The EMBO Journal vol.7 no.10 pp.3199 - 3209, 1988

Differential antero-posterior expression of two proteins encoded by a homeobox gene in *Xenopus* and mouse embryos

Guillermo Oliver, Christopher V.E.Wright, Jane Hardwicke and Eddy M.De Robertis

Department of Biological Chemistry, University of California School of Medicine, Los Angeles, CA 90024, USA

Communicated by E.M.De Robertis

The *X.laevis* XlHbox 1 gene uses two functional promoters to produce a short and a long protein, both containing the same homeodomain. In this report we use specific antibodies to localize both proteins in frog embryos. The antibodies also recognize the homologous proteins in mouse embryos. In both mammalian and amphibian embryos, expression of the long protein starts more posteriorly than that of the short protein. This difference in spatial expression applies to the nervous system, the segmented mesoderm and the internal organs. This suggests that each promoter from this gene has precisely restricted regions of expression along the anterior-posterior axis of the embryo. Because the long and short proteins share a common DNA-binding specificity but differ by an 82 amino acid domain, their differential distribution may have distinct developmental consequences.

Key words: embryogenesis/homeodomain/immunolocalization/ two promoters-two proteins/vertebrate development

Introduction

The best argument in favor of homeobox gene products specifying positional information in vertebrates, as they do in *Drosophila*, comes from studying the distribution of their transcripts in embryos. A number of homeobox mRNAs are expressed in restricted regions along the antero-posterior axis of mouse (Awgulewitsch *et al.*, 1986; Utset *et al.*, 1987; Gaunt *et al.*, 1986; Krumlauf *et al.*, 1987; Dony and Gruss, 1987; Toth *et al.*, 1987; Gaunt, 1987; Holland and Hogan, 1988; Breier *et al.*, 1988) and frog (Carrasco and Malacinski, 1987; Sharpe *et al.*, 1987; Condie and Harland, 1987) embryos. Antibody staining has been used to show that vertebrate homeodomain proteins are abundant nuclear proteins in some cultured cell lines and in adult tissues (Odenwald *et al.*, 1987; Kessel *et al.*, 1987).

XlHbox 1 (*Xenopus laevis* Homeobox gene 1) was the first homeobox gene to be isolated from vertebrate DNA (Carrasco *et al.*, 1984). In a previous paper (Cho *et al.*, 1988) we showed that XlHbox 1 produces two transcripts during early development which arise from the utilization of two promoters that are 9 kb apart in the genome. Use of promoter I (PR I) leads to a 2.2 kb transcript encoding a short open reading frame (ORF), while promoter II (PR II) gives rise to a 1.8 kb transcript with a long ORF. Both proteins share the same homeobox and differ only by an 82 amino acid extension present at the amino terminus of the long ORF.

In this paper we use anti-XlHbox 1 antibodies specific for various protein domains to show that both ORFs predicted by sequencing studies are indeed translated in embryos. Immunolocalization of both proteins in *Xenopus* as well as in mouse embryos was carried out. Mouse embryos were tested because the strong amino acid sequence conservation between homologous human and *Xenopus* cDNA clones indicated both that the antisera would cross-react between species (Simeone *et al.*, 1987; Cho *et al.*, 1988) and that mammals have a similar two-promoter/two-protein gene organization (Cho *et al.*, 1988). The homologous murine proteins proved to be easily detectable, and since mouse embryos have a greater degree of mesodermic differentiation in the prevertebral column of the early embryo, we found mice to be more suitable than *Xenopus* for the analysis of XlHbox 1 protein distribution in the mesoderm.

The XlHbox 1 proteins were localized in three main regions of the mouse embryo: the central nervous system (CNS), the segmented mesoderm and the visceral mesoderm of the internal organs. In all three areas the short protein (PR I type) region of expression starts more anteriorly than that of the long protein (PR II type). A similar distribution of the long protein relative to the short protein is found in *Xenopus* embryos. This suggests that the expression of each protein, which may have a distinct developmental role, is under very precise spatial regulation along the antero-posterior axis of the embryo. The antibodies permit the study of the homologous proteins in frogs and mice, and presumably in other chordates. Comparative studies on the expression of XlHbox 1 proteins in these embryos suggest that subtle differences may exist in the way the amphibian and the mammalian body plans are established.

Results

Antibodies specific for XlHbox 1

Figure 1 shows the fusion protein constructs used to raise rabbit polyclonal antibodies. The two non-overlapping fusion protein constructs A and B together encompass the entire ORF encoding the long protein (see Figure 3 of Cho *et al.*, 1988) except for its initiator codon. Antibody A was raised against, and purified with, a fusion protein containing the homeodomain and the carboxy-terminal 34 amino acids. Antibody B was raised against a fusion protein containing the amino-terminal 142 amino acids of the long XlHbox 1 ORF (Figure 1). A third specificity, called antibody C, was prepared by affinity-purifying antiserum B over a column of fusion protein C, which contains the first 73 amino acids of the extended ORF. Antibody C is therefore specific for the long version of XlHbox 1 protein, while both antibody A and B recognize regions common to the short and long versions. An antibody specific only for the short protein cannot be made because all of its amino acid sequence is contained within the long protein.

Figure 2, panel A, shows a 13-day-old mouse embryo

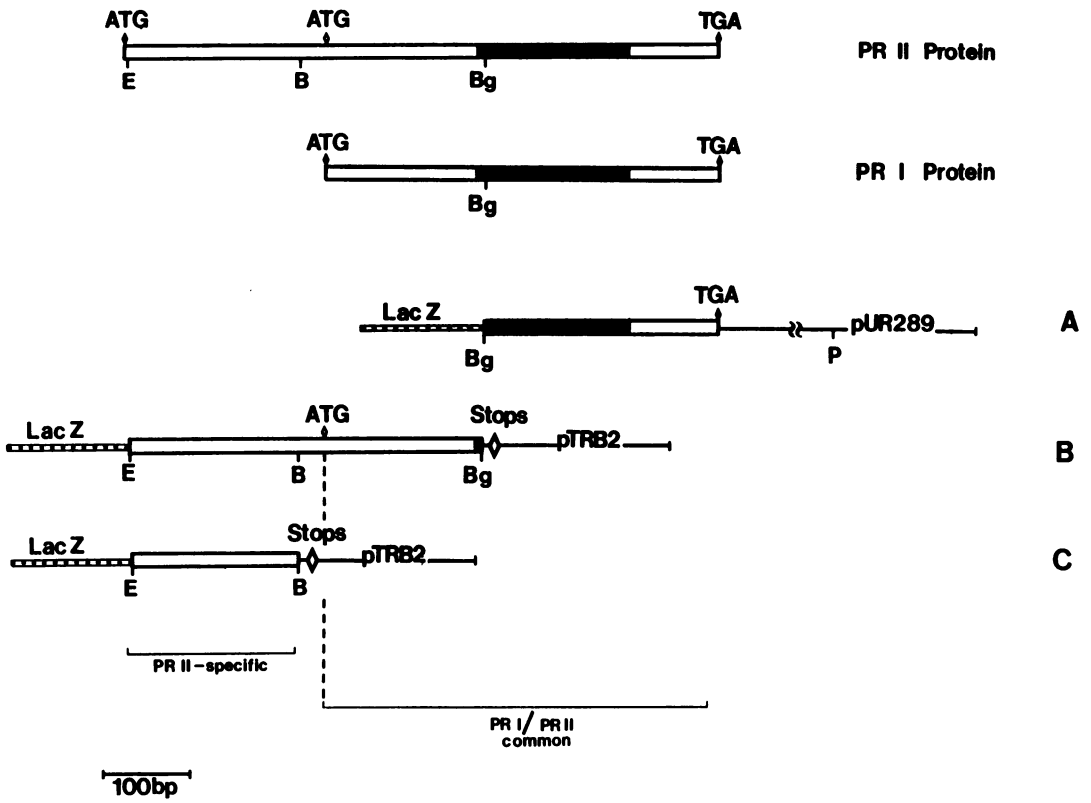

Fig. 1. Schematic representation of PR I- and PR II-specific proteins and of the fusion protein constructs. The structure of the XlHbox 1 proteins derived from promoter II mRNA (long ORF) and promoter I mRNA (short ORF), reported in Cho *et al.* (1988), are shown at the top of the figure. Protein-coding regions are indicated by the open boxes, and the homeodomain by the black box. Small filled diamonds labeled ATG and TGA indicate the natural start and stop codons, while the similarly marked small unfilled diamond is the second Met codon in PR II mRNA. Below this, lines A and B represent the two fusion protein constructs used to raise corresponding antibodies A and B, which detect both short and long XlHbox 1 proteins. Line C depicts the fusion protein used to affinity-purify long protein-specific antibodies from antiserum B. The hatched line represents the carboxy terminus of β-galactosidase. pUR289 and pTRB2 represent the fusion protein vectors (see Materials and methods). Large unfilled diamonds ('STOPS') indicate tandem vector translation termination codons. Restriction sites are E: *Eco*RI; B: *Bam*HI; Bg: *Bgl*II and P: *Pst*I.

