The EMBO Journal vol.9 no.4 pp.1209 – 1218, 1990

The segment-specific gene *Krox-20* encodes a transcription factor with binding sites in the promoter region of the *Hox-1.4* gene

Philippe Chavrier[1], Christine Vesque[1,2], Brigitte Galliot[3,4], Marc Vigneron[3], Pascal Dollé[3], Denis Duboule[1] and Patrick Charnay[1,2]

[1]European Molecular Biology Laboratory, Postfach 10.2209, 6900 Heidelberg, FRG, [2]Laboratoire de Génétique Moléculaire, CNRS D1302, Département de Biologie, Ecole Normale Supérieure, 46 rue d'Ulm, 75230, Paris Cedex 05 and [3]LGME du CNRS, Unité INSERM 184, Faculté de Médecine, 11 rue Humann, 67085 Strasbourg Cedex, France
[4]Present address: ZMBH, Universität Heidelberg, Postfach 106 249, 6900 Heidelberg, FRG

Communicated by P.Chambon

***Krox-20* is a mouse zinc finger gene expressed in a segment-specific manner in the early central nervous system, which makes it a potential developmental control gene. In this report, we show that the Krox-20 protein binds *in vitro* to two specific DNA sites located upstream from the homeobox containing gene *Hox-1.4*. The nucleotide sequence recognized by Krox-20 is closely related to the Sp1 target sequence, which is consistent with the similarity existing between the zinc fingers of the two proteins. In co-transfection experiments in cultured cells, Krox-20 dramatically activates transcription from the herpes simplex virus thymidine kinase promoter when an oligomer of its binding site is present in *cis* close to the promoter. Analysis of mutated binding sites demonstrates that the level of activation by Krox-20 correlates with the affinity of the protein for the mutant sequence. These data indicate that Krox-20 constitutes a sequence-specific DNA-binding transcription factor. Parallel analysis of the expression of *Krox-20* and *Hox-1.4* in the neural tube by *in situ* hybridization revealed no overlap, arguing against direct interactions between these two genes. The possible involvement of Krox-20 in the regulation of the transcription of other homeobox genes is discussed in view of their respective patterns of expression.**

Key words: zinc finger/homeobox/serum-inducible gene/DNA-binding protein/transcription control/central nervous system/rhombomere

Introduction

Intensive genetic analysis of the early aspects of *Drosophila* embryogenesis is providing deeper insights into the molecular events responsible for body plan formation. In particular, the process of segmentation has been shown to be controlled by a limited number of genes, many of them encoding established or putative DNA-binding proteins with transcription control properties (reviewed in Ingham, 1988; Levine and Hoey, 1988). The DNA-binding domains so far identified in these proteins belong to one of two categories: the homeodomain encoded by the homeobox (reviewed in Gehring, 1987) and the zinc fingers (Brown *et al.*, 1985; Miller *et al.*, 1985; Rosenberg *et al.*, 1986; Tautz *et al.*, 1987). In addition, genes involved in pattern formation appear to be organized in a complex regulatory network (reviewed in Ingham, 1988). Deciphering this network is critical to our understanding of insect development.

The genetic analysis of vertebrate development is much less advanced. Nevertheless, it is generally assumed that this process obeys principles similar to those involved in *Drosophila* development and that it is also controlled, at least in part, by a network of genes encoding transcription factors. Indeed, numerous genes with sequence similarity to *Drosophila* pattern formation genes have been identified in vertebrates, raising the possibility that these genes might also be involved in the regulation of developmental processes. Representative genes for both categories of putative transcription factors mentioned above have been found (McGinnis *et al.*, 1984; Miller *et al.*, 1985; Chowdhury *et al.*, 1987; Gehring, 1987; Chavrier *et al.*, 1988a; Köster *et al.*, 1988; Ruppert *et al.*, 1988), although the homeobox gene family has been better characterized. In the mouse, this family has been shown to be organized in several gene clusters. Sequence comparison allows alignment between mouse and *Drosophila* clusters (Duboule and Dollé, 1989; Graham *et al.*, 1989). Furthermore, as in *Drosophila*, expression of the genes along the antero-posterior body axis follows a positional hierarchy which reflects their respective positions within the cluster. This similarity in organization and expression of mouse homeobox genes with those of *Drosophila* strongly supports the idea of their involvement in body plan formation. Numerous mouse zinc finger genes have also been cloned (Chowdhury *et al.*, 1987; Chavrier *et al.*, 1988a; Lemaire *et al.*, 1988; Morishita *et al.*, 1988; Sukhatme *et al.*, 1988). However, no homologs of known *Drosophila* pattern formation genes have so far been identified. Nevertheless, several of these mouse genes are expressed in a tissue-specific manner (Chavrier *et al.*, 1988a,b; Chowdhury *et al.*, 1988; Christy *et al.*, 1988; Sukhatme *et al.*, 1988) and one of them, *Krox-20*, shows a segment-specific pattern of expression in the prospective hindbrain (Wilkinson *et al.*, 1989a). This latter observation raises the possibility of the involvement of *Krox-20* and eventually of other zinc finger genes in the control of the segmentation of this part of the early central nervous system (CNS).

Krox-20 was isolated as a serum-stimulated gene from a fibroblast cDNA library (Chavrier *et al.*, 1988b). The gene encodes a protein with three zinc fingers closely related to those of the transcription factor Sp1 (Kadonaga *et al.*, 1987; Chavrier *et al.*, 1988b, 1989). This suggests that the Krox-20 protein is able to bind DNA, possibly to a sequence related to the Sp1 binding site. This also raises the possibility that Krox-20 is a transcription factor, although the similarity between Krox-20 and Sp1 is limited to the zinc fingers (Chavrier *et al.*, 1988b). The human homolog of *Krox-20*

was independently isolated by Joseph *et al.* (1988) and named EGR-2. We have embarked on a detailed analysis of *Krox-20* and of its protein product. This analysis includes the demonstration of Krox-20 transcriptional control activity, the determination of Krox-20 DNA recognition sequence and the identification of putative target genes. Since Krox-20 may be involved in the regulation of segmentation of the mouse CNS, possible target genes could be other putative pattern formation genes expressed in the CNS. Several homeobox genes belong to this category and we have examined one of them, *Hox-1.4* (Duboule *et al.*, 1986; Wolgemuth *et al.*, 1986; Rubin, 1986). *Hox-1.4* was shown to be expressed during embryonic and fetal development as well as in adults according to temporal and tissue-specificities compatible with a functional interaction with *Krox-20* (Gaunt *et al.*, 1988, 1989; Galliot *et al.*, 1989). In addition, a promoter region of *Hox-1.4* contains several GC rich sequences similar to the Sp1 binding site; some of these sequences do bind the Sp1 transcription factor *in vitro*, while others bind factors which may be related to Sp1 (Galliot *et al.*, 1989). In the present study, we have investigated the possibility that Krox-20 could bind to the promoter region of *Hox-1.4*. This is indeed the case and this observation allowed the identification of the nucleotide sequence recognized by Krox-20. Furthermore, binding of Krox-20 to this sequence can lead to activation of a nearby promoter. These findings identify Krox-20 as a sequence-specific transcription factor, which might be involved in the regulation of the expression of *Hox-1.4*.

