Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 5306–5310, August 1987
Developmental Biology

Primary structure and nuclear localization of a murine homeodomain protein

(homeobox/hexapeptide/F9 cells/cDNA/immunofluorescence)

MICHAEL KESSEL*, FRANK SCHULZE*, MATHIAS FIBI*, AND PETER GRUSS*

Center for Molecular Biology of the University of Heidelberg (ZMBH), Im Neuenheimer Feld 282, 6900 Heidelberg, Federal Republic of Germany

Communicated by Manfred Eigen, March 13, 1987

ABSTRACT **The murine homeobox *Hox 1.1* (m6) is the first of a cluster of six boxes on chromosome 6. Using probes and synthetic peptides derived from the *Hox 1.1* sequence, we were able to isolate cDNAs and antibodies that allowed us to characterize the product of this homeobox-containing gene. From the open reading frame on the cDNA clone B21, a protein could be predicted, made up of 229 amino acids and having a calculated molecular weight of 25,740. A unique feature of this protein is that it has 15 glutamic acid residues as its carboxyl terminus, which gives it a very hydrophilic and acidic carboxyl terminal structure, most probably folding onto an α-helix. A second domain of six amino acids is present on the Hox 1.1 protein, which is conserved in other homeodomain proteins. Antibodies generated against synthetic peptides from the homeobox region were used in the immunoblotting procedure and revealed a major protein band of M_r 31,000 in extracts from 3T3 cells and F9 teratocarcinoma cells induced by retinoic acid and cAMP. The nuclear location of the protein was established by immunofluorescence. The presence of this protein in F9 cell nuclei is in faithful accordance with the kinetics established for the 2.4-kilobase Hox 1.1 transcript during differentiation into parietal endoderm cells.**

Homeoboxes are highly homologous regions of genes controlling *Drosophila* development (for a review, see ref. 1). They code for a strongly conserved domain of 60 amino acids showing a high proportion of basic residues, which can be predicted to fold into a helix-turn-helix secondary structure, as originally discovered for some prokaryotic DNA-binding proteins (2). In the fruit fly *Drosophila melanogaster*, the function of homeobox-carrying genes was first established by genetic analysis. In many cases it could be deduced from mutant phenotypes that homeobox genes are involved in the proper organization of the body plan, such as the determination of segment identity and number, compartmentalization, or polarity. Since the recognition of homeoboxes in the fruit fly genome, numerous homologous sequences have been found in DNA from other animals. Thus, earthworms, leeches, beetles, shrimps, sea urchins, tunicates amphioxus, frogs, chickens, mice, and humans reveal homeobox homology on genomic blots, whereas bacteria, fungi, sea anemonies, planaria, tapeworms, nematodes, and squid appear to lack such sequences (3).

The study of mammalian homeobox genes has been approached by applying the *Drosophila* boxes as molecular probes to identify and isolate genomic sequences that then could be used to analyze their expression in various cell lines and tissues. Two clusters of boxes (*Hox-1* and *Hox-2*) have been described on the murine chromosome 6 (4) and on chromosome 11 (5). In addition, there are apparently singular boxes on chromosome 1 (*en-1*; ref. 6), chromosome 5 (*en-2*), and chromosome 15 (*Hox-3*; refs. 7 and 8). The "*Hox-1*" locus on chromosome 6 contains in a region of about 70 kilobases (kb) a cluster of six homeoboxes designated *Hox 1.1–Hox 1.6* (refs. 4, 9, and 10; unpublished results). The expression of several boxes in this cluster has been studied in differentiating teratocarcinoma cells and in adult and embryonal tissues (4, 9, 10, 11).

Steps toward a functional analysis require the characterization of the gene products—i.e., the proteins with a homeobox-encoded domain. We approached the analysis of these proteins by manufacturing cDNA clones and antibodies against peptides deduced from the cDNA sequence. In this study we report the complete coding sequence of the gene containing the murine homeobox *Hox 1.1* (previously *m6*) and its splicing pattern in this region. We also describe the use of antibodies against Hox 1.1 peptides to identify the protein in cell extracts by immunoblotting, to establish its nuclear location by immunofluorescence, and to study its expression during F9 cell differentiation. A comparison of the predicted Hox 1.1 amino acid sequence with other available sequences from homeodomain-containing proteins reveals the presence of a second conserved domain.

MATERIALS AND METHODS

cDNA Cloning. RNA was prepared by the guanidinium thiocyanate method (12) and selected on oligo(dT)-cellulose. cDNA was prepared and cloned essentially as described (13).

