Vol. 190, No. 1, 1993 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
January 15, 1993 Pages 270-276

A NOVEL CHICKEN HOMEOBOX-CONTAINING GENE EXPRESSED IN NEURULATING EMBRYOS

José Serrano[2], Louis Scavo[3], Jesse Roth[4], Enrique J. de la Rosa[1] and Flora de Pablo[1,*]

[1] Centro de Investigaciones Biológicas, C.S.I.C., Velázquez 144, E-28006 Madrid, Spain

[2] East Carolina University School of Medicine, Greenville, NC 27858

[3] U.C.S.F., San Francisco, CA 94143

[4] Johns Hopkins University School of Medicine, Baltimore, MD 21224

Received November 12, 1992

We have isolated a cDNA clone from chicken embryo that contains a homeobox sequence (*CHox-cad2*). Analysis at the nucleotide and amino acid levels revealed closest similarity to the *Xenopus Xcad1* and to other homeoboxes related to *Drosophila caudal*. RNA blot analysis showed hybridization of *CHox-cad2* to two transcripts of 2.6 and 1.5 kb, present at day 1 of embryogenesis (E1). Using the highly sensitive polymerase chain reaction (PCR) to amplify cDNAs from embryonic RNAs from E0 (unincubated blastoderm) to E4, we confirmed the restricted expression of this homeobox sequence to the period of neurulation (E1).

The understanding of the molecular mechanisms that control morphogenesis has progressed enormously with the study of genes containing the conserved 180-bp homeobox first discovered in *Drosophila* (1). The hypothesis that the DNA-binding proteins encoded by homeobox-containing genes also regulate pattern formation in vertebrate embryos is supported by the recent finding of their restricted expression in time and space (2). The idea is also emerging that important steps of the body plan may be established by the interplay of peptide growth factors and homeobox-gene products (3,4).

Homeobox genes in vertebrates appear to be organized in highly conserved clustered arrangements (*Hox* complexes) plus a few apparently scattered genes. Typically, each group of vertebrate genes, including multiple genes in each species, is related to one of the homeobox genes in *Drosophila*. The *caudal* group appears to be one of the most conserved since genes with this type of homeobox have

*To whom correspondence should be addressed.

0006-291X/93 $4.00

been isolated from *C. elegans* (5), *Drosophila* (6), *Xenopus* (7), chicken (8) and mouse (9). Expression of *caudal*-related genes may be important for gastrulation and early induction events in vertebrates (7, 8, 9).

We have isolated a cDNA containing a homeobox similar to members of the *caudal* group (*CHox-cad2*). Its restricted pattern of expression in chicken embryos provides an additional evidence for a possible role of proteins of this group during gastrulation and neurulation. It also provides a tool to explore potential interactions with growth factors expressed in the early chicken embryo (10, 11).

MATERIALS AND METHODS

Chicken embryos. Fertilized White Leghorn eggs were incubated at 37.5°C, 60-90% relative humidity in an egg incubator. Embryos were staged, according to Hamburger and Hamilton (12), after dissection from membranes.

