ELSEVIER

Mechanisms of Development 63 (1997) 99–108

Transcription regulation and alternative splicing of an early zygotic gene encoding two structurally distinct zinc finger proteins in *Xenopus laevis*

Eric Bellefroid[a,1], Catherine Bourguignon[a], Tewis Bouwmeester[a,2], Oliver Rausch[a,3], Bruce Blumberg[b], Tomas Pieler[a,*]

[a]*Institut für Biochemie und Molekulare Zellbiologie, Humboldtallee 23, 37073 Göttingen, Germany*
[b]*The Salk Institute, PO Box 85800, San Diego, CA 92816, USA*

Received 6 January 1997; revised version received 27 February 1997; accepted 28 February 1997

Abstract

We describe the structural organization of a gene, termed XFDL 141/156, that is transiently activated during early *Xenopus* development. XFDL 141/156 is first transcribed at the midblastula transition (MBT) and during early gastrulation events. A roughly 200 nucleotide fragment immediately 5′ to the transcription start site is sufficient for transient, early zygotic activation of gene expression. The primary transcript is subject to alternative splicing. Corresponding cDNAs encode two structurally related but completely distinct C2H2-type zinc finger proteins of unknown biological function. © 1997 Elsevier Science Ireland Ltd.

Keywords: Zinc finger protein; *Xenopus laevis*; XFDL 141/156

1. Introduction

The C2H2-type zinc finger motif defines an extremely large superfamily of nucleic acid binding proteins with probably more than 1000 structurally distinct members encoded in a typical vertebrate genome. These proteins fall into two major groups which can be distinguished on the basis of several hallmark characteristics. Group 1 zinc finger proteins (ZFPs) are encoded in large genomic clusters, are found to be only moderately conserved in evolution, and their expression is largely ubiquitous. The function of these proteins is only poorly, if at all understood. Group 2 zinc finger proteins have been highly conserved in evolution, follow characteristic expression patterns during embryogenesis and in adult tissues, and several of these have been demonstrated to serve key regulatory functions during embryonic development and cell differentiation (reviewed in Pieler and Bellefroid, 1994).

ZFP gene activity can be regulated by alternative splicing; several vertebrate and invertebrate ZFP encoding transcription units have been demonstrated to encode multiple, structurally distinct proteins. Isoforms of the *Drosophila* zinc finger transcription factor CF2 differ in their zinc finger composition and bind to different DNA target sequences (Hsu et al., 1992). Similarly, alternative splicing of the WT1 gene, which is critically involved in mammalian kidney development, gives rise to different ZFP isoforms with distinct DNA binding specificities. Interestingly, these WT1 protein isoforms, which differ in their zinc finger clusters, localize to different subnuclear fractions, i.e. either to the splicing or to the transcription regulatory machinery (Bickmore et al., 1992; Larsson et al., 1995).

Transcription regulation of ZFP encoding genes has not been studied extensively. For the ubiquitous zinc finger transcription factor IIIA (TFIIIA) different levels of transcription in *Xenopus* oocytes are achieved via the activities of a negative and of three positively acting regulatory elements, all located within roughly 300 nucleotides 5′ to the transcription start site (Scotto et al., 1989). In early *Xenopus* embryogenesis, no zygotic transcription is detectable before

* Corresponding author. Tel.: +49 551 395683; fax: +49 551 395960; e-mail: tpieler@gwdg.de

[1] Present address: Wellcome Trust CRC Institute, Cambridge University, Cambridge, UK.

[2] Present address: School of Medicine, University of California, Los Angeles, CA, USA.

[3] Present address: Institute of Cancer Research, Chester Beatty Laboratories, London, UK.

0925-4773/97/$17.00 © 1997 Elsevier Science Ireland Ltd. All rights reserved
PII S0925-4773(97)00034-8

the midblastula transition (MBT), i.e. after 12 cleavage cycles (Newport and Kirschner, 1982). Genes activated at MBT include a selected number of protein-encoding transcription units, such as GS17 (Krieg and Melton, 1985) and EF-1α (Krieg et al., 1989). Regulatory elements responsible for zygotic gene activation have been identified by means of microinjection of cloned genes into fertilized eggs. In the case of GS17, sequences within a roughly 700 bp fragment upstream of the corresponding promoter have been demonstrated to contain all the essential information for MBT expression (Krieg and Melton, 1987), and they have been found to contain a CT-rich, 14 bp element that binds to the maternal OZ-1 transcription factor activity, and which is sufficient to drive MBT expression of a fused reporter gene (Ovsenek et al., 1992). For the second example of a gene that is strongly activated at MBT, EF-1α, identification of a G/C rich enhancer element that is located approximately 4.4 kb upstream of the transcription start site indicates that the zinc finger transcription factor SP1 may play a major role in the regulated expression of the EF-1α gene (Johnson and Krieg, 1995).

Here, we describe a novel ZFP encoding transcription unit in *Xenopus* that gives rise to two completely different, though structurally related ZFPs by means of alternative splicing. This gene is strongly and transiently activated at MBT and during early gastrula stages of *Xenopus* embryogenesis. A roughly 200 bp fragment that contains neither an OZ-1 nor an SP1 binding site is sufficient for transient transcription activation at MBT.

2. Results

2.1. Identification of two C2H2-type zinc finger protein encoding sequences overrepresented in a dorsal blastopore lip cDNA library

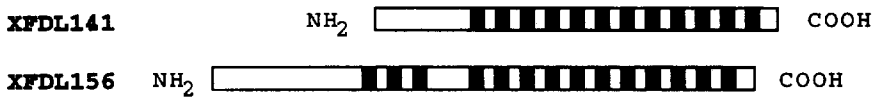
With the initial intention of identifying *Xenopus* ZFPs which would be specifically expressed in the Spemann organizer region, a cDNA library derived from the dorsal blastopore lip (Blumberg et al., 1991) was screened for C2H2-type ZFPs as described previously (Köster et al., 1988), resulting in the isolation of 50 independent, positive clones. Since numerous ZFP encoding genes have been described to be expressed in a non-localized fashion during gastrulation (Bouwmeester, 1993), this initial pool of 50 clones was screened by comparative Southern blot hybridization for sequences which would be the most abundant. One cDNA, designated XFDL (*Xenopus* Finger Dorsal Lip) 141, was represented by four independent clones, a second one, designated XFDL 156, was represented by seven independent clones. A different *Xenopus* ZFP, termed XFG 20, that is abundantly expressed during gastrulation in a non-localized fashion (Schäfer et al., 1994), was represented only once. Therefore, we concentrated on XFDL 141 and XFDL 156 in our further experimentation.