Antibodies. The two peptides Tyr-Gln-Thr-Leu-Glu-Leu-Glu-Lys-Glu-Phe-His-Phe-Asn-Arg-Tyr-Leu (peptide I) and Glu-His-Lys-Asp-Glu-Ser-Gln-Ala-Pro-Thr-Ala-Ala-Pro-Glu-Asp (peptide II) were synthesized by Nova Bachem (Babendorf, Switzerland) using the solution method. The synthetic peptides were coupled to the carrier protein keyhole limpet hemocyanin as described (14). Anti-peptide antisera were prepared from BALB/c mice using standard procedures.

For electrophoretic immunoblotting, cell lysates derived from equal numbers of cells were run on a 12.5% NaDodSO$_4$/polyacrylamide gel and then transferred to nitrocellulose filters as described (15). The filters were blocked with 0.2% Tween in phosphate-buffered saline (PBS), incubated with the first antibody, washed with 0.2% Tween in PBS and with 0.1% Triton and 0.2% Tween in PBS, and incubated with ^{125}I-labeled rabbit anti-mouse Ig antibodies. For immunofluorescence, cells were grown on gelatinized glass coverslips and fixed with 4% paraformaldehyde. The anti-peptide antisera were visualized with sheep anti-mouse Ig conjugated to Texas red, the anti-laminin antibodies with fluorescein isothiocyanate-coupled sheep anti-rabbit Ig antibodies. For quantification, the percentage of antigen-expressing cells was

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

*Present address: Department of Molecular Cell Biology, Max Planck Institute of Biophysical Chemistry, 3400 Göttingen-Nikolausberg, Federal Republic of Germany

calculated from 500 single cells counted under the microscope.

RESULTS

Primary Structure of the Hox 1.1 Protein as Deduced from cDNA. RNA from F9 cells grown for 24 hr in medium containing retinoic acid and cAMP had previously been shown to contain a considerable amount of the 2.4-kb Hox 1.1 transcript (9) and was used to prepare a cDNA library in phage λgt10. Clones were isolated that hybridized to two genomic fragments previously shown to pick up the Hox 1.1 transcript in blot-hybridization analysis of F9 cell RNA. Specifically, a genomic 1.3-kb *Sac* I–*Eco*RI fragment (Fig. 1), which is located 1.6 kb 5′ of the *Hox 1.1* box, and a second, box-specific fragment (*Eco*RI–*PVU* II; Fig. 1) that contains two-thirds of the box plus 25 base pairs (bp) downstream of it, were used as nick-translated probes. cDNA clone B21 contained the longest insertion with an internal *Eco*RI site expected for a *Hox 1.1* clone. The *Eco*RI fragments (613 bp and 349 bp) were mapped on the genomic phage λ clone λm6 (9) and sequenced by the dideoxy method. The sequence analysis revealed several features that show that the clone contains the coding information for the complete Hox 1.1 protein (Fig. 2). An intron of about 1.1 kb could be identified by comparison with the genomic sequence, indicating that the box is located on the last exon. Eight base pairs upstream of the box, perfect splice consensus (TTTCTGTTCCTAG↓G) is used, and also the splice donor on the second to last exon uses sequences of the splice consensus (CAG↓GT; ref. 16).