Results

Krox-20 binds to two specific DNA sequences 5′ to the Hox-1.4 gene

The pattern of *Krox-20* expression in the early CNS and the hierarchy existing in *Drosophila* between zinc finger genes and homeobox containing genes led us to look for possible Krox-20 target genes among mouse homeobox containing genes expressed in the prospective hindbrain between eight and ten days of development. *Hox-1.4* fulfills these criteria and in addition is expressed in adult testes like *Krox-20* (Duboule *et al.*, 1986; Rubin, 1986; Wolgemuth *et al.*, 1986; Chavrier *et al.*, 1988a,b). The organization of the gene, including the localization of its promoter, has recently been determined (Galliot *et al.*, 1989). Interestingly, the *Hox-1.4* 5′ flanking region contains six GC rich sequences, including some binding sites for Sp1 (Galliot *et al.*, 1989). Because of the similarity observed between Krox-20 and Sp1 zinc fingers (Chavrier *et al.*, 1988b), we decided to investigate possible binding of Krox-20 to the GC rich sequences located upstream to *Hox-1.4*. This was first analysed by DNase I footprinting experiments. A DNA fragment containing 360 bp of *Hox-1.4* 5′ flanking sequence was end-labelled with ^{32}P and incubated with *Escherichia coli* protein extracts from strains producing either Krox-20 or no eukaryotic protein, and the footprints were compared to those obtained with a HeLa cell nuclear extract (Figure 1). Two regions of protection were observed with extracts containing Krox-20, corresponding approximately to the GC rich regions previously named E and F (Galliot *et al.*, 1989), which are protected by the HeLa cell nuclear extract (Figure 1). Similar results were obtained when the other strand of the DNA probe was labelled (data not shown). The

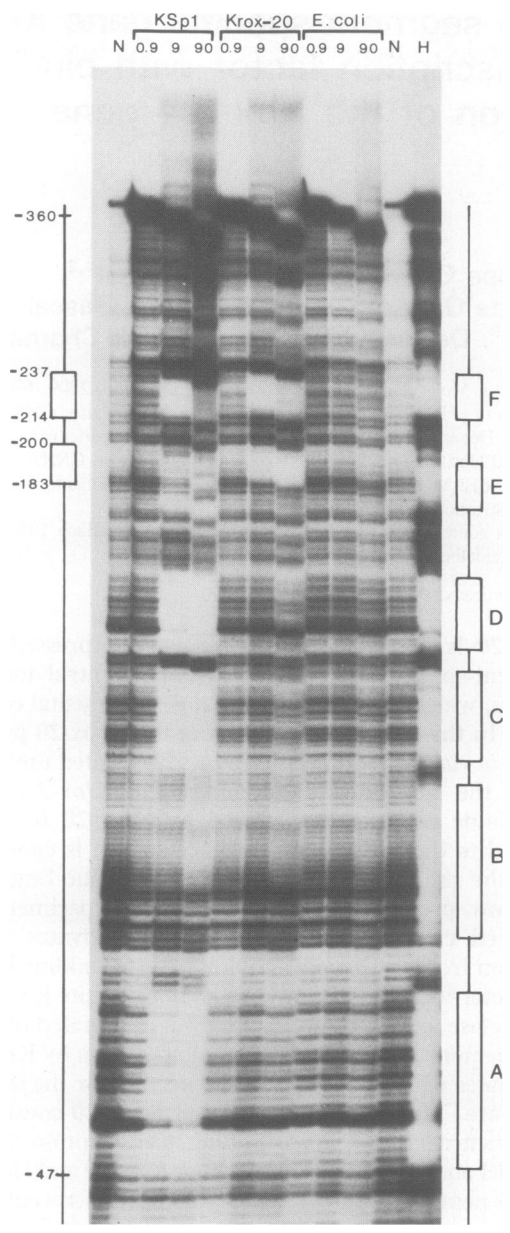

Fig. 1. DNase I footprinting analysis of the 5′ flanking region of *Hox-1.4*. A DNA fragment containing *Hox-1.4* 5′ flanking region up to position −360 was labeled at the 5′ end of the non-coding strand and subjected to DNase I treatment in absence of any extract (N), or in presence of extracts from *E.coli*, *E.coli* strains producing KSp1 or Krox-20, or of a HeLa cell nuclear extract (H). When different concentrations of extract were tested, the numbers above each lane indicate the amount of extract used in micrograms of protein. The positions of the two regions protected by Krox-20 are indicated on the left. The positions of the six major regions protected by the HeLa cell nuclear extract are indicated on the right.

protected regions will also be referred to in this paper as E and F, although the protection due to Krox-20 is slightly different from that observed with the HeLa cell nuclear extract (Figure 1 and Galliot *et al.*, 1989). Release of the protection by competition with oligonucleotides containing the sequences of the GC rich regions E or F, but not with irrelevant oligonucleotides indicated that the protection was specific (data not shown). In particular, absence of competition with an oligonucleotide containing an Sp1 binding site indicated that Krox-20 did not recognize such a site.

In addition to Krox-20, we tested the binding of KSp1 to the *Hox-1.4* 5′ flanking region. KSp1 is a hybrid protein identical to Krox-20, except for the zinc fingers and a few surrounding amino acids which are derived from Sp1 (Chavrier *et al.*, unpublished results). KSp1 was shown to bind to Sp1 target sequence with an affinity similar to that of Sp1 itself in a gel retardation assay (data not shown). This confirmed previous observations indicating that the specificity of DNA recognition is determined by the zinc finger domain (Kadonaga *et al.*, 1987, 1988; Courey and Tjian, 1988). In the DNase I footprinting assay, KSp1 protected four different regions, including the F region, but did not bind efficiently to the E region (Figure 1). The KSp1 footprint over the F region appeared to be displaced by 2 or 3 nucleotides in the 3′ direction compared to the Krox-20 footprint (Figure 1 and data not shown).

To more precisely localize the *Hox-1.4* sequences recognized by Krox-20, we carried out gel retardation experiments (Fried and Crothers, 1981; Garner and Revzin, 1981; Strauss and Varshavsky, 1984) using oligonucleotides corresponding to the two protected regions E and F (Figure 2A). Incubation of these oligonucleotides with a protein extract from the *E.coli* strain producing Krox-20 led to formation of specific complexes which were absent with a control *E.coli* extract (Figure 2B and data not shown). Formation of the complex was inhibited by addition of an excess of unlabeled E and F oligonucleotides, but not by oligonucleotides carrying binding sites for Sp1, AP2 or the serum response factor (Figure 2B and data not shown). The oligonucleotide corresponding to region F was chosen to define the Krox-20 recognition sequence using the methylation interference procedure. Each DNA strand was end-labeled, hybridized with the non-radioactive complementary strand and partially methylated on G residues with dimethylsulphate (DMS). After incubation with the extract containing Krox-20, complexed and free oligonucleotides were separated by gel electrophoresis and purified. Analysis of the pattern of G residues revealed that methylation of the F oligonucleotide at each of four positions on the same strand excluded the oligonucleotide from complex formation (Figure 3). Methylation of three other positions had milder effects (Figure 3). The DNase I footprinting and DMS interference results are summarized in Figure 4. They allow preliminary identification of the DNA binding site as including the sequence 5′-GCGGGGGCG-3′. Interestingly, the third G residue in this sequence does not appear to be closely contacted by Krox-20 (Figure 3) and the related sequence 5′-GCGCGGGCG-3′, where only the third G residue is modified, is observed in the center of the E footprint (Figure 4). Therefore this sequence is likely to represent Krox-20 recognition element in the E region.

