKSGRDQRH VTQSSGNRKF KCTECGKAFK YKHHLKEHLR IHSGEKPYEC PNCKKRFSHS GSYSSHISSK KCIGLMPVKG
HUMNIL2A  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

DEF1LS    RARSGLKTSQ CSSPSLSASP GSPARPQIRQ KIENKPLQEQ LPVNQIKTEP VDYEFKPIVV ASGINCSTPL QNGVFSGGSP LQATSSPQGV VQAVVLPTVG
HUMNIL2A  .......... .......... .......... .......... .......... .....KPIVV ASGINCSTPL QNGVFTGGGP LQATSSPQGM VQAVVLPTVG
                                                                                                      *

DEF1LS    LVSPISINLS DIQNVLKVAV DGNVIRQVLE NNHANLASKE QETISNASIQ QAGHSLISAI SLPLVDQDGT TKIIINYSLE QPSQLQVVPQ NLKKEHSVPT
HUMNIL2A  LVSPISINLS DIQNVLKVAV DGNVIRQVLE NNQANLASKE QETINASPIQ QGGHSVISAI SLPLVDQDGT TKIIINYSLE QPSQLQVVPQ NLKKENPVAT

DEF1LS    NSCKNEKLPE DLTVKSEKDK NFEGETNDST CLLCDDCPGD LNALQELKHY ETKNPPQLPQ SSGTEAEKPS SPAPSETGEN NLSPGQPPLK NLLSLLKAYY
HUMNIL2A  NSCKSEKLPE DLTVKSEKDK SFEGGVNDST CLLCDDCPGD INALPELKHY DLKQPTQPPP LPAAEAEKPE SSVSSATGDG NLSPSQPPLK NLLSLLKAYY

DEF1LS    ALNAQPSAEE LSKIADSVNL PLDVVKKWFE KMQAGQISVQ SSGPSSPEQV KISSPTDNDD QAATTNESEP QNSTNNSQNP ANTSKSQTSS GGSTQNGSRS
HUMNIL2A  ALNAQPSAQE LSKIADSVNL PLDVVKKWFE KMQAGQISVQ SSEPSSPEPG KVNIPAKNND QPQSANANEP QHSTVNLQSP SKMTNSPVLP VGSTTNGSRS

DEF1LS    STPSPSPLNL SSSRNSQGYT YTAEGVQEEP QMEPLDLSLP KQHGELLERS TITSVYQNSV YSVQEEPLNL TCAKKEPCKD NSITDSDPIV NVIPPSANPI
HUMNIL2A  STPSPSPLNL SSSRNTQGYL YTAEGAQEEP QVEPLDLSLP KQQGELLERS TITSVYQNSV YSVQEEPLNL SCAKKEPCKD SCVTDSEPVV NVIPPSANPI

DEF1LS    NIAIPTVTAQ LPTIVAIADQ NSVPCLRALA ANKQTILIPQ VAYTYSTTVS PAVQETPPKQ TQANGSQDER QDTSSEGVSN VEDQNDSDST PPKKKMRKTE
HUMNIL2A  NIAIPTVTAQ LPTIVAIADQ NSVPCLRALA ANKQTILIPQ VAYTYSTTVS PAVQEPPLKV IQPNGNQDER QDTSSEGVSN VEDQNDSDST PPKKKMRKTE

DEF1LS    NGMYACDLCD KIFQKSSSLL RHKYEHTGKR PHECGICKKA FKHKHHLIEH MRLHSGEKPY QCDKCGKRFS HSGSYSQHMN HRYSYCKREA EERDSTEQEE
HUMNIL2A  NGMYACDLCD KIFQKSSSLL RHKYEHTGKR PHECGICKKA FKHKHHLIEH MRLHSGEKPY QCDKCGKRFS HSGSYSQHMN HRYSYCKREA EERDSTEQEE

DEF1LS    VGQEVLSSEH AGARASPSQI DSDERESLTR EEEEDSEKEE EEEEEKDVEG LQEEKECRKL QDVEEEEE-- --VEEEEEEE EG------KT EGNKNDDVVN
HUMNIL2A  AGPEILSNEH VGARASPSQG DSDERESLTR EEDEDSEKEE EEE-DKEMEE LQEEKECEKP QGDEEEEEEE EEVEEEEVEE AENEGEEAKT EGLMKDDRAE