A single long open reading frame of 687 bases is present on B21 containing the homeobox in the 3′ portion of the cDNA. The first AUG codon after two in-frame stop codons is the most probable initiator codon. Like 75% of the eukaryotic mRNAs, the Hox 1.1 transcript also contains an adenosine at the position three bases upstream of the AUG, the only position showing a high degree of conservation around the initiator codon (17). The two AUG codons inside the reading frame contain a pyrimidine (T) at the equivalent position. A stop codon at position 787 terminates the open reading frame, leaving an untranslated region of 102 bp. A putative polyadenylylation signal (AATAAA; ref. 18) is located 5 bp before the site of linker addition. Clone B21 has no poly(A) tail; only a single adenosine residue is present at its 3′ end.

```
1   GGGTT TGGTG TAAAT CTGGG GGTGT AATGT TATCA TATAT CACGC TACCT

50  CGTAA AACCG ACACT GAAAG CTGCC GGACA ACAAA TCACA GGTCA AAATT

100 ATG AGT TCT TCG TAT TAT GTG AAC GCG CTT TTT AGC AAA TAT ACG
    Met Ser Ser Ser Tyr Tyr Val Asn Ala Leu Phe Ser Lys Tyr Thr

145 GCG GGG GCT TCT CTC TTC CAA AAT GCC GAG CCG ACT TCT TGC TCC
    Ala Gly Ala Ser Leu Phe Gln Asn Ala Glu Pro Thr Ser Cys Ser

190 TTT GCA CCC AAC TCG CAG AGA AGC GGC TAC GGG CCG GCG CCG GCC
    Phe Ala Pro Asn Ser Gln Arg Ser Gly Tyr Gly Pro Ala Pro Ala

235 GCC TTC GCC TCC ACT GTG CCG GGC TTA TAC AAT GTC AAC AGC CCC
    Ala Phe Ala Ser Thr Val Pro Gly Leu Tyr Asn Val Asn Ser Pro

280 CTC TAT CAG AGC CCC TTC GCG TCC GGC TAT GGC CTG GGA GCC GAC
    Leu Tyr Gln Ser Pro Phe Ala Ser Gly Tyr Gly Leu Gly Ala Asp

325 GCC TAC AAC CTG CCC TGC GCC TCC TAC GAC CAA AAC ATC CCC GGG
    Ala Tyr Asn Leu Pro Cys Ala Ser Tyr Asp Gln Asn Ile Pro Gly

370 CTC TGC AGT GAC CTC GCC AAA GGC GCC TGC GAC AAG GCG GAC GAG
    Leu Cys Ser Asp Leu Ala Lys Gly Ala Cys Asp Lys Ala Asp Glu

415 GGC GTG CTT CAC GGC CCG GCC GAA GCC AGT TTC CGC ATC TAC CCC
    Gly Val Leu His Gly Pro Ala Glu Ala Ser Phe Arg Ile Tyr Pro

460 TGG ATG CGC AGT TCA GGA CCC GAC AGG AAG CGG GGA CGC CAG ACC
    Trp Met Arg Ser Ser Gly Pro Asp Arg Lys Arg Gly Arg Gln Thr

505 TAC ACG CGC TAC CAG ACG CTG GAA CTG GAG AAG GAA TTC CAT TTC
    Tyr Thr Arg Tyr Gln Thr Leu Glu Leu Glu Lys Glu Phe His Phe

550 AAC CGC TAC CTG ACG CGG CGC CGC CGC ATC GAG ATC GCT CAC GCG
    Asn Arg Tyr Leu Thr Arg Arg Arg Arg Ile Glu Ile Ala His Ala

595 CTC TGC CTC ACT GAG CGC CAG ATC AAG ATC TGG TTC CAG AAT CGG
    Leu Cys Leu Thr Glu Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg

640 CGC ATG AAG TGG AAG AAA GAG CAT AAA GAT GAG AGC CAG GCT CCC
    Arg Met Lys Trp Lys Lys Glu His Lys Asp Glu Ser Gln Ala Pro

685 ACT GCA GCC CCG GAA GAC GCG GTG CCC TCC GTT TCC ACA GCT GCT
    Thr Ala Ala Pro Glu Asp Ala Val Pro Ser Val Ser Thr Ala Ala

730 GAC AAG GCG GAC GAG GAG GAA GAG GAG GAA GAG GAG GAA GAA GAA
    Asp Lys Ala Asp Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu

775 GAG GAA GAG GAG TAAAG GGCCA GGCAC AGGAC CCTGG CTGCA CAGGA
    Glu Glu Glu Glu

822 CAGTT GGAAA AGCGT CTTTA AGAGA CTCAT TGATT TTAGT TACAA AAATG

872 GGGGG AAATA AAGTG TA
```

FIG. 2. Primary structure of cDNA clone B21 containing the murine homeobox Hox 1.1. The amino acid sequence deduced from the open reading frame including the homeobox is given below the DNA sequence. The start and stop codons and the termination signal are underlined, the splice site is indicated by a triangle, and the homeobox sequence is boxed.

The deduced protein sequence indicates a protein consisting of 229 amino acids—128 amino acids at the N terminus,