The 1.9 kb XFDL 141 cDNA sequence contains a 411 amino acid open reading frame with multiple 5′-in frame stop codons and a poly(A) tail. The predicted protein sequence defines a 47 kDa protein with a cluster of 12 C2H2 zinc finger units in tandem repeat (Fig. 1). The 2.1 kb XFDL 156 sequence is a composite cDNA clone; the cDNA fragment isolated from the blastopore lip library overlaps with two different, incomplete 3′-cDNA fragments termed HC 14 and XFG 32-1, that we had described previously (Köster et al., 1988; Nietfeld et al., 1989). The composite XFDL 156 sequence contains a 573 amino acids long open reading frame that encodes for a 66 kDa protein with two clusters of three and 11 C2H2 zinc finger units in tandem repeat, separated by an unrelated spacer element (Fig. 1).

Primary sequence comparison of the two proteins reveals extensive sequence similarity, but only within the large zinc finger clusters; sequence identity within the carboxy-terminal portion of XFDL 141 and XFDL 156 including the entire zinc finger cluster of the former protein is 78% (Fig. 1). On the nucleotide level, there exist two blocks of 91% and 94% sequence identity in the region corresponding to zinc fingers 1–5 and 7–11 in XFDL 141. The 62 amino acids long amino-terminal region of the XFDL 141 protein is only 37% identical to the corresponding sequence element in XFDL 156. Finally, the extreme 5′-end of the XFDL 141 cDNA is found to contain a 76 nucleotides long sequence element that is 93% identical to the 5′-portion of the untranslated region in XFDL 156 (see below). In summary, XFDL 141 and XFDL 156 cDNAs encode two closely related C2H2-type zinc finger proteins which are both overrepresented in a blastopore lip specific cDNA library.

2.2. Temporal and spatial expression characteristics of XFDL 141 and XFDL 156

XFDL 141 and XFDL 156 gene expression was examined by means of Northern blot analysis with total embryonic RNA preparations (Fig. 2A). During embryogenesis, both genes are similarly expressed. Using the two different cDNAs as a probe, very low levels of maternal transcripts can be detected in the egg and in all embryonal stages prior to the mid-blastula-transition (MBT). Upon zygotic activation of gene transcription a strong, but transient increase in the XFDL 141/156 RNA concentrations can be detected (stages 9 and 10). During gastrulation, RNA levels decline to a low level, which is similar to the level of the maternal transcripts. Due to the extreme degree of sequence identity within certain elements of the zinc finger encoding portion of the two cDNAs (see above), crosshybridization with the two probes appears to be likely. Indeed, if a cDNA probe is utilized that corresponds to the less conserved N-terminal portion of XFDL 141, no maternal and late embryonic mRNA can be detected (Fig. 2A). Therefore, although the quantity of both transcripts is strongly increased at MBT,




```
XFDL141   1        ▼                MSKMKKISKDRSE..NIQASLSQADSDNELCRLPNKKDLPQMSQESCADG
XFDL156   1 MMHLIRNWPEKEKQNLNGSE--R----IEVG--RE--WGN---TQTHK-SY-T--QNYTSP-P--YSYN-

XFDL141  49 KSPALRKRKSQNTK.....
XFDL156  71 -LF--QRE-LH--NVQVIRMPNIAKQPKFNREIRSDSPKRKKAKSSEKSHVANISHQGTKINKFSKAQSA

XFDL141     .....
XFDL156 141 LKSFICPICKRRFCNARNLLTHQRIHTGRRFVCSKCGKCFSNKKILIAHRWLHTGRKPFTCTKCNKGFLW

XFDL141  63 ..........DIRLSFDCRECGQSFRKKSKLKTHFLCHTGEKPC
XFDL156 211 NRDLQKHEKSHKEMRNKSLGLVHRHDKLREVSTDPCKREK----T---K--KR----------------F

XFDL141  97 VCVHCGKRFADKYKLSLHLRIHTGENLSVCPHCGKSYRDKTKLTVHLRIHTGETPFVCPECGKGFRDASF
XFDL156 281 -------G-R-N-------------------D-----T--N--I--M------K.-M-S------S-FYN

XFDL141 167 LKSHLSIHTGEKPFVCTECGKSFRKKCELKSHLQVHTGVKQFVCTECGKSFRDRSELNIHINVHTGKTFP
XFDL156 350 -----Q----------P------LS-Y-F----LI---KYP-I-S------M-KRY-K--S---S-N---

XFDL141 237 CTECGKSFAAKKNLKRHQMIHTGEKPHECTECGKQFLEKSKLKRHHLSHTGVKPFSCFECGEQFTWKHLL
XFDL156 420 ----------Q------K--------K---------Q-N--D----------------------------

XFDL141 307 QNHQLSHTGERPFVCSECGKSFKTKASLALHCHIHTGERPFVCPECGKGFRDPTSLNSHVRRHTGEKPFA
XFDL156 490 -Y-----------------GY----------------K--------NSY-SNGA-RY-LQT-

XFDL141 377 CAECGKTYTDSSTLKNHLKLHVKGKYFAHSECGES*          411aa
XFDL156 553 .....I---------H---DKSETNE*   573aa
```

Fig. 1. Primary sequence comparison of the XFDL 141 and XFDL 156 zinc finger proteins. (A) Structural architecture of XFDL 141 and XFDL 156 proteins. Bars indicate the position of C2H2 zinc finger units. (B) Amino acid sequence comparison of XFDL 141 and XFDL 156. Dots indicate amino acids present in one but not the other sequence; hyphens reflect positions of sequence identity. C2H2 zinc finger module encoding sequence elements are indicated by lines.

