

The homeo domain of a murine protein binds 5' to its own homeo box

(mouse development/protein-DNA binding/DNA binding protein blots/gel retardation/fusion protein)

ABRAHAM FAINSD*¹, LEONARD D. BOGARAD*², TARMO RUUSALA¹, MARTIN LUBIN², DONALD M. CROTHERS¹, AND FRANK H. RUDDLE*

Departments of *Biology and ¹Chemistry, Yale University, New Haven, CT 06511; and ²Department of Microbiology, Dartmouth Medical School, Hanover, NH 03755

Contributed by Frank H. Ruddle, August 29, 1986

ABSTRACT Nuclear protein extracts from day 12.5 mouse embryos were used to study protein binding to DNA sequences 5' of the Hox 1.5 homeo box. Embryos of this developmental stage are known to express this gene. DNA binding protein blotting and retardation gel techniques show that murine embryonic nuclear proteins specifically bind a 753-base pair (bp) DNA fragment from the region upstream of the Hox 1.5 homeo box. A fusion protein containing the Hox 1.5 homeo domain constructed in λ gt11 also binds the same 753-bp DNA fragment. Specific binding of the fusion protein to the upstream DNA fragment shows that the homeo box contains the sequences required for specific protein-DNA interactions, and the 753-bp fragment contains a homeo domain binding site. These results support the hypothesis that murine homeo boxes are DNA binding domains of proteins involved in the regulation of embryonic development.

Several *Drosophila* homeotic and segmentation genes involved in embryonic development contain an 180-bp (base-pair) conserved sequence known as the homeo box (1-4). These homeo box sequences exist in about 20 loci in the *Drosophila* genome. *Xenopus*, chicken, mouse, and man were also found to contain approximately 20 homeo box sequences (5). The homeo box portion of the coded protein, termed the homeo domain (3), possesses regions of homology with well characterized prokaryotic DNA binding proteins, such as the λ repressor, the *lac* repressor, *cro*, and eukaryotic DNA binding proteins, such as the yeast mating type proteins (6-9). It is believed that these proteins all share a common helix-turn-helix DNA binding motif. One helix (helix 2) interacts electrostatically with the phosphate backbone of DNA, while the other (helix 3) makes specific hydrogen bonds with the bases of the major groove of B-DNA (10). This helix-turn-helix region of the homeo domains shows the highest degree of conservation of the homeo box sequence. It has been predicted that the homeo domain of the protein specifically binds DNA on the basis of its sequence conservation, structural properties, and similarities to known DNA binding proteins (2).

The protein products of several *Drosophila* homeo box genes have been localized in the nucleus (11-14). The protein coded by the *Drosophila* engrailed (*en*) gene has been reported to specifically bind to the upstream region of its own gene as well as to the upstream region of the fushi tarazu (*ftz*) gene (15). The homeo domain of the *en* gene product mediates this function. In mammalian systems, the lack of mutations affecting the homeo box genes limits their developmental analysis. However, the function of the homeo domain can be studied on a molecular level. We present here the results of

protein-DNA binding experiments using the mouse Hox 1.5 (*Mo10*) (16) gene that demonstrate the existence of (i) a specific DNA binding property of the Hox 1.5 protein product and (ii) a binding site for this protein in the 5' region of Hox 1.5.

MATERIALS AND METHODS

Enzymes. Restriction enzymes, *Escherichia coli* DNA polymerase large fragment (Klenow fragment) and T4 DNA ligase were purchased from commercial sources [New England Biolabs, Promega Biotec (Madison, WI), and International Biotechnologies (New Haven, CT)].

Breeding of Mice and Dissection of Embryos. CD1 outbred mice (Charles River Laboratories) were used in this study. A successful mating was determined the next morning by the presence of a vaginal plug. This was considered day 0.5 of gestation. Dissection of embryos was done as described (17).