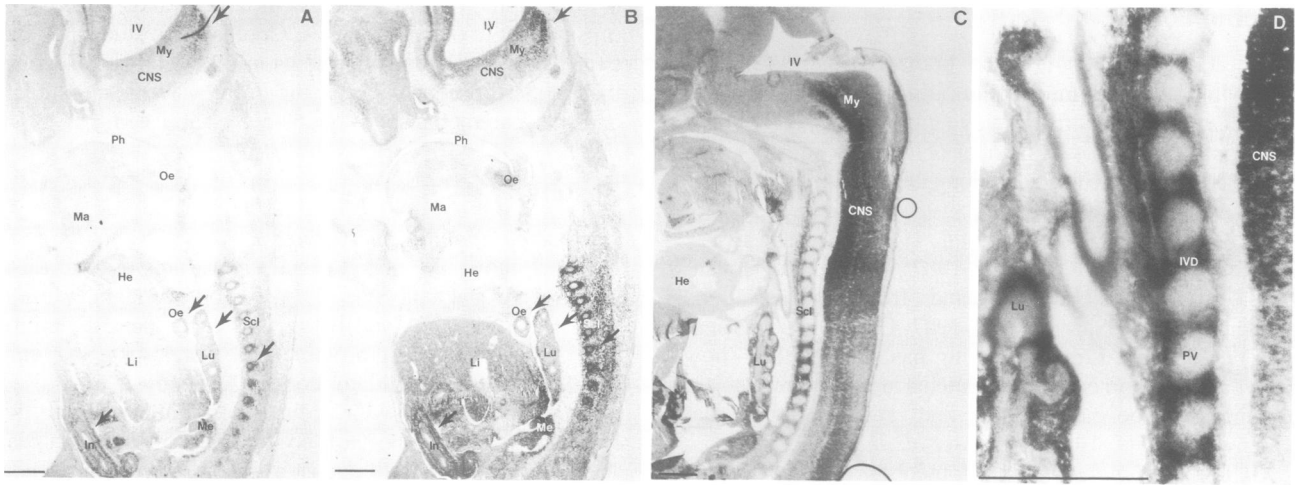

Fig. 2. Detection of XlHbox 1 proteins in mouse embryos by anti-*Xenopus* and anti-human antibodies. **Panels A** and **B** depict adjacent parasigittal sections of a 13-day embryo stained with antibody A and antibody B, respectively. Arrows point out regions immunopositive with both antibodies, specifically the myelencephalon (My), sclerotomes (Scl), lung (Lu), oesophagus (Oe), and intestine (In). **Panels C** and **D** show another day-13 mouse embryo stained with an XlHbox 1 antibody derived from a human antigen. For orientation within the embryo, other organs are also labeled in abbreviated form: CNS, central nervous system; IV, fourth brain ventricle; He, heart; Li, liver; Ph, pharynx; Ma, mandible; Me, mesonephric tubules; IVD, intervertebral disc; PV, prevertebrae. The scale bar represents 0.4 mm.

stained with antibody A. Staining was found in the nuclei of cells in the CNS, starting in the hindbrain (myelencephalon), continuing over the cervical spinal cord (this region is not present in this section), and gradually disappearing in the thoracic spinal cord (see Figure 3A). As discussed later, the protein distribution in *Xenopus* embryos

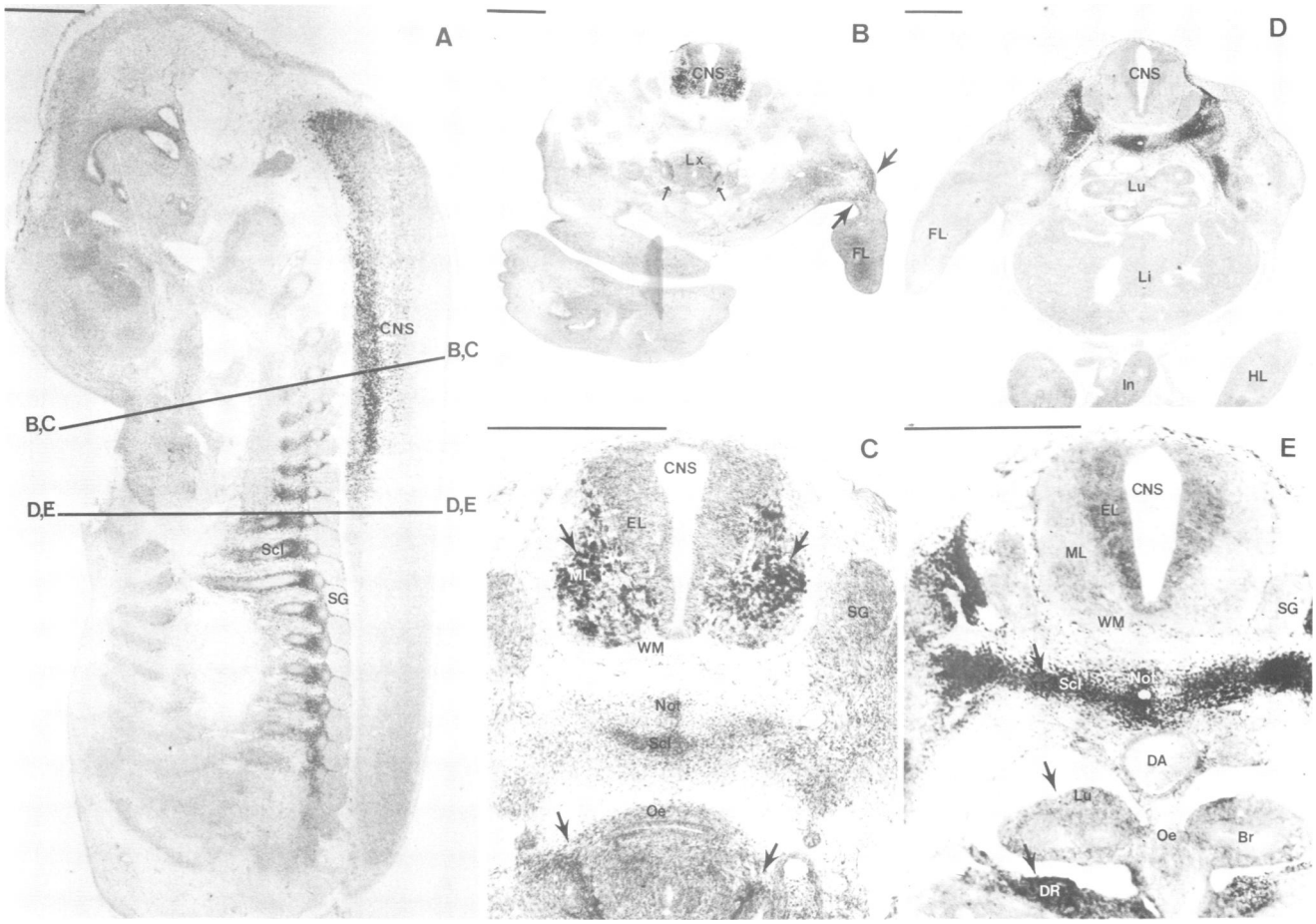

Fig. 3. Distribution of XlHbox 1 proteins in the day-12 mouse embryo. **Panel A** shows a parasagittal section of a day-12 embryo stained with antibody B, which detects both short and long forms of XlHbox 1 protein. Positive nuclei are seen in the CNS. Additional staining is seen in some mesodermally derived cells such as the sclerotomal derivatives that will give rise to vertebrae and ribs (Scl) and mesenchymal cells sheathing the spinal ganglia (SG). **Panels B** and **C** show two magnifications of a transverse section taken along the plane B,C−B,C shown in panel A. Immunopositive cells (arrowed) are seen in the mantle layer (ML) of the spinal cord and in cells surrounding the digestive tract at the level of the larynx (Lx) and the oesophagus (Oe). Note that the cells in the border of the growing forelimb bud (FL; indicated by large arrows on panel B) are also stained. **Panels D** and **E** show two magnifications of a section taken in the D,E−D,E plane of panel A. In this region of the mouse embryo, the CNS is no longer stained, while the sclerotomes, lung (Lu), oesophagus, diaphragmatic region (DR) and intestine (In) are stained. Other abbreviations are: EL, ependymal layer; WM, white matter; Not, notochord; Li, liver; HL, hindlimb; Br, bronchi; DA, dorsal aorta. The scale bar represents 0.4 mm.

is very similar, in agreement with *in situ* hybridization studies showing XlHbox 1 transcripts mostly in the anterior spinal cord of 3-day-old *Xenopus* embryos (Carrasco and Malacinski, 1987). Staining was exclusively nuclear (coincident with Hoechst 33258 staining) for all the antibodies used in this study. In some tissues background staining was observed (see for example liver in Figure 2), but was readily distinguishable from 'positive' cells at higher magnification because this staining was not nuclear.