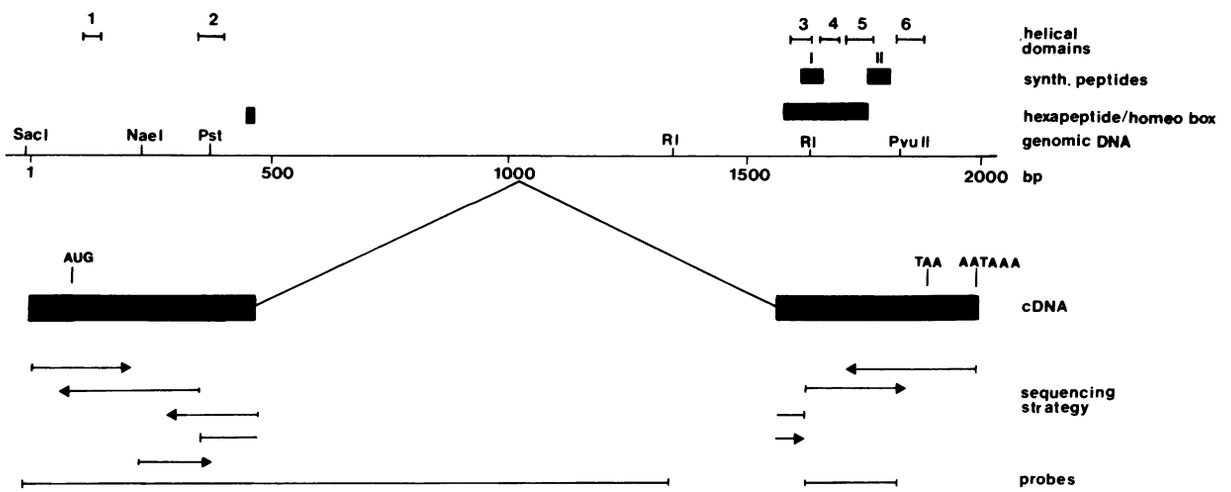

FIG. 1. Schematic representation of the described structural elements. From top to bottom the following structural features are outlined: the helical domains of the predicted Hox 1.1 protein, the location of the synthetic peptides used for immunization, the location of the Hox 1.1 homeobox and the hexapeptide on the genomic DNA and the cDNA, the genomic structure as deduced from the phage λ clone m6 (9), the structure of the F9-derived cDNA clone B21 showing the two exons as black bars, the sequencing strategy, and the probes used to isolate the cDNA. The base pairs are numbered starting with nucleotide one from the cDNA clone B21 and continuing through the intron.

followed by the homeobox (60 amino acids), followed by 41 amino acids of which the 3'-terminal portion consists of a peptide containing 15 glutamic acid residues. The calculated molecular weight is 25,740.

Immunological Detection of the Hox 1.1 Protein. To assay the protein predicted by means of the cDNA, two polyclonal antisera were generated by immunization of BALB/c mice with synthetic peptides coupled to keyhole limpet hemocyanin. The sequence of peptide I was deduced from parts of the homeobox, while peptide II was selected because it is translated from unique sequences located 3' of the conserved box region prior to the stop codon (Fig. 1). When tested in an ELISA assay, neither antiserum cross-reacted with bovine serum albumin, and both were only specific for the injected peptide. The serum specific for peptide I proved to be applicable to the detection of a protein by immunofluorescence or blotting; the data we show and discuss were obtained with this serum. On the other hand, several independently prepared sera against peptide II gave only low intensity of fluorescence and negative results in the electrophoretic immunoblot analysis. One possible explanation for this result is that the antigenic peptide is embedded between the two α-helices 5 and 6 (Fig. 1), which may explain its lower accessibility in the protein.

To identify the Hox 1.1 protein, we applied the immunoblotting technique to F9, differentiated F9, and NIH 3T3 (hereafter 3T3) whole-cell protein lysates. A single protein was only detected in extracts from cells known to produce Hox 1.1 RNA (that is, in differentiated F9 and 3T3 cells), while the F9 stem cells did not contain this protein (Fig. 3). The molecular weight determined from the gel is 31,000. In this analytical system, no difference could be detected in the protein from F9 and that from 3T3 cells (Fig. 3).

RNA analysis had established the time course of *Hox 1.1* expression in the murine teratocarcinoma cell line F9 after induction of differentiation by retinoic acid and cAMP (8, 9). We applied our sera to F9 cells at various stages of differentiation and stained the cells with a second antibody coupled to Texas red. As a control, laminin-detecting antibodies were chosen, which gave rise to a characteristic fluorescence only in the cytoplasm of differentiated F9 cells. The fluorescence in F9 cells could be detected 24 hr after induction in about 80% of the cells, increased up to 48 hr in both percentage of positive cells and intensity of fluorescence and decreased quite remarkably in all cells 74 hr and 96 hr after induction (Fig. 4). The anti-Hox 1.1 activity of the sera could be eliminated by preadsorption with specific peptides. In comparison, the laminin signal rose from zero in undifferentiated cells to a stable plateau after growth for 48–96 hr in medium containing retinoic acid and cAMP. The time course of expression closely followed the data obtained on the RNA level (8, 9). The immunofluorescence analysis also clearly showed that the Hox 1.1 protein was detected exclusively in the cell nucleus of both differentiated F9 and 3T3 cells (Fig. 4).