Preparation of Nuclear Extracts. Mouse embryos (10-13) were homogenized in 5 ml of buffer 1 (0.25 M sucrose/50 mM Tris Cl, pH 7.5/25 mM KCl/5 mM MgCl₂), and the cells were collected at 500 \times g in a refrigerated centrifuge. The cells were resuspended and lysed in 5 ml of buffer 2 (10 mM Hepes, pH 8.0/50 mM NaCl/0.5 M sucrose/0.1 mM EDTA/0.5% Triton X-100/1 mM dithiothreitol/5 mM MgCl₂). The nuclei were pelleted at 500 \times g as above. The pellet was resuspended in buffer 2 at 7 \times 10⁷ nuclei per ml. Spermidine was added to a concentration of 5 mM, and NaCl was added to 0.5 M. The suspension was incubated on ice for 45 min and then centrifuged for 10 min in the microcentrifuge. The supernatant was dialyzed against buffer 3 (10 mM Hepes, pH 8.0/1 mM MgCl₂/1 mM dithiothreitol/50% (vol/vol) glycerol/50 mM NaCl) overnight. The dialysate was spun again for 10 min in the microcentrifuge, and the supernatant was divided into aliquots and stored at -20°C.

Protein Blotting and DNA Binding Protein Blots. Protein blots were prepared with 50 μ g of mouse embryo nuclear extracts as previously described (18)—the only modification being that 10% NaDodSO₄/polyacrylamide gels were routinely used. For the detection of DNA-binding proteins, the procedure of Miskimins *et al.* (18) was followed with the omission of nonfat dry milk in the binding buffer. Washes were performed in 50 mM NaCl.

Construction of the Fusion Protein. A 92-amino acid open reading frame was inserted into the expression vector λ gt11 as described. The construct contains 4 amino acids upstream of the homeo box, the complete Hox 1.5 homeo box, and 28 amino acids downstream of the homeo box. The plasmid pMo10 (16) was digested with *Bst*NI, and the ends were end-filled with Klenow fragment in the presence of all four

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.

Abbreviations: kb, kilobase(s); bp, base pair(s); HDBS, protein homeo domain binding site(s); gtHB, β -galactosidase-Hox 1.5 homeo domain fusion protein.

nucleotides (19). *EcoRI* linkers (CGGAATTCGG, New England Biolabs) were ligated and then digested with *EcoRI*. The homeo box-containing fragment with *EcoRI* ends was gel-purified and ligated to *EcoRI*-digested λ gt11 DNA (20) (Promega Biotec; see Fig. 4). After *in vitro* packaging, initial infections were performed on *E. coli* Y1090 r^- (Promega Biotec), and lysogens were finally prepared in *E. coli* Y1089 r^- . Screening of the clones was performed by protein blotting and immunodetection of the fusion proteins with antibodies directed against β -galactosidase followed by Southern blotting and probing with the antennapedia gene homeo box to confirm the identity and orientation of the insert.

Preparation of Fusion Protein. Lysogens of the construct λ gtHB were used to prepare crude fusion protein extracts as described by Huynh *et al.* (21) with the only modification that 5 mM isopropyl β -D-thiogalactopyranoside (Sigma) was used.

Retardation Gels. Gel-purified DNA fragments were end-labeled using Klenow fragment of DNA polymerase and [α - 32 P]dCTP (3000 Ci/mmol; 1 Ci = 37 GBq; Amersham). Retardation gel analysis was performed as previously described (22, 23) with the following modifications: The binding buffer used was 50 mM NaCl/10 mM MgCl₂/10 mM Tris Cl, pH 8.0/0.5 mM EDTA/0.1 mM dithiothreitol/4% (vol/vol) glycerol. The 4% acrylamide gels were run in Tris/borate buffer (19).

RESULTS

The restriction map of the Hox 1.5 genomic region is shown in Fig. 1. The direction of transcription has been deduced from the coding sequence and the genomic orientation of the Hox 1.5 homeo box. As indicated in the *Introduction*, previous studies by Desplan *et al.* (15) have shown that the 5'-flanking DNA sequences of the *Drosophila* homeo box gene engrailed (*en*) contain a binding site for the encoded engrailed protein. Assuming the adjacent presence of a binding site to be a general feature of homeo box genes, we screened the 5' region of the murine homeo box Hox 1.5 gene for a Hox 1.5 protein binding site. To accomplish this, we divided the 5' region of the Hox 1.5 gene, and then tested the DNA subfragments for binding to (i) mouse embryo nuclear proteins and (ii) to a fusion protein containing the Hox 1.5 homeo domain.