Because fusion protein A contains 54 out of the 60 amino acids in the homeodomain, it was necessary to show that this reflected XlHbox 1-specific staining and not cross-reaction with otherwise unrelated homeodomain proteins, although antibody A stained the predicted region of XlHbox 1 expression. Accordingly, we immunostained adjacent parasagittal mouse or *Xenopus* sections with either antibody A or antibody B, which are directed against two non-overlapping portions of the XlHbox 1 protein. As can be seen from Figure 2 panels A and B, both antibodies bind to the same regions, including some areas outside of the CNS (sclerotomes, oesophagus, lung and intestine; arrowed in Figure 2). A number of competition experiments were also performed. The antibodies were incubated in the presence of the corresponding fusion protein during the immunocytochemical reaction and, as expected, staining was abolished. When fusion proteins containing different frog (or human) homeodomains or the conserved hexapeptide region found in many homeobox genes (Fritz and De Robertis, 1988; Fibi *et al.*, 1988) were used, no competition was detected (data not shown). Furthermore, to support the specificity of the staining, another fusion protein antibody was raised against a partial cDNA clone carrying the last 106 amino acids of the human homolog of the XlHbox 1 gene. This clone was isolated from a neuroblastoma cDNA library and is identical in sequence to the previously described HHO.c8 (Simeone *et al.*, 1987). Immunolocalization with this antibody (Figure 2C and 2D) confirmed the protein localization found with the anti-*Xenopus* antibodies A and B. In addition, antibodies (made in our laboratory) against different homeodomain-containing proteins (XlHbox 2 and XlHbox 6) immunostain entirely different regions of the embryo (unpublished observations). We therefore conclude that these

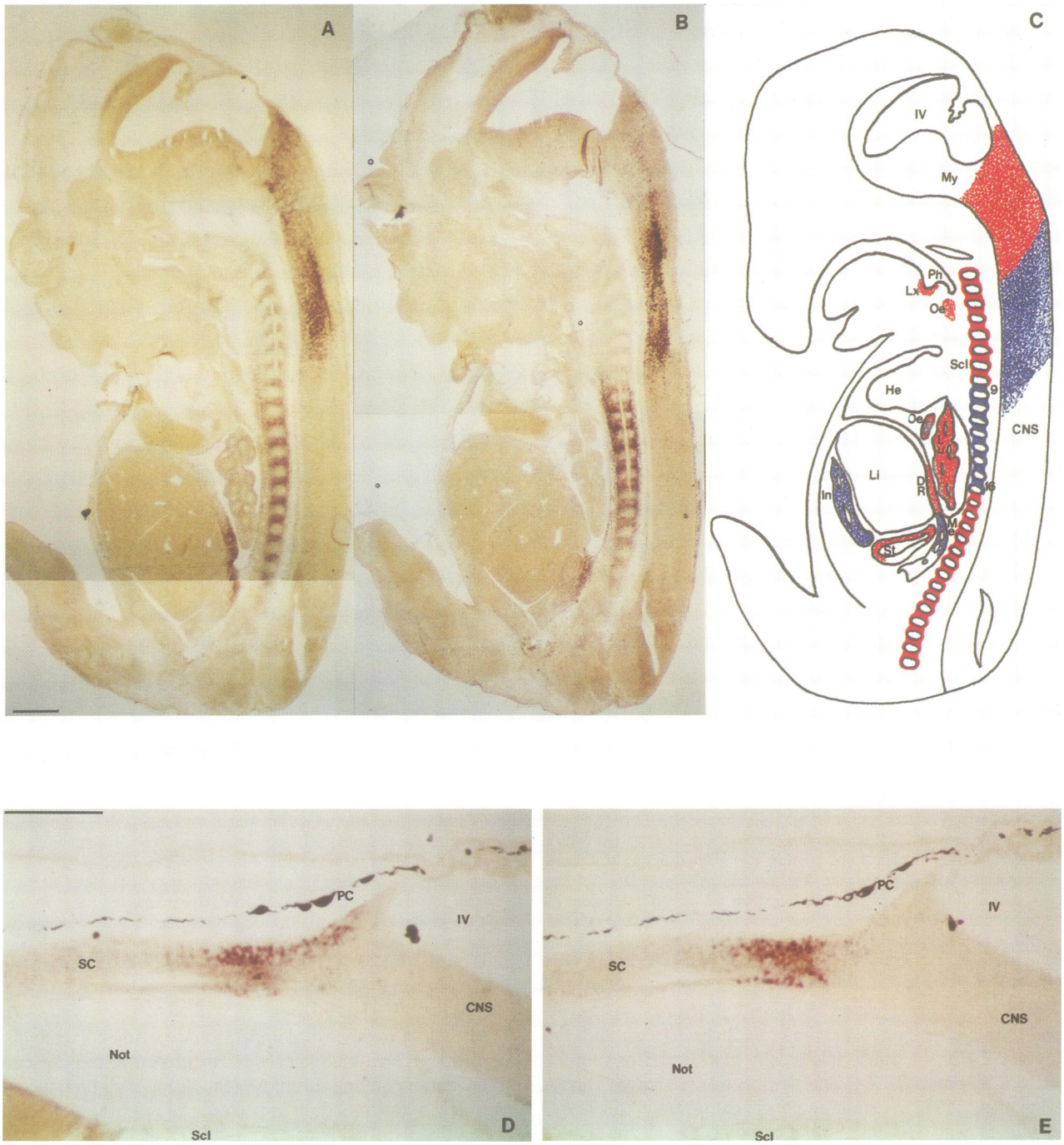

Fig. 4. Differential localization of the long and short XlHbox 1 proteins in mouse and frog embryos. **Panels A** and **B** are adjacent parasagittal sections of a day-13 mouse embryo stained with antibody B and C, respectively, showing the main differences in the localization of the short and long version of XlHbox 1 protein. **Panel C** presents an idealized day-13 mouse embryo, summarizing the different regions of expression of the long and short proteins: red indicates areas expressing only the short protein (positive only with antibody B) and blue those regions reacting with the long protein-specific antibody (antibody C) as well as with the 'common' antibody B. Antibody B (panels A and C) reacts with nuclei in the myelencephalon and cervical spinal cord. All sclerotomal prevertebrae, including the most anterior, are stained. The lung, the diaphragmatic region, mesonephric tubules and anterior regions of the digestive tract (pharynx and oesophagus) are also stained. Antibody C-positive nuclei (panels B and C) are restricted to the more positive regions: in the CNS the myelencephalon is negative but the cervical spinal cord is stained. Only sclerotomal prevertebrae 9–16 are stained, and the only internal organ that binds antibody C is the degenerating mesonephric region (in other sections the intestine is also stained). **Panels D** and **E** show that a similar differential distribution (although only sagittal sections through the central nervous system are shown here) occurs in stage 45 *Xenopus* embryos. Antibody B reacts with nuclei in both myelencephalon and cervical spinal cord (panel D), whereas antibody C only reacts in cervical spinal cord (panel E). Abbreviations used: CNS, central nervous system; Scl, sclerotomes; Lu, lung; DR, diaphragmatic region; Me, mesonephric tubule region; Ph, pharynx; Oe, oesophagus; Li, liver; In, intestine; St, stomach; Lx, larynx; My, myelencephalon; IV, fourth ventricle; PC, pigment cells; SC, spinal cord; Not, notochord. The scale bar in Panel A represents 0.4 mm while in Panel D it represents 0.1 mm.

antibodies are specific for XlHbox 1 antigens. Attempts to detect XlHbox 1 antigens on Western blots of *Xenopus* and mouse embryos were unsuccessful, perhaps due to the low amounts of these proteins in whole embryo extracts.