Further delineation of the Krox-20 binding specificity was carried out in gel retardation experiments using various mutated F oligonucleotides as competitors. The different mutated oligonucleotides are described in Figure 2A. They define the region recognized by Krox-20 since Fm1 and Fm6 bind the protein as well as the wild-type sequence, while Fm2 to Fm5 have seriously decreased affinities (Figure 2A and B). This delimitation of the target sequence is consistent with the data obtained by methylation interference and in addition implicates an additional G residue evidenced by the effect of mutation m5. Therefore, the Krox-20 recognition sequence appears to be 5′-GCGGGGGCGG-3′. The E

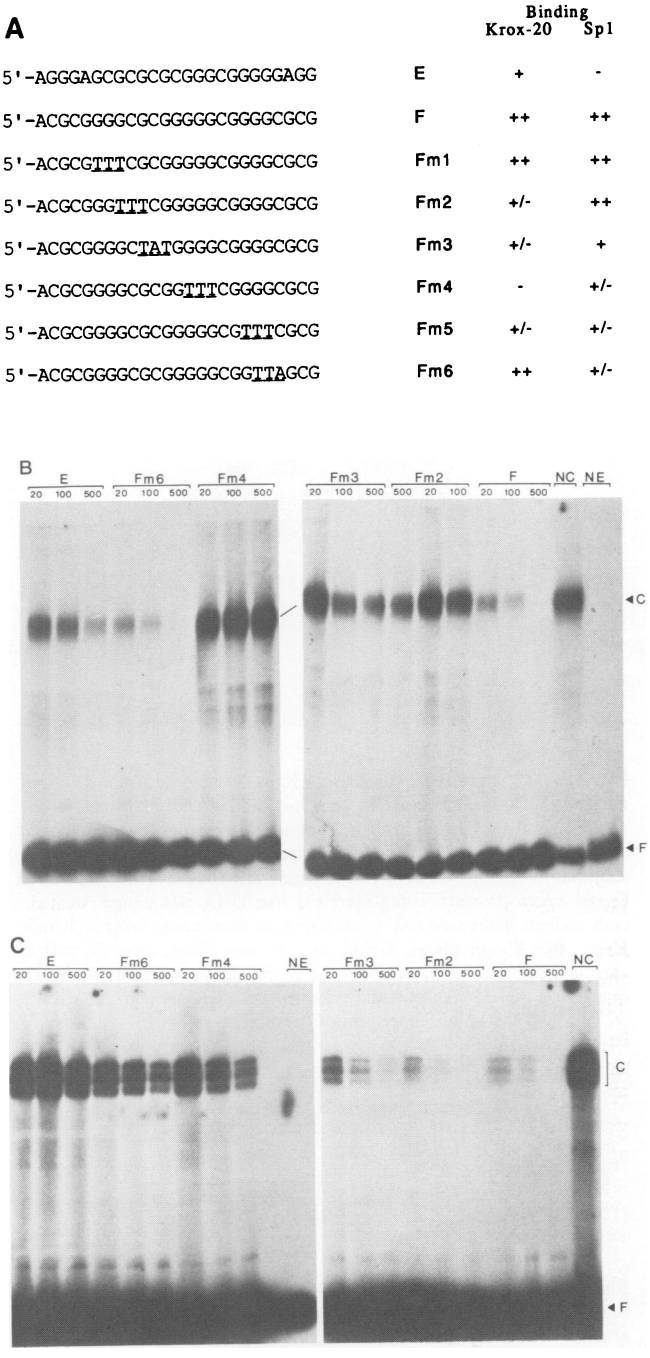

Fig. 2. Gel retardation analysis of Krox-20 and Sp1 binding.
(**A**) Sequence of the central part of the coding strands of the wild type and mutant F and E oligonucleotides used in these experiments and relative affinities for Krox-20 and Sp1. The complete description of the oligonucleotides is given in the Materials and methods section. The mutated nucleotides are underlined. Relative affinities for Krox-20 or Sp1 were estimated by competition as described below: ++ (competes like the F oligonucleotide within a range of concentration of 5-fold); + (requires 5- to 25-fold higher concentration to compete like the F oligonucleotide); +/− (requires >25-fold higher concentration); − (no competition). (**B**) Gel retardation analysis of Krox-20 binding to F and E oligonucleotides and to some of the mutants of the F oligonucleotide. The assays were set up with ~1 ng of labeled F oligonucleotide and 2 μg of protein of bacterial extract containing Krox-20. Above each lane is indicated the identity of the competitor oligonucleotide and its molar excess over probe DNA. Lanes: NE, no extract added; NC, no competition. Positions of complexed (C) and free (F) probe are indicated. (**C**) Gel retardation analysis of Sp1 binding to F and E oligonucleotides. Conditions are the same as above.

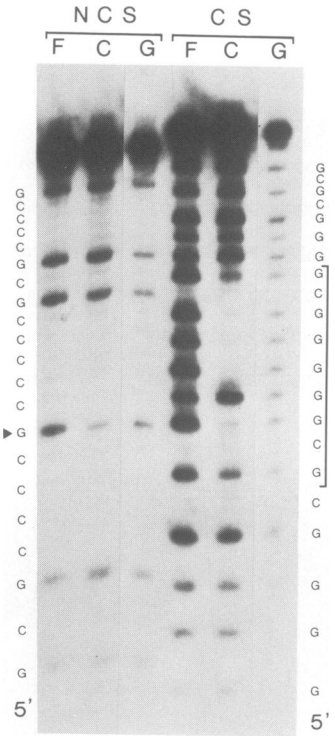

Fig. 3. DMS interference analysis of Krox-20 binding to the F oligonucleotide. Coding strand (CS) or non-coding strand (NCS) oligonucleotides were labeled at their 5′ ends and reassociated with the complementary strands. Double stranded oligonucleotides were subjected to DMS treatment and exposed to Krox-20. After separation by gel electrophoresis, complexed and free DNA was eluted, treated with sodium hydroxide and fractionated on sequencing polyacrylamide gels. Lanes C, complexed DNA; lanes F, free DNA; lanes G, partial chemical degradation products of the probe cleaved at guanosines. The bracket on the CS indicates the region of interference and the arrow on the NCS the unique interfering G residue.

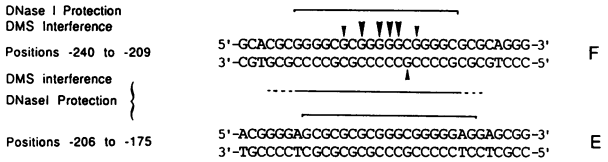

Fig. 4. Summary of data obtained by DNase I footprinting and methylation interference. The sequences of *Hox-1.4* F and E regions are shown together with the limits (brackets) of detected variation in DNase I reactivity after exposure to an *E.coli* extract containing Krox-20. Dashed lines in the case of the non-coding strand indicate that the limits of the footprints were not determined precisely. Positions of methylation that interfere with Krox-20 binding in the F region are indicated by arrows. The symbol size is indicative of the magnitude of the effect.

sequence contains a G to C transversion compared with the F sequence within the binding site and shows a slightly reduced affinity for Krox-20 (Figure 2A and B).

The F oligonucleotide also contained a sequence, 5′-GGGGCGGGG-3′, close to the consensus sequence for the Sp1 binding site (Figure 2A). In addition, KSp1 protected the F region in DNase I protection experiments (Figure 1). We therefore analyzed the binding of the 516 amino acid carboxy-terminal part of Sp1, containing the zinc fingers, to the F oligonucleotide in gel retardation experiments (Figure 2C) and by methylation interference (Lemaire *et al.*, submitted). We found that the consensus sequence indeed

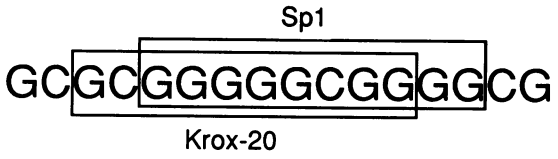

Fig. 5. Overlapping nucleotide sequences recognized by Krox-20 and Sp1 on the F oligonucleotide.

constituted an Sp1 binding site (Lemaire *et al.*, submitted). Competition gel retardation experiments indicated that mutations m3, m4 and m5 seriously affected both Krox-20 and Sp1 binding, while mutations m2 and m6 affected only Krox-20 or Sp1 binding, respectively (Figure 2). These results are consistent with the pattern of methylation interference observed for each protein (Figure 4 and Lemaire *et al.*, submitted). Finally, the E oligonucleotide did not bind the Sp1 protein, as predicted from the DNase I protection experiment. This is likely to be due at least in part to the G to C transversion noticed above between F and E sequences.

In conclusion, Krox-20 and Sp1 bind to two overlapping sequences on the F oligonucleotide (Figure 5), but data from both methylation interference and competition with mutant oligonucleotides clearly indicate that the two sites are distinguishable.