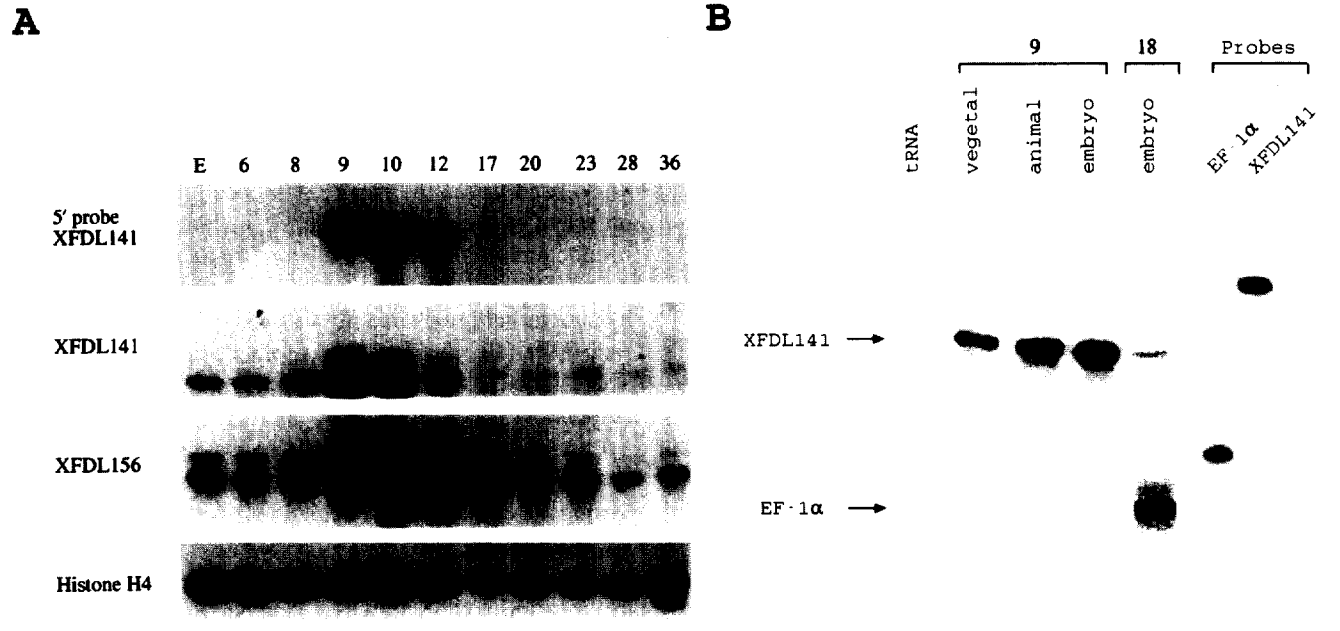


Fig. 2. XFDL 141/156 expression characteristics during *Xenopus* embryogenesis. (A) Northern blot analysis. Four egg/embryo equivalents of total RNA preparations were loaded per lane; embryonic stages are indicated according to Nieuwkoop and Faber, 1967. Probing was with either a 0.4 kb cDNA fragment derived from the 5′-region of the XFDL 141 cDNA or the entire XFDL 141 and XFDL 156 cDNAs isolated in this study. The histone H4 cDNA insert was used as a control. (B) RNase protection analysis of XFDL 141 mRNA with total RNA preparations (10 μg each) from vegetal and animal halves of stage 9 embryos. Control RNA preparations include tRNA, as well as total stage 9 and 18 RNA preparations (as indicated).

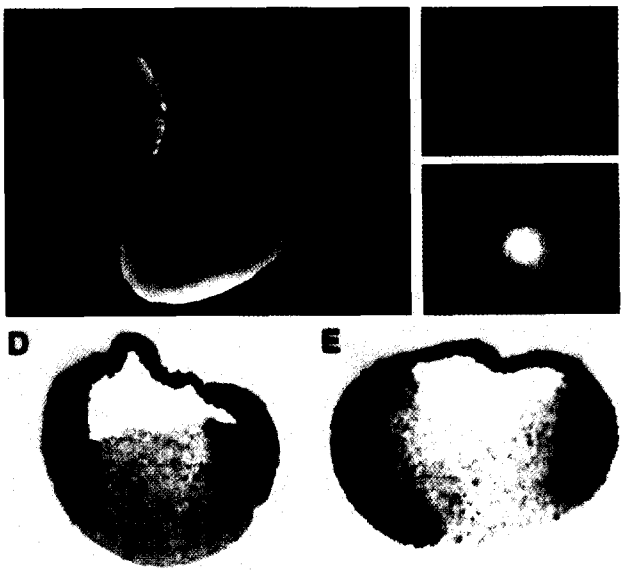


Fig. 3. In situ hybridization analysis of XFDL 141/156 expression during early *Xenopus* embryogenesis. *Xenopus* embryos (staged according to Nieuwkoop and Faber, 1967) were stained for XFDL 141/156 gene transcription making use of the XFDL 141 cDNA as a probe. (A) Stage 9.5 blastula embryos viewed from the vegetal half, the animal half and from the side (clockwise). Note that the staining appears to be much stronger in the animal half than in the vegetal half. (B) Stage 12 late gastrula embryo; anterior is to the right hand side. (C) Stage 14 neurula embryo; anterior is to the right. Note the speckled pattern of cells on both sides along the dorsal midline. (D) Transverse section through a stage 10 + early gastrula embryo stained for XFDL 141 expression post-sectioning. (E) Transverse section through a midgastrula stage 11.5 embryo stained as described in (D).

low level maternal and late embryonic expression may be specific to XFDL 156. Finally, results obtained with the XFDL 141-derived 5′-probe also reveal a slight increase of RNA mobility correlating with the progress of development, raising the possibility that the rapid disappearance of the XFDL 141 mRNA could be caused by an active degradation mechanism that is initiated by a shortening of the poly(A) tail.