DNA binding protein blotting technique was first used to identify the specific DNA binding region. Nuclear protein extracts were prepared from day-12.5 mouse embryos, which by RNA analysis had been previously shown to express the Hox 1.5 gene (24). DNA binding protein blot analysis of DNA fragments from phage λ Mo10 revealed a 1.37-kb (kilobase) *HindIII* fragment that was reproducibly bound by four mouse embryo proteins of approximately 39, 45, 68, and 95 kDa in molecular mass (Fig. 2, lane A). Nonspecific binding by histones served as a quantitative control. No other DNA fragments tested from phage λ Mo10 showed specific binding to mouse embryo nuclear proteins. To refine the location of the DNA binding site(s) the 1.37-kb *HindIII* DNA fragment was digested with the restriction enzyme *Ava I*, resulting in

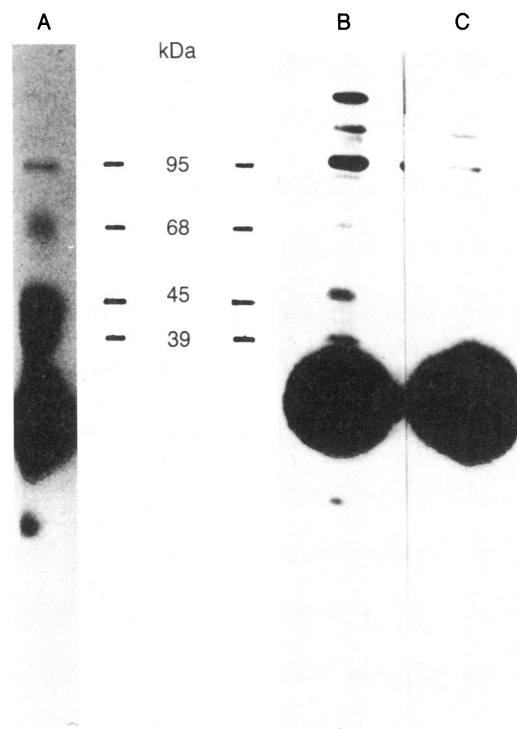


FIG. 2. Southern/immunologic blot analysis of DNA fragments upstream from the Hox 1.5 homeo box. Nuclear protein extracts from day-12.5 mouse embryos (50 μ g) were fractionated on NaDodSO₄/polyacrylamide gels and transferred onto nitrocellulose. The different strips were tested for specific DNA binding with the 1.37-kb *HindIII* fragment (lane A), the 753-bp (lane B), and the 616-bp (lane C) *Ava I-HindIII* fragments. The nonspecific binding by histones served as a quantitative control.

two fragments of sizes 753 and 616 bp. Analysis of the two fragments by DNA binding protein blotting is shown in Fig. 2. The same four proteins specifically and reproducibly bind to the 753-bp fragment (lane B), while the 616-bp fragment showed no specific binding by these proteins (lane C).

The binding specificity was further studied by gel retardation of protein-DNA complexes. The results of using the 753-bp *Ava I-HindIII* DNA fragment as the target DNA, together with nuclear proteins from day-12.5 mouse embryos are shown in Fig. 3. At low protein concentrations only one complex is formed (C1). As the protein concentration is increased, a larger complex appears (C2). Note that several complexes are formed, which suggests that more than one protein or proteins are binding simultaneously to the fragment. The degree of DNA binding specificity was measured competitively by comparing the amounts of poly[d(I-C)] \cdot poly[d(I-C)], or unlabeled 753-bp fragment necessary to reduce the radioactive complex by one-half. Competition by the nonspecific polymer was 10³- to 10⁴-fold weaker compared with the specific fragment (data not shown).