Krox-20 is a transcriptional activator

The determination of Krox-20 binding site allowed to test whether the protein was a transcriptional activator. For this purpose, we made use of the transient co-transfection assay developed by Courey and Tjian (1988). The Krox-20 coding sequence was placed under the control of the *Drosophila* actin 5C promoter. The construction was co-transfected into Schneider line 2 (SL2) cells along with reporter constructs and the plasmid pPadh-βgal, which contained the *E.coli lacZ* gene under the control of the *Drosophila* Adh promoter and was used for normalization of the experiments. The reporter constructs were derived from ptkCAT, which consisted of the herpes simplex virus (HSV) thymidine kinase (tk) gene promoter driving the chloramphenicol acetyl transferase (CAT) gene (Figure 6A). This promoter contained a weak proximal Sp1 binding site as well as a moderate affinity distal one (Jones *et al.*, 1985). Wild-type E and F Krox-20 binding sites, as well as mutant derivatives of the F binding site, were polymerized in a head to tail configuration and inserted in front of the tk promoter (Figure 6B). The *Drosophila* SL2 cells were chosen because they have been shown to be devoid of endogenous Sp1-like activity (Courey and Tjian, 1988), which could have interfered with the assay, since the F oligonucleotide contained an Sp1 binding site. CAT activity determined 48 h after transfection was taken, after normalization with the β-galactosidase activity, as a measure of the capacity of the Krox-20 protein to modulate transcription of the reporter gene. Presence of the Krox-20 plasmid led to very strong activation of CAT expression (>100-fold) from a construct, p4F−, containing four F binding sites (Figure 7 and Table I). In contrast, only marginal induction was observed with ptkCAT (1.5-fold). Transactivation occurred independently of the orientation of the binding sites respective to the promoter (compare p3F+ and p3F−, Table I) and its level correlated with the number of F binding sites present on the plasmid (compare p4F−, p3F− and

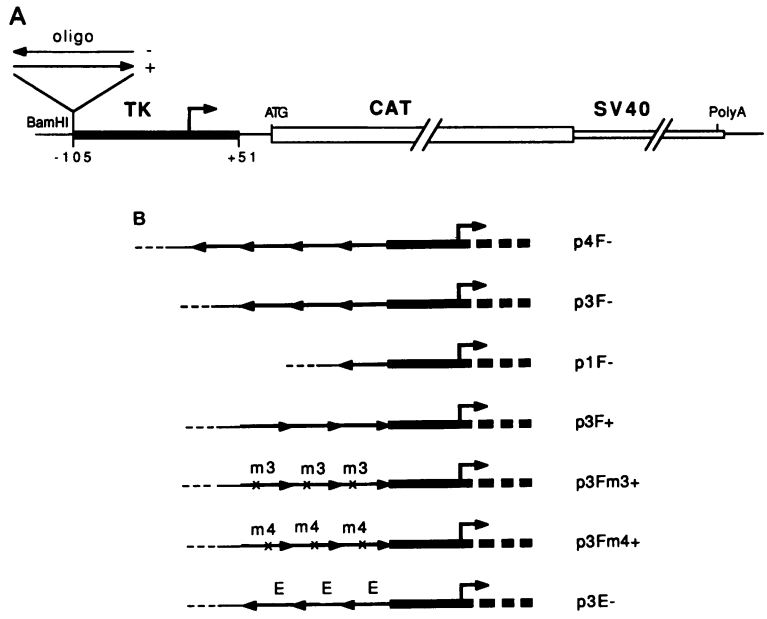

Fig. 6. Structure of the reporter genes. (**A**) Structure of the ptkCAT plasmid. The hybrid tkCAT gene containing the HSV tk promoter region from positions −105 to +51 was used as a recipient for the introduction into its *Bam*HI site of oligomers of the F and E Krox-20 binding sites and of mutant derivatives of F. The transcription initiation site is indicated by an arrow. (**B**) Schematic representation of the constructs containing various oligomers of the F or E oligonucleotides. For each construct, the first number indicates the number of copies of the oligonucleotide and the mutation is indicated when applicable (m3 or m4). The '+' and '−' signs refer to the respective orientations of the oligonucleotide and of the transcription, '+' indicating same relative orientation as in *Hox-1.4*.

p1F−, Table I). The E binding site was slightly less effective than the F one, an observation consistent with the reduced affinity of Krox-20 for E compared to F (Table I, Figures 2 and 7). In addition, for the mutants of the F oligonucleotide examined in the transactivation assay, the level of CAT expression correlated well with their respective affinity for Krox-20 determined *in vitro* (Figure 2 and Table I). These different data strongly support the idea that transactivation of the tkCAT derivatives by Krox-20 involves binding of the protein to the target sequence.

Additional co-transfection experiments included KSp1 and Sp1 expression plasmids as well as a reporter construct (SV6−tkCAT) in which the six GC boxes from the SV40 promoter have been brought upstream of the tk TATA box (Courey and Tjian, 1988). This construction was strongly activated by Sp1 and moderately by KSp1. Surprisingly, it was also slightly activated by Krox-20 (Figure 7 and Table I). This effect could be explained by a low affinity of Krox-20 for isolated Sp1 binding sites, not detectable in gel retardation assays, and a higher cooperative binding of the protein to multiple tandem GC boxes, sufficient to allow transactivation. Indeed, we have observed that a DNA fragment containing the six SV40 GC boxes competes for Krox-20 binding with an authentic Krox-20 binding site in the gel retardation assay (data not shown). Finally, the level of ptkCAT transactivation by Sp1 (~27-fold, Table I) was higher than that reported by Courey and Tjian (1988) for a similar construct, −105tkCAT. This might be due to larger amounts of pPacSp1 expression plasmid used in the present study.

To prove that transactivation by Krox-20 occurred at the RNA level, we carried out an RNase mapping experiment to measure the level of tkCAT mRNA in presence or in absence of Krox-20. Two major bands were observed after co-transfection with the Krox-20 expression plasmid, which were not detected after co-transfection with the control

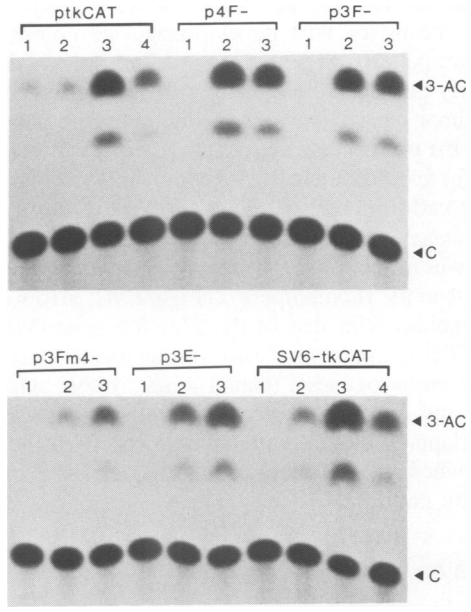

Fig. 7. *Trans*-activation assay. SL2 cells were co-transfected with the expression vector pPacU-NdeI (**1**) or Krox-20 (**2**), Sp1-516C (**3**) or KSp1 (**4**) expression plasmids on one side and different reporter plasmids on the other, as indicated. CAT activity was assayed 48 h after transfection. The structure of the ptkCAT derivatives containing Krox-20 binding sites are described in Figure 6. 'C' corresponds to the unreacted substrate [^{14}C]chloramphenicol and '3-AC' to the major acetylated form (in position 3). The quantification of the results is presented in Table I.

expression vector (Figure 8). The lower band (5′ tk) revealed a protected fragment of ~225 nucleotides, likely to correspond to a transcript correctly initiated from the tk promoter (expected size 218 nucleotides). A fragment of the same size was observed with the SV6−tkCAT construct

Table I. Levels of *trans*-activation of the different reporter constructs by Krox-20, Sp1 or KSp1