Since XFDL 141 and 156 were overrepresented in the dorsal blastopore lip specific cDNA library, one expectation was that corresponding transcripts might be localized to the Spemann organizer. However, the experiment illustrated in Fig. 3 shows that this is not the case. Using a XFDL 141-derived probe in whole mount in situ hybridization experiments, expression is first detected in blastula (stage 9.5) embryos. Staining is found to be most intense within the animal half of the embryo and can only be detected in a few isolated cells of the vegetal half (Fig. 3A). At stage 12.5, XFDL141 is only weakly expressed in a group of cells which define a thin line along what appears to be the anterior, dorsal-ventral boundary (Fig. 3B). At neurula stage (stage 14), XFDL 141 remains weakly expressed in a group of cells localized to the anteriormost portion of the folding neural tube (Fig. 3C). In order to ensure equal access of the probe to all regions of the embryo (Lemaire and Gurdon, 1994), we have also carried out in situ hybridizations on 12 μm sections. XFDL 141 is found to be very strongly expressed within all cell layers of the animal half of late blastula and early gastrula stage embryos. No, or only very weak expression is detected in endodermal cells (Fig. 3D,E). RNase protection analyses with RNA preparation from the animal half and from the vegetal half of stage 9 embryos reveal that the signal originating from RNA preparations of the vegetal half is much lower than with corresponding samples representing the animal half (Fig. 2B). In RNA preparations from embryos which were either ventralized by UV treatment or dorsalized by LiCl treatment, no effect on XFDL 141 or XFDL 156 expression could be detected (data not shown). In conclusion, XFDL 141 is strongly activated at the mid-blastula transition and transcripts are only present for a brief period, primarily during early gastrula stages.

2.3. Genomic organization of the XFDL 141 and XFDL 156 encoding transcription units

Preliminary genomic Southern blot analysis had revealed that probes specific for XFDL 141 and XFDL 156 generate an overlapping pattern of fragments, indicating that the corresponding transcription units colocalize (data not shown). The successful identification and structural characterization of a genomic DNA fragment that contains the sequence information for both of these proteins revealed that this is indeed the case (Fig. 4A). Interestingly, this analysis also demonstrated that the XFDL 141 and 156 cDNAs correspond to mRNAs which are generated by alternative splicing. Both mRNAs share the same 5′-exon, which is part of the non-translated region (as pointed out above). The second and third exons define the remainder of the XFDL 156 mRNA; the fourth exon contains the entire structural information for XFDL 141. Exon/intron boundaries were mapped by DNA sequence analysis. A comparison of cDNA and genomic primary sequence information reveals an extent of diversity (98% identity) which is typical for the pseudotetraploid species *Xenopus laevis* in alignments of pseudoallelic DNA sequence information.

The transcription start site was mapped by RNase protection (Fig. 4B). This analysis reveals that the transcription start is only eight nucleotides upstream of the position corresponding to the 5′-end of the XFDL 156 cDNA. A TATAAAA element is found to be present at position −48 to −41. Interestingly, a DNA sequence element that is 88% identical to the first 252 nucleotides upstream of the XFDL 141/156 transcription unit is present in close proximity to the 3′-end of the gene (Fig. 4A,B). This observation raises the possibility that this DNA element is part of a second, closely spaced and structurally related transcription unit; however, the genomic fragment characterized is devoid of further zinc finger protein encoding elements that might be under the control of this putative, second promoter (see below).

