xlgv7: a maternal gene product localized in nuclei of the central nervous system in *Xenopus laevis*

Mill Miller,1,4 Malgorzata Kloc,1,4 Bramham Reddy,1 Eric Eastman,1 Christine Dreyer,2 and Laurence Etkin1,3

1Department of Molecular Genetics, The University of Texas, M.D. Anderson Cancer Center, Houston, Texas 77030 USA; 2Max Planck Institute for Developmental Biology, Department of Cell Biology, D-7400 Tubingen, FRG

The *Xenopus* oocyte nucleus (GV) is a storehouse for a large number of proteins that are used during early development. We have cloned and characterized a cDNA coding for a maternal gene product that is localized in the GV and then becomes highly enriched in the nuclei of the central nervous system (CNS) of tadpoles and adult frogs. This cDNA (xlgv7) is 2.1 kb and hybridizes to a 2.4-kb RNA species on Northern blots. Southern blots of genomic DNA suggest that this gene is a member of a multigene family. The cDNA sequence reveals a long open reading frame (ORF) of 1773 nucleotides, with a putative nuclear targeting signal (Glu Arg Arg Lys Lys Lys Thr) at the extreme carboxyl terminus and an internal histidine (His)-rich region with a repeated conserved amino acid sequence between His pairs. The significance of this region is unclear, but the protein is a DNA-binding protein, and it is possible that this region is involved in this function. The xlgv7 protein also possesses a putative nucleotide-binding consensus sequence that is similar to the bacterial RecA and RecB and yeast RAD proteins. Protein xlgv7 exists as several isotypes that exhibit developmental and cell-specific changes during development. Northern blot analysis of the abundance of the xlgv7 mRNA shows an accumulation following neural induction at stages 15–16. There is a transient expression of the mRNA in the gut of tadpoles. In the adult, the mRNA is highly enriched in the brain and is absent or in very low abundance in other tissues. Immunohistochemical analysis of the protein shows that the protein is localized in the nuclei of the brain cells. We conclude that the xlgv7 gene product is a maternal protein that may serve several important functions, one of which may be in the development and maintenance of the CNS.

[Key Words: *Xenopus*, nuclear protein; CNS; differentiation; development]

Received August 17, 1988; revised version accepted February 22, 1989.

A central question in developmental biology concerns the mechanism by which the embryo specifies cell fates during development and cell differentiation. Evidence suggests that the process of cell specification may be initiated very early in development by localized cytoplasmic determinants or morphogens that establish asymmetry in the embryo. This initial asymmetry which may bias cell fates, may be amplified by inductive cell–cell interactions [Gurdon 1985; Davidson 1986].

The existence of such factors has been analyzed both biochemically, with the discovery of localized mRNAs for growth factor-like molecules [Weeks and Melton 1987], and genetically, by the analysis of maternal-effect mutations that influence the synthesis or patterns of localization of morphogens [e.g., Driever and Nüsslein-Volhard 1988]. In fact, there are numerous examples of maternal-effect mutations that effect early development and cell commitment, as well as body pattern [for review, see Davidson 1986; Lawrence 1988]. It is likely that maternal products also have pronounced effects on the developmental events much later during development.

In an attempt to identify such maternal factors, Dreyer et al. [1981, 1982] produced a series of monoclonal antibodies against maternal gene products, which were localized in the nucleus (GV) of the *Xenopus laevis* oocyte. The fate of these antigens was followed using immunohistochemical staining of sectioned embryos, larvae, and adults. Upon GV breakdown at maturation, the antigens are released into the cytoplasm [Dreyer et al. 1983; Dreyer et al. 1985; Hausen et al. 1985]. During development, groups of these antigens reaccumulate within the embryonic nuclei at characteristic developmental stages [Dreyer and Hausen 1983]. Some proteins reenter the nuclei during early cleavage, whereas others reenter the nuclei at the blastula stage or later. During early embryogenesis, most of the antigens are distributed equally to all nuclei. Interestingly, some of these GV proteins become localized in the nuclei of specific cell types later during development.

One of these antigens, referred to as protein 7, reenters the nuclei at the early gastrula stage of development and eventually becomes localized predominantly in the nuclei of the central nervous system (CNS) in the tadpole.

572 GENES & DEVELOPMENT 3:572–583 © 1989 by Cold Spring Harbor Laboratory Press ISSN 0890-9369/89 $1.00
Xenopus primer extension using a radiolabeled synthetic oligonucleotide. Analysis of the cDNA extension products of this precise length of the 5' end of the mRNA, we carried out it codes for a protein that is comparable to the molecular weight of protein xlgv7, which was determined by gel electrophoresis [Dreyer et al. 1982]; [2] no other ORF can encode a polypeptide of this size; and [3] RNA blots probed with single-stranded probes demonstrate that this reading frame is represented on the sense strand. There was no significant match to any protein or DNA sequences in the GenBank, EMBL, or NBRF data bases.