DEF1LS    RASNAEPEVI QSNG-QVSEE KTNKA
HUMNIL2A  SQASSLGQKV GESSEQVSEE KTNEA
```

Fig. 10. Alignment of amino acid sequences of EF1 and Nil-2-a. Putative DNA-binding domains, zinc finger clusters and the homeobox are boxed, and identical amino acid residues on EF1 and Nil-2-a sequences are stippled. The first methionine residue of Nil-2-a sequence is marked by an asterisk.

mouse, and hence likely in the human: the most N-proximal methionine codon of Nil-2-a is probably located in the middle of a large exon.

δEF1 as a ZFH family DNA-binding protein

An intriguing feature of EF1 protein as a DNA-binding protein is that it has a multiplicity of potential DNA-binding domains as combinations of zinc-fingers and a homeodomain in a way analogous to *Drosophila zfh* gene products (Fortini et al., 1991) and an -fetoprotein enhancer binding factor (Morinaga et al., 1991). A striking similarity is found between EF1 and ZFH-1 in the amino acid sequence of the DNA-binding domains and in their organization, which suggests that these two proteins have properties in common as DNA-binding proteins. It is interesting to note that EF1 is expressed in most of the mesodermal lineages and, in *Drosophila*, *zfh-1* expression is also primarily mesoderm-specific (Lai et al., 1991).

How the separate DNA-binding domains interact with genomic DNA sequences poses an interesting question concerning not only EF1 itself but also proteins having multiple DNA-binding domains in general. The JF12 portion of EF1 carrying only the C fingers had the same binding specificity as the native EF1 when the DC5 fragment of the enhancer core was used as the probe of binding, indicating that C fingers alone account for the binding to the enhancer core region. In the case of PRDII-BF1, two zinc finger clusters were separately located close to the N and C termini, and these clusters had the same binding specificity (Fan and Maniatis, 1990). As shown in Fig. 5C, N fingers 3-4 had significant sequence similarities to C fingers 6-7, which may suggest that N fingers bind to nucleotide sequences having some similarity to C-finger-binding sites. It has been demonstrated that joining of zinc fingers of different binding specificity results in combined plural specificities (Keller and Maniatis, 1992), which may suggest that EF1 has multiple sets of sequence specificities. It is not certain whether the homeodomain of EF1 binds DNA in a sequence-specific manner, but the homeodomain of ZFH-1 does have such an activity (Fortini et al., 1991). Alternatively, the homeodomain of EF1 may participate in protein dimer formation, as is suggested for POU homeodomains (Treacy et al., 1992) to which EF1 homeodomain has an appreciable similarity.

δEF1 as a transcriptional regulator

An intriguing feature of EF1 expression in the lens is that it begins after -crystallin expression is turned on and that it appears to increase along with -crystallin synthesis.

Cotransfection experiments using a EF1 expression vector and HN enhancer fragment-carrying reporters demonstrated that EF1 acts as a repressor of the -crystallin enhancer. By detailed mutational analysis of the HN fragment, we have shown that EF1 competes with an activator in lens cells through overlapping binding sites (Kamachi and Kondoh, 1993). In embryonic lens cells

where a high activator level is achieved, the repressor action of EF1 is probably counteracted. However, in a variety of non-lens cells where the activator is not in effect, EF1 likely represses -crystallin expression so as to make lens specificity very stringent. In fact, in a number of primary cultures and cell lines tested, EF1 binding of HN fragment always resulted in repression of reporter genes (Kamachi and Kondoh, 1993).

How, then, can we reconcile the persistent and even augmented expression of EF1 in developing lens with increase of -crystallin expression which seem to take place in parallel? It should be noted that there are a remarkable increase of crystallin synthesis and a substantial loss of non-crystallin expression, which occur simultaneously in maturating lens cells. Perhaps EF1 acts to repress non-crystallin genes as long as they possess a EF1-binding site but lack proper activating factors in lens cells.