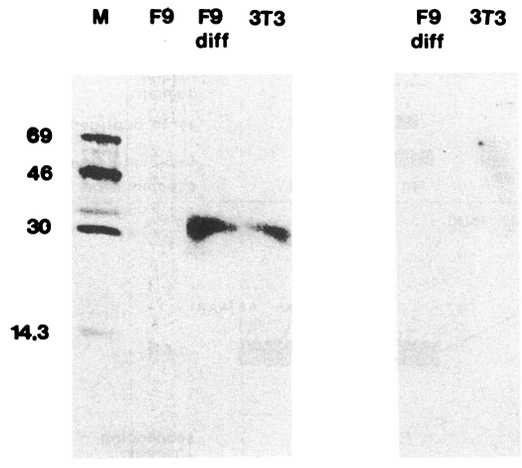

FIG. 3. Detection of the Hox 1.1 homeobox protein in F9 and 3T3 cells. Protein extracts from identical numbers of F9 cells, F9 cells grown in retinoic acid/cAMP-supplemented medium for 24 hr (lanes F9 diff), and 3T3 cells were separated on a 12.5% $NaDodSO_4$/polyacrylamide gel and analyzed by using polyclonal antibodies specific for peptide I. The two lanes on the right were treated with preimmune serum. ^{14}C-labeled protein markers (indicated as $M_r \times 10^{-3}$) were run and transferred in parallel (lane M).

DISCUSSION

The availability of antibodies against synthetic peptides with sequences deduced from the *Hox 1.1* gene has made possible the detection of a protein in expressing cells by immunofluorescence as well as in protein extracts by an immunoblotting procedure. Peptide I antibodies are directed against a region that is highly conserved, also in other homeoboxes. However, in the F9 cell system, the Hox 1.1 transcript seems to be the most dominant RNA, at least when compared with Hox 1.2 (m5), Hox 1.5 (mo10), and Hox 3.1 (hox 3, m31). Another quite abundant homeobox RNA in these cells is the 1.9-kb Hox 1.3 (m2) transcript, which was originally assigned to the *Hox 1.2 (m5)* box (4) and which shares 15 out of 16 amino acids with peptide I. However, this transcript is expressed in F9 cells also after differentiation for 72 or 96 hr, when the Hox 1.1 RNA is already downregulated. While cross-reactivity cannot be excluded, we assume that the immunological detection is directed against the Hox 1.1 protein. The kinetics of the expression in F9 cells and the singular protein detected in the Western blot corroborate this assumption. In this context it is of interest that peptide I antibodies detect antigens in those embryonal mouse tissues in which also a Hox 1.1-specific RNA can be detected by *in situ* hybridization (F.S., unpublished data; C. Dony, personal communication).

The staining by immunofluorescence demonstrated the location in the nucleus of the homeodomain protein. This observation represents yet another step in establishing a relationship to genes in the *Drosophila* system, such as Ultrabithorax (*Ubx*), fushi tarazu (*ftz*), and engrailed (*en*), where protein products have also been shown to reside in the nucleus (19–21). Although it seems likely that the Hox 1.1 protein has DNA-binding capacity, this ability still remains to be demonstrated.

The calculated size of the Hox 1.1 protein (M_r 25,740) is remarkably smaller than found for *Drosophila* homeodomain proteins, which have $M_r > 40,000$ in the case of Antennapedia (*Antp*), *Ubx*, and *ftz* (2, 22, 23). It is apparent in both systems, however, that a relatively large RNA codes for the respective protein, thus leaving room for long leader or trailer regions at the 5' and 3' ends. For example, The *Antp* RNAs (3.2, 3.4, 4.6, and 4.8 kb) possess an open reading frame of only 1042 bases (22), and the *Ubx* transcript of 3.2 kb contains 1100 bases in the open frame (23). In comparison, the 2.4-kb Hox 1.1 transcript contains 687 bp without a stop codon. Thus, at most only one-third of the RNA is used for coding purposes. The remaining sequences may have regulatory functions, perhaps including differential splicing [which has already been shown for *Antp* and *Ubx* (22, 23)] or RNA stability.

