XIHbox 8: a novel *Xenopus* homeo protein restricted to a narrow band of endoderm

CHRISTOPHER V. E. WRIGHT, PATRICK SCHNEGELSBERG and EDDY M. DE ROBERTIS

*Department of Biological Chemistry, University of California, Los Angeles, CA 90024-1737, USA*

Summary

We report the isolation of a new homeobox gene from *Xenopus laevis* genomic DNA. The homeodomain sequence is highly diverged from the prototype *Antennapedia* sequence, and contains a unique histidine residue in the helix that binds to DNA. The homeodomain is followed by a 65 amino acid carboxy-terminal domain, the longest found to date in any vertebrate homeobox gene. We have raised specific antibodies against an XIHbox 8-β-gal fusion protein to determine the spatial and temporal expression of this gene. The nuclear protein first appears in a narrow band of the endoderm at stage 33 and develops into expression within the epithelial cells of the pancreatic anlagen and duodenum. Expression within the pancreatic epithelium persists into the adult frog. This unprecedented restriction to an anteroposterior band of the endoderm suggests that vertebrate homeobox genes might be involved in specifying positional information not only in the neuroectoderm and mesoderm, but also in the endoderm. Our data suggest that XIHbox 8 may therefore represent the first member of a new class of position-dependent transcription factors affecting endodermal differentiation.

Key words: *Xenopus laevis*, homeodomain protein, endodermal differentiation, embryogenesis.

Introduction

The majority of homeobox-containing genes in *Drosophila* specify or interpret positional information along the anteroposterior axis of the fly (Gehring, 1987; Akam, 1987; Ingham, 1988). Vertebrate homeobox-containing genes share a number of structural features with those of *Drosophila* (Cairasco et al. 1984; Fienberg et al. 1987; Dressler & Gruss, 1988; Boncinelli et al. 1988; Scott et al. 1988). One of the most compelling arguments in favour of vertebrate homeobox genes having a role similar to those of *Drosophila* is that many of them are expressed in restricted and precisely defined regions along the anteroposterior axis of the embryo (Awgulewitsch et al. 1986; Holland & Hogan, 1988; Oliver et al. 1988), rather than in a tissue-specific or cell type-specific manner. To date, all vertebrate homeobox genes have been found to be expressed in the neuroectoderm and sometimes also in the mesoderm (e.g. Utset et al. 1987; Dony & Gruss, 1987; Toth et al. 1987; Graham et al. 1988), but none of the genes reported are expressed in the endoderm. In *Drosophila*, a small amount of transient expression of *fushi tarazu*, *engrailed*, *Ultrabithorax* and *caudal* has previously been noted in small areas of the embryonic gut (Akam & Martinez-Arias, 1985; DiNardo et al. 1985; Fjose et al. 1985; Ingham et al. 1985; Kornberg et al. 1985; Mlodzik et al. 1985; Krause et al. 1988). However, in all of these cases, the analysis did not extend to showing whether the stained cells are endodermal in origin or derived from another germ layer. Whatever the case, it is definite that no *Drosophila* or vertebrate gene isolated thus far is expressed exclusively in the endoderm.

In this paper, we report the isolation of a new homeobox gene, called XIHbox 8, which is expressed solely in a narrow band of the endoderm in early *Xenopus* embryos. As development proceeds XIHbox 8 protein is restricted to the nucleus of endodermal cells of the duodenum and the developing pancreas. The homeodomain sequence and the unique pattern of expression suggest that previously undiscovered classes of homeodomain protein may be involved in the specification or interpretation of anteroposterior position within the vertebrate endoderm. In vertebrates coculturing of tissue explants has shown that the differentiation pathway followed by the embryonic endoderm (e.g. choice between lung or intestinal epithelium) is induced by the underlying mesoderm (reviewed by Gurdon, 1987; Wessels, 1977). Thus we believe that the gene isolated
here promises to be a valuable marker for studying endodermal differentiation along the anteroposterior axis and its earliest induction by mesoderm.