However, considering the dynamic expression of EF1 in early embryogenesis, it is tempting to speculate a positive function for EF1 in a non-lens context. The domains rich in prolines in the middle and rich in glutamic acids near the C terminus resemble identified activation domains of various activators (Mitchell and Tjian, 1989; Seipel et al., 1992; and references therein). In a variety of cell-type-specific enhancers of non-crystallin origin, EF1-binding sites have been identified. Thus, putative dual functions and heterogeneous DNA-binding domains associated with single EF1 protein raise the possibility that this protein factor acts in a variety of ways depending on the cell type and the gene to be regulated.

Dynamic expression of δEF1 during embryogenesis

In the embryo, the primary site of EF1 expression is the mesoderm. Expression of EF1 begins when gastrulation is over and organogenesis has just begun. Most of the mesoderm derivatives express EF1: the notochord first, and then somites, nephrotomes and lateral plates. It has been demonstrated that at stages 12 to 14 when all notochord nuclei contain a significant level of EF1, the notochord is active in organizing dorsoventral polarity of the neural tube (Yamada et al., 1991). In the somite, EF1 is turned on first in the external epithelial part leaving the cell mass in the core negative, and when three somitic compartments are differentiated, the myotome becomes and remains the site where EF1 level is the highest. In the nephrotome, EF1 is most conspicuous where tubular assembly is very actively going on. This dynamic change of expression strongly suggests that EF1 has a significant role in the early histogenesis of mesodermal tissues.

In conjunction with mesodermal expression of EF1, it is interesting to note that mesodermal expression of *zfh-1* in *Drosophila* is dependent upon the activity of the *snail* and *twist* genes (Lai et al., 1991). Mouse homologues of the latter two, *Sna* (Smith et al., 1992) and *M-twist* (Wolf et al., 1991) have been cloned and their expression in the mid- to postgastrulation mouse embryos examined. Both *Sna* and *M-twist* are expressed in wide variety of mesodermal tissues and mesenchymal components of neural crest derivatives. Extrapolating the mouse data to chicken embryos, *snail* and *twist* homologues are expressed earlier than EF1. It is possible that a hierarchy of mesodermal regulatory genes analogous to *Drosophila* exists in vertebrates.

Thus, expression pattern of EF1 poses interesting problems of how this transcription factor gene is regulated in the spatiotemporal order of embryonic development, and how EF1 exerts its effects on diverse cell types. Since genomic clones of the chicken and mouse EF1 genes are already available (R. Sekido, M. Okanami, T. Takagi and Y. Higashi, unpublished), the problems can be solved if advantage is taken of transgenic and gene targeting technologies.

We thank S. Subramani for providing firefly luciferase gene, M. Okanami for assistance in cDNA analysis, and H. Nakamura, Y. Wakamatsu and Y. Higashi for discussion. This work was supported by grants from the Ministry of Education, Science and Culture of Japan to J. F., Y. K. and H. K., and by Special Coordinating Funds for Promoting Science and Technology from the Science and Technology Agency of Japan to H. K. In addition, J. F. and Y. K. are recipients of Fellowships from the Japan Society for the Promotion of Science for Japanese Junior Scientists.

REFERENCES

Chen, C. and Okayama, H. (1987). High efficiency transformation of mammalian cells by plasmid DNA. *Mol. Cell. Biol.* **7,** 2745-2752.

deWet, J. R., Wood, K. V., Deluca, M., Helinski, D. R. and Subramani, S. (1987). Firefly luciferase gene: structure and expression in mammalian cells. *Mol. Cell. Biol.* **7,** 725-737.

Fan, C-M. and Maniatis, T. (1990). A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. *Genes Dev.* **4,** 29-42.

Fortini, M. E., Lai, Z. and Rubin, G. M. (1991). The Drosophila zfh-1 and zfh-2 genes encode novel proteins containing both zinc-finger and homeodomain motifs. *Mech. Dev.* **34,** 113-122.