Expression plasmids	Reporter plasmids													
	ptkCAT		SV6-CAT		p4F−		p3F−		p1F−		p3F+		p3Fm3+		p3Fm4−		p3E−	
Krox-20	1.5	(1−1.9)	5	(3.5−7.5)	117	(101−134)	34	(31−37)	6.5	(6−7)	41	(25−57)	2.7	(2.6−2.9)	3.9	(2.3−6.3)	22	(16−28)
Sp1	27	(17−36)	109	(52−203)	32	(9−56)	22	(12−31)	15	(9−21)	44	(27−61)	19	(15−22)	26	(10−51)	45	(17−73)
KSp1	5	(3−6.8)	14	(8.5−20)	ND		ND		ND		ND		ND		ND		ND	

The level of transactivation of the reporter construct is defined as the ratio of normalized CAT activities measured after co-transfection with the expression plasmid (pPacKrox-20, pPacSp1 or pPacKSp1) or the expression vector pPacU-Ndel. The figures shown are the means of the values obtained in at least two independent experiments. The figures in parenthesis are the extreme values obtained. Denominations of the reporter plasmids containing wild-type or mutant Krox-20 binding sites are described in Figure 6. ND, not done.

trans-activated by Sp1 (data not shown). The slower-migrating protected fragment (upstream) revealed the presence of transcripts initiated upstream of the correct initiation site, since the size of the protected fragment corresponded to the region of homology between the RNA and the probe. In conclusion, we find that co-transfection with the Krox-20 expression plasmid leads to an increase in the level of correctly initiated tkCAT mRNA (Figure 8), which is consistent with the effect observed on the enzymatic activity. Our data suggest that Krox-20 acts by stimulating transcription. It is therefore a sequence-specific DNA binding transcriptional activator.

Expression of Krox-20 and Hox-1.4 in the developing hindbrain

To test the possibility of an interaction between Krox-20 and *Hox-1.4* during *in vivo* development, we compared the expression patterns of both genes to find out whether they temporally and spatially overlapped. *In situ* hybridization on neighbor serial sections using antisense riboprobes revealed the presence of Krox-20 transcripts within rhombomeres 3(a) and 5(b) in an 8.25 day old embryo (Figure 9A), as anticipated from earlier studies (Wilkinson *et al.*, 1989a). At this stage, the anterior border of *Hox-1.4* transcript domain within the hindbrain neuroectoderm was not more anterior than the rhombomere 7 (Figure 9C, arrow c). This limit coincides with that of the *Hox-2.6* gene (Wilkinson *et al.*, 1989b), a paralog of *Hox-1.4*, as it is the case at later stages of embryogenesis (Gaunt *et al.*, 1989). Therefore, *Hox-1.4* and *Krox-20* transcripts were present in separate, non-overlapping regions within the prospective hindbrain at the time when *Krox-20* exhibited its characteristic segmental expression pattern.

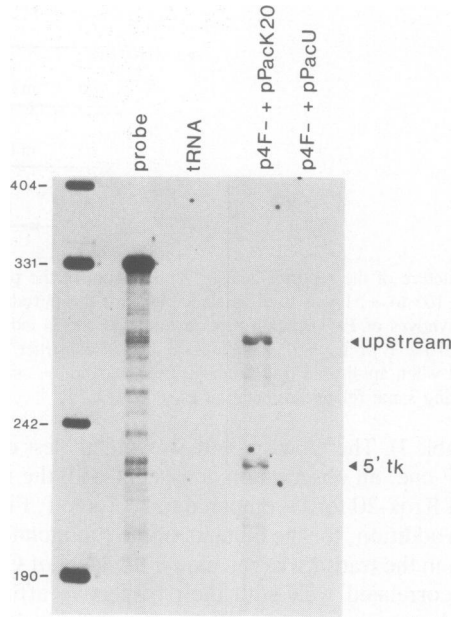

Fig. 8. Krox-20 is a transcriptional activator. A tkCAT derivative containing four copies of the F site (4F−) was co-transfected into SL2 cells with the expression vector (pPacU) on Krox-20 expression plasmids. Total RNA was extracted 40 h after transfection and the level of tkCAT RNA was measured using an RNase protection assay (see Materials and methods for details). Protected fragments were analysed by electrophoresis on a 6% polyacrylamide sequencing gel. Indicated are protected fragments corresponding to correctly initiated tkCAT RNA (5′ tk) and to transcripts initiated upstream and marking the position of divergence with the probe (upstream). The first lane corresponds to the size marker consisting of pUC19 DNA digested with *Hpa*II, the second lane to the undigested probe and the third lane to an assay carried out in presence of tRNA only.

Discussion

Krox-20 is a sequence-specific transcription factor

In this study, we have shown that Krox-20 binds specifically to the sequence 5′-GCGGGGGCGG-3′ *in vitro* and is able to activate transcription from the HSV tk promoter when oligomers of this sequence are placed in the vicinity of the promoter. Furthermore, mutations in the binding site altering its affinity for the protein affect Krox-20 dependent transcriptional activation in a parallel manner. These data strongly suggest that Krox-20 acts directly by binding to the promoter region and that, therefore, Krox-20 is a sequence-specific transcription factor. In this study, we have used high levels of *Krox-20* expression plasmid (5 μg), since we knew that the Krox-20 protein was very unstable in eukaryotic cells (Chavrier *et al.*, unpublished results). Nevertheless, co-transfection experiments performed with 100-fold less expression plasmid gave *trans*-activation levels only ~4-fold lower than those obtained with 5 μg of plasmid (data not shown). The activation property of Krox-20 is not restricted to *Drosophila* cells, since we have demonstrated recently that the protein can also activate transcription in human HeLa cells (Vesque *et al.*, unpublished result).

In recent years, analysis of a number of transcription factors from higher and lower eukaryotes has demonstrated that DNA recognition and transcriptional activation are generally performed by different domains of the protein (see Mitchell and Tjian, 1989, for a recent review). This is also likely to be the case for Krox-20, since the hybrid protein KSp1 is a transcriptional activator (Table I), while the Sp1 zinc fingers do not contain a transcription activation domain (Courey and Tjian, 1988). This observation indicates that such a domain has to be located within Krox-20, outside of the zinc finger region. So far, three types of transcription