In summary, structural analysis of a genomic DNA frag-

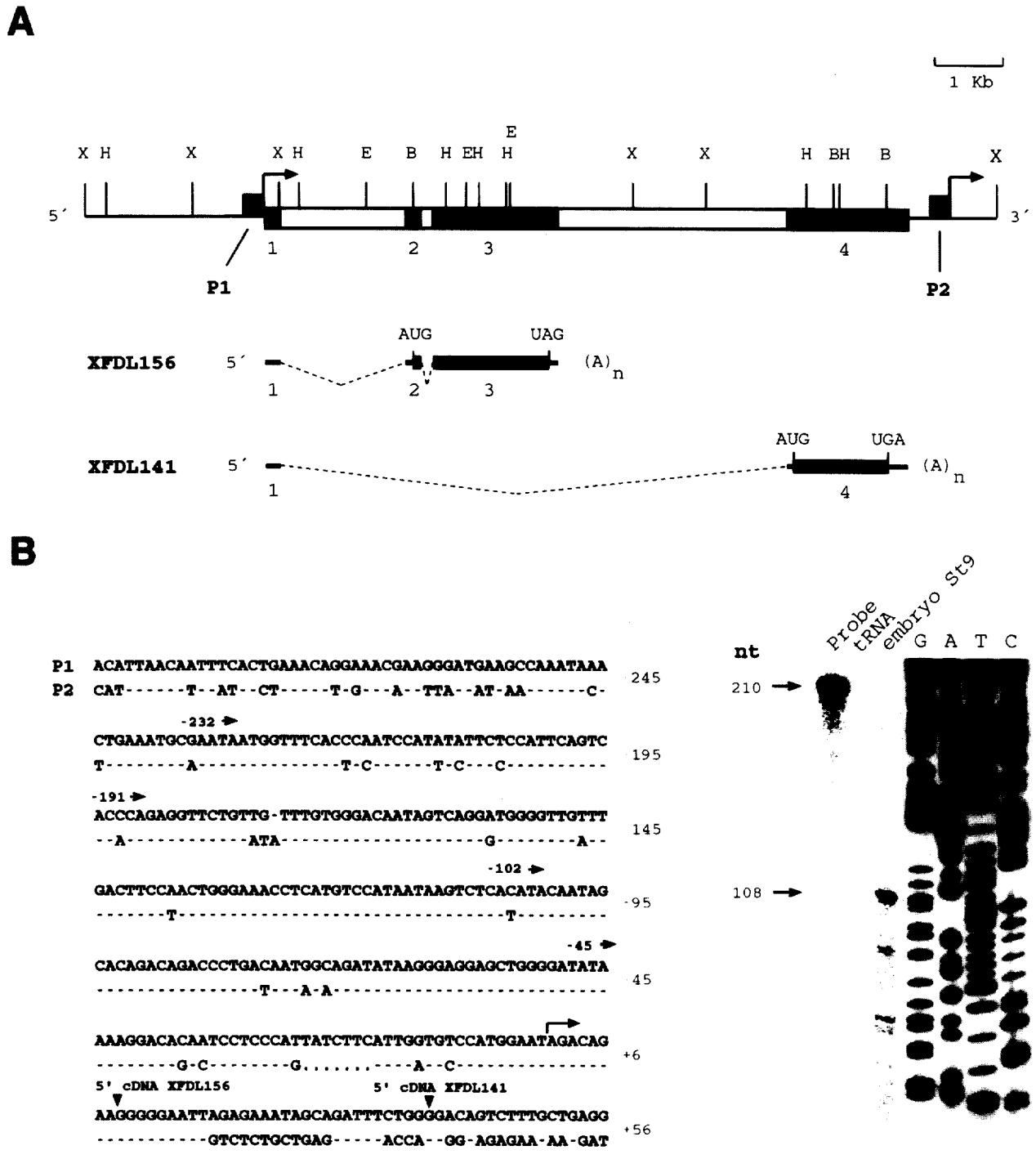


Fig. 4. Genomic organization of the XFDL 141/156 transcription unit. (A) Functional units within a *Xenopus laevis* genomic fragment termed XFG 141/156 that contains the XFDL 141/156 transcription unit. Restriction sites are indicated as: X, XbaI; H, HindIII; E, EcoRI; and B, BamHI. Filled and open boxes reflect exons and introns, respectively. The exons corresponding to either the XFDL 156 or the XFDL 141 cDNAs are indicated in the lower portion of the figure. The minimal promoter of the XFDL 141/156 transcription unit is indicated as P1, a duplication of the same element as P2. (B) Primary sequence comparison of the P1 and P2 promoter elements (left hand panel). Dots indicate nucleotides present in one but not the other sequence; hyphens reflect positions of sequence identity. Residues corresponding to the 5′-ends of the XFDL 141 and XFDL 156 cDNAs are indicated. The transcription start site as mapped by RNase protection analysis is indicated by an arrow. Determination of the transcription start site of the XFDL 141/156 transcription unit by RNase protection analysis (right hand panel). A genomic DNA fragment that extends the region corresponding to the 5′-ends of the XFDL 141/156 cDNAs was used to generate a ^{32}P-labelled antisense RNA. This antisense RNA was hybridized to total RNA preparations from stage 9 *Xenopus* embryos and subjected to RNase digestion. The size of the protected antisense RNA fragment was determined in correlation with a DNA sequencing ladder of an unrelated DNA fragment.

ment reveals that the XFDL 141 and XFDL 156 mRNAs are generated by alternative splicing from the same primary transcript. The XFDL 141/156 transcription unit could be part of a gene cluster with duplicated promoter structures.

2.4. Functional analysis of the XFDL 141/156 promoter

Early embryonic expression of the XFDL 141/156 transcription unit is characterized by two prominent features, namely early zygotic and transient activation (as detailed above). In order to map promoter elements responsible for these properties, the 2.7 kb 5′-upstream portion of the genomic clone was fused to a sequence encoding chloramphenicol acetyl transferase (CAT). Injection of this reporter gene construct into fertilized eggs of *Xenopus laevis* results in the induction of strong CAT activity in extracts derived from stage 10 embryos (Fig. 5B). The level of activity achieved is similar to what is obtained upon injection of a CAT reporter construct driven by a strong, constitutively expressed viral promoter (pSV2CAT).

In order to map the essential regulatory elements within the XFDL 141/156 promoter, progressive 5′-deletions were generated and the resulting constructs assayed for transcription activity by embryo microinjection as above. It turns out that a fragment containing only the first 191 nucleotides of the 5′-upstream region is sufficient to generate a transcription response in early embryos that is comparable to the one achieved with the entire upstream fragment. Further deletions result in a loss of CAT reporter activity (Fig. 5B). As pointed out above (Fig. 4B), a DNA sequence element with a high degree of sequence identity in comparison to this minimal promoter structure is localized immediately downstream of the XFDL 141/156 transcription unit. Fusion of this putative second promoter to the CAT encoding region equally results in induction of strong CAT expression in microinjected embryos (Fig. 5A and data not shown).

In order to analyze if the transcription activation obtained with the deletion constructs is transient, in a way comparable to the early zygotic expression of endogenous gene, we made use of the RNase protection assay. As one would predict from the enzymatic activities produced with the microinjected reporter constructs, significant amounts of CAT mRNA were only detected with deletion mutants up to position 192. More importantly, RNA produced from these transcriptionally active microinjected reporter constructs was readily detected in blastula (stage 10) *Xenopus* embryos, whereas RNA preparations from later stages of embryogenesis, such as from stage 18 or stage 32 embryos, reveal only strongly reduced or background levels of CAT mRNA. In contrast, microinjection of a reporter construct driven by the viral promoter (pSV2CAT) results in high levels of CAT mRNA also in RNA preparations from stage 18 and 32 embryos (Fig. 5C).

