Isolation and Characterization of a Xenopus cDNA Which Encodes a Homeodomain Highly Homologous to Drosophila Distal-less

(Received for publication, November 18, 1991) Misaki Asano, Yasufumi Emori, Kaoru Saigo, and Koichiro Shikawa

From the Laboratory of Molecular Embryology, Zoological Institute, Faculty of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan, and the Department of Biophysics and Biochemistry, Faculty of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

A novel homeobox gene of Xenopus was isolated from the ovary cDNA library. The homeodomain of the encoded protein was homologous to that of Drosophila Distal-less (DII). The mRNA exists in a large amount in ovary, and in a small amount in testis, but was not detected in muscle, kidney, gut, and liver. The mRNA also occurs in a large amount in oocytes and is maintained in unfertilized eggs and cleavage stage embryos as a maternal mRNA at a low but distinctly detectable level. The amount of the mRNA per embryo increases gradually in later stages by zygotic expression. Embryo dissection experiment revealed that the transcript is abundant in the anterior part of the embryo at the neurula stage, suggesting that Xdll may play a role in the establishment of the structures in the anterior part of the embryo.

It has been well established in Drosophila that homeobox-containing genes play many important roles in early embryonic development (1, 2). Also in vertebrates, a number of homeobox genes have been isolated and shown to play important roles in various developmental processes (3). In Xenopus laevis, about 20 homeobox genes have been isolated, which are expressed in localized patterns. For instance, Xlbbox1 is expressed exclusively in the anterior region (4), whereas Xxho3 (5, 6) and Xlhox6 (7) are in the posterior region, and Mix.I (8) in the vegetal region of the embryo. Most of these homeobox genes are not expressed maternally, and appear to be involved in the establishment of body axis at specific stages of development. It has been determined that a few homeobox genes like Xlbbox2 (9) and Xxho2A (10) are maternally transcribed, although at very low levels.

In Xenopus laevis embryos, the determination of the body axis starts relatively early during development; the animal-vegetal axis is established during oogenesis (11), and the dorso-ventral axis secondarily determined through the sperm-induced subcortical rotation (12). It may, therefore, be interesting to search for other maternally expressed homeobox genes in Xenopus and clarify their functions, since studies of such early expressed homeobox genes may add useful information for the understanding of the mechanism of body axis determination in more details.

In the present experiment, we isolated a cDNA of a novel homeobox-containing gene which is abundantly expressed in ovary and in oocytes. We describe here the structural features of the gene and some characteristics of the expression pattern in embryos and in adult tissues.

MATERIALS AND METHODS

Isolation and Sequencing of Homeobox-containing cDNA of Xenopus—A mixed degenerated oligonucleotide probe which has the sequence of 5'-AA(AG)AT(AC)TGTTT(CT)(AC)AGAACG(CG- (AGT)(AC)G-3' (25-mers, 192-fold) was synthesized based on an eight-amino acid motif, KIWF(K)INRR, of the helix 3 region of the homeodomain (13) according to the codon usage frequencies of X. laevis.1.2 X. laevis ovary cDNA library constructed in agt10 vector (14, 15) was screened with this probe by the plaque hybridization at 42°C.

To obtain the 5'-region of the cDNA clone, a primer extension cDNA library was constructed in agt10 with ovarioly(A)+ RNA, (15) using a specific primer of 23 bases (Fig. 1a), and then screened with the cDNA probe (the Psfl fragment of X914) (Fig. 1a). DNA fragments were subcloned into pUC19 or BluescriptII SK+ and sequenced by the dideoxy method (16).

Oocytes, Embryos, and Dissection of Embryos—An adult female that had not ovulated for several months was sacrificed, and the ovary was obtained. The ovary was treated with 0.5 mg/ml collagenase for 15 h at 21°C, and then oocytes from stage I to stage VI (17) were collected. X.laevis embryos obtained by injection of a gonadotropic hormone (18) were dejellied and cultured in 1/10 × Steinberg solution. Embryos at the stage 24 (19) were dissected into head, trunk, and tail parts in complete Stearn’s medium which contained 10 µg/ml each penicillin and streptomycin (20).