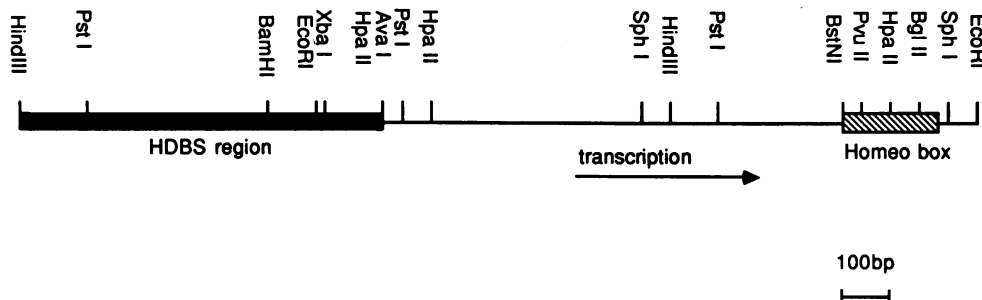


FIG. 1. Restriction map of the Hox 1.5 homeo box region. The homeo box (hatched bar) and the DNA fragment containing the homeo domain binding site (HDBS) (solid bar) are shown. The direction of transcription is indicated by the arrow.

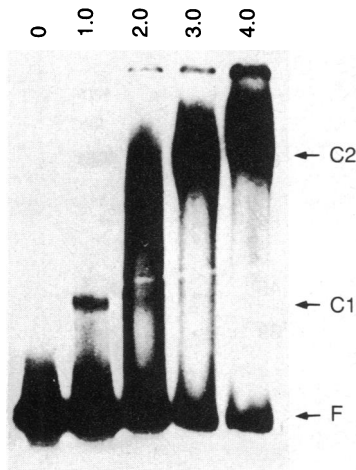


FIG. 3. Polyacrylamide retardation gel analysis of protein-DNA complexes between the 753-bp fragment and mouse embryo nuclear extracts. The 753-bp DNA fragment was incubated with 0, 1.0, 2.0, 3.0, or 4.0 μ g of day-12.5 mouse nuclear proteins in the presence of poly[d(I-C)] \cdot poly[d(I-C)] and then loaded directly onto a nondenaturing acrylamide gel. The arrow labeled F indicates the position of free DNA, C-1 is the first protein-DNA complex, and C-2 is the second protein-DNA complex.

Moreover, mouse embryo extracts show no specific binding to the 616-bp fragment, and nuclear mouse extracts derived from tissues that do not express the Hox 1.5 gene—i.e., adult liver, show no specific binding to either the 753- or the 616-bp fragment (data not shown). In agreement with the DNA binding protein blotting results only the 753-bp fragment

shows the formation of specific DNA-protein complexes. Both series of experiments taken together map the binding site for the embryo proteins to the 753-bp fragment indicated in Fig. 1.

In the described experiments, mouse embryo nuclear extracts were used under the assumption that the Hox 1.5 gene product would be among the proteins specifically binding to Hox 1.5 DNA. A β -galactosidase-Hox 1.5 homeo domain fusion protein (gtHB) was constructed to examine this possibility (Fig. 4A). The construct was shown to contain the homeo box in the correct orientation and reading frame by immunologic analysis using antibodies directed against β -galactosidase (Fig. 4B) and Southern transfer analysis (data not shown).

To test whether the upstream DNA fragment does indeed contain the homeo domain binding site (HDBS) for the Hox 1.5 protein homeo domain, the fusion protein was tested by the retardation gel technique. Fig. 5A shows the results of such an experiment using the two *Ava* I-*Hind*III fragments. The gtHB protein binds only the upstream fragment. To ensure that the homeo domain of fusion protein gtHB is responsible for this binding the upstream fragment was tested for binding with gt11 extract without the homeo box insert (Fig. 5B); three-fold excess of this λ gt11 extract had no binding activity, which shows that the homeo domain is necessary for specific DNA binding. These results map the HDBS to the upstream *Ava* I-*Hind*III fragment and support the model that DNA binding is the basis of the biochemical function of the protein's homeo domain.