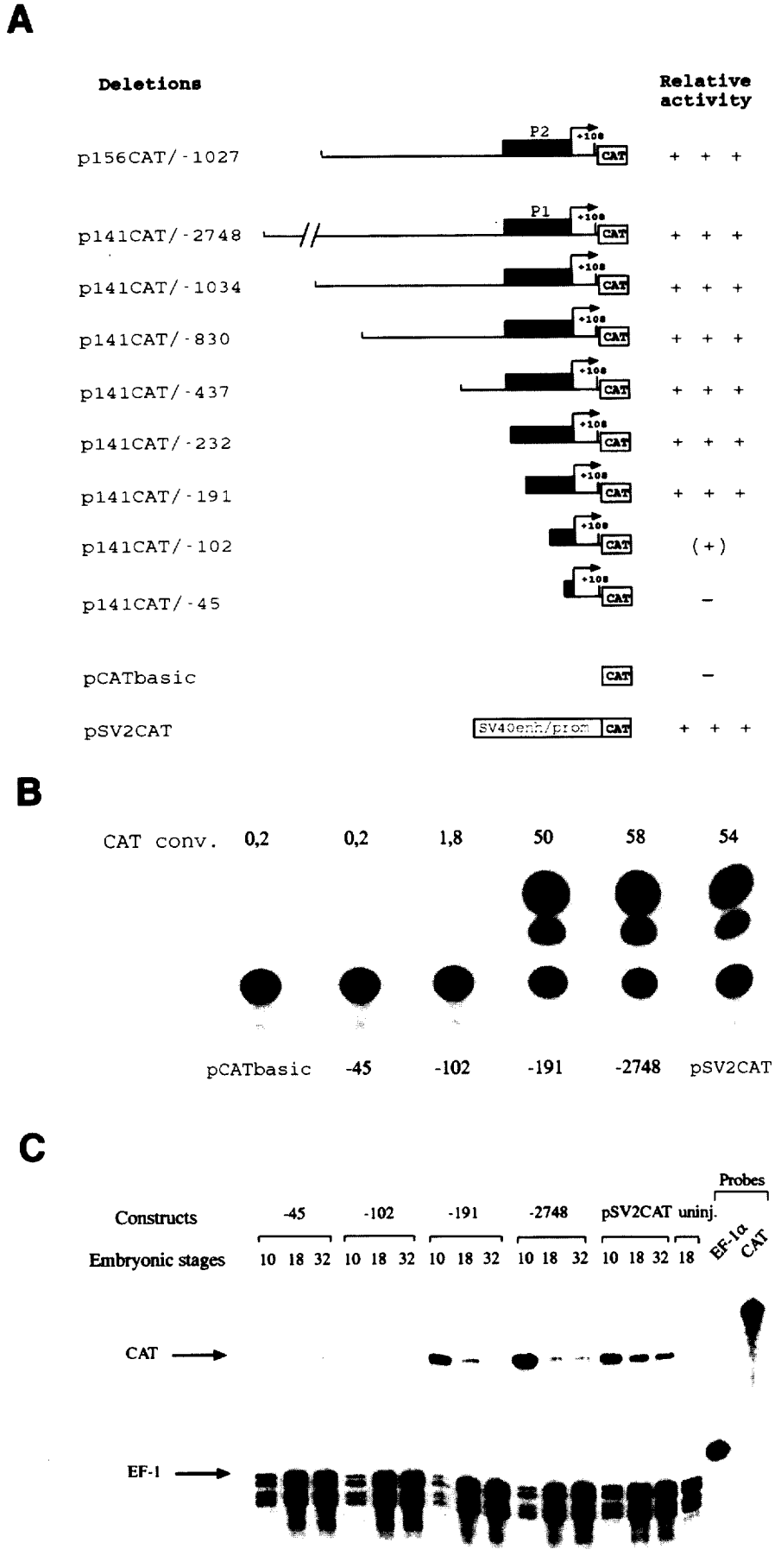
We conclude that regulatory elements contained within only 191 nucleotides of 5′-flanking region in the XFDL 141/156 transcription unit are sufficient for the early zygotic and transient expression of this gene. A functional duplication of this promoter is localized in the immediate 3′-flanking region of the XFDL 141/156 gene.

3. Discussion

We have resolved structural and regulatory elements of a transcription unit that, by means of alternative splicing, gives rise to two different C2H2-type zinc finger proteins in *Xenopus*. During early embryogenesis, expression of this gene is transiently activated at the mid blastula transition (MBT).

Transcription regulatory elements of another gene that shares these expression characteristics and is termed GS17 has been characterized in detail. A 74 bp enhancer element that is sufficient to confer these embryonic transcription properties to a heterologous promoter-reporter construct was identified and found to contain a crucial binding site for the maternal transcription factor OZ-1 (Krieg and Melton, 1987; Ovsenek et al., 1992). Corresponding sequence elements have also been identified in a number of other *Xenopus* embryonic genes; however, the XFDL 141/156 transcription regulatory region is devoid of a consensus binding site for the OZ-1 protein. Therefore, in this case, transcription activation at MBT appears to depend on one or more factors distinct from OZ-1. Transcription regulation of the gene encoding elongation factor 1α is similarly activated at MBT but, in contrast to the transcription units discussed above, continues to be expressed constitutively in the embryo and in all somatic tissues. These related transcription properties appear to require the zinc finger factor SP1 for maximal levels of activity (Johnson and Krieg, 1995). Again, the XFDL 141/156 transcription control region does not contain an SP1 binding site. Thus, the expression profile of the XFDL 141/156 gene must be under the control of different transcription factors active at MBT; it will be of interest to identify these regulators and the sequence ele-

Fig. 5. Determination of a minimal promoter fragment for transient zygotic gene activation. (A) Schematic representation of reporter constructs containing either fragments from the XFDL 141/156 transcription unit or a control promoter. Progressive 5′-deletions of a fragment containing the region upstream to the XFDL 141/156 transcription start site, or the P2 element (compare Fig. 4) were fused to a CAT-encoding cDNA sequence. The pSV2CAT and pCATbasic plasmids were used as positive and negative controls, respectively. (B) CAT assays with extracts from microinjected embryos. *Xenopus* embryos were microinjected with the reporter constructs shown in (A). The activity is given as percent chloramphenicol conversion. (C) RNase protection analysis of CAT mRNA levels from microinjected embryos. *Xenopus* embryos were microinjected with the reporter constructs shown in (A) and allowed to develop to the embryonic stages as indicated. RNase protection assays were performed with total RNA preparations from such embryos and a radioactive antisense RNA probe corresponding to a portion of the CAT mRNA. The EF1α probe was used as a control.


A
Deletions
Relative activity
p156CAT/-1027
P2
+1
CAT
+ + +
p141CAT/-2748
P1
p141CAT/-1034
p141CAT/-830
p141CAT/-437
p141CAT/-232
p141CAT/-191
p141CAT/-102
(+)
p141CAT/-45
–
pCATbasic
pSV2CAT
SV40enh/prom
B
CAT conv.
0,2
0,2
1,8
50
58
54
pCATbasic
-45
-102
-191
-2748
pSV2CAT
C
Probes
Constructs
-45
-102
-191
-2748
pSV2CAT
uninj.
Embryonic stages
10 18 32 10 18 32 10 18 32 10 18 32 10 18 32 18
EF-1α
CAT
CAT
EF-1

ments within the XFDL 141/156 control region they recognize, which will be different from what has been identified previously to be important for early zygotic gene activation in *Xenopus laevis*.