Concerning the nomenclature of the mouse gene, we realize that since the mouse homolog of the frog XlHbox 1

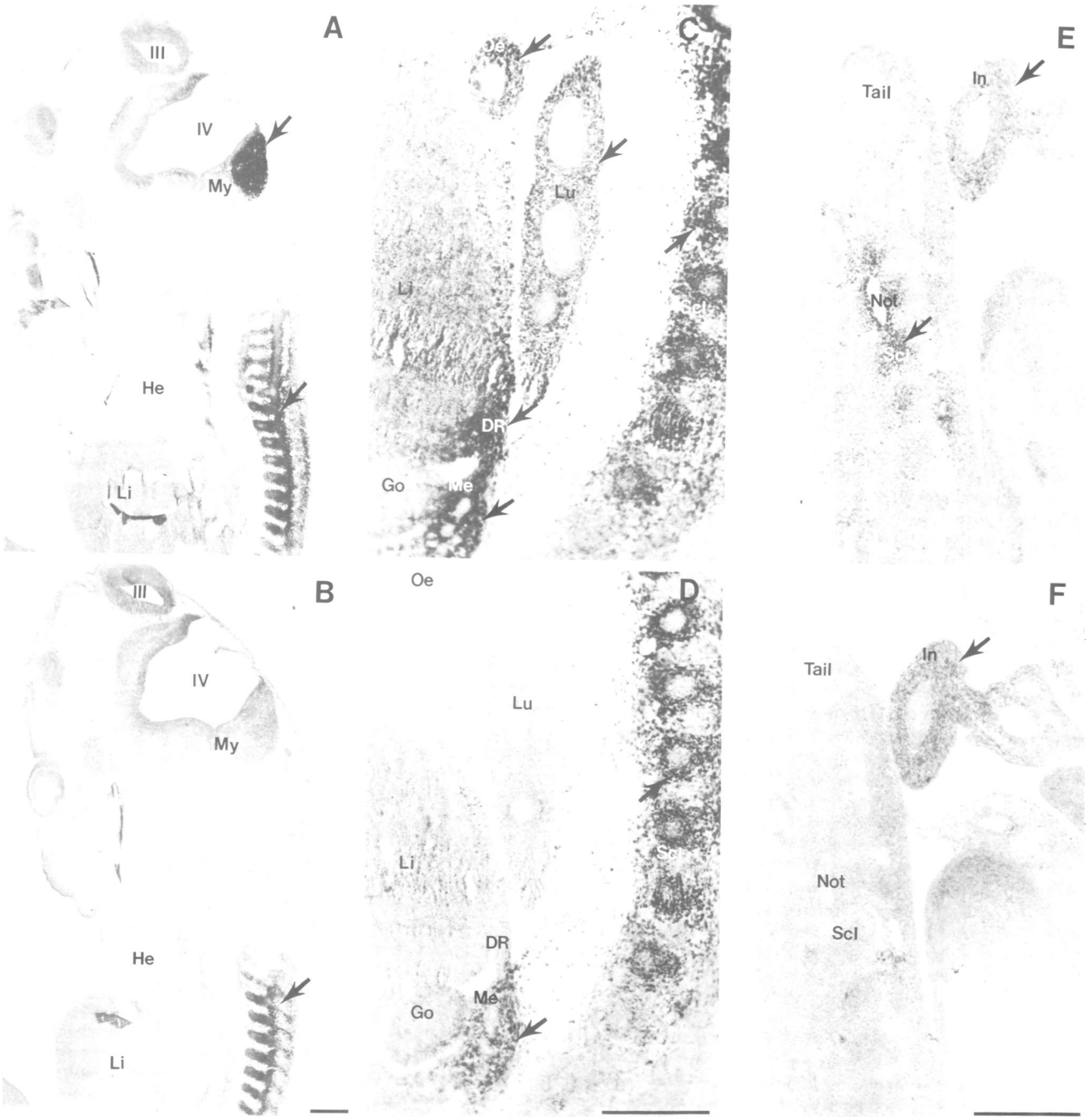

Fig. 5. Distribution of the long and short XlHbox 1 proteins in internal organs of day-13 mouse embryos. **Panels A, C** and **E** are immunolocalizations with antibody B (which is specific for the long protein). Comparison of panels A and B shows that antibody C does not stain the myelencephalon (My) and the anterior sclerotomes (Scl), which are stained by antibody B. The cervical spinal cord is absent in these particular preparations because of the plane of section. Panel C represents a higher magnification of the lung region of the embryo: antibody B stains lung (Lu), sclerotomes (Scl), oesophagus (Oe), diaphragmatic region (DR) and mesonephric tubule region (Me). In contrast, only the sclerotomes and mesonephric tubule region are stained with antibody C (panel D). Panel E shows that antibody B stains the intestinal mesenchyme and most posterior sclerotomes of the tail. Conversely, panel F shows that antibody C only stains the intestine. Note that in the lungs and digestive tract the mesenchymal cells, but not the epithelium, are stained. The liver sometimes displays a variable degree of cytoplasmic background staining (panel C). Arrows indicate immunopositive staining. Additional abbreviations are: III, third brain ventricle; IV, fourth brain ventricle; He, heart; Li, liver; Go, gonad; Not, notochord. The scale bar represents 0.4 mm.

has not been isolated or named, it would be strictly correct to call this antigen 'mouse homolog of frog XlHbox 1'. However, for brevity we will sometimes refer to it in this paper as XlHbox 1 antigen, whether or not we are dealing with frog or mouse embryos. The human homolog has been mapped to a region of chromosome 12 that is close to the Hox-3 gene complex (Cannizzaro *et al.*, 1987). Once a chromosomal walk has confirmed that this gene is indeed part of the Hox-3 complex, a definitive numbering according to the accepted nomenclature (Martin *et al.*, 1987) will be possible, which will supercede our present designation.

Localization of antigens common to the long and short ORF

We first analyzed the distribution of antigens which are common to the short and long XlHbox 1 homeodomain-

containing proteins. Figures 2 and 3 illustrate the expression of XlHbox 1 antigen in day-13 and day-12 mouse embryos. Using both antibodies A and B, the protein was determined to be exclusively nuclear (coincident with Hoechst 33258 staining), and localized in the following regions:

The nervous system. Expression starts in the myelencephalon, continues over the cervical spinal cord and disappears gradually over the upper thoracic level (Figures 2C and 3A). Transverse sections show that staining is restricted to nuclei of the mantle layer (arrows in Figure 3C), which is formed by postmitotic neuroblasts that will differentiate into neurons and glia (Gilbert, 1985). XlHbox 1 protein is still present in restricted parts of the CNS in 18-day-old mouse embryos and in adult mice (not shown). From immunolocalization experiments using the well-characterized adult cat spinal cord we conclude that these XlHbox 1 antigens are present in nuclei of motor, intermediary and sensory neurons but not glial cells (data not shown).

In mouse, the dorsal root spinal ganglia of the cervical region were usually negative (labeled SG in Figures 3A and 3E) or weakly positive (Figure 3C). Distinctly positive spinal ganglia nuclei were found in earlier mouse embryos (day 10.5) and in adult cat, while in *Xenopus* the cervical spinal ganglia were clearly stained at all embryonic stages.

The segmented mesoderm. The prevertebrae (derived from the sclerotomes) of the 12-day (Figure 3, panels A, D and E) and 13-day (Figure 2) embryo also stain heavily with anti-XlHbox 1 antibodies. Staining is most intense in the condensed cells of the intervertebral discs (Figures 2C, 2D, 4A and 4B) and in the cartilage-forming cells that surround the prevertebral bodies and the rib and neural arch anlagen (Figure 5, panels A, C and E). A striped appearance (Figure 3A) results from the staining of the segmentally repeated sclerotome derivatives, which will ultimately become the axial skeleton (vertebrae and ribs). Immunostaining with antibody A or B reveals that expression of the XlHbox 1 protein definitely begins in the first cervical prevertebra (e.g. Figure 4A), becomes maximal over the thoracic segments and continues, with some attenuation, into the most posterior sclerotomes of the tail (Figure 5E). Use of the anti-human XlHbox 1 antibody confirmed expression starting in the first prevertebra (Figure 2C).

XlHbox 1 antigens are also found in the mesenchyme surrounding a group of mesonephric tubules located near the gonad (labelled Me in Figures 2, 4 and 5). At this stage the mesonephric tubules (which originate from the segmentally repeated nephrotomes) are undergoing extensive rearrangements, giving rise later to the excretory ducts of the gonad (Rugh, 1968). Nuclei are also stained in the nearby posterior diaphragmatic region (DR, Figures 4C and 5C).

The unsegmented mesoderm. The lateral plate mesoderm (which will form the visceral mesoderm and the body wall) does not become segmented, but still expresses XlHbox 1 protein. Mesenchymal nuclei, but not the endodermally derived epithelia, are stained in the larynx (Lx, Figure 3B and 3C), oesophagus (Oe, Figures 2A, 2B and 5C), lung (Lu, Figures 2A, 2B, 2D, 3E, 4A and 5C), stomach (not shown) and small intestine (In, Figure 5E). Expression over mesenchyme, but not over the endodermal epithelium, has also been noted in mice for Hox 2.1 in the lung (Krumlauf *et al.*, 1987), and for Hox 1.3 in the lung, stomach and gut (Dony and Gruss, 1987).

The limb buds. Staining is observed in the mesenchymal nuclei of cells at the base of the forelimb bud (arrowed in Figure 3B). XlHbox 1 may therefore provide a useful marker for analysis of positional specification during limb bud development.