cDNA library screening and sequence analysis. A stage HH10 (30-36 h of development) cDNA chicken embryo library (kindly provided by Dr. R. M. Grainger, University of Virginia, Charlottsville, VA) had been prepared by cloning the cDNAs from the embryonic mRNAs into the *Eco* RI site of *lambda* gt10 bacteriophage. A total of 10^6 phages were plated on 150X25 mm dishes (Falcon, Lincoln Park, NJ) and double replica filters were prepared (Nitrocellulose BA85, Schleicher and Schuell, Keene, NH). The filters were treated sequentially with i) 0.5 M NaOH, 1.5 M NaCl for 2 min; ii) 0.5 M Tris-HCl, pH 8.0, 1.5 M NaCl for 3 min; iii) 2X SSC for 5 min (1X SSC consists of 0.15 M NaCl, 15 mM sodium citrate); followed by baking at 80°C for 20 min. Prehybridization was carried out at 50°C for 2 h in 5X SSPE (1X SSPE consists in 0.18 M NaCl, 10 mM sodium phosphate, pH 7.7, 1 mM EDTA), 0.5X Denhardt's solution (1X Denhardt's solution consists in 0.2% polyvinylpyrrolidone, 0.2% Ficoll, 0.2% bovine serum albumin), 0.1% SDS and 100 μg/ml ssDNA. Subsequently, 10^6 cpm/ml of a random-primed, [^{32}P]-labeled probe was added and the filters were hybridized at 50°C overnight. The filters were then washed twice for 15 min at 50°C in 5X SSPE containing 0.1% SDS, followed by two additional washes for 20 min at 50°C in 2X SSPE with 0.1% SDS. Under these non-stringent conditions, the cDNA fragment of rat insulin-like growth factor I (kindly provided by Dr. Derek LeRoith, National Institutes of Health, Bethesda, MD) hybridized to the homeobox-containing cDNA. The reason for cross-hybridization with the homeobox clone is not clear, but we suspect that it was due to the presence of multiple short stretches of sequence similarity (there is no known ancestral relationship between homeobox and peptide growth factor genes). After secondary screening, the strongly hybridizing clones were subjected to restriction analysis, and the appropriate fragment of one of them (named p3) was partially characterized and found to contain a homeobox motif (to be reported elsewhere). We rescreened then the same library using p3 as a probe, and obtained a second cDNA clone, p*CHox-cad2*, which was subcloned into the *Eco* RI site of p*Gem 4z* (Promega, Madison, WI). Recombinant plasmids containing the approximately 1-kb insert cDNA were sequenced as dsDNA by the dideoxynucleotide method (13) using commercial kits (K/RT, Promega; Sequenase, USBC, Cleveland, OH). The sequence obtained was compared to those in the EMBL bank using the GCG programs.

RNA isolation from embryonic tissues and Northern blot analysis. Total RNA was prepared by the method of Cathala et al. (14). Polyadenylated RNA was prepared by affinity chromatography using poly(dT)-cellulose (BRL, Gaithersburg, MD), according to the manufacturer's protocol. RNA was resolved by

electrophoresis in 1% agarose gels in the presence of formaldehyde and transferred to nylon filters (Gene Screen, Dupont, Boston, MA) by capillary blotting. The filters were probed with a [^{32}P]-end labeled 48-base oligonucleotide (5'TTGTTTACACAGACCATCAGAGACTAGAATTAGAGAAGGAATTTCACT3') (primer U, corresponding to nucleotides 17 to 64 of the homeobox in Fig. 1) or with a random-primed [^{32}P]-labeled cDNA from p*CHox-cad2*. The filters were prehybridized and hybridized at 37°C overnight in 50% formamide, 4X SSPE, 1X Denhardt's solution, 1% SDS, followed by sequential washes in i) 4X SSPE, 0.1% SDS; ii) 2X SSPE, 0.1% SDS; iii) 1X SSPE, 0.1% SDS; 15 min each at 37°C. The filters were exposed to XAR-2 film (Eastman Kodak, Rochester, NY) with intensifying screens (Cronex lighting-plus, Dupont) at -70°C, for approximately 24 h.

Reverse transcription-polymerase chain reaction (RT-PCR). Total or poly(A)$^+$ RNAs from chicken embryos at different stages of development were analyzed for expression of *CHox-cad2* by using the polymerase chain reaction with reverse transcriptase-synthesized cDNA as described (15). In short, cDNAs were synthesized by reverse transcriptase (30 U/μl, Boehringer Mannheim, Indianapolis, IN) in the presence of RNAsin (Promega), 5X RT buffer and deoxinucleotides, using a 48mer (primer D) (5'AGATACTGCTGAAGATTCAAGCAAACCTAATTCTAGTCTGCTCATGTA3') complementary to a 3' untranslated region of the *CHox-cad2*, 422 bases downstream of the homeobox. One third of the cDNA mixture was processed for amplification, using the kit GeneAmp (Perkin-Elmer, Cetus Corp., Emeryville, CA), following the protocol of the manufacturer. The oligonucleotide primers used were the above described, U and D. Thirty cycles of denaturation (90°C, 1 min), annealing (50°C, 1.5 min) and extension (70°C, 2 min) were performed. One tenth of the amplified product was electrophoresed on a 2% agarose gel in Tris-borate-EDTA buffer. The DNA was transferred then to Nytran filters (Schleicher and Schuell) and it was hybridized with the [^{32}P]-labeled random-primed insert of p*CHox-cad2*.