The predicted amino acid sequence of the xlgv7 protein exhibits several interesting features. The first is the presence of a basic peptide sequence located between amino acids 579 and 583 at the carboxyl terminus of the protein. This sequence [Glu Arg Arg Lys Lys Thr] closely resembles the nuclear targeting sequences known to confer nuclear compartmentalization of other nuclear proteins (Table 1). For example, the Xenopus histone-binding proteins N1/N2 have a targeting sequence, Val Arg Lys Arg Lys Thr, which is almost identical to that found in protein xlgv7. Both are variations of the SV40 T antigen nuclear targeting signal sequence. This is consistent with the nuclear localization of this protein.

A second feature of the amino acid sequence is the presence of a His-rich region (amino acids 101–437). The His residues are separated by variable numbers of amino acids. Within several of the paired His residues, there is a conserved motif. This consists of a core Glu-Lys-Asn-Gln, which is located between the His pairs at amino acids 170–177, 182–191, 222–229, and 229–240. The functional significance of these repeated sequences is not clear.

A third feature is the presence of a putative nucleotide-binding site. Walker et al. (1982) defined two conserved motifs associated with nucleotide-binding proteins. Motif A consists of Leu Gly n Gly Gly Gln, whereas motif B consists of a single-stranded probe demonstrates that this reading frame is represented on the sense strand. There was no significant match to any protein or DNA sequences in the GenBank, EMBL, or NBRF data bases.

The predicted amino acid sequence of the xlgv7 protein exhibits several interesting features. The first is the presence of a basic peptide sequence located between amino acids 579 and 583 at the carboxyl terminus of the protein. This sequence [Glu Arg Arg Lys Lys Thr] closely resembles the nuclear targeting sequences known to confer nuclear compartmentalization of other nuclear proteins (Table 1). For example, the Xenopus histone-binding proteins N1/N2 have a targeting sequence, Val Arg Lys Arg Lys Thr, which is almost identical to that found in protein xlgv7. Both are variations of the SV40 T antigen nuclear targeting signal sequence. This is consistent with the nuclear localization of this protein.

A second feature of the amino acid sequence is the presence of a His-rich region (amino acids 101–437). The His residues are separated by variable numbers of amino acids. Within several of the paired His residues, there is a conserved motif. This consists of a core Glu-Lys-Asn-Gln, which is located between the His pairs at amino acids 170–177, 182–191, 222–229, and 229–240. The functional significance of these repeated sequences is not clear.

A third feature is the presence of a putative nucleotide-binding site. Walker et al. (1982) defined two conserved motifs associated with nucleotide-binding proteins. Motif A consists of Leu Gly n Gly Gly Gln, whereas motif B consists of an aspartate [Asp] residue that binds ATP. In most cases, the Asp residue is preceded by four hydrophobic amino acids. In xlgv7, the A motif is located at amino acid residues 120–129. This sequence, Leu Gly n Gly Gly Gly Lys Thr, is very similar to the sequences found on the yeast RAD18 and RAD3 proteins [Reynolds et al. 1985; Jones et al. 1988] and several bacterial proteins such as RecA and RecB, which bind and hydrolyze ATP (see Table 2). The B motif is located 196 amino acids away from the A motif [315–327 see Table 2] and consists of Lys n n n Lys Leu Tyr Lys Asp. Table 2 shows a comparison of this sequence with other known nucleotide-binding proteins.

The xlgv7 cDNA may encode the nuclear protein 7

To determine the relationship between the clone xlgv7 and nuclear protein 7, we devised the following experimental strategy relying on the production of specific chymotryptic digestion fragments based on the conceptual amino acid sequence. This was necessary because it was difficult to obtain unequivocal hybrid select or hybrid arrest data.

From the amino acid sequence we predicted that the proteolytic enzyme chymotrypsin should produce 11 polypeptide fragments, ranging in size from 590 to...
33,447 daltons (Table 3). One of these fragments should be detectable on Western blots of the digested protein, assuming that the chymotrypsin did not digest in the middle of the epitope and destroy its reactivity. Therefore, we carried out peptide mapping experiments on total proteins isolated from GVs and on total protein isolated from a bacterial expression system (M. Kloc and L. Etkin, unpubl.) in which the xlgv7 clone was ex-

Figure 1. Nucleotide sequence of the xlgv7 cDNA clone and the deduced amino acid sequence of the encoded protein. (*) The ATG initiation codon of the long ORF; (**) stop codons. The putative nuclear targeting sequence is underlined. The conserved sequences located between the His residues are indicated with dashed lines. The site of the putative nucleotide-binding sequence is indicated by double underline.