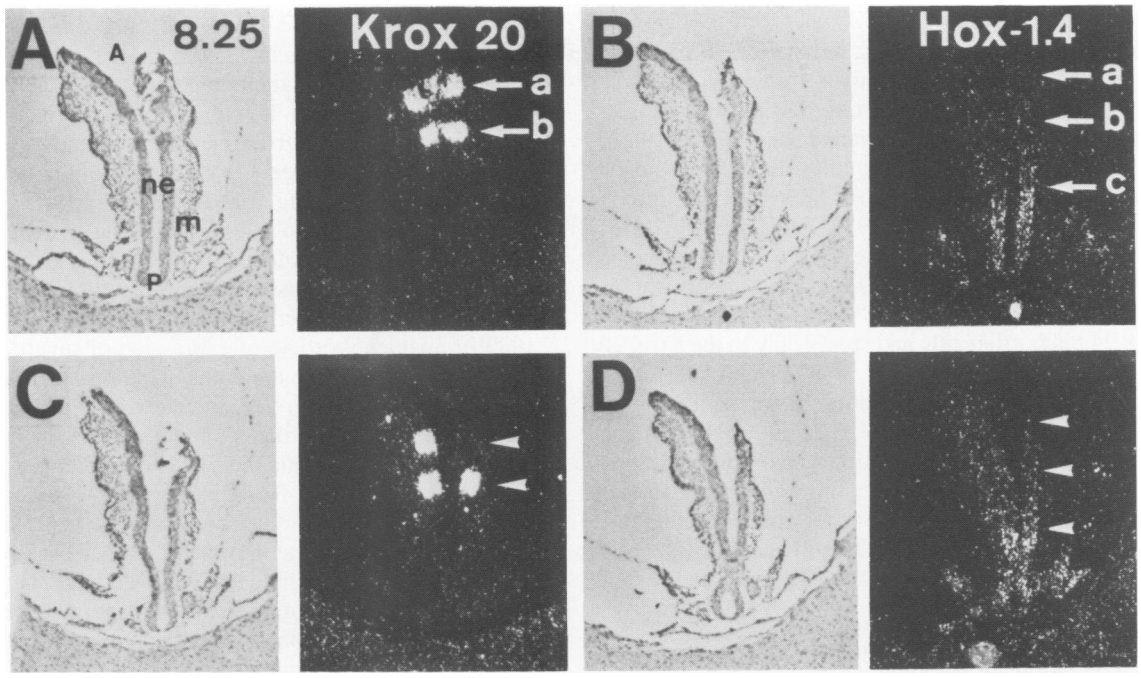

Fig. 9. *Krox-20* and *Hox-1.4* expression domains in the developing hindbrain of a 8.25 day old embryo. Four serial sections are shown hybridized successively with the *Krox-20* probe (**A,C**) or the *Hox-1.4* probe (**B,D**). The arrows in (A) and (B) and arrowheads in (C) and (D) indicate domains of expression of *Krox-20* (rhomobomere 3: a; rhombomere 5: b) as well as the anterior limit of the *Hox-1.4* expression domain (c). In addition to expression in the hindbrain, the presence of *Hox-1.4* transcripts within the mesodermic lineage is clearly seen (bottom of D). A: anterior; P: posterior; ne: neuroectoderm (from the hindbrain); m: mesoderm.

activation domains have been identified in transcription factors: acidic domains, first identified within yeast transcriptional activators (Ma and Ptashne, 1987a,b; Hope and Struhl, 1986; Struhl, 1987), glutamine rich regions, recognized in the case of Sp1 (Courey and Tjian, 1988) and proline-rich domains, found more recently in CTF (Mermod *et al.*, 1989). Unlike Sp1, Krox-20 does not contain any glutamine-rich domain (Chavrier *et al.*, 1988b, 1989); however, two relatively acidic regions can be defined within the protein: from positions 23 to 63 (net charge −7) and from positions 160 to 184 (net charge −4) (Chavrier *et al.*, 1989). In addition, Krox-20 contains a high proportion of proline, with one particularly rich region (34% proline between positions 204 and 264). Reverse genetics experiments will be required to determine whether any of these three regions constitutes a transcriptional activation domain.

Krox-20 DNA binding domain

The studies carried out so far do not allow us to determine whether the sequence 5′-GCGGGGGCGG-3′ constitutes the optimal recognition sequence of the protein. This will require the analysis of a number of natural binding sites or a more detailed mutagenesis of the sequence available. Nevertheless, comparison of the DNA-binding activity of extracts from bacteria producing Krox-20 or KSp1 by gel retardation assay and of the relative amounts of Krox-20 and KSp1 present as estimated by Western blotting indicates that the affinities of the two proteins for their respective binding sites are of the same order of magnitude (data not shown). Since Krox-20 and KSp1 have the same type of DNA-binding domain and might therefore have a similar affinity for their respective DNA-binding sites, our data suggest that the 5′-GCGGGGGCGG-3′ sequence is close to the optimal Krox-20 recognition sequence.

Determination of the respective DNA target sequences of Krox-20 and KSp1 has shown that the specificity of DNA recognition is governed, at least in first approximation, by the amino acid sequence of the zinc fingers. This is consistent with the previous analysis of Sp1 indicating that the zinc fingers constitute the DNA binding domain (Courey and Tjian, 1988). However, regions located outside of the zinc fingers were found to subtly modulate the specificity of DNA recognition (Vigneron, unpublished results).

The nucleotide sequences recognized respectively by Krox-20 and Sp1 in the F region of *Hox-1.4* are similar, both in terms of length and sequence: the two binding sites consist of a GC rich sequence 9−10 nucleotides long (Figure 5). These close specificities are likely to be related to the similarity that we noted earlier between the zinc fingers of the two proteins (Chavrier *et al.*, 1988b): each protein contains three zinc fingers and in the region surrounding the conserved leucine residue (position 23 in our nomenclature, Chavrier *et al.*, 1988b), the first and third finger of Krox-20 are closely related to the second finger of Sp1, while the second finger of Krox-20 is more similar to the third finger of Sp1. This region of the fingers is of particular importance since it is organized into an α-helix (Lee *et al.*, 1989) which, according to the models proposed by Berg (1988) and by Gibson *et al.* (1988), contains the amino acids involved in base-specific contacts with DNA and therefore responsible for the sequence specificity of DNA recognition. The comparative study of the interactions of Sp1 and Krox-20 with their respective DNA binding sites is therefore likely to yield general information on the process of DNA recognition by zinc finger proteins. In particular, the comparison of the amino acid sequences suggests that some amino acid positions are critical for specific DNA recognition. We are currently testing such predictions by *in vitro* mutagenesis.

The similarity between Krox-20 and Sp1 recognition sequences and DNA binding domains suggests that the two proteins may belong to a new family of transcription factors with related DNA binding specificities. Several such families have already been identified (Mitchell and Tjian, 1989). Indeed, several laboratories, including ours, have cloned another gene belonging to the Sp1/Krox-20 family (Milbrandt, 1987; Christy *et al.*, 1988; Lemaire *et al.*, 1988; Sukhatme *et al.*, 1988). This gene, which we have named *Krox-24*, encodes a protein with three zinc fingers almost identical to those of Krox-20 and we have recently shown that Krox-24 binds efficiently and specifically to the Krox-20 recognition sequence (Lemaire *et al.*, submitted). An interesting aspect of this transcription factor family is that although Sp1 and Krox-20 have similar DNA-binding domains, they may not share the same type of transactivation domain, as discussed above. If true, this will provide further illustration of the flexibility of the transcriptional apparatus: DNA-binding and transactivation domains can be effectively assembled in various combinations (Mitchell and Tjian, 1989).

Possible physiological significance of Krox-20 interaction with Hox-1.4

The fact that Krox-20 binds *in vitro* to two sites within the promoter region of *Hox-1.4* raises the possibility that it may also be involved in its regulation *in vivo*. We have tested whether Krox-20 was able to activate a construction consisting of the Hox-1.4 promoter driving the CAT coding sequence in the co-transfection assay in SL2 cells. Presence of Krox-20 led to only marginal induction (2-fold) of the *Hox-1.4* promoter (data not shown). Nevertheless, this result does not preclude an involvement of Krox-20 in the activation of *Hox-1.4* in other cell types: Krox-20 might require to interact with other transcription factors absent from SL2 cells in order to act on the *Hox-1.4* promoter. We also considered a possible interaction in the hindbrain where *Krox-20* might play a role in setting up a segmental pattern at ~8.5 days of development. However, our parallel analysis of the expression of the two genes in this region does not favor such a possibility: the patterns of mRNA hybridization do not overlap during this period. Since the Krox-20 protein is located within the nucleus (Chavrier *et al.*, unpublished results), it is expected to be present in the cells containing the RNA. This suggests that Krox-20 is not directly involved in *Hox-1.4* activation in this region of the embryo. Nevertheless, the possibility of an involvement of Krox-20 in *Hox-1.4* repression cannot be excluded, since some transcription factors can act both as activator and as repressor (Meisfeld *et al.*, 1987; Akerblom *et al.*, 1987). In addition, our data do not preclude an involvement of Krox-20 in the regulation of *Hox-1.4* in other tissues or at other stages of development. For instance, although their respective patterns of expression are quite restricted in adult animals, both genes are expressed in testes (Duboule *et al.*, 1986; Rubin *et al.*, 1986; Wolgemuth *et al.*, 1986; Chavrier *et al.*, 1988a,b). In the case of *Hox-1.4*, the expression appears to be germ cell-specific and restricted to cells that have entered into and progressed beyond the meiotic prophase stage of differentiation (Wolgemuth *et al.*, 1987). We are currently investigating the possibility that *Krox-20* is also expressed in these cells.