In this context, it is important to know whether the end of the cDNA clone B21 does indeed represent the 3' end of the message. We have analyzed other clones showing more than the one adenosine residue present on B21, but because on the genomic sequence 12 contiguous adenosine residues are present, it cannot be determined at present whether the adenosines on B21 are transcribed or polyadenylylated. The

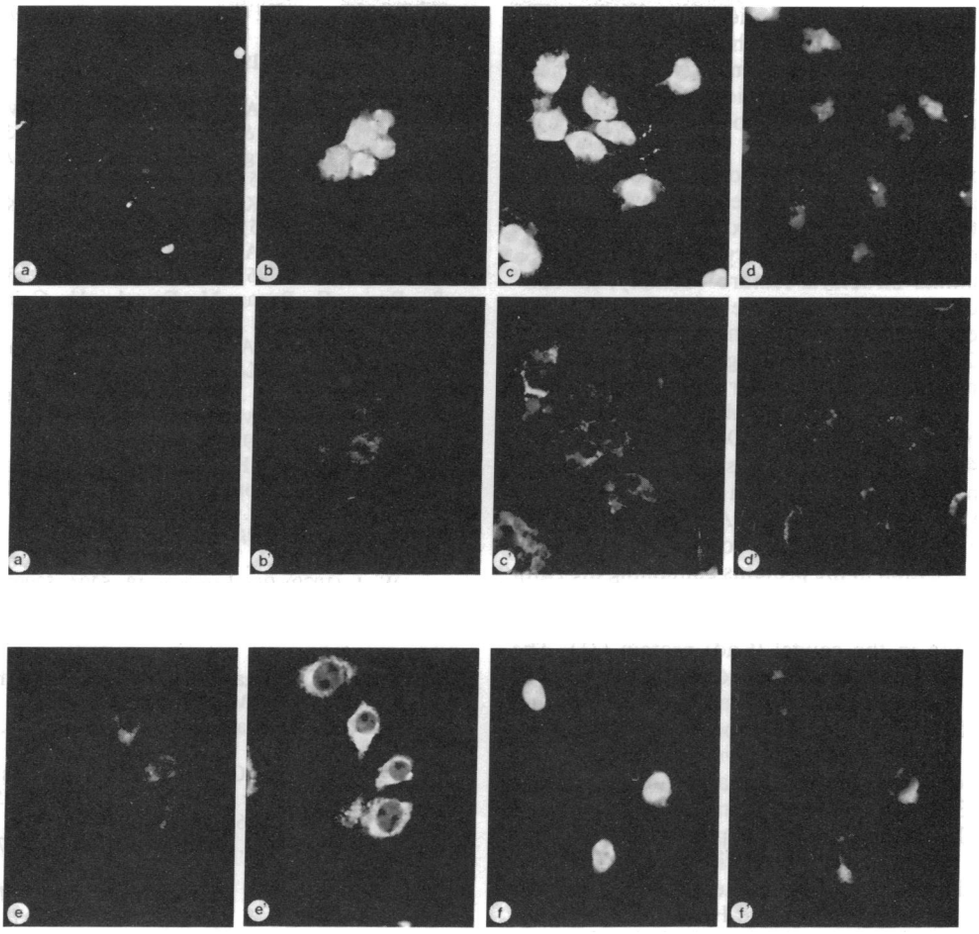

FIG. 4. Detection of the Hox 1.1 protein and laminin in 3T3 cells and in F9 cells during differentiation using an immunofluorescence double-staining method. The cells were fixed before staining with paraformaldehyde and permeabilized by 0.1% Triton X-100. The staining was performed by incubating the cells with peptide I-specific polyclonal mouse antibodies, which were detected by Texas red-coupled sheep anti-mouse Ig antibodies. The second staining of the identical cells was performed with rabbit anti-laminin antiserum and anti-rabbit Ig antibodies labeled with fluorescein isothiocyanate. The photographs of the second stain are marked with a prime. (*a*) Hox 1.1 expression in undifferentiated F9 cells. (*a'*) Laminin expression in undifferentiated F9 cells. (*b*) Hox 1.1 expression in 24-hr differentiated F9 cells. (*b'*) Laminin expression in 24-hr differentiated F9 cells. (*c*) Hox 1.1 expression in 48-hr differentiated F9 cells. (*c'*) Laminin expression in 48-hr differentiated F9 cells. (*d*) Hox 1.1 expression in 96-hr differentiated F9 cells. (*d'*) Laminin expression in 96-hr differentiated F9 cells. (*e*) 3T3 cells incubated with preimmune serum. (*e'*) Laminin expression in 3T3 cells. (*f*) Hox 1.1 expression in 3T3 cells. (*f'*) Laminin expression in 3T3 cells.

cDNA cloning procedure used may preferentially prime internally at these adenosines, or the nuclease S1 may find here a suitable structure that is cleaved and where linkers are inserted. Nuclease S1 analysis of F9 RNA (data not shown) shows a protected fragment mapping the 3′ end to the position indicated by the cDNA clone. In contradiction to this finding, we could detect the transcript also by downstream genomic probes.