RESULTS AND DISCUSSION

Isolation and characterization of the homeobox of *CHox-cad2*.

We have cloned a homeobox-containing cDNA from a stage HH10 chicken embryo library that, by comparison with previously identified genes, shows the

```
AAA GAG AAG TAC CGG GTT GTT TAC ACA GAC CAT CAG AGA CTA GAA[45]
Lys Glu Lys Tyr Arg Val Val Tyr Thr Asp His Gln Arg Leu Glu

TTA GAG AAG GAA TTT CAC TGC AAC AGA TAC ATT ACA ATA AGG AGA[90]
Leu Glu Lys Glu Phe His Cys Asn Arg Tyr Ile Thr Ile Arg Arg

AAG TCA GAA CTT GCA GTA AAC CTT GGA CTA TCT GAA AGA CAG GTA[135]
Lys Ser Glu Leu Ala Val Asn Leu Gly Leu Ser Glu Arg Gln Val

AAA TCC TGG TTC CAG AAT CGC CGA GCC AAA GAA AGA AAA ATA ATC[180]
Lys Ser Trp Phe Gln Asn Arg Arg Ala Lys Glu Arg Lys Ile Ile
```

Figure 1. Nucleotide and derived amino acid sequence of the homeobox of *CHox-cad2*.

Caudal homeoboxes

```
Chicken CHox-cad2      KEKYRVVYTDHQRLELEKEFHCNRYITIRRKSELAVNLGLSERQVKSWFQNRRAKERKII
Xenopus Xcad-1         *D********Q**********YS********A**********T****I**********N
Mouse Cdx-1            *D*S*****************YS**********A*****T*****I**********VN
Chicken CHox-cad       *D*****************YS********A***AA***T*****I**********VN
Xenopus Xcad-2         *D*****************YS********A***AA***T*****I**********VN
Drosophila caudal      *D********F********YCTS**********QT*S*******I**********TSN
C. elegans ceh-3       AD***M**S*Y**********TSPF**SD***Q*STM*S*T***I*I********D*RDK
```

Figure 2. Alignment of derived amino acid sequences of the *caudal*-related homeoboxes of several species. The single letter code is used. * indicates identity.

highest similarity to the group of *caudal*. Since it is different from another chicken *caudal* homologue recently reported (*CHox-cad*), we have named it *CHox-cad2*. The nucleotide and derived amino acid sequences of the homeobox of *CHox-cad2* are shown in Fig. 1. The open reading frame in p*CHox-cad2* continued 90 amino acids upstream of the homeobox. An in-frame translational stop codon was found 35 amino acids downstream of the homeobox (data not shown).

At both the nucleotide and amino acid levels, *CHox-cad2* homeobox was remarkably similar to the *caudal*-related homeoboxes, with 78% of identity in amino acids with the *Drosophila* gene (Figs. 2 and 3). The closest similarity was found with the *Xenopus Xcad-1* (87% of amino acids identical), while the most distant was the *C. elegans ceh-3* (62%). The similarity was less with any other group of homeoboxes. Chicken is the only species besides *Xenopus* (that has a tetraploid genome and, thus, many duplicated genes) in which two *caudal*-related cDNAs have been reported to date. This fact may suggest that there are several groups of *caudal* homeobox genes in vertebrates. This phenomenon of multiplicity of vertebrate homeobox-containing genes relative to the *Drosophila* counterpart is well established for the *Hox*-complexes (1, 2).