Funahashi, J., Kamachi, Y., Goto, K. and Kondoh, H. (1991). Identification of nuclear factor EF1 and its binding site essential for lens-specific activity of the 1-crystallin enhancer. *Nucl. Acids Res.* **19,** 3543-3547.

Goto, K., Okada, T. S. and Kondoh, H. (1990). Functional cooperation of lens-specific and nonspecific elements in the 1-crystallin enhancer. *Mol. Cell. Biol.* **10,** 958-964.

Grainger, R. M. (1992). Embryonic lens induction shedding light on vertebrate tissue determination. *Trends Genet.* **8**, 349-355.

Hamburger, V. and Hamilton, H. L. (1951). A series of normal stages in the development of the chick embryo. *J. Morphol.* **88,** 49-92.

Hayashi, S., Goto, K., Okada, T. S. and Kondoh, H. (1987). Lens-specific enhancer in the third intron regulates expression of the chicken 1-crystallin gene. *Genes Dev.* **1,** 818-828.

He, X., Treacy, M. N., Simmons, D. M., Ingraham, H. A., Swanson, L. W. and Rosenfeld, M. G. (1989). Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* **340,** 35-42.

Kamachi, Y. and Kondoh, H. (1993). Overlapping positive and negative regulatory elements determine lens-specific activity of the 1-crystallin enhancer. *Mol. Cell. Biol.,* in press.

Kato, K., Kanamori, A. and Kondoh, H. (1990). Rapid and transient decrease of N-myc expression in retinoic acid-induced differentiation of OTF9 teratocarcinoma stem cells. *Mol. Cell. Biol.* **10,** 486-491.

Keller, A. D. and Maniatis, T. (1992). Only two of the five zinc fingers of the eukaryotic transcriptional repressor PRDI-BF1 are required for sequence-specific DNA binding. *Mol. Cell. Biol.* **12**, 1940-1949.

Lai, Z., Fortini, M. E. and Rubin, G. M. (1991). The embryonic expression patterns of zfh-1 and zfh-2, two Drosophila genes encoding novel zinc-finger homeodomain proteins. *Mech. Dev.* **34**, 123-134.

Mitchell, P. J. and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

Morinaga, T., Yasuda, H., Hashimoto, T., Higashio, K. and Tamaoki, T.

(1991). A human -fetoprotein enhancer-binding protein, ATBF1, contains four homeoproteins and seventeen zinc fingers. *Mol. Cell. Biol.* **11**, 6041-6049.

Seed, B. (1987). An LFA-3 cDNA encodes a phospholipid-linked membrane protein homologous to its receptor CD2. *Nature* **329**, 840-842.

Seipel, K., Georgiev, O. and Schaffner, W. (1992). Different activation domains stimulate transcription from remote ('enhancer') and proximal ('promoter') positions. *EMBO J.* **11**, 4961-4968.

Singh, H., LeBowitz, J. H., Baldwin, Jr. A. S. and Sharp, P. A. (1988). Molecular cloning of an enhancer binding protein, isolation by screening of an expression library with a recognition site DNA. *Cell* **52**, 415-423.

Smith, D. B. and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67**, 31-40.

Smith, D. E., del Amo, F. F. and Gridley, T. (1992). Isolation of *Sna*, a mouse gene homologous to the *Drosophila* genes *snail* and *escargot*: its expression pattern suggests multiple roles during postimplantation development. *Development* **116**, 1033-1039.

Treacy, M. N., Neilson, L. I., Turner, E. E., He, X. and Rosenfeld, M. G. (1992). Twin of I-POU: a two amino acid difference in the I-POU homeodomain distinguishes an activator from an inhibitor of transcription. *Cell* **68**, 491-505.

Vinson, C. R., LaMarco, K. L., Johnson, P. F., Landschulz, W. H. and McKnight, S. L. (1988). In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. *Genes Dev.* **2**, 801-806.

Williams, T. M., Moolten, D., Burlein, J., Romano, J., Bhaerman, R., Godillot, A., Mellon, M., Rauscher, F. J. and Kant, J. A. (1991). Identification of a zinc finger protein that inhibits IL-2 gene expression. *Science* **254**, 1791-1794.

Wolf, C., Thisse, C., Stoezel, C., Thisse, B., Gerlinger, P. and Perrin-Schmitt, F. (1991). The *M-twist* gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus X-twi* and the *Drosophila twist* genes. *Dev. Biol.* **143**, 363-373.

Yamada, T., Placzek, M., Tanaka, H., Dodd, J. and Jessel, T. M. (1991). Control of cell pattern in the developing nervous system: polarizing activity of the floor plate and the notochord. *Cell* **64,** 635-647.

(Accepted 23 June 1993)