RNA Extraction and Northern Blot Analysis—Oocytes, embryos, various adult organs, and dissected embryos were crushed in frozen state and mixed with the acid guanidium thiocyanate solution. Total RNAs were prepared as described previously (15). Poly(A)+ RNA was purified by oligo(dT)-cellulose chromatography.

RNA blot analyses were carried out (15) using 10 µg of the total RNA (15). The nylon filters were hybridized in rapid hybridization buffer (Amersham Corp.) with P-labeled X914 insert (Fig. 1a) as a probe at 65°C overnight and finally washed in 0.2 × SSC containing 0.1% sodium dodecyl sulfate at 65°C. As the internal standard to monitor the amount of the applied RNAs, histone H4 cDNA was used as described by Shikawa et al. (21).

RESULTS AND DISCUSSION

Molecular Cloning of a Xenopus Homeobox Gene, Xdll—We used a mixed oligonucleotide probe corresponding to the consensus sequence of homeobox helix 3 region to screen a cDNA library of X. laevis oocytes and obtained several classes of cDNA clones (data not shown). One of these clones, designated X914, was found to include the homeobox sequence, and this was subjected to further analyses. The analysis of nucleotide sequence revealed that the clone encodes a homeodomain similar to that of Drosophila Distal-less (DII) as shown in Fig. 1b, and this Xenopus gene was termed Xdll.

1 K. Brian, personal communication.
2 L. B. Dawid, personal communication.

* This study was supported in part by research grants from the Ministry of Education, Science, and Culture of Japan and a grant from The Fujisawa Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBankTM/EMBL Data Bank with accession number(s) D 10859.
**Figure 1.** Restriction map and primary sequence of Xdll. a, schematic representation and restriction map of the Xdll cDNA. The box indicates the coding region, the shaded area the homeobox, and the asterisks the oligonucleotide used as the probe on the screening (*) and as the primer of the primer extension library (**). The lines indicate the obtained cDNA clone; λB14, λA-7, and λA-8. b, Nucleotide sequence and deduced amino acid sequence of the Xdll cDNA. The amino acid sequence is provided in the single-letter amino acid code. Four possible translation initiation codons are overlined and numbered. Amino acid residues are numbered from the third ATG sequence which is most probable (see text). Distinctive features of the ORF are a homeodomain (shaded box), M-repeat (box), and the histidine-rich region (thin underline). In-frame stop codon (TGCA, thick overline) which precedes the translation initiation site and the conserved eight amino acid motif used for screening (thick underline, see text) are indicated.

(Xenopus homolog of Distal-less).

Since the Xdll clones lacked the 5' part of the coding region, the primer extension cDNA library (see "Materials and Methods" and Fig. 1a) was constructed and screened. Two clones, λA-7 and λA-8, were found to contain the 5'-terminus of the open reading frame preceded by an in-frame termination codon (Fig. 1b). Then, the nucleotide sequence of 1625 base pairs covering the whole coding sequence was determined as presented in Fig. 1b.

**The Structural Features of the Xdll—Sequence analysis (Fig. 1b) revealed that the Xdll cDNA contains a single long open reading frame with 4 ATG codons near its 5' end. The real initiation point has not yet been determined explicitly, but we assume that it may be the third ATG (sequence no. 439), because its surrounding sequence (TGACCATGA) matches the consensus of the translation initiation region of vertebrate genes (23). Thus, the encoded protein would consist of 250 amino acid residues. The size of the Xdll protein deduced here is relatively small among the Xenopus homeobox proteins so far reported but is comparable to the XhoxlA protein (232 amino acids).

The amino acid sequence was examined for its homology to the known sequences of the NBRP protein sequence database (version 29.0) using the FAST-P computer program. Over 150 sequences showed significant homologies to Xdll, and most of them were homeodomain-containing proteins. Drosophila Distal-less (Dll) protein showed the highest homology score, and the sequence is aligned in Fig. 2 together with recently reported sequence of murine Dll gene product (24). In the homodeomain, only 8 and 4 out of 61 amino acid residues in our sequence are different from those in Drosophila Dll and mouse Dll, respectively. Moreover, these differences are all included in conservative changes in the Dayhoff's criteria and are mostly located outside the helices 2 and 3, which are reportedly critical for the specificity of DNA binding.