Localization of antigens unique to the long XlHbox 1 protein

Having defined the regions in which antigens common to the short and long types of XlHbox 1 homeodomain protein are expressed, we now turn to the region of expression of the long protein, derived from promoter II transcription. Immunostaining with antibody C, as described above and illustrated in Figure 1, is specific for the extended amino terminus of the long protein. The results are summarized in Figure 4C, in which regions stained by antibody C are indicated in blue, while additional regions stained with antibody A or B are indicated in red. It is not possible to conclude from our experiments whether there are nuclei in regions of the embryo that express exclusively the long protein, or if such regions express both short and long proteins in the same cells (Figure 4C, blue area). Cells expressing only the short protein, however, can be identified because they are stained by the 'common' antibodies A and B, but not by the long protein-specific antibody C (Figure 4C, red area).

In all regions of the embryo that express XlHbox 1 protein, the anterior border of expression of the long (PR II-type) protein is more posterior than that of the short (PR I-type) protein:

The CNS. The long protein is expressed more posteriorly (compare Figure 4A and B). The myelencephalon is stained with antibody B but not with antibody C, as is clearly seen by comparing Figure 5A and B.

The segmented mesoderm. Expression of the long protein is restricted to a thoracic domain, starting in prevertebra 9 and ending in prevertebra 16 (Figures 4B and 5B). Antibodies A or B (which recognize the short protein in addition to the long protein) immunostain more anterior regions, starting from prevertebra 1 (compare panel A with panel B in both Figure 4 and Figure 5). The same result is seen with a new antibody against the human homolog (Figure 2C). The staining with antibody A or B extends more posteriorly into the tail sclerotomes (arrowed in Figure 5E), although it gradually decreases in intensity.

The internal organs. The long (PR II-type) protein is expressed only in those internal organs that arise from more posterior regions of the embryo: the mesonephric region (Figure 5D) and the intestine (Figure 5F). More anterior organs, such as larynx, oesophagus, lung and stomach do not express the long protein (Figure 5D).

From these studies in mouse we conclude that the two types of translation product that arise from the murine homolog of XlHbox 1 are differentially expressed along the antero-posterior axis of the embryo.

As expected, this spatial displacement of the two XlHbox 1 proteins is also found in *Xenopus* embryos. In Figure 4 a

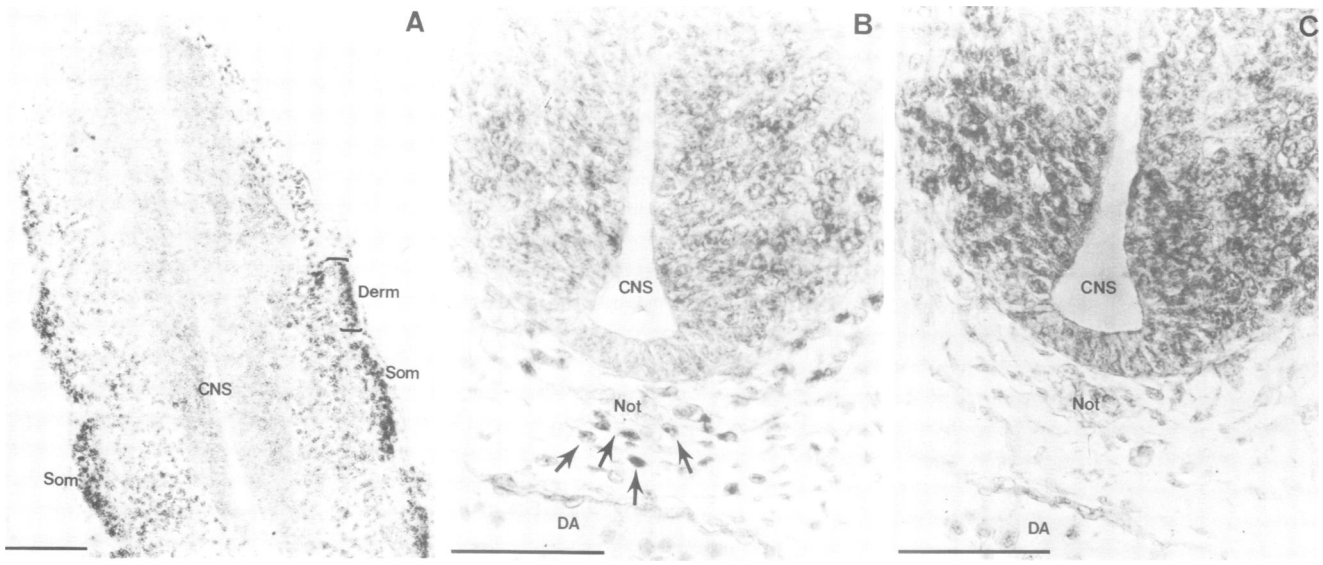

Fig. 6. Distribution of XlHbox 1 proteins in somite-derived cells of the day 9.5 mouse embryo. **Panel A** shows a 9.5-day embryo sectioned transversely through the somites and spinal cord (CNS) and immunostained with antibody B, which detects both short and long XlHbox 1 proteins. **Panels B** and **C** are adjacent transverse sections, at higher magnification, of a different region of the same embryo immunostained with antibody B or antibody C, respectively. In this region the cells immediately surrounding the notochord (Not) are immunopositive for antibody B (indicated by arrows) but not for antibody C. Som, somite; Derm, dermatome; DA, dorsal aorta. The scale bar represents 0.1 mm.

comparison of panels D and E reveals a more posterior distribution of the long protein in the CNS of stage 46 *Xenopus* embryos. At this stage mesodermal expression is weaker than in mouse embryos, being barely detectable in the sclerotomes (Figure 4D and E). We believe this merely reflects a difference in the timing of sclerotomal differentiation in the prospective vertebral column. In the mouse embryo XlHbox 1 staining is maximal at the time in which chondrocyte formation starts in the future vertebrae. In *Xenopus* the differentiation of cartilage is delayed until the third week of development when tadpoles of this stage indeed show clear staining of a subset of sclerotomes (data not shown).

Long and short protein distribution in earlier mouse embryos

Earlier embryos were analyzed to establish whether the apparent areas of expression of XlHbox 1 protein result from differences in the maturation state of somites (the anterior segments form first) or are indeed characteristic of a specific region of the embryo. Analysis of day-9 mouse embryos, in which the somites have not yet started to disaggregate, shows nuclear staining over the entire somite, with the distribution of the two proteins being similar to that in day-13 embryos (data not shown). This indicates that each protein is indeed a region-specific marker, and that the localized expression in the mesoderm is not simply because of differences in somite age.

An intriguing result was obtained when slightly later (9.5-day) embryos were analyzed: at this stage the somites disaggregate, giving rise to dermatomes (the most external part, Figure 6A), myotomes and sclerotomes (the part innermost to the embryo). All somite-derived nuclei are still weakly stained, but the staining is most prominent in the dermatomes, perhaps because dermatome nuclei are more tightly packed (Figure 6A). At this stage in the development of the vertebral column, cells from the ventromedial angle of the sclerotome migrate medially towards the notochord forming the transverse commissure. These cells surround the notochord and will give rise to the perichordal tube (Verbout, 1985). Perichordal mesodermic cells provide the vertebral body anlagen, which are formed later in development after the appearance of the intervertebral discs (Verbout, 1985). Unexpectedly, some sections contained a small group of highly stained nuclei immediately surrounding the notochord (indicated by arrows in Figure 6B), in what will later become the perichordal tube. These cells are probably the first sclerotomal cells to arrive in the vicinity of the notochord. In more anterior sections of the same embryos (more mature regions) this strong staining extends gradually to neighboring cells so that the transverse commissure and eventually, in older embryos, the whole of the sclerotomal derivatives including neural arches and rib anlagen are stained (Figure 3E). The results suggest that the notochord may have a late inductive effect, activating XlHbox 1 expression in mesodermal derivatives (see Discussion).

CNS and mesodermal expression of XlHbox 1 proteins are out of register in the mouse, but not in Xenopus

Immunostaining of mouse embryos with XlHbox 1 antibodies shows that the anterior border of expression in the mesoderm is more posterior than in the CNS, as best illustrated by the longitudinal sections of Figure 4B. Out-of-register expression in CNS and mesoderm was first noted by Utset *et al.* (1987) for mouse homeobox gene Hox 3.1. This phenomenon also applies to mouse homeobox genes Hox 1.3 (Dony and Gruss, 1987) and Hox 1.2 (Toth *et al.*, 1987). The question then arises of whether this is a peculiarity of mouse development or a property intrinsic to all homeobox genes in vertebrates.

We approached this by analyzing the distribution of XlHbox 1 proteins in *Xenopus* embryos. A detailed study has been carried out (Hardwicke *et al.*, manuscript in preparation), but here the salient points are summarized.