Alternative splicing is a mechanism that generates structurally and functionally distinct protein isoforms from the same transcription unit. Several examples for alternatively spliced zinc finger protein encoding transcripts have been described; often, proteins with distinct sets of zinc finger modules and therefore distinct DNA binding specificities are generated (Bordereaux et al., 1990; Bickmore et al., 1992; Hsu et al., 1992). In the case of the XFDL 141/156 gene described here, the two alternative protein products, though structurally related, are completely distinct. In respect to the two corresponding mRNAs, structural identity is restricted to the first exon, i.e. the 5′-non coding region. The high degree of structural homology observed in a primary sequence comparison of the two zinc finger proteins generated from these RNAs suggests that an internal sequence duplication event may have given rise to this particular gene structure. Also, the fact that the immediate 3′-flanking region of the XFDL 141/156 transcription unit contains a duplicated, functional copy of the transcription control region indicates that this gene is part of a genomic cluster, similar to what has been described for other *Xenopus* and human ZFP genes (Bellefroid et al., 1993; Nietfeld et al., 1993). While the structural data obtained on predicted protein sequence and genomic organization of the numerous vertebrate zinc finger protein encoding genes illustrate evolutionary aspects, the biological function for the vast majority of these proteins remains highly enigmatic.

4. Materials and methods

4.1. Isolation and structural analysis of the XFDL 141 and 156 cDNAs

A 0.6 kb cDNA fragment derived from the XlcNF1 cDNA containing a cluster of zinc finger repeats (Köster et al., 1988) was used to screen a dorsal blastopore lip specific IZAPII cDNA library (Blumberg et al., 1991) under reduced stringency. Fifty positive, single clones were isolated, out of which two groups of overrepresented clones were identified by restriction digest and Southern blot analysis. The XFDL (*Xenopus* Finger Dorsal Lip) 141 clone were represented four times and XFDL 156 seven times in this primary selection of 50 cDNAs. Those clones containing the longest inserts, for XFDL 141 a 1.9 kb EcoRI-XhoI fragment and for XFDL 53 a 1.7 kb EcoRI-XhoI fragment, were sequenced using a Sequenase II kit (USB). The sequence of the partial XFDL 53 cDNA clone was found to overlap with previously isolated cDNAs, termed HC 14 and XFG 32-1 (Köster et al., 1988; Nietfeld et al., 1989). The resulting composite sequence constitutes a 2.1 kb cDNA, referred to as XFDL 156.

4.2. Isolation and structural analysis of a XFDL 141/156 encoding genomic fragment

A *Xenopus laevis* genomic library in λFIXII (Stratagene) was screened with a 0.4 kb EcoRI-PstI fragment corresponding to the 5′ end of the XFDL 141 cDNA. Two positive clones were isolated, which, according to restriction enzyme mapping, appeared to be identical. Southern blot analysis using different parts of the XFDL 141 and XFDL 156 cDNAs as probes indicated that the genomic clone contains the sequence information for both cDNAs. Subclones were generated which cover the entire genomic fragment. Sequencing of these genomic subclones was performed on an Applied Biosystems sequencing system, using Taq dye terminator cycle sequencing. The sequence analysis for the genomic clone was completed with the exception of part of the large intron separating the last exon of XFDL 156 from the second exon of XFDL 141.

4.3. RNA isolation, Northern blot and RNase protection analysis

RNA was isolated from embryonic tissues by SDS/proteinase K digestion and selective precipitation with lithium chloride (Krieg et al., 1989). Northern analysis was performed as described in El Baradi et al., 1991. For XFDL 141, either the full length 1.9 kb EcoRI-XhoI cDNA or a 0.4 kb EcoRI-PstI fragment from the 5′ end of the cDNA were used as probes. For XFDL 156, a 0.5 kb EcoRI-EcoRI cDNA fragment encompassing the zinc finger repeats was used as probe. A histone H4 probe was used as a control. RNase protection assays were carried out basically as described by Melton et al. (1984). The probe for mapping the transcription start site of the XFDL 141/156 gene was generated from a 210 bp genomic fragment covering 70 bp and 100 bp of the most 5′ portion of the XFDL141 and 156, respectively, and extending further upstream. The CAT probe was generated from an XbaI-EcoRI fragment of the pCATbasic vector (nucleotides 2266–2528) (Promega). The EF1a probe has been described previously (Ferreiro et al., 1994).

4.4. In situ hybridization and histology

In situ hybridization to sections was performed as described (Lemaire and Gurdon, 1994). Whole mount in situs were carried out according to Chitnis et al. (1995). The full length XFDL141 sequence was subcloned in pBluescript SK(−) between the EcoRI and XhoI sites. To prepare the antisense probe, this plasmid was linearized at the EcoRI site and transcribed using T7 polymerase in the presence of digoxygenin-11-UTP (Boehringer 1209 256). The DIG-labelled hybrids were visualized using alkaline phosphatase-conjugated anti-dig Fab fragments and NBT and BCPIP substrates (Boehringer 1383 321 and 1383 213). Images from embryos and mounted sections were

visualized using Nomarski interference optics on a Zeiss Axioskop and were acquired with a CCD camera (Sony).

4.5. Construction of CAT-reporter plasmids

The XFDL 141/156 promoter, corresponding to positions −2748 up to +108 relative to the start site of transcription as determined on the genomic fragment, was cloned into the HindIII site of the promoterless pCAT-Basic vector (Promega). A genomic fragment encoding the P1-related sequence as identified in close proximity to the 3′ end of the transcription unit was similarly cloned 1135 nucleotide fragment using PstI and SalI restriction sites. The 5′ deletions of the XFDL 141/156 promoter were obtained using appropriate restriction sites within the promoter region (XbaI, −1034; SphI, −830), or by PCR using specific primers located at positions −437, −232, −191, −102 and −45. The pSV2CAT plasmid contains the CAT gene under the control of SV40 promoter and enhancer sequences.

4.6. Embryo microinjection and CAT assays

Embryos were obtained from *Xenopus laevis* adult frogs by hormone-induced egg-laying and in vitro fertilization using standard methods. Embryos were staged according to Nieuwkoop and Faber (1967). Embryos were injected in the animal region in each blastomeres of the two cell stage with 100 pg of supercoiled plasmid DNA in 0.5 × MBS, 1% Ficoll and were then transferred to 0.1 × MBS at 1 h after injection. For each assay, eight embryos grown up to the desired stage were homogenized in 0.25 M Tris (pH 7.8) and equal amounts of protein were used in the CAT assays. All experiments were repeated at least three times. Levels of CAT activity were quantified from TLC plates by use of a Phosphorimager (Molecular Dynamics).