The N-terminal region of Xdll protein (residues 1-124) contains a short M-repeat (25) and a stretch rich in histidine, 

FIG. 2. Alignment of the amino acid sequences of the homoeodomains from Xdll, Dlx, and Dll. The dashes indicate positions where amino acids are shared with Xdll. The regions of the α-helices of the homoeodomain are indicated.
which are both general features of homeodomain-containing proteins (Fig. 1b). High Ser content (14%) in Xdill is also a common feature of some homeodomain-containing proteins, as has been reported for FoxH1 (26). Xdill, however, does not have other characteristic features such as the PRD repeat (27), the long acidic region (28), and IYPWM pentapeptide in the N-terminal border of the homeodomain (14).

Expression of Xdill Transcript in Oocytes, Embryos, and Adult Tissues—The expression of the Xdill gene was examined by Northern blot analysis. Fig. 3 shows that the size of the mRNA obtained is ~1.8 kb irrespective of the stage of the eggs and embryos. This size is in a good agreement with the size deduced from the sequence analysis (1.6 kb, Fig. 1b) if we take the poly(A) sequence into account. This finding supports that the sequence obtained in Fig. 1b may cover nearly the full length of the Xdill mRNA.

Based on the results of the sequence analysis (Fig. 1b) and the Northern blotting (Fig. 3), we reach the conclusion that the sizes of the mRNAs of three Dll-related genes are significantly different from one another; that for Xdill may be 1.8 kb, whereas that for Dll is reported to be 4.5 kb, and that for Dlx 2.5 and 3.4 kb. This suggests that the structures of these proteins are considerably different in the regions outside the homeodomains.

Fig. 3a shows that Xdill is expressed at a high level in oocytes, and the mRNA persists at low but distinct levels in unfertilized eggs and in early embryos. The level of histon H4 mRNA, which was examined at the same hybridization experiment and serves as the internal standard of the added RNA, is more or less constant throughout the stages (Fig. 3a). The amount of Xdill mRNA per embryo appears to increase gradually in later stages (about twice at the tadpole stage compared with that in unfertilized eggs), probably by zygotically expression. We have also examined with poly(A)+ RNA extracted from embryos of these stages and got almost the same pattern of expression (data not shown).

The spatial distribution of the Xdill transcript was studied at the neurula stage (Fig. 3b). Late neurulae were dissected into three parts (head, trunk, and tail), and the RNAs were subjected to Northern blot analysis. As shown in Fig. 3b, the Xdill transcript existed predominantly in the head part. It is possible that the anterior-abundant expression may be due to the new expression of Xdill gene in the central nervous system, just as was the case with Dlx (24), although further analyses will be needed.

Fig. 3c shows the results of the examination of the expression of Xdill in adult tissues. It is clearly shown here that transcription of the Xdill gene mainly occurs in the ovary. A small amount of mRNA was detected also in testis, but not in muscle, kidney, gut, or liver. The almost exclusive expression of Xdill in gonads in adult animals suggests that Xdill may have some important function in germ cell development, especially in female animals.

The expression of Xdill in gonads observed here is quite unique for Xdill, since neither Dll nor Dll is expressed maternally or in oocytes. It has been reported that Dll is expressed zygotically in the limb primordia and in the optic center of the larval brain of Drosophila embryos (24, 29), whereas Dlx is expressed in the restricted region of the brain in early mouse embryos (24). Thus, Xdill appears to have some special functions in the process of oogenesis and/or very early stage of development, in addition to the common features of temporal and spatial expression of Dll and Dlx. We assume that the elucidation of these unique aspects of Xdill gene expression may help us understand the mechanism which controls the early phases of Xenopus embryogenesis.

Acknowledgments—We thank Drs. Igor B. Dawid and Kay Brian for kindly providing the codon usage tables for Xenopus laevis. We also thank Dr. Rosuke Tashiro, Satoshi Ichimaru, and Gen Fujiwara for their warm encouragement throughout the work.

REFERENCES
Xenopus Homeobox Gene Related to Drosophila Distal-less