As mentioned above, there is at least one other transcription factor, Krox-24, with the same DNA specificity as Krox-20 (Lemaire *et al.*, 1988; Lemaire *et al.*, submitted). It constitutes another potential candidate for the regulation of *Hox-1.4*. It will therefore be interesting to determine whether *Krox-24* or another putative gene with closely related zinc fingers is expressed in the neural tube in a pattern overlapping with that of *Hox-1.4*.

Recent *in situ* hybridization studies have indicated that several homeobox-containing genes are expressed in segment-specific patterns in the developing mouse hindbrain (Gaunt *et al.*, 1986; Murphy *et al.*, 1989; Wilkinson *et al.*, 1989b). Segment borders of expression of several of these genes are common with the limits of expression of *Krox-20*. For two of them, *Hox-2.7* and *Hox-2.8*, the domain of transcription includes the rhombomeres 3 or 5, where *Krox-20* is expressed, and their level of expression is higher in these rhombomeres (Wilkinson *et al.*, 1989a,b). Since expression of *Krox-20* in the rhombomeres seems to precede the appearance of homeobox gene transcripts (Wilkinson *et al.*, 1989b), it is possible that Krox-20 might be involved in the definition of the expression domain of some of these genes. In such a case, one would anticipate to find Krox-20 binding sites in the vicinity of the genes, like in the case of *Hox-1.4*.

In conclusion, the possible involvement of Krox-20 in the regulation of the expression of homeobox containing genes is particularly attractive in view of the parallelism observed in the organization and expression of homeobox genes between insects and vertebrates (Duboule and Dollé, 1989; Graham *et al.*, 1989). It raises the possibility that, like in *Drosophila*, mouse zinc finger genes might control aspects of important developmental processes by modulating the expression of homeobox genes.

Materials and methods

Oligonucleotides

The different oligonucleotides were obtained by annealing of two chemically synthesized strands as follows. E: 5′-CCGAGGGAGCGCGCGCGGGCGGGGGAGGA-3′ and 5′-TCGGTCCTCCCCCGCCCGCGCGCTCCCC-3′; F: 5′-CCGACGCGGGGCGCGGGGGCGGGGCGCGCA-3′and 5′-TCGGTGCGCGCCCCGCCCCCGCGCCCCGCG-3′; Sp1 binding site: 5′-GATACGCGTATCGGGGCGGAGAAACACTGC-3′ and 5′-GCAGTGTTTCTCCGCCCCGATACGCGTATC-3′; AP2 binding site: 5′-GAACTGACCGCCCGCGGCCCGTGT-3′ and 5′-ACACGGGCCGCGGGCGGTCAGTTC-3′; serum response factor-binding site: 5′-AATTCCTCAGTCCATATATGGGCAGCGACGTCACGGG-3′ and 5′-AATTCCCGTGACGTCGCTGCCCATATATGGACTGAGG-3′. The mutant derivatives of the F oligonucleotide were obtained by introducing the complementary changes described in Figure 3A in both strands of the F oligonucleotide.

Plasmid constructions

ptkCAT reporter plasmids containing oligomers of the E and F oligonucleotides and of their derivatives were obtained in the following way: the oligonucleotides were phosphorylated at their 5′ ends with T4 polynucleotide kinase and ATP, annealed and subsequently ligated into head to tail tandem repeats by virtue of the presence of asymmetric complimentary AvaI extremities. The ligation products were separated by electrophoresis on 5% polyacrylamide gel, trimers and tetramers were purified and cloned into the *Ava*I site of the plasmid pV2 (Fromental *et al.*, 1988). The oligomers were subsequently excised from the pV2 derivatives by restriction with *Bgl*II and cloned into the BamHI site of ptkCAT (generous gift of R.Miksicek and G.Schütz). The orientation and the sequence of the inserts in ptkCAT were determined after cloning into M13 derived vectors. Single-stranded DNA was prepared (Messing, 1983) and the nucleotide sequence was established using the Sequenase™ procedure (US Biochemical, Cleveland,

Ohio). ptkCAT derivatives containing either a monomer or a dimer of the F oligonucleotide were obtained by partial digestion of the derivative containing the tetramer by *Bss*HII, which cuts at a unique position within the F oligonucleotide. Plasmid pET-Krox-20 was constructed by inserting a 1.63 kb *Acc*I−*Bam*HI fragment encoding the C-terminal 463 amino acids of Krox-20 (Chavrier *et al.*, 1989) and a synthetic *Nde*I−*Acc*I adaptor encoding the N-terminal seven amino acids of Krox-20 into the *Nde*I and *Bam*HI sites of plasmid pET3a (Studier *et al.*, 1986; Rosenberg *et al.*, 1987). Plasmid pET-KSp1 was constructed as follows. A 399 bp *Hgi*AI−*Hin*dIII fragment encoding the C-terminal 43 amino acids of Krox-20 was ligated to a 243 bp *Bst*XI−*Sau*961 fragment from plasmid pSp1-516C (Kadonaga *et al.*, 1987) encoding the 82 amino acids zinc finger domain of Sp1, using a *Sau*96I/*Hgi*AI synthetic adaptor. A 1060 bp fragment encoding the 333 N-terminal amino acids of Krox-20 was then added using a *Apa*I/*Bst*XI synthetic adaptor. The resulting 1757 bp *Eco*RI−*Hin*dIII fragment encoding the 472 amino acids of KSp1 was subsequently cloned into plasmid pGEM1 (Promega Biotech). The complete coding sequence of KSp1 was verified by double-stranded DNA sequencing (Chen and Seeberg, 1985). The subsequent cloning of KSp1 coding sequence into the expression plasmid pET3a, resulting in plasmid pET-KSp1, was performed as described in the case of pET-Krox-20. The constructions of the expression plasmids derived from pPacU-NdeI (generous gift of A.Courey and R.Tjian) were carried out as follows. A 550 bp *Nde*I−*Bam*HI fragment derived from pET-Krox-20 and encoding the N-terminal 184 amino acids of Krox-20 was cloned into pPacU-NdeI digested by *Nde*I and *Bam*HI. Fragments encoding the C-terminal parts of Krox-20 or KSp1 were subsequently inserted into the unique *Bam*HI site to yield plasmids pPac-Krox-20 and pPac-KSp1, respectively. Plasmid pPacSp1 (Courey and Tjian, 1988) was generously provided by A.Courey. Plasmid pHox1.4-CAT was constructed as follows. A 360 bp *Apa*I−*Sac*I fragment containing the promoter region of *Hox-1.4* from plasmid pMT5AS360a (Galliot *et al.*, 1989) was converted into a *Xho*I fragment by addition of synthetic linkers. It was subsequently inserted into the *Xho*I site upstream of the CAT coding sequence in a plasmid derived from pGEM7ZF (Promega Biotech) by cloning of a *Hin*dIII−*Bam*HI fragment containing the CAT sequence from pSV2-CAT (Gorman *et al.*, 1982). The orientation of the insert and sequence of the junction were determined by double-stranded DNA sequencing (Chen and Seeberg, 1985). The plasmid pPadh−βgal, susceptible of expressing the *E.coli lacZ* gene in *Drosophila* cells, was constructed by insertion into the *Bam*HI site of pPadh (Courey and Tjian, 1988) of a *Bam*HI fragment carrying the coding sequence of the *lacZ* gene and derived from the plasmid pSPTlacZ1 (generous gift of P.Gruss). The construction resulted in the formation of a hybrid gene with conservation of the reading frame.

Cell lines and DNA transfection

D.melanogaster Schneider 2 cells (Schneider, 1972) were grown in Schneider medium (Gibco) supplemented with L-glutamine, 1% penicillin−streptomycin solution (Gibco) and 10% fetal calf serum. They were transfected at densities ranging from 0.5 to 5 × 10^6 cells/ml, using the calcium-phosphate procedure as described (Di Nocera and Dawid, 1983; Courey and Tjian, 1988). Each plate received a total amount of plasmid DNA of 20 μg, consisting of 10 μg of reporter DNA, 5 μg of expression plasmid and 5 μg of a control plasmid, pAdh-βgal, susceptible of expressing the *E.coli lacZ* gene under the control of the *Drosophila* Adh promoter.