Table 1. Summary of nuclear targeting sequences

<table>
<thead>
<tr>
<th>Nuclear protein</th>
<th>Nuclear targeting sequence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mat α₂</td>
<td>lys ile pro ile lys</td>
<td>Hall et al. (1984)</td>
</tr>
<tr>
<td>large T antigen</td>
<td>pro lys lys lys arg lys val</td>
<td>Lanford and Butel (1984)</td>
</tr>
<tr>
<td>Lamin L1*</td>
<td>gln arg lys arg ile asp</td>
<td>Krohne et al. (1987)</td>
</tr>
<tr>
<td>nucleoplasmin sequence A</td>
<td>gln arg lys arg ile asp</td>
<td>Dingwall et al. (1987)</td>
</tr>
<tr>
<td>Bb</td>
<td>gly glut lys lys lys leu asp</td>
<td>Burglin and DeRobertis (1987)</td>
</tr>
<tr>
<td>C</td>
<td>pro thr lys lys gly gly</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>arg lys pro ila lys</td>
<td></td>
</tr>
<tr>
<td>N1N2</td>
<td>val arg lys lys lys thr</td>
<td></td>
</tr>
<tr>
<td>nucleoprotein 7-2*</td>
<td>glu arg lys lys lys thr</td>
<td>Kleinschmidt et al. (1986)</td>
</tr>
</tbody>
</table>

* These sequences are putative nuclear targeting signals and have not been tested in a functional assay.

** A functional analysis has indicated that this is the primary nuclear targeting signal in nucleoplasmin.
Table 2. Homology between putative nucleotide-binding sequence of protein xlgv7 and other known ATPases

<table>
<thead>
<tr>
<th>Protein</th>
<th>Residues</th>
<th>Sequence</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motif A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAD 3</td>
<td>40–49</td>
<td>L E M P S S T G K T</td>
<td>Reynolds et al. (1985)</td>
</tr>
<tr>
<td>Rec A</td>
<td>64–73</td>
<td>I Y G P E S S S T G K T</td>
<td>Bullock et al. (1987)</td>
</tr>
<tr>
<td>Rec B</td>
<td>21–30</td>
<td>I E A S A G T G K T</td>
<td>Finch et al. (1986)</td>
</tr>
<tr>
<td>xlgv7</td>
<td>120–129</td>
<td>L P I P A P R G K T</td>
<td>this paper</td>
</tr>
<tr>
<td>Motif B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xlgv7</td>
<td>315–327</td>
<td>K D Q G P E F L L Y K D</td>
<td>this paper</td>
</tr>
<tr>
<td>bovine ATPase</td>
<td>246–258</td>
<td>R D Q E G Q D V L L F I D</td>
<td>Walker et al. (1982)</td>
</tr>
<tr>
<td>E. coli ATPase</td>
<td>232–244</td>
<td>R D Q E G R D V L L F I D</td>
<td>Saraste et al. (1981)</td>
</tr>
</tbody>
</table>

Underlined amino acids represent identical or similar sequences shared by all proteins.

The xlgv7 gene may be a member of a multigene family

Figure 3 shows an autoradiograph of a Southern blot in which genomic DNA from *Xenopus* liver was digested with several different restriction enzymes and probed blot to nitrocellulose, and probed with mAb 37-1A9. Figure 2A (lanes 3 and 4) shows that a peptide fragment of ~7400 daltons was detected on Western blots from both the GV proteins and bacterially produced protein. This is similar in size (7381 daltons) to one of the chymotryptic digestion fragments predicted from the amino acid sequence [fragment 9, Table 3]. Also, Figure 2A (lane 2) shows a partial chymotryptic digest of the GV protein revealing a ladder of fragments, along with the complete digestion product of the proper size. We have never detected any reaction with bacterial extracts using mAb 37-1A9, except when the xlgv7 gene has been expressed [Fig. 2A, lane 5]. We believe that the common antigenicity of the GV protein and the bacterially produced protein, as well as the detection of a predicted chymotryptic digestion fragment from both the bacterially produced and endogenous GV protein, suggests that the xlgv7 clone codes for the GV protein 7 or a closely related polypeptide. Interestingly, in another expression system that produces a fusion protein, there are differences in the chymotryptic digestion patterns. Currently, we are analyzing this phenomenon.

Table 3. Predicted chymotrypsin digestion fragments from protein 7

<table>
<thead>
<tr>
<th>Fragment</th>
<th>Location [nucleotide]</th>
<th>Amino acids</th>
<th>Size [daltons]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24–65</td>
<td>14</td>
<td>2,209</td>
</tr>
<tr>
<td>2</td>
<td>66–120</td>
<td>18</td>
<td>2,148</td>
</tr>
<tr>
<td>3</td>
<td>121–240</td>
<td>39</td>
<td>5,145</td>
</tr>
<tr>
<td>4</td>
<td>241–606</td>
<td>121</td>
<td>15,381</td>
</tr>
<tr>
<td>5</td>
<td>607–660</td>
<td>17</td>
<td>2,457</td>
</tr>
<tr>
<td>6</td>
<td>661–691</td>
<td>9</td>
<td>1,347</td>
</tr>
<tr>
<td>7</td>
<td>692–757</td>
<td>21</td>
<td>2,673</td>
</tr>
<tr>
<td>8</td>
<td>758–811</td>
<td>17</td>
<td>2,431</td>
</tr>
<tr>
<td>9</td>
<td>812–984</td>
<td>57</td>
<td>7,381</td>
</tr>
<tr>
<td>10</td>
<td>985–996</td>
<td>4</td>
<td>590</td>
</tr>
<tr>
<td>11</td>
<td>996–1794</td>
<td>266</td>
<td>33,447</td>
</tr>
</tbody>
</table>