The product of the *Hox 1.1* gene reveals several characteristics that seem unusual. We assume that one or more of these structural features may explain also the difference between the experimentally determined M_r 31,000 and the calculated M_r 25,740. On the other hand, post-translational modification might explain the size difference. The region preceding the box is rich (10%) in proline residues; therefore, no α-helices can be formed, except possibly from regions 1 and 2 (Fig. 1); also the hydropathy plot shows a rather hydrophobic character. It may be of significance that the

Hexapeptide		Homeo domain
Ile-Tyr-Pro-Trp-Met-Arg	--- 5 amino acids ---	Hox 1.1
Leu-Tyr-Pro-Trp-Met-Arg	--- 8 amino acids ---	Antp
Ile-Tyr-Pro-Trp-Met-Lys	--- 16 amino acids ---	Dfd
Val-Tyr-Pro-Trp-Met-Lys	--- 15 amino acids ---	Xhox-1
Trp-Met-Lys	--- 15 amino acids ---	Cad

FIG. 5. The conserved hexapeptide and its location in relation to the homeodomain. The sequences for the different proteins were taken from this paper and refs. 22, 24, 30, and 31. For Xhox-1 it is not proven that a splice occurs upstream of the box.

Antp protein (22), the ftz protein (2), and the Xhox-1 protein (24) also contain about 10% prolines. This region is followed by the hydrophilic homeobox with homology to the typical helix-turn-helix structure of prokaryotic DNA-binding proteins and a predominance of basic amino acids (17 arginine or lysine residues out of 60 amino acids). One striking feature marks the carboxyl terminus of the Hox 1.1 protein in the form of a stretch of 15 glutamic acid residues, which add a very acidic counterpart to the basic box region and, furthermore, fold into another helical secondary structure (region 6, Fig. 1). A very similar structural feature—namely, six consecutive glutamic acid residues at the end of the homeobox protein—has been deduced both from the human cDNA clone HHO.c1 (25) and the *Xenopus* genomic clone MM3 (26). Regions consisting of only one, or predominantly one, amino acid are described for a number of homeotic or homeodomain-containing proteins in *Drosophila* and other organisms (22, 24, 27, 28, 29). Possibly such a feature relates to a common function in the regulation of development.

A hexapeptide (Ile-Tyr-Pro-Trp-Met-Arg) located near the end of the first exon of the Hox 1.1 protein is conserved in sequence and localization in the proteins containing the Antp (22), the Dfd (30), and the Xhox-1 homeobox (ref. 24; Fig. 5). A possible relic of this peptide (Trp-Met-Lys) lies at a comparable location in the caudal (Cad) protein (31). The consensus emerging from this sequence comparison postulates an aliphatic amino acid (Ile/Val/Leu) in the first position, then a strictly conserved tetrapeptide (Tyr-Pro-Trp-Met) and a basic residue (arginine or lysine) at position six. Of the seven protein sequences available for comparison [Antp (22), Dfd (30), Cad (31), ftz (2), engrailed (29), Xhox-1 (24), and Hox 1.1], the hexapeptide is present in five; thus, the conservation spans the large evolutionary distance from the fruit fly over the frog to the mouse. Recent results indicate that this peptide is also present in the Hox 1.3 (m2) protein (data not shown). The conservation also includes the localization; it is always encoded at the end of the second-to-last exon, always on a different exon than the homeobox. The only two proteins lacking this sequence, ftz (2) and engrailed (29), have in common that they are encoded by segmentation genes. It may also be of importance that engrailed and Cad are homeoboxes only about 50% homologous to the Antp prototype.

At least on the basis of the Drosophila proteins, it can be assumed that the hexapeptide is present in the products of truly homeotic genes with a prototype (Antp) homeobox. We speculate that the conserved peptide represents a small homeotic domain either cooperating with or fulfilling its function independently of the homeobox.

After submission of this manuscript, a paper by Mavilio *et al.* (32) described the presence of the conserved hexapeptide in the human homeodomain proteins c13 (Val-Tyr-Pro-Trp-Met-Lys), c1 (Ile-Tyr-Pro-Trp-Met-Arg), and c8 (Ile-Tyr-Pro-Trp-Met-Gln).