	CHox-cad2	Xcad-1	Cdx-1	CHox-cad	Xcad-2	caudal	ceh-3
CHox-cad2	*						
Xcad-1	87	*					
Cdx-1	85	88	*				
CHox-cad	83	90	95	*			
Xcad-2	83	90	95	100	*		
caudal	78	82	82	82	82	*	
ceh-3	62	63	67	67	67	67	*

Figure 3. Percentage of identity between derived amino acid sequences from *caudal*-related homeoboxes.

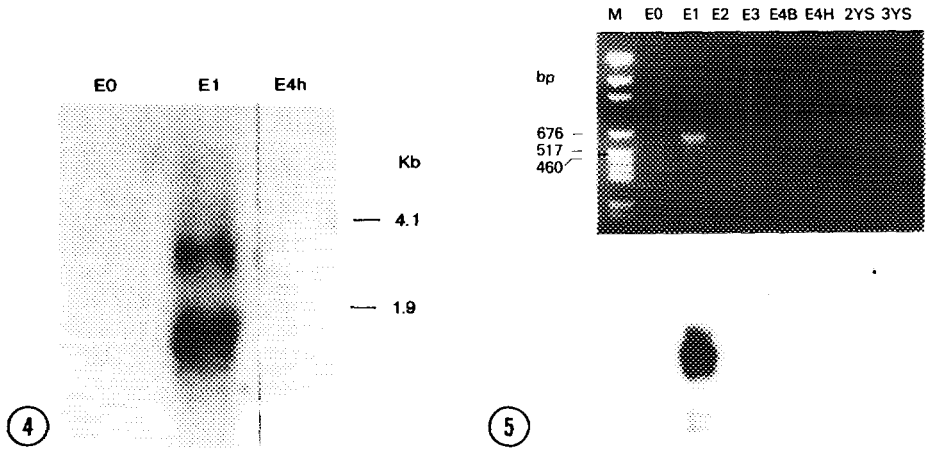

Figure 4. Expression of *CHox-cad2* during early embryogenesis. Four μg of poly(A)+ RNA from whole chicken embryos before incubation (E0), after 1 day (E1), and from heads of E4 embryos (E4h) was fractionated in a gel and transferred to a nylon filter as described in Materials and Methods. The blot was hybridized to the [^{32}P]-labeled 48-base oligonucleotide U, described in Materials and Methods.

Figure 5. *CHox-cad2* expression analysis by RT-PCR. Five μg of poly(A)+ RNA from E0 and E1 whole embryos and 5 μg of total RNA from E3 whole embryos, from E4 heads (E4h) and bodies (E4b), and from E2 yolk sac (2YS) and E3 yolk sac (3YS), was used for reverse-transcription followed by PCR amplification of the cDNAs. The amplified fragments were photographed under UV light (upper panel) and transferred to a nylon membrane. This blot was hybridized to an *Eco* RI fragment of p*CHox-cad2* as described in Materials and Methods.

Expression of *CHox-cad2* in early chicken development is restricted to E1 embryos.

Since the cDNA had been isolated from a library of stage HH10 (30-36 h) chicken embryos, we knew that the mRNA was expressed very early in embryogenesis. An RNA blot hybridized with a 48mer complementary to the p*CHox-cad2* sequence (described in Materials and Methods as primer U) showed transcripts of 1.5 and 2.6 kb in E1 embryos. No expression was detected at the time the egg is laid (E0, blastoderm of a few thousand cells) or in heads of E4 embryos (Fig. 4). Other stages of embryogenesis (whole embryo RNA from E2, E3, E4 and E8) did not show any expression when RNA blots were hybridized with the full-length insert of p*CHox-cad2* (data not shown). To confirm this restricted temporal expression with a technique of much higher sensitivity than RNA blot hybridization, we used RT-PCR with RNAs extracted from embryos from E0 to E4 and from yolk sac membranes of E2 and E3. After hybridization of the amplified products, only the E1 embryo showed expression of *CHox-cad2* (Fig. 5). The integrity of the RNAs was confirmed by the positive amplification of unrelated genes, CTK-1 and CTK-2, in all samples (16).