Figure 2. Western blots of chymotryptic digests of GV protein xlgv7. (A) Proteins were extracted from oocyte GV's or bacteria containing a plasmid expression vector with the xlgv7 cDNA. The protein was either used undigested or digested with chymotrypsin (see Materials and methods) and separated on a 17% polyacrylamide gel. (Lane 1) 20 µg of total GV protein (undigested); (lane 2) 20 µg of total GV protein partially digested with chymotrypsin [10 mg/ml]; (lane 3) complete digestion of bacterially produced protein [20 mg/ml]; (lane 4) 20 µg of total GV protein completely digested with chymotrypsin [20 mg/ml]; (lane 5) proteolytic fragment predicted from the amino acid sequence. (B) Chymotrypsin digestion of proteins isolated from nuclei of stage 24 [Lane 1] Undigested protein [120 µg of total protein]; [lane 2] digested protein [120 µg of total protein digested with chymotrypsin [20 mg/ml].
Miller et al.

Figure 3. Southern blots of genomic DNA digested with various restriction enzymes and probed with a random primed probe from xlgv7 [A] and with a probe of another maternal single copy gene [B] [M. Miller, E. Eastman, L. Etkin, unpubl.]. Each lane contains 10 μg of DNA. (Lanes 1–4 contain genomic DNA digested with EcoRI, BamHI, HindIII, and PstI, respectively.

Loci have been identified [Bisbee et al. 1977] or, possibly, a large number of introns in the chromosomal gene. It should be pointed out, however, that the Xenopus genome has undergone substantial diploidization. We are currently cloning and characterizing the genomic members of the xlgv7 gene family.

The xlgv7 protein exists as several developmentally expressed isotypes

Figure 4 shows a Western blot of protein xlgv7 extracted from brain, oocytes, and early embryos. It is clear that there is a difference in mobility of the protein isolated from each sample. The xlgv7 protein isolated from oocytes consists of two major species, whereas that of brain consists of a single major component. In the early embryo there are two major isotypes that migrate more slowly than either the oocyte or the brain isotypes. A third isotype in the embryo appears to comigrate with the more slowly migrating isotype in the oocyte sample. The brain isotype migrates differently than those from either oocytes or early embryos. These proteins are related to one another because chymotryptic digestion followed by Western blotting showed a similar sized internal antigenic fragment [data not shown]. We have not yet tested the proteins from tissues in which the gene is transiently expressed [i.e., gut].

Protein xlgv7 binds to DNA cellulose

Some cellular proteins have the capacity to bind specifically to DNA and to regulate processes such as replication and transcription. As an initial step to examine the possible function of the nuclear protein xlgv7, we attempted to determine the extent of binding of this protein from crude GV extracts to DNA cellulose. This assay has been used to establish the DNA-binding nature of proteins [Alberts et al. 1968] and as a step to purify transcription factors [Rosenberg and Kelly 1986]. Figure 5 shows that protein xlgv7 binds to a column containing double-stranded DNA cellulose in a low-salt buffer [50 mM NaCl in buffer A]. The protein was eluted from the column with buffer A containing 0.3 M NaCl, 0.5 M NaCl. In this experiment, ~95–98% of the bulk protein was in the flowthrough [data not shown]. We estimate that 70–80% of xlgv7 protein was bound to the DNA cellulose column. This result suggests that protein xlgv7 is a DNA-binding protein. The flowthrough fraction did not bind to the column upon repassage.

xlgv7 is expressed at the neurula stage and in adult brain

The GV protein xlgv7 is first detectable during oogenesis in X. laevis and is accumulated in the GV. After oocyte maturation, the protein is dispersed into the cytoplasm
and remains there during early development through the gastrula stage. At this time, protein xlgv7 reenters all nuclei of the embryo [Dreyer et al. 1982]. It is detected in all nuclei until late in development (stages 51 and onward) when it becomes localized in the nuclei of specific regions of the tadpole. Figure 6 shows immunohistochemical staining of different regions of stage 51 and 55 tadpoles with mAb 37-1A9, which recognizes GV protein xlgv7. Clearly, the nuclei in both the secretory epithelial cells of the stomach and gut, as well as the nuclei of the neural tube and spinal ganglia stain positive, whereas the nuclei of the pharyngobranchial cells show no detectable staining. The lack of staining with the antibody in the pharyngobranchial cells in the section shown in Figure 6E is not due to an artifact because these sections can be counterstained with a different antibody [see Dreyer et al. 1982]. The presence of protein xlgv7 in the stomach and gut is a transient phenomenon. In adult frogs, the antigen is highly enriched in the cells of the CNS with some staining of the epidermis and lymphocytes (Wedlich and Dreyer 1988).