We are grateful to Sabine Quick for excellent technical assistance, Rosemary Franklin for help with the manuscript, Uwe Drescher for computer analysis, Anamaris M. Colberg-Poley for discussions, and Eberhard Pfaff for helpful advice on the protein-coupling technique. The work was supported by the Deutsche Forschungsgemeinschaft (SFB 229-86).

1. Gehring, W. J. (1985) *Cell* **40,** 3–5.
2. Laughon, A. & Scott, M. P. (1984) *Nature (London)* **310,** 25–31.
3. McGinnis, W. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 263–270.
4. Colberg-Poley, A. M., Voss, S. D., Chowdhury, K., Stewart, C., Wagner, E. & Gruss, P. (1985) *Cell* **43,** 39–45.
5. Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. (1985) *Cell* **43,** 9–18.
6. Joyner, A. L., Kornberg, T., Coleman, K. G., Cox, D. R. & Martin, G. (1985) *Cell* **43,** 29–37.
7. Awgulewitsch, A., Utset, M. F., Hart, C. P., McGinnis, W. & Ruddle, F. (1986) *Nature (London)* **320,** 328–335.
8. Breier, G., Bucan, M., Francke, U., Colberg-Poley, A. M. & Gruss, P. (1986) *EMBO J.* **5,** 2209–2215.
9. Colberg-Poley, A. M., Voss, S., Chowdhury, K. & Gruss, P. (1985) *Nature (London)* **314,** 713–718.
10. Wolgemuth, D. J., Engelmyer, E., Duggal, R. N., Gizang-Ginsberg, E., Mutter, G. L., Ponzetto, C., Vivano, C. & Zakeri, Z. F. (1986) *EMBO J.* **5,** 1229–1235.
11. Ruddle, F. H., Hart, C. P., Awgulewitsch, A., Fainsod, A., Utset, M., Dalton, D., Kerk, N., Rabin, M., Ferguson-Smith, A., Fienberg, A. & McGinnis, W. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 277–284.
12. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
13. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning: A Practical Approach*, ed. Glover, D. (IRL, Oxford), pp. 49–78.
14. Pfaff, E., Mussgay, M., Böhm, H. O., Schulz, G. E. & Schaller, H. (1982) *EMBO J.* **2,** 869–874.
15. Towbin, H., Staehelin, J. & Gordon, J. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 4350–4354.
16. Padgett, R. A., Grabowski, P. J., Konarska, M. M. & Sharp, P. (1985) *Trends Biochem. Sci.* **10,** 154–157.
17. Kozak, M. (1986) *Cell* **44,** 283–292.
18. Proudfoot, N. J. & Brownlee, G. G. (1976) *Nature (London)* **263,** 211–214.
19. White, R. A. H. & Wilcox, M. (1985) *EMBO J.* **4,** 2035–2043.
20. Carroll, S. B. & Scott, M. P. (1985) *Cell* **43,** 47–57.
21. DiNardo, S., Kuner, J. M., Theis, J. & O'Farrell, P. H. (1985) *Cell* **43,** 59–69.
22. Schneuwly, S., Kuroiwa, A., Baumgartner, P. & Gehring, W. J. (1986) *EMBO J.* **5,** 733–739.
23. Hogness, D. S., Lipshitz, H. D., Beachy, P. A., Peattie, D. A., Saint, R. B., Goldschmidt-Clermont, M., Gavis, E. R. & Helfand, S. L. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 181–194.
24. Harvey, R. P., Tabin, C. J. & Melton, D. A. (1986) *EMBO J.* **5,** 1237–1244.
25. Bonicinelli, E., Simeone, A., LaVolpe, A., Faiella, A., Fidanza, V., Acampora, D. & Scotto, L. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 301–306.
26. Mueller, M., Carrasco, A. E. & De Robertis, E. M. (1984) *Cell* **39,** 157–162.
27. Wharton, K. A., Johansen, K. M., Xu, T. & Artavanis-Tsakonas, S. (1985) *Cell* **43,** 567–581.
28. Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. & McGinnis, W. (1985) *Cell* **43,** 71–80.
29. Poole, P. J. W., Kauvar, L. M., Drees, B. & Kornberg, T. (1985) *Cell* **40,** 37–43.
30. Laughon, A., Carrol, S. B., Storfer, F. A., Riley, P. D. & Scott, M. P. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50,** 253–262.
31. Mlodzik, M., Fjose, A. & Gehring, W. (1985) *EMBO J.* **4,** 2961–2969.
32. Mavilio, F., Simeone, A., Giampaolo, A., Faiella, A., Poiana, G., Russo, G., Peschle, C. & Boncinelli, E. (1986) *Nature (London)* **324,** 664–668.