DISCUSSION

The sequence homology found between the homeo domains and several DNA binding proteins suggests that several genes involved in *Drosophila* development encode DNA binding

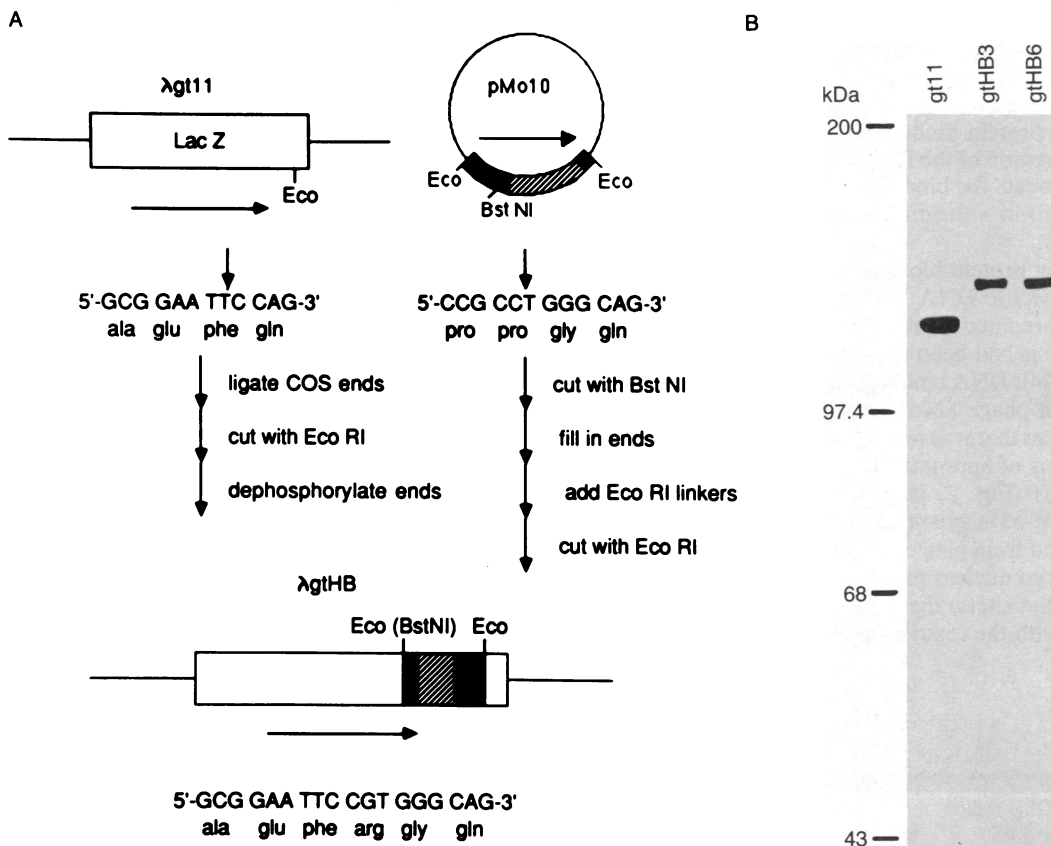


FIG. 4. The galactose-Hox 1.5 homeo domain fusion protein (gtHB). (A) Schematic summary of the construction of the Hox 1.5 fusion protein. (B) Immunoblot analysis of the fusion protein with antibodies directed against β -galactosidase. Crude extracts of λ lysogens were used (5 μ g). The extract of a λ gt11 lysogen and two identical independent homeo domain- β -galactosidase fusion protein constructs (gtHB3 and gtHB6) are shown.