Figure 7A shows a Northern blot containing RNAs from embryos and tadpoles from stage 3–48 of X. laevis development probed with xlgv7. A single RNA species,
Figure 7. Northern blot analysis of xlgv7 mRNA during development in X. laevis and X. borealis. RNA was extracted from different stage embryos, separated by gel electrophoresis under denaturing conditions, and blotted to GeneScreen (see Materials and methods). (A) RNA from different staged X. laevis and X. borealis embryos. The numbers refer to the developmental stage (Nieuwkoop and Faber 1967); the arrow indicates 2.4-kb xlgv7 mRNA in both X. laevis and X. borealis. (B) A graphic representation of A, representing a compilation of data from several experiments. The statistical analysis (t-test) was based on a subset of that data, as all developmental stages were not represented in each experiment. The decrease in RNA titer from stage 3 to 10 is significant with a confidence level of \( p < 0.05 \). From stage 10 to 18, the increase is significant with >95% confidence. The differences between stages 32 and 34/35 are not statistically significant \( p > 0.1 \).

− 2.4 kb in length, was detected, and the size of the transcript appeared the same throughout development (Fig. 7A). The data are plotted in Figure 7B. The points are based on densitometric scans of autoradiographs from several experiments and are the abundance relative to stage 12. During early cleavage stages, the xlgv7 mRNA declines ~5- to 10-fold by the gastrula stage of development. The first marked accumulation of xlgv7 mRNA occurred after the presumptive mesoderm had nearly completed involution and neural induction had begun. At the neural fold/neural closure stage (stage 15–16; Nieuwkoop and Faber 1967), there was a 7- to 10-fold increase in the amount of mRNA than was present during blastopore closure at the end of gastrulation (stage 12–13). Following neurulation, the abundance of xlgv7 mRNA declines throughout the remainder of development. The small increase in xlgv7 mRNA at stages 32–34 is not statistically significant based on analysis with the t-test. Reprobing of this blot with a probe for \( \text{Xenopus H4} \) showed normal abundance levels of this mRNA, suggesting that changes detected for the xlgv7 mRNA are not due to differences in RNA loading, etc. (data not shown).

Also, we analyzed the expression of xlgv7 in a closely related species, \( \text{Xenopus borealis} \). The \( \text{X. laevis} \) probe recognized an RNA species of similar size (~2.4 kb) to the \( \text{X. laevis} \) mRNA on Northern blots of \( \text{X. borealis} \) RNA. Figure 7A shows that there is a similar increase of the xlgv7 mRNA at the time of neurulation in developing \( \text{X. borealis} \) embryos.

Figure 8 shows the distribution of the xlgv7 mRNA in various adult tissues. Equivalent amounts of RNA were loaded onto each lane. The xlgv7 mRNA was most abundant in the brain. In some instances, low levels were detected in the gut (data not shown). There was no signal detected in either muscle or skin. The presence of a low level of xlgv7 mRNA in the adult gut is consistent with the transient expression in the gut of the tadpole. Also shown in Figure 8 is RNA from total ovary, which was separated into poly(A)+ and poly(A)− fractions. The xlgv7 mRNA was detected only in the poly(A)+ fraction.

It was of interest to determine the similarity or difference between the protein product present in the oocyte and the protein found in the nuclei of later developmental stages (e.g., stage 24). The protein isolated from oocytes and stage 24 embryos appeared to be of similar size, as determined by gel electrophoresis and Western blot data (Fig. 2B; Dreyer and Hausen 1983). We extended these studies by comparing the presence of the diagnostic chymotryptic digestion fragment in protein
Here, we report the cloning of a cDNA that encodes a maternal gene product exhibiting a unique pattern of nuclei in all regions of the embryo at the late blastula-gastrula stage. Because the amount of synthesis of the xlgv7 protein is relatively low during early embryogenesis (Dreyer et al. 1983), it is probable that the protein entering the early embryonic nuclei originates from the maternal pool. At the late tadpole stage, it becomes predominantly localized in nuclei of the CNS. Also, it is detected transiently in other regions such as secretory cells in the gut and stomach.

Northern blot analysis of xlgv7 transcripts reveals a single RNA species ~2.4 kb in size that exhibits modulations during development similar to those seen for the protein (Dreyer et al. 1983). It is probable that the protein entering the early embryonic nuclei originates from the maternal pool. At the late tadpole stage, it becomes predominantly localized in nuclei of the CNS. Also, it is detected transiently in other regions such as secretory cells in the gut and stomach.

Discussion

Here, we report the cloning of a cDNA that encodes a maternal gene product exhibiting a unique pattern of temporal and spatial distribution later during development of the frog, *X. laevis*. Following oocyte maturation, the xlgv7 protein is cytoplasmic, reentering the somatic nuclei in all regions of the embryo at the late blastula-gastrula stage. Because the amount of synthesis of the xlgv7 protein is relatively low during early embryogenesis (Dreyer et al. 1983), it is probable that the protein entering the early embryonic nuclei originates from the maternal pool. At the late tadpole stage, it becomes predominantly localized in nuclei of the CNS. Also, it is detected transiently in other regions such as secretory cells in the gut and stomach.