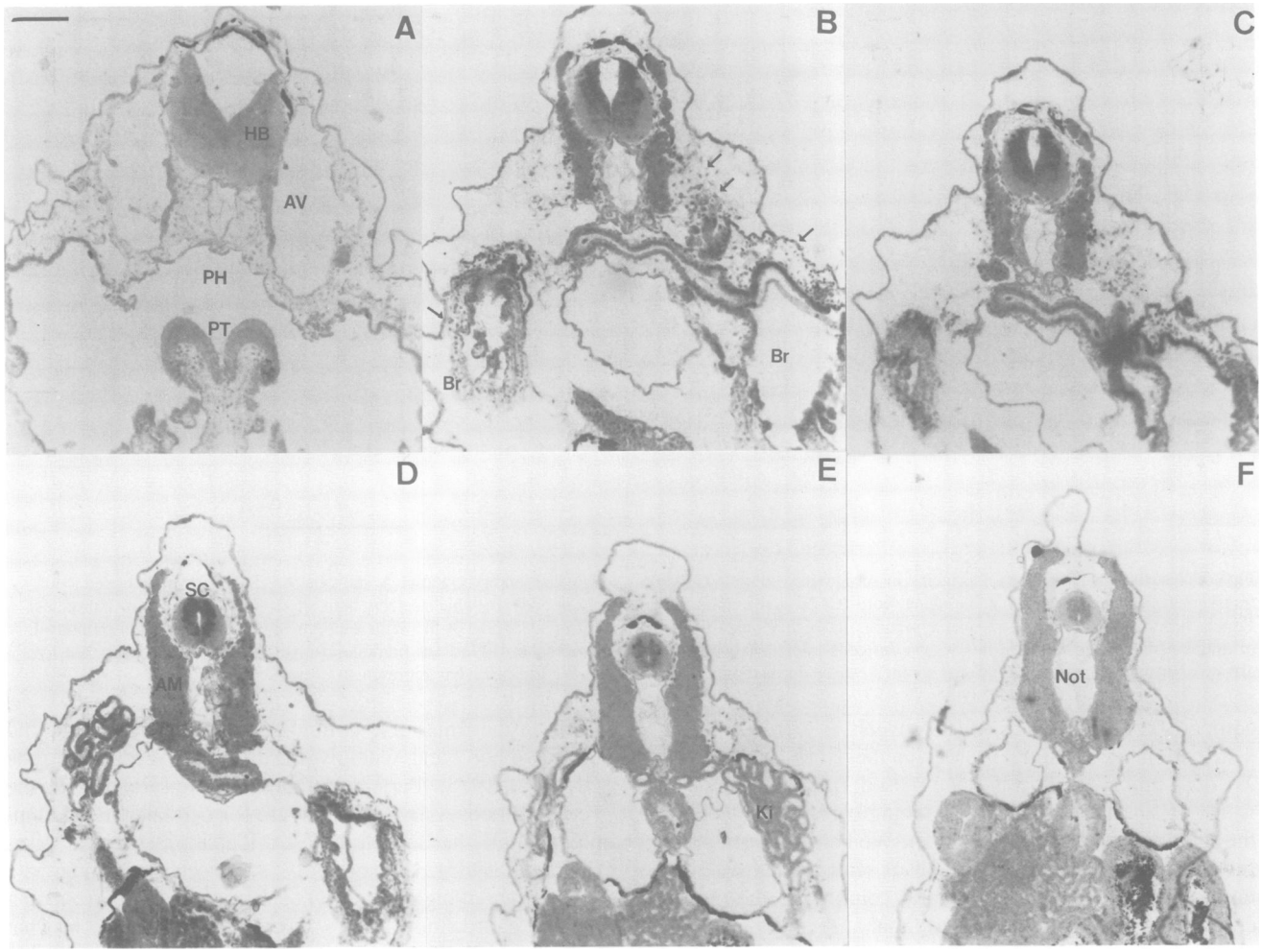

Fig. 7. The domains of expression of XlHbox 1 protein in mesoderm and CNS of the *Xenopus* stage 46 tadpoles are in register. **Panels A–F** are six representative transverse sections taken in an anterior to posterior direction from a serially sectioned stage 46 *Xenopus* tadpole. All sections were immunostained with antibody A, which detects both long and short versions of the XlHbox 1 protein. **Panels B–E** show immunostaining in spinal cord nuclei. Note that it is only over this same region that mesodermally derived cells are also stained (indicated by small arrows in panel B). Abbreviations used: AM, axial muscle; AV, auditory vesicle; HB, hindbrain; PH, oro-pharyngeal cavity; PT, primitive tongue; Br, branchial cavity; SC, spinal cord; Ki, kidney; Not, notochord. The scale bar represents 0.2 mm.

At 6 days of development, stage 46 *Xenopus* tadpoles clearly show a localized domain of expression of XlHbox 1 protein in the CNS, and we have shown that the short protein product of promoter I is expressed more anteriorly than the long protein from promoter II (Figure 4, panels D and E). The nuclei on the ventral side of the notochord, which are part of the presumptive sclerotome, definitely stain although too weakly to be reproduced in Figure 4D and E. This staining was in register with the expression in the CNS for both antibodies. In 3-week-old tadpoles, which already have some cartilage laid down by the chondrocytes which have differentiated from the sclerotome, the cells surrounding the notochord are much more strongly stained and staining is still in register with the CNS (data not shown). Figure 7 shows mesodermal nuclei immunostained with antibody A in transverse sections of a 6-day-old tadpole. Many nuclei distributed throughout the interstitial spaces express XlHbox 1 protein (e.g. the nuclei between the myotomes and the skin, arrowed in Figure 7B). Furthermore, staining in these mesodermal regions is restricted to those serial sections where the CNS is also positive (Figure 7, panels B, C, D and E) and is not present on sections where the CNS is negative (Figure 7, panels A and F).

Transverse sections of mouse embryos, using an antibody reacting with both long and short XlHbox 1 proteins clearly show that mesodermal expression extends much more posteriorly than that in CNS, as depicted in Figure 3, panel E. The out-of-register expression of mouse neuroectoderm and segmented mesoderm is even more striking when the long protein-specific antibody is used (antibody C, see longitudinal sections in Figures 4B and 5B). When serial longitudinal sections of stage 40 *Xenopus* embryos were similarly stained with antibody C, expression was in register in the CNS, sclerotomal mesoderm and lateral plate mesoderm (data not shown).

We conclude from this that in *Xenopus*, in contrast to mouse, the domains of expression of XlHbox 1 protein are in register in the mesoderm and CNS, even at late embryonic stages.

Discussion

Two proteins from XlHbox 1

Using specific affinity-purified polyclonal antibodies we have

shown that the two proteins predicted by Cho *et al.* (1988) to arise from the two major embryonic transcripts of XlHbox 1 are indeed translated in frog (and mouse) embryos. The two proteins have the same DNA-binding specificity (Cho *et al.*, 1988) but presumably differ in biochemical functions provided by the amino-terminal extension present only in the long protein. Lambert *et al.* (1987) have described a gene from bovine papilloma virus in which two DNA-binding proteins are encoded by the same ORF, with the long protein being an activator and the short version a repressor of transcription. The long and short XlHbox 1 proteins, which may have different developmental functions, are expressed in distinct regions of the embryo with the long protein being expressed more posteriorly than the short one. If translation mimicks transcription, then the results would also suggest that the two XlHbox 1 promoters are under precise spatial regulation along the antero-posterior axis of the embryo.

Antibodies detect homologous proteins in different vertebrates

In this study we have been able to use mouse, as well as *Xenopus* embryos, to analyze the expression of the two protein products of XlHbox 1. The sequence data obtained by Cho *et al.* (1988) suggested that our antibodies would crossreact between species and allow the study of the homologous gene in mammals. The reason for this is that Simeone *et al.* (1987) have reported the complete sequence of a human cDNA clone (called HHO.c8) which is the homolog of the frog PR I transcript encoding the short protein. Out of 152 amino acids, 128 are identical between frog and man, and many of the 24 amino acid changes are conservative (Cho *et al.*, 1988). Furthermore, comparison of the 5′ leader of clone HHO.c8 to *Xenopus* sequences shows that it also contains a substantial amount of the ORF utilized in the long version of the XlHbox 1 protein. Of 67 long protein-specific amino acids that we can compare, 57 are identical in frog and man, and of the 10 amino acid changes, five are conservative (Cho *et al.*, 1988). We therefore expected the anti-*Xenopus* and the anti-human antibodies to crossreact with the homologous mouse proteins. In addition, the sequence comparison suggests that in humans the organization of this gene is similar to that of *Xenopus*, with two promoters driving the production of a long and a short version of a homeodomain protein from the same ORF.