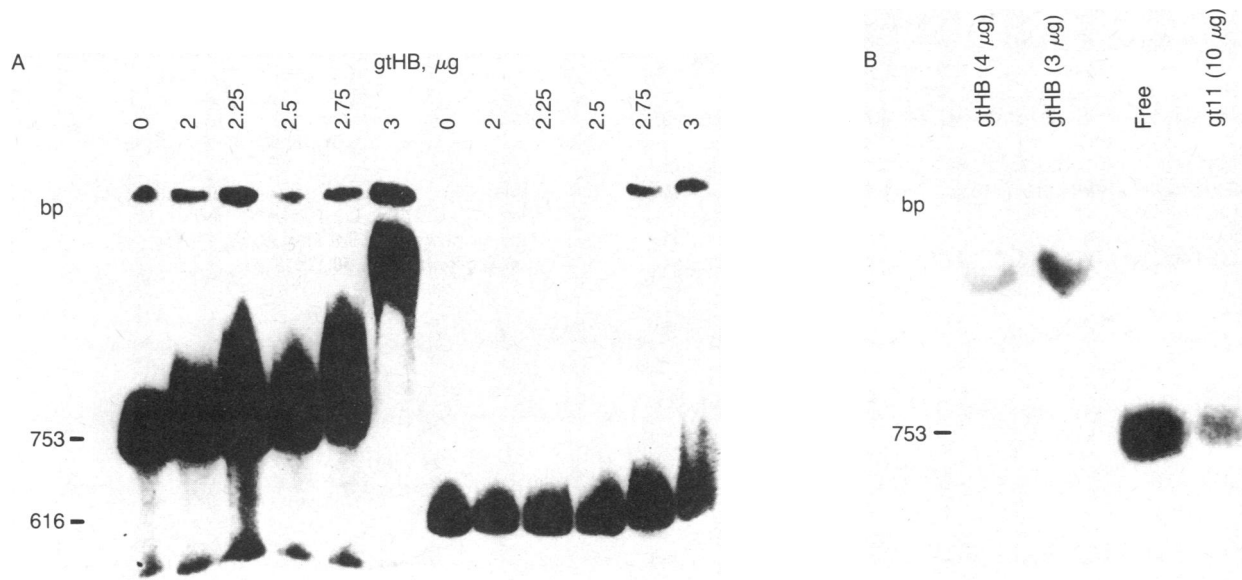


FIG. 5. DNA binding by the Hox 1.5 homeo domain fusion protein. (A) The two *Ava* I-*Hind*III DNA fragments (753 and 616 bp) were analyzed by the gel retardation technique in the presence of different amounts of gtHB protein. (B) Comparison between gtHB and gt11 extracts in binding of the 753-bp fragment. The track labeled "Free" represents the unbound DNA fragment; the other tracks contain different amounts of gtHB or gt11 extracts as labeled.

proteins. Moreover it has been demonstrated by Desplan *et al.* (15) that the homeo domain of the engrailed (*en*) gene product shows specific DNA binding affinity to upstream sequences of its own and another homeo box containing the gene, *fushi tarazu*. Here we have described DNA-protein interaction experiments performed with embryonic mouse nuclear proteins and a homeo domain fused to β -galactosidase. We find that the region upstream of the Hox 1.5 homeo box binds four nuclear proteins (39-kDa, 45-kDa, 68-kDa, and 95-kDa) from day-12.5 mouse embryos. The binding specificity is demonstrated by the inability of any other DNA fragments from this region to bind to these or other nuclear proteins (histones excluded) and by the fact that only four nuclear proteins reproducibly bind to this fragment. To determine if any of the protein-DNA complexes found was specific to a Hox 1.5 homeo domain-DNA interaction, a fusion protein between the Hox 1.5 homeo domain and β -galactosidase was constructed. Protein extracts from the λ lysogen (gt11) could specifically bind the 753-bp fragment only when it contained the Hox 1.5 homeo domain fusion gene. In other experiments not described here antibodies raised against the Hox 1.5 homeo domain have been shown to recognize a protein of 39 kDa in mouse embryo nuclear extracts (A.F., M.L., and F.H.R., unpublished results). These results taken together strongly argue for specific binding between the Hox 1.5 homeo domain and the upstream sequence of the Hox 1.5 gene and that the 39-kDa protein is one of the Hox 1.5 gene products.