Northern blot analysis of xlgv7 transcripts reveals a single RNA species ~2.4 kb in size that exhibits modulations during development similar to those seen for the protein (Dreyer et al. 1982). The xlgv7 mRNA decreases in abundance from fertilization through the late gastrula stage of development. It increases in abundance during the period of neural induction. In the adult, the transcripts for xlgv7 are highly enriched in the brain and are absent or at very low abundance levels in other somatic tissues. It is likely that the protein localized in the adult brain is newly synthesized because there is a high abundance of xlgv7 mRNA in this tissue.

The presence of a single band on Northern blots suggests that there is a single mRNA, even though the xlgv7 genes may be present in multiple copies. It is possible that the mRNAs detected at different developmental stages or in different tissues are not identical and may differ in sequence, even though they migrate to a similar position on gels under denaturing conditions. The determination of this possibility will have to await the characterization of the genomic clones and the production of gene-specific probes.

The sequence of the xlgv7 cDNA does not have a significant match in the Genbank, NBRL, and EMBL databases. There are several salient structural features of the conceptual protein, which include the presence of a putative nuclear targeting signal (Glu Arg Arg Lys Lys Thr), very similar to the nuclear targeting signal observed on SV40 T antigen and various other *Xenopus* nuclear proteins (Table 1). The xlgv7 protein does not reenter the embryonic nuclei until the gastrula stage of development, even though other nucleoproteins present in the same cells reenter during early cleavage. It is possible that the nature and position of the targeting signal may be important in the timing of reentry into embryonic nuclei.

Another interesting feature is a region containing an unusual arrangement of His residues. Several of the His pairs contain a conserved sequence or a close variation of the sequence Glu Lys Asn Gln, located between the His residues. Presently, we do not know the function of this domain. An intriguing possibility is that this region may be involved in the interaction of this protein with the DNA or its association with another protein.

We have also detected a homology at the amino acid level to nucleotide-binding sequences that are present in a number of proteins binding and hydrolyzing ATP, such as the RAD18 and RAD3 proteins from yeast (Jones et al. 1988) and the *Escherichia coli* proteins UvrA (Hussain et al. 1986), UvrD (Finch and Emmerson 1984), RecA (Walker et al. 1982), and RecB (Finch et al. 1986). Many of these proteins are involved in DNA repair, helicase activity, or recombination.

Although we do not know the function of the xlgv7 gene product, we assume that it must have a nuclear function because it is found localized in the nucleus. Protein xlgv7 appears to be a DNA-binding protein based on its ability to bind to DNA cellulose and its appearance on loops of lampbrush chromosomes (C. Dreyer, pers. comm.). The xlgv7 protein could function as a structural protein in chromatin, as a transcription factor, or in RNA processing. Also, it is possible that the products of the different xlgv7 gene family members may function in different capacities at various times during development or that the proteins do not function until a specific time and in a specific cell type. The high abundance of the protein and mRNA in the CNS of both the tadpole and adult suggest that the xlgv7 may play an important role in the development and maintenance of the CNS.
We have detected at least three different isotypes of protein xlgv7, which appear during early embryogenesis and in the adult brain. The shift detected from the oocyte stage to the early cleavage stage embryo is due to phosphorylation because alkaline phosphatase treatment reduces the isotypes into one species, and $^{32}$P is incorporated into the protein during maturation (M. Miller and L. Etkin, in prep.). Phosphorylation is a common modification occurring during oocyte maturation in Xenopus. Perhaps the state of modification of the protein is related, in some way, to its function. It is known that phosphorylation affects the DNA-binding ability and biological activity of numerous cellular components such as maturation promoting factor (Lohka et al. 1988), and transcription factors such as the heat shock element-binding protein and cAMP response element (CRE)-binding protein (Sorger et al. 1987; Yamamoto et al. 1988). The protein from the brain appears to migrate differently from either the oocyte isotypes or the early embryonic isotypes. We are currently attempting to delineate the precise relationship between the different isotypes and the xlgv7 cDNA.

It is interesting that the gene is expressed transiently in a variety of cell and tissue types during development. The presence of the protein in secretory epithelial cells of the gut, the CNS, and the female germ line does not correspond to any common embryological or cell lineage derivation. It is possible that the presence of the antigen in specific cell types may be indicative of the state of differentiation of these cells or the necessity for these cells to carry out a common cellular function. The transient pattern of expression and localization of xlgv7 is reminiscent of the expression of several of the homeo box genes such as fushi tarazu (ftz) and even-skipped (eve) in Drosophila (Doe et al. 1988a,b). The ftz and eve genes are segmentation genes that are expressed in striped patterns during early embryogenesis at the blastoderm stage. Both genes encode nuclear proteins. Their role at the blastoderm stage is in the establishment of the normal pattern of segmentation. Recently, it has been found that both the ftz and eve genes are expressed transiently a second time in different but overlapping sets of neuronal precursor cells in the developing CNS. Analysis of mutants suggests that the products of these genes also may function in determination of neuronal cells (Doe et al. 1988a,b). Another example of a diverse pattern of expression of a homeo box gene is observed in the HHO.c13 gene of the human (Mavilio et al. 1986). No Northern blots of human RNA probed with this gene show a pattern of three different size transcripts, which are expressed differentially in the spinal cord, brain, backbone rudiments, limb buds, and the heart throughout embryogenesis. This suggests that this gene may exert a wide spectrum of functions in the nervous system, as well as other organs or body parts.