Materials and methods

DNA sequence

An XlHbox 8 genomic clone was isolated by screening a Charon 4A library (Wahl & Dawid, 1980) with Drosophila homeobox probes as described previously (Carrasco et al. 1984). The EcoRI fragments were subcloned and a 1·8 kb insert that showed the only weak hybridization with the probe was cut separately with Alul and HaeIII and the EcoRI ends filled in, which allowed subcloning into M13 mp8 restricted with Smal. Two subclones were used further: an 860 bp Alul fragment and a 350 bp EcoRI/ HaeIII fragment (see Fig. 1A). The Alul and EcoRI/ HaeIII fragments are inserted in M13 in the opposite orientation relative to each other. Use of the M13 universal and homeobox-specific primers allowed the determination on both strands of the sequence between the homeobox EcoRI and the HaeIII site downstream of the translation stop codon. Sequencing was by the dideoxynucleotide chain-terminator method of Sanger et al. (1977). After the subcloning of all detectable EcoRI fragments, the ends of which were all double-strand sequenced in a search for the 5' portion of the homeodomain sequence, the original genomic phage stock was inadvertently lost in a trans-

Antibody procedures

Antisera were raised in NZW female rabbits using the protocol given in Oliver et al. (1988), using approximately 1·5–2 mg of fusion protein per inoculation. The first immune serum of reasonable titre was taken after boosting the animal twice. Subsequent boosts yielded antisera of higher titre and specificity. The anti-XlHbox 8 antibodies were purified by depletion and affinity purification as described by Oliver et al. (1988). In most cases, the XlHbox 8 antiseraum was additionally depleted against a β-gal fusion protein matrix containing an Antennapedia-like vertebrate homeodomain (X1Hbox 1 fusion protein A in Oliver et al. 1988). This step removed residual anti-E. coli/β-gal antibodies and nonspecific cross-reaction of the antibody with other homeodomains. Immunoblotting experiments with several homeodomain-containing fusion proteins (XlHbox 1, XlHbox 2, human Hox 5.3) and fusion protein competition experiments during the immunostaining of sections (data not shown) clearly show that the staining patterns shown in this paper are specific for XlHbox 8 proteins, and not due to cross-reaction with the homeodomain of other proteins. It should be noted that the distribution of the protein antigen coincides with that of the XlHbox 8 RNA (Fig. 2). Sections were prepared and stained with affinity-purified antibodies and alkaline phosphatase-coupled second antibodies (Promega) as described in Oliver et al. (1988).

RNAs protection

RNA was isolated from embryos by a modified proteinase K/SDS method. Stage 38 embryos were anaesthetized in 0·02% MS222 (Sigma) and dissected as shown in Fig. 2. The dissected pieces (Fig. 2A) from thirty stage 38 embryos were pooled. RNA was also extracted from twenty whole tadpoles of the same stage. A pooled sample was immediately added to 400 μl of 40 mM-Tris-HCl pH 7·5, 4 mM-

Results

Isolation of a highly divergent Xenopus homeobox gene

A genomic library of X. laevis DNA was screened with an Antennapedia homeobox probe as described previously (Carrasco et al. 1984). A 1·8 kb EcoRI fragment of the XlHbox 8 genomic clone which hybridized weakly with the probe contains the last 126 nucleotides of the homeobox and sequences 3' to it.
(Fig. 1A). This fragment was subcloned (see Materials and methods) and most of the resulting sequence is shown in Fig. 1B. The homeodomain is followed by an additional 65 amino acids before the translation termination codon is reached. This is the longest carboxyl terminal domain yet reported in any vertebrate homeodomain protein, and its unusual length permitted the generation of high-titre specific antibodies from a β-galactosidase fusion protein derived from this genomic subclone (Fig. 1A).

The only similarity to other proteins arising from a computer search using the carboxy-terminal domain lies in a string of eight amino acids (underlined in Fig. 1B) that is also present in the 'potentiator' region described for Drosophila. One amino acid change that is particularly noteworthy is a change in amino acid 44 of the homeodomain (Gln to His), which is located in the 'recognition helix' (Fig. 1C) that determines DNA-binding specificity (Desplan et al. 1988; Hoey & Levine, 1988). No other homeodomain isolated in any organism has histidine in this position (see Scott et al. 1988, for a compilation of 87 homeodomain sequences). Therefore, it is possible that this change may endow the XlHbox 8 homeodomain protein with unique DNA sequence recognition properties.

We conclude that XlHbox 8 is a novel type of homeobox gene that differs from those characterized...
markers fragments

previously in several respects. When the pattern of XlHbox 8 expression was analysed in detail, its uniqueness became even more apparent.

**XlHbox 8 is expressed in a narrow endodermal band**

We analysed the spatial expression of XlHbox 8 initially by using a sensitive RNase protection assay (Materials and methods). Fig. 2 shows an experiment in which hatching tadpoles (stage 38) were dissected into five fragments. XlHbox 8 RNA expression was restricted to a single fragment (Fig. 2, lane 4) containing the anterior digestive tract, branchial arches, developing heart, and lateral body wall. No transcripts are detected in fragments containing central nervous system. The distribution of XlHbox 8 transcripts determined by RNase protection is in strict agreement with the protein distribution pattern determined by immunolocalization which is described below.