The interpretation of immunolocalization results may be complicated by potential crossreactivity with other homeodomain proteins. This problem is alleviated by comparing results obtained with several non-overlapping antibodies, and by competition experiments with fusion proteins derived from other homeobox genes. Despite these potential problems, antibody probes have the great advantage that they permit one to cross evolutionary boundaries and to compare the expression of the homologous genes in embryos of diverse organisms because of the high conservation of the proteins during evolution. The anti-*Xenopus* and anti-human XlHbox 1 antibodies reported here recognize similar regions in vertebrate embryos as different as frogs, mice, zebra fish and chickens (unpublished observations). This will be useful in studies comparing the strategy of development in different chordates. Much of our present knowledge on the vertebrate body plan has come from comparative anatomical and embryological studies (Romer and Parsons, 1986), and we believe that the type of molecular comparison reported here will further extend our understanding.

The long protein is expressed more posteriorly

Antibodies that recognize both the short and the long XlHbox 1 proteins bind to nuclei in three main regions of the mouse embryo: (i) the CNS (myelencephalon and cervical spinal cord), (ii) the segmented mesoderm (sclerotomes, dermatomes) and (iii) the visceral unsegmented mesoderm (mesenchyme of the larynx, oesophagus, lung, stomach and intestine). Other immunopositive regions (spinal ganglia, mesonephros, posterior diaphragmatic region, limb buds), have been discussed in the results. The antibody specific for the long XlHbox 1 protein reacts with nuclei of cells that occupy more posterior locations in all of the above three regions. Regions stained with long protein-specific antibody are indicated in blue in Figure 4C; in the CNS, expression starts in the spinal cord (posterior to the myelencephalon), in the segmented mesoderm it is restricted to prevertebrae 9−16, and in the unsegmented mesoderm it is confined to the intestine. Expression of the long protein is also detectable in mesonephric tubules and limb buds. The gene is also under temporal regulation, so that by day 18 of mouse development the mesoderm is negative and XlHbox 1 protein is only detectable in the neurons of the CNS. Taken together, the results show that both XlHbox 1 gene products are exquisitely regulated, spatially and temporally, along the antero-posterior axis of the embryo.

The antibody localization results reported here are in general agreement with mRNA distribution studies carried out by Simeone *et al.* (1987), who found two transcripts corresponding to the human cDNA HHO.c8 on Northern blots from human embryo spinal cord, backbone and limb buds, which were absent from brain and liver. The detailed *in situ* hybridization studies carried out in *Xenopus* by Carrasco and Malacinski (1987) with a 3′ XlHbox 1 probe do not distinguish between PR I and PR II type transcripts, but are entirely consistent with our immunolocalizations in *Xenopus* tadpoles using antibodies that do not distinguish between the short and long XlHbox 1 protein. Recently the cloning of a partial mouse cDNA clone with sequence homology to XlHbox 1 and HHO.c8 has been reported (Sharpe *et al.*, 1988), but because the chromosome mapping does not agree with a gene belonging to the Hox 3 complex, and because sufficient upstream sequence is not yet available, it is not clear whether this mouse gene is the murine homolog of XlHbox 1, or rather a related gene.

Concerning XlHbox 1 expression in the internal organs (the long protein is expressed in the mesenchyme of the intestine, while the short protein is present in that of oesophagus, lung and stomach), it is perhaps relevant that Slack (1985) has suggested that metaplasia in the human digestive tract may be considered a type of homeotic transformation. A substantial number of adult humans have patches of gastric epithelium in the oesophagus, or of intestinal epithelium in the stomach. These have been noted by pathologists because they sometimes lead to clinical disorders (Slack, 1985; Robbins *et al.*, 1984). It may be that the disruption of homeodomain protein function in the underlying mesoderm plays a role in such metaplasia.

XlHbox 1 in vertebral column development

The availability of antibody probes for XlHbox 1 proteins

may provide unexpected insights into the mechanisms of vertebrate development. Analysis of staining patterns in 9.5-day mouse embryos suggests that the notochord may have an inductive effect late in embryogenesis, in addition to its early role in neural tube formation. In this model, induction causes increased expression of XlHbox 1 protein in newly arrived migratory sclerotomal cells, when they reach the perichordal zone (arrows in Figure 6B). While the morphological data do not prove the existence of this late inductive effect, this proposition could be tested experimentally by culturing embryonic fragments in close contact.

The pattern of expression of some other Hox genes in the mouse vertebral column has been previously studied in detail. Hox 1.3 is expressed in sclerotomes 8–22, predominantly in the rib anlagen (Dony and Gruss, 1987). Hox 2.1 is not expressed in sclerotome-derived mesoderm, although it is expressed in nephrotome and lateral plate mesoderm (Holland and Hogan, 1988). Hox 3.1 is expressed in prevertebrae 6–10 and not in intervertebral discs (Breier *et al.*, 1988). Finally the murine paired-box homolog PAX-1 is initially expressed in the sclerotomal cells of the differentiated somites, which then continue to express this gene as they migrate medially to surround the notochord and form the perichordal tube (Deutsch *et al.*, 1988). It therefore appears that XlHbox 1 differs from the above mentioned genes in its expression during development of the vertebral column. A possible induction of homeobox gene products in the immediate vicinity of the notochord (Figure 6B) has not been reported in any of the previous studies.

Alignment of expression in CNS and mesoderm

In mouse, the anterior border of expression of XlHbox 1 proteins in the mesoderm is more posterior than in the CNS (e.g. Figure 4B). This 'out-of-register' expression has previously been noted to occur in three other mouse homeobox genes (Utset *et al.*, 1987; Dony and Gruss, 1987; Toth *et al.*, 1987) analyzed by *in situ* hybridization. Surprisingly, in *Xenopus* embryos XlHbox 1 protein expression in the CNS and the mesoderm is in register along the antero-posterior axis. This points to a difference in the developmental plan adopted by mouse and frog. It will be interesting to analyze earlier mouse embryos to determine whether the neuroectodermal and mesodermal domains of expression of both promoters are initially in register when positional values are first established in the embryo and then slide apart as the individual develops.

Despite the differences mentioned above, the similarities in the expression of XlHbox 1 in *Xenopus* and of its mouse homolog are striking. In both embryos the short protein is expressed more anteriorly than the long protein. The long and short proteins can bind to the same (or similar) target sequences in DNA (Cho *et al.*, 1988), but they presumably differ in the biochemical function provided by the amino-terminal domain present only in the long protein. We therefore envisage that the difference in spatial expression of these two proteins could have distinct consequences on embryonic cell determination.

Materials and methods

Preparation of mouse and frog embryos for antibody staining

Pregnant mice were purchased from Simmonsen Labs Inc. (California) and their embryos staged according to Rugh (1968). *Xenopus laevis* albino tadpoles were staged according to Nieuwkoop and Faber (1967). Mouse and frog embryos were fixed in Bouin's solution (75 parts saturated picric acid, 25 parts paraformaldehyde and 5 parts glacial acetic acid), embedded in paraffin, microtome sections (8 μm) mounted on 8-well histological slides (Roboz Surgical Instruments), and subjected to immunostaining essentially as described by Zeller *et al.* (1983), except that we used alkaline phosphatase-labeled second antibody as described below.

Construction and isolation of β-galactosidase-XlHbox 1 fusion proteins

Fusion protein A (Figure 1) was constructed by ligating a *Bgl*II–*Pst*I fragment of genomic DNA carrying 54 amino acids of the XlHbox 1 homeodomain and the downstream 34 amino acid region, in-frame with the carboxy terminus of β-galactosidase in the pUR289 vector (Ruther and Muller-Hill, 1983). As an intermediate step in the preparation of fusion proteins B and C (Figure 1), the entire cDNA clone p15 (Cho *et al.*, 1988) was blunt-ended and fused in-frame to the carboxy terminus of β-galactosidase in pTRB2, a modified version of pUR292 (Burglin and De Robertis, 1987). Construct B is a deletion of all sequences downstream of the *Bgl*II site at the start of the homeodomain, and construct C is a deletion of all sequence downstream of the *Bam*HI site (Figure 1). The human antibody was raised against a fusion of the last 106 amino acids of the human homolog of XlHbox 1 to β-galactosidase.

Fusion proteins were induced and isolated by the method of Ullman (1984), with some modifications. Briefly, mid-log phase bacterial cultures were induced by addition of IPTG (final concentration 0.5 mM) for 2 h. The cells were harvested and resuspended (10 ml per liter-equivalent of cells) in 50 mM Tris–HCl pH 7.9, 25% sucrose, 1% NP-40, 0.5% sodium desoxycholate, 2 mM dithiothreitol, 5 mM EDTA (Bikel *et al.*, 1983). The mixture was sonicated with a Branson Sonifier microprobe with four 30 s bursts, cooling on ice. After centrifugation to remove insoluble material, the supernatant was diluted 5-fold with 2 M NaCl, 20 mM Tris–HCl pH 7.5, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol, left for 15 min at room temperature and recentrifuged. The supernatant was loaded on a Sepharose 4B-TPEG column, previously equilibrated and washed in 1.6 M NaCl, 20 mM Tris–HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM β-mercaptoethanol. After washing with a large volume of 1.6 M NaCl buffer, the bound fusion protein was eluted with 1 ml fractions of 0.1 M boric acid–NaOH pH 10, which were immediately neutralized with 0.4 ml each of 2 M Tris–HCl pH 7.