Although xlgv7 is not a homeo box gene and may not be involved in pattern formation it is quite possible that it may have multiple functions during development, one of which may be associated with the developing CNS. We anticipate that the elucidation of the function of the xlgv7 gene product(s) will help us to understand the role of maternal information in the regulation of development and cell differentiation.

### Materials and methods

#### Animals

Embryos were obtained from spawnings of adult *X. laevis* and were reared in charcoal-filtered water (Gurdon 1967). For biochemical analyses of protein, RNA, or DNA, embryos were de-jellied with 2.5% cysteine HCl in 0.6 mM Tris (pH 8.0) and washed extensively with modified Barth’s solution (MBS, Etkin et al. 1984).

#### Protein blotting

Germinal vesicles were manually isolated in MBS (4°C), homogenized, boiled for 3 min in 1.0% SDS in MBS, and stored at –20°C. Proteins were separated by electrophoresis on 12.5% polyacrylamide gels according to Dreyfuss et al. (1984) and transferred to 0.45-µm nitrocellulose (Schleicher and Schuell), using the method of Towbin et al. (1979). Blots were washed 10-15 min in TBS (10 mM Tris-HCl [pH 8.0], 150 mM NaCl) with 0.05% Tween 20 and blocked for 1 hr with 20% calf serum, 0.05% Tween 20 in TBS. Blots were washed twice for 5 min in TBS–Tween 20 and 5 min in TBS alone. Antibodies were bound at room temperature for 1 hr to the blot in 1 : 500 dilutions made in TBS with 0.1% BSA and washed as above. Sequentially, blots were incubated with the mouse mAb 37-1A9 (against GV protein 7), goat antimouse IgG (Cappel), and mouse alkaline phosphatase antialkaline phosphatase [APAAP] (Boehringer–Mannheim). Following a final wash, blots were soaked for 5 min in AP buffer [0.1 M Tris-HCl (pH 9.5), 0.1M NaCl, and 5 mM MgCl$_{2}$] and reacted in 50 ml of AP buffer with 330 µl nitroblue tetrazolium [50 mg/ml in 100% DMF] and 165 µl 5-bromo-4-chloro-3-indoly phosphate [50 mg/ml in 100% DMF]. Reactions were stopped by dilution.

#### Nucleic acid isolation and Northern and Southern blotting

Total cytoplasmic RNA was extracted from different stage embryos and various frog tissues using either a guanidinium isothiocyanate/hot phenol procedure or phenol extractions (Etkin and Maxson 1980). All samples were treated with 1 unit of RQ1 DNase (Promega) per microgram of total nucleic acid in 40 mM Tris-HCl [pH 8.0], 0.5 mM MgCl$_{2}$, and 10 mM NaCl for 15 min, reextracted with phenol and chloroform/isoamyl alcohol (24 : 1 vol/vol), and precipitated with ethanol. intact RNA was glyoxylated (McMaster and Carmichael 1977), separated electrophoretically through 1.1% agarose gels, blotted to GeneScreen Plus nylon membranes (NEN), according to the manufacturer’s specifications, and subsequently hybridized overnight at 65°C according to Amasino (1986). Randomly primed $^{32}$P-labeled probes were synthesized from the EcoRI inserts of xlgv7 clone according to the method of Feinberg and Vogelstein (1984) and hybridized to each blot with $1 \times 10^{6}$ to $2 \times 10^{6}$ cpm/ml [sp. act. >$10^6$ cpm/µg].

Genomic DNA was isolated from *X. laevis* liver, which was frozen and pulverized under N$_{2}$ with mortar and pestle and resuspended in DNA extraction media [50 mM Tris-HCl (pH 8.0), 10 mM NaCl, and 10 mM EDTA] with 0.5% SDS and 100 µg/ml proteinase K. After a 3-hr incubation at 50°C, the slurry was extracted several times with phenol and dialyzed against the
extracted DNA. Samples were phenol/chloroform extracted, precipitated, and resuspended in 0.5% SDS for 1 hr and reextracted with an equal volume of phenol and chloroform. The remaining aqueous layer was dialyzed and stored at 4°C.

Genomic DNAs were digested with different restriction enzymes to completion, as judged by parallel digestions with λ DNA. Samples were phenol/chloroform extracted, precipitated, and resuspended with HCl, and blotted to GeneScreen Plus nylon membrane according to manufacturer's specifications (NEN). Hybridizations were performed as described above.

Cloning xlgv7

A λ gt11 cDNA library was constructed from RNA isolated from intact X. laevis ovary. The library, grown in Y1088 cells, had 2.0 × 10^9 independent recombinants and was screened as follows. Tittered phage were plated on Y1090 bacterial lawns, incubated at 42°C for 3 hr, and overlaid with nitrocellulose filters presoaked in 10 mM isopropyl-B-D-thiogalactopyranoside (IPTG). After 4 hr at 37°C, the filters were blocked with 20% horse serum in TBS and reacted successfully to the mouse mAb 37-1A9 and goat antismouse IgG conjugated either to alkaline phosphatase or horseradish peroxidase. From 1.0 × 10^9 plaques screened, three immunoreactive clones were identified and plaque purified for three cycles. Purified phage cDNA inserts subsequently were subcloned into the EcoRI site of the pGem-1 (Promega Biotec).