Acknowledgements

We are grateful to Regina Buhl and Michaela Beuße-Jendraszek for expert technical assistance. EJB is recipient of a long term EMBO fellowship. This work was supported by funds from the Deutsche Forschungsgemeinschaft to TP.

References

Bellefroid, E.J., Marine, J.-C., Ried, T., Lecocq, P.J., Rivière, M., Amemiya, C., Poucelet, D.A., Coulie, P.G., de Jong, P., Szpirer, C., Ward, D.C. and Martial, J. (1993) Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. EMBO J. 12, 1363–1374.

Bickmore, W.A., Oghene, K., Little, M.H., Seawright, A., Van Heyningen, V. and Hastie, N.D. (1992) Modulation of DNA binding specificity by alternative splicing of the Wilms tumor WT1 gene transcript. Science 257, 235–237.

Blumberg, B., Wright, C.V.E., De Robertis, E.M. and Cho, K.M.Y. (1991) Organizer specific homeobox genes in *Xenopus laevis* embryos. Science 253, 194–196.

Bordereaux, D., Fichelson, S., Tambourin, P. and Gisselbrecht, S. (1990) Alternative splicing of the Evi-1 zinc finger gene generates mRNAs which differ by the number of zinc finger motifs. Oncogene 5, 925–927.

Bouwmeester, T. (1993) The molecular and functional characterization of zinc finger and Myb-related proteins during early embryonic development in *Xenopus laevis*. PhD thesis, Free University of Amsterdam, Holland, pp. 35–44.

Chitnis, A., Henrique, D., Lewis, J., Ish-Horowicz, D. and Kintner, C. (1995) Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene Delta. Nature 375, 761–766,.

El Baradi, T., Bouwmeester, T., Giltay, R. and Pieler, T. (1991) The maternal store of zinc finger protein encoding mRNAs in fully grown *Xenopus* oocytes is not required for early embryogenesis. EMBO J. 10, 1407–1413.

Ferreiro, B., Kintner, C., Zimmerman, K., Anderson, D. and Harris, W. (1994). *XASH* gene promotes neurogenesis in *Xenopus* embryos. Development 120, 3649–3655.

Hsu, T., Gogos, J.A., Kirsh, S.A. and Kafatos, F.C. (1992) Multiple zinc finger forms resulting from developmentally regulated alternative splicing of a transcription factor gene. Science 257, 1946–1950.

Johnson, A.D. and Krieg, P.A. (1995) A *Xenopus laevis* gene encoding EF-1 alpha S, the somatic form of elongation factor 1 alpha: sequence, structure and identification of regulatory elements required for embryonic transcription. Dev. Genet. 17, 280–290.

Köster, M., Pieler, T., Pöting, A and Knöchel, W. (1988) The finger motif defines a multigene family represented in the maternal mRNA of *Xenopus* oocytes. EMBO J. 7, 1735–1741.

Krieg, P.A. and Melton, D.A. (1985) Developmental regulation of a gastrula-specific gene injected into fertilized *Xenopus* eggs. EMBO J. 4, 3463–3471.

Krieg, P.A. and Melton, D.A. (1987) An enhancer responsible for activating transcription at the mid-blastula transition in *Xenopus* development. Proc. Natl. Acad. Sci. USA 84, 2331–2335.

Krieg, P.A., Varnum, S.M., Wormington, W.M. and Melton, D.A. (1989) The mRNA encoding elongation factor-1a (EF-1a) is a major transcript at the mid-blastula transition in *Xenopus*. Dev. Biol. 133, 93–100.

Larsson, S.H., Charlieu, J.-P., Miyagawa, K., Engelkamp, D., Rassoulzadegan, M., Ross, A., Cuzin, F., Van Heyningen, V. and Hastie, N.D. (1995) Subnuclear localization of WT1 in splicing of transcription factor domains is regulated by alternative splicing. Cell 81, 391–401.

Lemaire, P. and Gurdon, J.B. (1994) A role for cytoplasmic determinants in mesoderm patterning: cell-autonomous activation of the goosecoid and Xwnt-8 genes along the dorsoventral axis of early *Xenopus* embryos. Development 120, 1191–1199.

Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12, 7035–7056.

Newport, J. and Kirschner, M. (1982) A major developmental transcription in early *Xenopus* embryos: I. Characterization and timing of cellular changes at the midblastula stage. Cell 30, 675–686.

Nietfeld, W., el Baradi, T., Mentzel, H., Pieler, T., Köster, M., Pöting, A and Knöchel, W. (1989) Second-order repeats in *Xenopus laevis* finger proteins. J. Mol. Biol. 208, 639–659.

Nietfeld, W., Courad, S., van Wijk, I., Giltay, R., Bouwmeester, T., Knöchel, W. and Pieler, T. (1993) Evidence for a clustered genomic organization of Fax-zinc finger protein encoding transcription units in *Xenopus laevis*. J. Mol. Biol. 230, 400–412.

Nieuwkoop, P.D. and Faber, J. (1967). Normal Table of *Xenopus laevis*. North Holland, Amsterdam.

Ovsenek, N., Zorn, A.M. and Krieg, P.A. (1992) A maternal factor, OZ-1,

activates embryonic transcription of the *Xenopus laevis* GS17 gene. Development 115, 649–655.

Pieler, T. and Bellefroid, E. (1994) Perspectives on zinc finger protein function and evolution: an update. Mol. Biol. Rep. 20, 1–8.

Schäfer, U. Rausch, O., Bouwmeester, T. and Pieler, T. (1994) Sequence specific recognition of a repetitive DNA element by a C2H2 zinc finger protein in *Xenopus*. Eur. J. Biochem. 226, 567–576.

Scotto, K.W., Kaulen, H. and Roeder, R.G. (1989) Positive and negative regulation of the gene for transcription factor IIIA in *Xenopus laevis* oocytes. Genes Dev. 3, 651–662.