Analysis of protein-DNA interaction on polyacrylamide gels showed that the embryo extract gave rise to several protein-DNA complexes (Fig. 3). The large mobility shift from C1 to C2 and the disappearance of C1 with only a 2-fold increase in protein concentration indicates that the binding is cooperative. The DNA binding protein blots identify four different proteins that can independently bind to this 753-bp fragment (Fig. 3). The formation of C2 may represent the binding to C1 of either more of the same protein(s) in C1, or of the other proteins (or both). A more detailed analysis of the binding site(s) and the interacting proteins will be necessary to resolve these possibilities.

This study supports the specific DNA-binding function of the Hox 1.5 homeo domains as predicted from its sequence

and shows binding of this conserved peptide to DNA upstream from the Hox 1.5 homeo box. Our results taken together with the genetic evidence on the developmental role of homologous proteins in *Drosophila* suggests that DNA binding is the function through which homeo proteins regulate their host cells. We hope these results and subsequent studies will aid in understanding the complex molecular processes that regulate murine development.

We thank Dr. Keith Miskimins and David Straney for their technical advice and helpful discussions. A.F. is a National Cancer Cytology Center Postdoctoral Fellow. T.R. is an EMBO Postdoctoral Fellow. L.D.B. is a National Institutes of Health Predoctoral Trainee (GM07223-11). This work was supported in part by National Institutes of Health Grant GM09966 to F.H.R.

1. McGinnis, W., Levine, M. S., Hagen, E., Kuroiwa, A. & Gehring, W. J. (1984) *Nature (London)* **308**, 428-433.
2. Scott, M. P. & Weiner, A. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4115-4119.
3. Regulski, M., Harding, K., Kostriken, R., Karch, F., Levine, M. & McGinnis, W. (1985) *Cell* **43**, 71-80.
4. Gehring, W. J. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50**, 243-251.
5. McGinnis, W., Garber, R. L., Wirz, J., Kuroiwa, A. & Gehring, W. J. (1984) *Cell* **37**, 403-408.
6. Laughon, A. & Scott, M. P. (1984) *Nature (London)* **310**, 25-31.
7. Shepherd, J. C. W., McGinnis, W., Carrasco, A. E., De Robertis, E. M. & Gehring, W. J. (1984) *Nature (London)* **310**, 70-71.
8. Laughon, A., Carroll, S. B., Storfer, F. A., Riley, P. D. & Scott, M. P. (1985) *Cold Spring Harbor Symp. Quant. Biol.* **50**, 253-262.
9. Porter, S. D. & Smith, M. (1986) *Nature (London)* **320**, 766-768.
10. Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53**, 293-321.
11. White, R. A. H. & Wilcox, M. (1984) *Cell* **39**, 163-171.
12. Beachy, P. A., Helfand, S. L. & Hogness, D. S. (1985) *Nature (London)* **313**, 545-551.
13. Carroll, S. B. & Scott, M. P. (1985) *Cell* **43**, 47-57.
14. DiNardo, S., Kuner, J. M., Theis, J. & O'Farrell, P. H. (1985) *Cell* **43**, 59-69.
15. Desplan, C., Theis, J. & O'Farrell, P. H. (1985) *Nature (London)* **318**, 630-635.

16. McGinnis, W., Hart, C. P., Gehring, W. J. & Ruddle, F. H. (1984) *Cell* **38**, 675-680.
17. Hart, C. P., Awgulewitsch, A., Fainsod, A., McGinnis, W. & Ruddle, F. H. (1985) *Cell* **43**, 9-18.
18. Miskimins, W. K., Roberts, M. P., McClelland, A. & Ruddle, F. H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6741-6744.
19. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
20. Young, R. A. & Davis, R. W. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1194-1198.
21. Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in *DNA Cloning: A Practical Approach*, ed. Glover, D. M. (IRL, Arlington, VA), Vol. 1, pp. 49-78.
22. Fried, M. & Crothers, D. M. (1981) *Nucleic Acids Res.* **9**, 6505-6525.
23. Carthew, R. W., Chodosh, L. A. & Sharp, P. A. (1985) *Cell* **43**, 439-448.
24. 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.