DNA sequencing

Alkali-denatured double-stranded pGem-1 plasmid subclones of xlgv7 cDNA served as a template, on which a dyeoxy sequencing procedure was carried out (Sanger et al. 1977). Both strands of the xlgv7 cDNA were sequenced independently using the GemSeq K/RT system and Sequenase system (Promega Biotec and USB, respectively).

Primer extension

Total RNA from manually defolliculated stage VI Xenopus oocytes (Dumont 1972) was prepared according to Etkin and Maxson (1980). A primer complementary to the 5' end of the coding strand of the xlgv7 cDNA (from -8 to +1) was labeled using T4 polynucleotide kinase (Pharmacia) and [γ-32P]ATP (3000 Ci/mmole, Amersham). The primer extension reaction was done according to McKnight et al. (1981). The products of reaction were run on an 8% polyacrylamide gel with 8 μM urea and exposed to XAR-5 Kodak film at −80°C.

Expression of xlgv7 in bacteria and peptide mapping

The transcription and translation of the xlgv7 cDNA were carried out using the T7 RNA polymerase/promoter system (gift from Stan Tabor), which consists of two plasmids: pT7-7 and pGP1-2 (Tabor and Richardson 1985). pT7-7 contains a T7 RNA polymerase promoter, ribosome binding site, and translation start codon located at the 5' end of a multiple cloning site, which permits cloning an insert into one of three different reading frames. Vector pGP1-2 is an M13 phage harboring the T7 RNA polymerase gene under the control of heat-inducible lac promoter.

The xlgv7 cDNA was recloned from the pGem vector into the EcoRI site of pT7-7. This places the xlgv7 cDNA under the control of the T7 polymerase promoter and in the proper reading frame relative to the ATG codon. K38 E. coli cells harboring the pGP1-2 plasmid were transformed with pT7-7-xlgv7 clone, and production of protein 7 was initiated by induction of T7 RNA polymerase lac promoter using a temperature shift from 30°C to 42°C. As a control, we have used K38 pGP1-2 cells transformed with pT7-7 vector alone, xlgv7 cDNA inserted into pGem vector, and xlgv7 cDNA inserted into pT7-7 vector, but in the opposite orientation. Two hours from the induction, bacterial cells expressing xlgv7 cDNA (from 1 ml culture) and control cells were harvested, and the pellet was dissolved in 120 μl of cracking buffer [60 mM Tris-[pH 6.8], 1% SDS with glycerol, 1% SDS with glycerol, and bromophenol blue]. One milliliter of cells yielded -250 μg of total proteins. Proteins were separated on SDS-polyacrylamide gels and detected on Western blots with the 37-1A9 antibody.

For analysis of chymotrypsin digestion products, total GV proteins isolated from manually dissected Xenopus GV's and/or proteins synthesized in the bacterial expression system were suspended in the buffer [60 mM Tris-HCl (pH 6.8), 1% SDS, 1% β-mercaptoethanol, 1% glycerol] at concentration 160-660 μg/ml, heated at 100°C for 2 min, and digested with chymotrypsin (type 1-S from bovine pancreas, Sigma) at a concentration of 4-2000 μg/ml for 30 min at 37°C. After digestion, samples were boiled for 2 min before loading on SDS-polyacrylamide gels. Separation was performed on 12.5-17% gels in different experiments. The markers consisted of prestained protein molecular weight standards [both high and low range, Sigma, St. Louis, Missouri and BRL, Bethesda, Maryland].

Histological staining of tadpoles

All sectioning and staining of Xenopus tadpoles was performed according to Hausen and Dreyer (1981).

DNA cellulose chromatography

Approximately 250 μg of protein extracted from GVs was loaded onto a 1-ml column containing double-stranded DNA cellulose (Pharmacia) in buffer A [15 mM Tris-HCl (pH 8.0), 15% glycerol, 1 mM DTT, 0.05% NP-40, 5 mM EDTA, 1 mM PMSF, and 50 mM NaCl]. The column was washed with buffer A, and the proteins were eluted with buffer A containing 0.3, 0.5, and 1.5 M NaCl. The fractions were collected and proteins separated on SDS-polyacrylamide gels and blotted to nitrocellulose. Protein 7 was detected on the Western blots by reaction with mAb 37-1A9, which is specific for protein 7.

Acknowledgments

This work was supported by a grant (DCB-860869) from the National Science Foundation to L.D.E. and a postdoctoral training fellowship (CA09299) from the National Institutes of Health to M.M. We acknowledge the critical comments of Dr. Andres Carrasco.

References


Miller et al.


xlgv7: a maternal gene product localized in nuclei of the central nervous system in Xenopus laevis.


Genes Dev. 1989, 3:
Access the most recent version at doi:10.1101/gad.3.4.572