Northern blots of RNA from various stages of development failed to detect XlHbox 8 transcripts specifically presumably because of the small number of cells that express this gene. However, we analysed the spatial and temporal expression of the XlHbox 8 homeodomain protein more precisely by collecting embryos at various stages of development, serially sectioning the entire animal, and immunostaining with anti-XlHbox 8 antibodies. Antibodies were raised in rabbits using a β-gal fusion protein (Fig. 1A) as antigen. Antibodies reacting with *E. coli* proteins or with other Antennapedia-type homeodomains were removed by depletion over *E. coli*-Sepharose and XlHbox 1 homeodomain-Sepharose columns. The antiserum was then affinity-purified over an XlHbox 8 fusion protein column (Oliver et al. 1988).

By several criteria, the resulting antibodies react specifically with the carboxy-terminal region of the XlHbox 8 protein (see Materials and methods).

**Immunolocalization of XlHbox 8 protein in Xenopus embryos of various stages (staged according to Nieuwkoop & Faber, 1967)** is summarized in Fig. 3. At all stages of development, the protein is nuclear and localized exclusively in cells of endodermal origin. During early development immunostaining is undetectable until stage 33 (tailbud) when a narrow band of endodermal nuclei becomes weakly stained.
Fig. 3. Immunolocalization of XlHbox 8 protein. *Xenopus* embryos staged according to Nieuwkoop & Faber (1967) were serially sectioned and stained with anti-XlHbox 8 antibody. (A) Frontal section of a stage 33 embryo. A narrow band of expression of the XlHbox 8 protein (bracketed) is faintly visible in the endoderm slightly posterior to the pharynx. (B) Sagittal section through the left half of a stage 38 embryo. In this section the anterior border where expression of the XlHbox 8 antigen starts is clearly visible. The dorsal and ventral pancreatic rudiments are positive for the antigen as indicated. The stained endodermal tube in the centre will later mature into duodenum (note the cavity at its anterior tip). The liver is entirely unstained. (C) Sagittal section through a stage 41 embryo. Note the band of expression in the endoderm (dark stain) while the remainder of the embryo is unstained. (D) Transverse section (slightly oblique angle) through a stage 41 embryo at low magnification. Only endodermal cells are stained; neuroectoderm and mesoderm are unstained. (E) Same section as in D but at higher magnification. Note XlHbox 8 expression in the gut epithelium as well as in the pancreatic excretory duct. All nuclei in the endodermal band are stained; the lower right corner appears unstained because of the slightly oblique plane of sectioning. (F) The same section as in E counterstained with the fluorescent DNA stain Hoechst 33258. (G) Section through a stage 49 embryo. All nuclei in the epithelium of the duodenum and pancreatic ducts are stained, but only a small percentage of cells within the pancreas proper contain XlHbox 8 protein at this stage. No XlHbox 8 protein is expressed in liver, stomach, or small intestine. (H) Same section shown in G at higher magnification. (I) Section through an adult *Xenopus* pancreas. All the nuclei in the pancreatic duct system are stained with anti-XlHbox 8 antibody. The nuclei of cells in the acini are weakly stained. Abbreviations; ph, pharynx; en, endoderm; p, pancreas; dp, dorsal pancreatic anlage; vp, ventral pancreatic anlage; l, liver; pd, pancreatic excretory duct; s, stomach; in, intestine; cns, central nervous system; pn, pronephros.
(bracketed in panel A). Examination of serial sections showed that this band spans the region of the embryonic endoderm located between the third and fifth postotic somites. The unstained endodermal regions anterior to this band will give rise to pharynx, oesophagus and stomach, while the stained region will give rise to duodenum (see below). At stage 38 (hatching tadpole, panel B) and stage 41 (swimming tadpole, panels C to F) the narrow band of nuclear staining, which spans the entire endoderm dorsoventrally, is more intensely stained. Examination of sections counterstained with the DNA stain Hoechst 33258 reveals that all endodermal nuclei within this narrow region contain the XIHBox 8 protein (panels D and F). Intense nuclear staining is also detected in the dorsal and ventral pancreatic anlagen which evaginate from this endodermal region during development. The anatomical location of the *Xenopus* pancreas can be determined readily in serial sections thanks to the detailed description of normal development provided by Nieuwkoop & Faber (1967). Whole-mount staining of tadpoles with XIHBox 8 (see Fig. 4) confirms this assignment. The dorsal and ventral pancreas can be seen best in Fig. 3, panel B. The liver, which is localized more anteriorly, is entirely negative (panels B, C, G and H). The pancreatic anlagen are connected to the duodenum via an excretory duct system, the epithelium of which strongly expresses XIHBox 8 throughout its length (panels D, E and G).