Antibody preparation

For each fusion protein, one 6-week-old female New Zealand White rabbit was immunized subscapularly and intramuscularly (four injection points) with 3 mg (total) of fusion protein emulsified by sonication with Freund's complete adjuvant (1:1 v/v). Booster immunizations with 2 mg of fusion protein emulsified in incomplete adjuvant (1:1 v/v) were given 4, 8, 12 and 16 weeks after the first inoculation.

Rabbits were bled (50 ml) via the marginal ear vein 10 days after each inoculation, beginning after the second boost. Antisera were first depleted of anti-β-galactosidase and anti-*E.coli* antibodies. A bacterial lysate of an induced culture containing the pTRB0 vector without any insertion of foreign DNA was coupled to CNBr-activated Sepharose 4B (Pharmacia; 8–10 mg of protein per ml of swollen Sepharose). Approximately 4 ml of antiserum was mixed at room temperature with 2 ml of this depletion matrix for 2 h with end-over-end rotation. The supernatant was taken and similarly incubated with a fresh 2 ml aliquot of the same matrix. The supernatant was then affinity-purified by incubation of ~4 ml of depleted antisera with 0.5 ml of CNBr-activated Sepharose 4B matrix coupled with the corresponding fusion protein (2–4 mg of protein per ml of swollen Sepharose) for a period of 2 h at room temperature with end-over-end rotation. This mixture was then transferred to a column constructed from a 1 ml plastic hypodermic syringe and allowed to settle. The matrix was washed with 25–50 ml of TBST (10 mM Tris pH 8, 150 mM NaCl, 0.05% Tween-20). Antibodies were eluted with 0.2 ml aliquots of 0.15 M glycine pH 2.5 which were collected directly into tubes each containing 0.2 ml of 2 M Tris pH 8. The fractions containing the affinity-purified antibody were detected by immunostaining strips cut from a Western blot of IPTG-induced bacteria harboring the appropriate fusion protein construct. The best three fractions were pooled and stored in aliquots at −70°C, usually stabilized by the addition of fetal calf serum (10% final). Approximately 75–100 μg of specific antibody at 50 μg/ml was recovered. The antibodies were named A, B or C, according to the fusion protein used in the affinity purification. Note that antibodies B and C were made by affinity-purifying antiserum B using a matrix coupled with either fusion protein B or C (Figure 1).

Immunostaining analysis

Sections were dewaxed by passage through xylene and graded ethanol baths,

and excess non-specific protein-binding sites on the sections were saturated by incubation with 30 μl per well of blocking solution (10 mM Tris pH 7.4, 100 mM $MgCl_2$, 0.5% Tween-20, 1% BSA, 5% fetal calf serum) as described by Zeller *et al.* (1983), except that the urea incubation step was omitted since we found that this resulted in variable and unacceptable loss of antigen. Sections were incubated overnight with 1:100 dilutions of antibody A, B or C in TBST. Excess antibody was removed by rocking (three times, 5 min each wash) in a large volume of TBST. The second antibody (anti-rabbit IgG conjugated with alkaline phosphatase; Promega-Biotec) was diluted 1:1000 in TBST and 30 μl per section applied for 3 h at room temperature. Excess second antibody was removed as above. The bound antibodies were visualized by the addition of color development substrate solution (BCIP and NBT; as described by Promega Biotec) to each section for ~30 min. The reaction was stopped with 20 mM Tris pH 8, 5 mM EDTA. Sections were dried by aspiration and rapidly mounted in glycerol.

Acknowledgements

We thank Drs Larry Zipursky, David Meyer, Kathryn Calame and Ken Cho for critical reading of the manuscript. This work was supported by NIH grant HD 21502-02 and a grant of the California Institute for Cancer Research. C.V.E.W. was the recipient of an EMBO Long Term Fellowship.

References

Awgulewitsch,A., Utset,M.F., Hart,C.P., McGinnis,W. and Ruddle,F.H. (1986) *Nature*, **320**, 328–335.
Bikel,I., Roberts,T.M., Bladon,M.T., Green,R., Amann,E. and Livingston,D.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 906–910.
Breier,G., Dressler,R. and Gruss,P. (1988) *EMBO J.*, **7**, 1329–1336.
Burglin,T.R. and De Robertis,E.M. (1987) *EMBO J.*, **6**, 2617–2625.
Cannizzaro,L.A., Croce,C.M., Griffin,C.A., Simeone,A., Boncinelli,E. and Huebner,K. (1987) *Am. J. Hum. Genet.*, **41**, 1–15.
Carrasco,A.E., McGinnis,W., Gehring,W.J. and De Robertis,E.M. (1984) *Cell*, **37**, 409–414.
Carrasco,A.E. and Malacinski,G.M. (1987) *Dev. Biol.*, **121**, 69–81.
Cho,K.W., Goetz,J., Wright,C.V., Fritz,A., Hardwicke,J. and De Robertis, E.M. (1988) *EMBO J.*, **7**, 2139–2149.
Condie,B.G. and Harland,R.M. (1987) *Development*, **101**, 93–105.
Deutsch,U., Dressler,G.R. and Gruss,P. (1988) *Cell*, **53**, 617–625.
Dony,C. and Gruss,P. (1987) *EMBO J.*, **6**, 2965–2975.
Fibi,M., Zink,B., Kessel,M., Colberg-Poley,A.M., Labeit,S., Lehrach,H. and Gruss,P. (1988) *Development*, **102**, 349–359.
Fritz,A. and De Robertis,E. (1988) *Nucleic Acids Res.*, **16**, 1453–1469.
Gaunt,S.J. (1987) *Development*, **101**, 51–60.
Gaunt,S.J., Miller,J.R., Powell,D.J. and Duboule,D. (1986) *Nature*, **324**, 662–664.
Gilbert,S.F. (1985) *Developmental Biology*. Sinauer Associates, Inc., Sunderland, MA.
Holland,P.W.H. and Hogan,B.L.M. (1988) *Development*, **102**, 159–174.
Kessel,M., Schulze,F., Fibi,M. and Gruss,P. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5306–5310.
Krumlauf,R., Holland,P.W.H., McVey,J.H. and Hogan,B.L.M. (1987) *Development*, **99**, 603–617.
Lambert,P.F., Spalholz,B.A. and Howley,P.M .(1987) *Cell*, **50**, 69–78.
Martin,G., Boncinelli,E., Douboule,D., Gruss,P., Jackson,I., Krumlauf,R., Lonai,P., McGinnis,W., Ruddle,F. and Wolgemuth,D. (1987) *Nature*, **325**, 21–22.
Nieuwkoop,P.D. and Faber,J. (1967) *Normal Table of Xenopus laevis (Daudin)*, 2nd edition. North Holland Publ. Co., Amsterdam.
Odenwald,W.F., Taylor,C.F., Palmer-Hill,F.J., Friedrich,V., Tani,M. and Lazzarini,R.A. (1987) *Genes Dev.*, **1**, 482–496.
Robbins,S.L., Cotran,R.S. and Kumar,V. (1984) *Pathologic Basis of Disease*. Saunders, Philadelphia, PA.
Romer,A.S. and Parsons,T.S. (1986) *The Vertebrate Body*. Saunders, Philadelphia, PA.
Rugh,R. (1968) *The Mouse*. Burgess, Minneapolis.
Ruther,U. and Muller-Hill,B. (1983) *EMBO J.*, **2**, 1791–1794.
Sharpe,C.R., Fritz,A., De Robertis,E.M. and Gurdon,J.B. (1987) *Cell*, **50**, 749–758.
Sharpe,P.T., Miller,J.R., Evans,E.P., Burtenshaw,M.D. and Gaunt,S.J. (1988) *Development*, **102**, 397–407.
Simeone,A., Mavilio,F., Acampora,D., Giampaolo,A., Faiella,A., Zappavigna,V., D'Esposito,M., Pannese,M., Russo,G., Boncinelli,E. and Peschle,C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4914–4918.
Slack,J.M.W. (1985) *J. Theor. Biol.*, **114**, 463–490.
Toth,L., Slawin,K.L., Pintar,J.E. and Nguyen-Huu,M.C. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 6790–6794.
Ullman,A. (1984) *Gene*, **29**, 27–31.
Utset,M.F., Awgulewitsch,A., Ruddle,F.H. and McGinnis,W. (1987) *Science*, **235**, 1379–1382.
Verbout,A.J. (1985) *Adv. Anat. Embryol. Cell Biol.*, **90**, 1–121.
Zeller,R., Nyffenegger,T. and De Robertis,E. (1983) *Cell*, **32**, 425–434.

Received on March 14, 1988; revised on June 26, 1988