The expression of *CHox-cad2* corresponds temporally to the period of late gastrulation and early neurulation (our E1 RNA was extracted from embryos at stages HH6 to HH10). The time of expression partially overlaps, but is more restricted than the expression of both *Xenopus Xcad-1* and *Xcad-2* (7) and the other chicken *caudal*-related gene, *CHox-cad* (8). These three homeobox genes have been shown to be highly expressed in the embryonic Spemann's organizer (the blastopore lip in *Xenopus* and Hensen's node in chicken). The developmental stages at which we detect expression of *CHox-cad2* are compatible with its possible expression in the embryonic organizer, although future studies using "in situ" hybridization will be needed.

Interestingly, despite the finding that *caudal* transcripts in *Drosophila* are first detected as maternal mRNA that forms a gradient from the posterior pole (6), no maternal mRNA from either *Xcad-1* or *Xcad-2* has been found in the unfertilized *Xenopus* egg (7). In the chicken embryo, *CHox-cad2* transcripts found in E1 are not likely to be maternal, since they are not detected at E0.

The reported expression for another member of the group, mouse *Cdx-1*, first detected at E14 by "in situ" hybridization in the epithelial lining of the intestine (9), overlaps neither with *CHox-cad2* expression nor with the expression patterns of other vertebrate *caudal*-related homeobox genes (7, 8). This suggests, again, the existence of multiple *caudal*-related genes in vertebrates, which may have evolved from an ancestral gene by duplication and divergence. Each new gene within a species may have restricted its expression and/or its role to a narrower time and region.

ACKNOWLEDGMENTS

We thank Dr. Robert M. Greinger for providing the chick embryo library and Drs. Catherine McKeon and Margarita del Val for useful comments on the manuscript. These studies were partially supported by Fondo de Investigaciones Sanitarias (Spain, grant 91/218). The experimental part was done in the Diabetes Branch, NIDDK, National Institutes of Health, Bethesda, Maryland. The sequence is accessible in GenBank with the number L06621.

REFERENCES

1 McGinnis, W., and Krumlauf, R. (1992) Cell 68, 283-302.
2 Wright, C. V. E. (1991). Current Opinion Cell Biol. 3, 976-982.
3 Ruiz i Altaba, A., and Melton, D. A. (1989) Nature 341, 33-38.
4 Jessell, T. M., and Melton, D. A. (1992) Cell 68, 257-282.
5 Bürglin, T. R., Finney, M., Coulson, A., and Ruvkun, G. (1989) Nature 341, 239-243.
6 Mlodzid, M., and Gehring, W. J. (1987) Cell 48, 465-478.

7 Blumberg, B., Wright, C. V. E., De Robertis, E. M., and Cho, K. W. Y. (1991) Science 253, 194-196.
8 Frumkin, A., Rangini, Z., Ben-Yehuda, A., Gruenbaum, Y., and Fainsod, A. (1991) Development 112, 207-219.
9 Duprey, P., Chowdhury, K., Dressler, G. R., Balling, R., Simon, D., Guenet, J.- L.,and Gruss, P. (1988) Genes Dev. 2, 1647-1654.
10 De Pablo, F., and Roth, J. (1990) TIBS 15, 339-342.
11 Mitrani, E., Ziv, T., Thomsen, G., Shimoni, Y., Melton, D. A., and Bril, A. (1990) Cell 63, 495-501.
12 Hamburger, V., Hamilton, H. L. (1951) J. Morphol. 88, 49-92.
13 Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
14 Cathala, G., Savouret, J. F., Mendez, B., Wets, B. L., Karin, M., Martial, J. A., and Baxter, J. D. (1983) DNA 2, 329-335.
15 Serrano, J., Shuldiner, A. R., Roberts, C. T. Jr, Le Roith, D., and De Pablo, F. (1990) Endocrinology 127, 1547-1549.
16 Scavo, L. M., Serrano, J., Roth, J., and De Pablo, F. (1991) Biochem. Biophys. Res. Commun. 1393-1401.