XIHBox 8 antigen distribution in the digestive tract was also analysed in serial sections of later embryos (stages 45, 47, 48 and 49 – which correspond to 4, 5, 7 and 12 days of development) and of adult pancreas. By day 12 the endoderm in the duodenal region has evolved into a single layer of columnar epithelial cells all of which strongly express XIHBox 8 (Fig. 3, panels G and H), except for the initial third of the duodenum (ventral horizontal segment – see Nieuwkoop & Faber, 1967) which is unstained (not shown). At this stage, the dorsal and ventral pancreatic anlagen have fused to form a single organ in which the number of strongly stained nuclei has decreased substantially. Less than 10% of the cells express XIHBox 8 strongly (Fig. 3, panels G and H), in contrast to earlier stages when the vast majority of presumptive pancreatic cells were stained (panel B). All the epithelial nuclei of the pancreatic excretory duct system are strongly stained. XIHBox 8 protein is not expressed detectably in stomach, liver, or small intestine (panel G), or in gall bladder or bile duct (not shown).

Nuclear XIHBox 8 protein is also detectable in the pancreas of adult frogs. Fig. 3, panel I, shows that the most intense staining is found in the nuclei of the excretory tubules. The nuclei of cells in the exocrine acini are also stained above background, although much more weakly. The cells in the Islets of Langerhans are not stained (data not shown).

Fig. 4 depicts a stage 40 *Xenopus* tadpole stained for XIHBox 8 in a whole-mount procedure. This is presented to help in reconstructing the pattern of XIHBox 8 protein expression determined from the section stainings shown in Fig. 3. This type of staining confirms the view that this homeo protein is exclusively localized in epithelial nuclei of the region of the posterior foregut.

We conclude that XIHBox 8 homeodomain-containing protein is found only in nuclei of cells lying across a narrow band of the endoderm in early *Xenopus* embryos. These cells continue to express the antigen as they mature into the epithelial lining of the duodenum, the pancreatic epithelium and the epithelium of the interconnecting duct system.
Discussion

We report here the isolation of a new vertebrate homeobox-containing gene. Of the nine homeobox genes isolated to date from *Xenopus* (sequences listed in De Robertis et al. 1988), the XlHbox 8 homeodomain is most diverged from the prototype Antennapedia homeodomain. XlBox8 has an unusual Glu to His change in the proposed DNA recognition helix. A histidine residue in this position is not found in a homeodomain from any organism (Scott et al. 1988). In addition, XlBox8 has a carboxy-terminal domain of 65 amino acids downstream of the homeodomain, the longest of any vertebrate homeodomain protein.

The most exceptional feature of XlBox 8 is that it is the only homeobox gene known to be expressed exclusively in the endoderm. The XlBox 8 protein is found only in the nuclei of a narrow band of endodermal cells as early as stage 33. Our detailed analysis demonstrates that XlBox 8 is expressed in the epithelium of the duodenum and of the pancreas, an endodermal organ that evaginates from this region. Expression in the pancreas and its excretory duct system continues into the adult. The distributions of gene products from a large number of vertebrate homeobox genes have been described. These were always found to be expressed in the neuroectoderm and/or in the mesoderm, usually in derivatives of both of these germ layers (for example, Oliver et al. 1988) but never in the endoderm.

What might be the function of the long carboxy-terminal domain of XlBox 8? We might draw a comparison between this protein and the even-skipped protein of *Drosophila*. The eve protein has a long carboxy-terminal domain which has been found to change the DNA-binding specificity of the homeodomain lying upstream of it (Hoey & Levine, 1988; Hoey et al. 1988). It may be that the carboxy-terminal region of XlBox 8 acts synergistically with the His change in the recognition helix to give XlBox 8 proteins unique DNA-binding characteristics and hence unique target gene specificities.

XlBox 8 is expressed exclusively in a very narrow slice of the gut endoderm – the posterior part of the foregut. One might speculate that other homeobox genes may be expressed in bands in other regions along the anteroposterior axis of the endoderm. Why have these putative homeobox genes not been detected yet? Hybridization of the standard Antennapedia-type DNA probes to the XlBox 8 homeobox sequence is weak. There is only 67% nucleotide sequence homology between the region of XlBox 8 reported here and Antp, and is probably even less in the first third of the homeobox (not sequenced here) because this region is usually less conserved than the region of the DNA-binding helices. Perhaps an XlBox 8 homeobox probe will allow the isolation of other genes contained within an XlBox 8-type family, expressed in endoderm, by low stringency screening of gene libraries. One potentially useful application of XlBox 8 probes may be in the diagnosis of the site of origin of intestinal tract carcinomas.

Vertebrate endoderm/mesoderm coculture experiments have shown that the type of epithelial differentiation undergone by many portions of the digestive tract is ‘instructed’ by induction by the underlying mesenchyme (Gurdon, 1987; Saxen, 1977; Wessels & Rutter, 1969; Wessels & Cohen, 1967; Rutter et al. 1964). For example, chick allantois (which originates from the posterior endoderm) is capable of differentiating into gizzard-, intestine- or lung-type epithelium when cultured in close contact with mesenchymes from early organs of these types. Similarly, lung epithelium that has not yet been instructed by mesoderm can be induced to form stomach or intestinal epithelium if apposed in culture together with mesoderm obtained from these regions of the gut (Wessels, 1977). XlBox 8 expression is detected in a narrow band of the gut long before the pancreatic anlagen are formed (Fig. 3A). If its expression is triggered by mesodermal induction, then its activation precedes any sign of pancreatic differentiation. After the pancreas is formed, its continued growth requires growth factors (‘permissive induction’) from mesoderm, but these can be provided by a variety of mesenchymes or even cell extracts (Wessells, 1977). Using XlBox 8 probes as sensitive markers should allow experiments addressing the question of whether duodenum differentiation requires ‘instructive induction’ from a specific type of mesoderm, as is the case in other parts of the endodermal tube.

Expression of several homeobox genes in the mesoderm that surrounds the endodermal epithelium has been observed in vertebrate embryos (Holland & Hogan, 1988; Dony & Gruss, 1987; Oliver et al. 1988). Clearly, different homeodomain proteins are expressed along the anteroposterior axis of the gut mesoderm (Oliver et al. 1988; Graham et al. 1988). It is also interesting to note that Slack (1985, 1986) has suggested that the most frequent homeotic transformation in man is digestive tract metaplasia. It is not uncommon, for example, for adult humans to have patches of intestinal epithelium in the stomach or patches of gastric epithelium in the esophagus. XlBox 8 is expressed exclusively in the endoderm, conceivably is involved in the earliest specification of epithelial identity, and might even be a target for an epithelial instruction signal from the mesoderm. This proposition might be tested experimentally by micro-injecting XlBox 8 protein or synthetic mRNAs into the parts of *Xenopus* embryos (vegetal blastomeres) that give rise to endodermal structures, in an attempt to change the epithelial fate. Markers are available (e.g. specific digestive enzymes) which can be used to test the efficacy of such a transformation. Preliminary experiments in which XlBox 8 antibodies were injected into embryos in the hope of causing a phenotypic defect, as described for XlBox 1 antibodies by Cho et al. (1988), have so far failed to produce significant changes in the organization of the tadpole gut.
Endoderm-specific homeodomain protein

XIIHbox 8 may provide an excellent marker for the re-examination and characterization of the mesenchymal/endodermal induction process at the molecular level. Faced with such an intriguing distribution and potentially powerful ways to explore and manipulate the expression of this gene we are now undertaking a complete characterization of the XIIHbox 8 gene, beginning with the cloning of full-length cDNAs.

We acknowledge Jane Hardwicke for expert technical assistance during the sectioning, staining and photographic work, and Clyde Luhum (BioRad Microsciences Division) for image analysis of whole-mount embryos. We thank Larry Zipursky and John Merriam for comments on the manuscript. We recognize the invaluable contribution of the laboratory of M. Klymkowsky for the Xenopus whole-mount procedure. This work was supported by NIH grant HD 21502-03. CVEW is an ACS (California Division) Senior Fellow.

References


(Accepted 16 December 1988)

Note added in proof
While this paper was in press, the distribution of mouse gene Cdx-1, which is most similar to the Drosophila gene caudal, was reported (Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J.-L. & Gruss, P., 1988; A mouse gene homologous to the Drosophila gene caudal is expressed in epithelial cells from the embryonic intestine. Genes Dev. 2, 1647–1654.). The endoderm-specific expression of the Cdx-1 gene is different from that of XlHbox 8 in at least two respects. First, XlHbox 8 is expressed more anteriorly within the gastrointestinal tract than the Cdx-1 gene. Second, Cdx-1 mRNA is first found quite late in intestinal differentiation, when the villi are just about to form. In contrast, XlHbox 8 protein is already found in a restricted endodermal band long before morphogenesis and differentiation of the duodenum and pancreatic anlagen.