Protein extracts and DNA footprints

The expression system and protocols of Studier and collaborators were used to produce the different proteins in *E.coli* (Studier and Moffatt, 1986; Rosenberg *et al.*, 1987). Protein extracts were prepared according to Kadonaga *et al.* (1987). After dialysis against buffer B containing 20% (v/v) glycerol, 20 mM Tris−HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 1 mM EDTA, 10 μM $ZnSO_4$, 1 mM DTT, 0.2 mM PMSF (phenylmethylsulphonyl fluoride) and 1 mM $NaHSO_3$, the protein concentration of the extract was adjusted to 8 mg/ml. These extracts were tested when applicable for the presence of Krox-20 or KSp1 by Western blotting with a rabbit antibody raised against a β-galactosidase−Krox-20 fusion protein (Zerial *et al.*, unpublished) and were used directly in DNA footprinting experiments and DNA-binding gel electrophoresis assays. Control HeLa cell nuclear extracts were prepared as described (Wildeman *et al.*, 1984). Footprinting experiments were performed as described (Galliot *et al.*, 1989).

DNA-binding gel electrophoresis assay and methylation interference

For DNA binding gel electrophoresis assays, either strand of the F oligonucleotide (50 ng) was labeled at its 5′ end using T4 polynucleotide kinase and [γ-^{32}P]ATP (Amersham, 3000 Ci/mmol). It was subsequently annealed with a 10-fold excess of the complementary strand by incubation at 65°C for 5 min and slow cooling to room temperature. The bacterial extract (2 μg of protein) was first pre-incubated on ice for 10 min in a volume of 20 μl in buffer C (20 mM Tris−HCl, pH 7.5, 50 mM KCl, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM DTT) supplemented with 6 μg of poly(dI−dC) and 8% Ficoll. A 1 ng sample of end-labeled F oligonucleotide, previously mixed with cold competitor oligonucleotide if required, was subsequently added and the incubation was pursued for an additional 20 min on ice. The mixture was then loaded on a 5% polyacrylamide gel and electrophoresis was performed for 2 h at 14.5 V/cm in 0.5 × TBE buffer (Maniatis *et al.*, 1982). The gels were dried and autoradiographic exposures were performed at −70°C with an intensifying screen. Methylation interference experiments were carried out on the F oligonucleotide labeled at the 5′ end of sense or antisense strand and subsequently partially methylated as described (Maxam and Gilbert, 1980). Free and complexed oligonucleotide were recovered by preparative gel electrophoresis after exposure of 2 ng of oligonucleotide to 16 μg of protein extract and subsequently subjected to alkali treatment (Raymondjean *et al.*, 1988). After three ethanol precipitations using poly dIdC as carrier, the cleaved oligonucleotides were analyzed by electrophoresis on 18% polyacrylamide sequencing gel. Autoradiographic exposures were performed at −70°C with an intensifying screen.

CAT assay and RNase mapping

For CAT assays, cells from one 9 cm plate were harvested 48 h after transfection, washed once with phosphate-buffered saline (PBS) and resuspended into 200 μl of 0.25 M Tris−HCl, pH 7.8. They were lysed by addition of 0.5% Triton X-100 and the debris were removed by centrifugation in an Eppendorf microfuge for 10 min. CAT activity was assayed in aliquots of the supernatant as described (Gorman *et al.*, 1982). β-galactosidase activity was measured as described (Herbomel *et al.*, 1984). Quantification of the experiments was obtained by cutting out the parts of the chromatograms corresponding to each spot and measuring the radioactivity. The CAT results were normalized with the levels of β-galactosidase. For RNase mapping, total cellular RNA was extracted from the cells 40 h after transfection, using the guanidinium procedure (Chirgwin *et al.*, 1979). RNase mapping was performed as described (Chavrier *et al.*, 1989). The RNA probe was derived from the plasmid pSP-TK-CAT (generous gift from G.Schütz) linearized with *Eco*RI. Transcription with the SP6 polymerase generates a 347 nucleotide probe. Hybridization with correctly initiated tkCAT RNA is expected to protect a 218 nucleotide fragment.

In situ hybridization

^{35}S-Labeled antisense RNA probes were synthesized from the same DNA templates as previously used (Gaunt *et al.*, 1988; Galliot *et al.*, 1989 for *Hox-1.4* and Wilkinson *et al.*, 1989 for *Krox-20*). Embryo recovery, embedding, sectioning and *in situ* hybridization were performed as previously described (Gaunt *et al.*, 1986; Dollé and Duboule, 1989) with no prehybridization step.

Acknowledgements

We thank Professor P.Chambon for his support, Drs R.Tjian and A.Courey for various plasmid constructs and the communication of unpublished information, Dr F.W.Studier for his *E.coli* expression system, Drs R.Miksicek, G.Schütz and P.Gruss for plasmids, Dr Echalier for the SL2 cells and Drs S.A.Ness and D.Bohman for critical reading of the manuscript. This work was supported by grants from the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale (grant no. 881005), the Association pour la Recherche sur le Cancer, the Ligue Nationale Française Contre le Cancer and the Fondation pour la Recherche Médicale. P.C. and C.V. were supported by fellowships from the European Molecular Biology Organization and the Ligue Nationale Française Contre le Cancer, respectively.

References

Akerblom,I.E., Slater,E., Beato,M., Baxter,J.D. and Mellon,P.L. (1987) *Science*, **241**, 350−353.

Berg,J.M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 99−102.

Brown,R.S., Sander,C. and Argos,P. (1985) *FEBS Lett.*, **186**, 271−274.

Chavrier,P., Lemaire,P., Bravo,R. and Charnay,P. (1988a) *Mol. Cell. Biol.*, **8**, 1319−1326.

Chavrier,P., Zerial,M., Lemaire,P., Almendral,J., Bravo,R. and Charnay,P. (1988b) *EMBO J.*, **7**, 29−35.

Chavrier,P., Janssen-Timmen,U., Mattéi,M.-G., Zerial,M., Bravo,R. and

Charnay,P. (1989) *Mol. Cell. Biol.*, **9**, 787–797.
Chen,E.Y. and Seeberg,P.H. (1985) *DNA*, **4**, 165–170.
Chirgwin,J.J., Przbyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Chowdhury,K., Deutsch,U. and Gruss,P. (1987) *Cell*, **48**, 771–778.
Chowdhury,K., Dressler,G., Breier,G., Deutsch,U. and Gruss,P. (1988) *EMBO J.*, **7**, 1345–1353.
Christy,B.A., Lau,L.F. and Nathans,D. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7857–7861.
Courey,A.J. and Tjian,R. (1988) *Cell*, **55**, 887–898.
Di Nocera,P.P. and Dawid,I.B. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7095–7098.
Dollé,P. and Duboule,D. (1989) *EMBO J.*, **8**, 1507–1515.
Duboule,D. and Dollé,P. (1989) *EMBO J.*, **8**, 1497–1505.
Duboule,D., Baron,A., Mähl,P. and Galliot,B. (1986) *EMBO J.*, **5**, 1973–1980.
Fried,A. and Crothers,D.M. (1981) *Nucleic Acids Res.*, **9**, 6505–6525.
Fromental,C., Kanno,M., Nomiyama,H. and Chambon,P. (1988) *Cell*, **54**, 943–953.
Galliot,B., Dollé,P., Vigneron,M., Featherstone,M.S., Baron,A. and Duboule,D. (1989) *Development*, **107**, 343–359.
Garner,M.M. and Revzin,A. (1981) *Nucleic Acids Res.*, **9**, 3047–3059.
Gaunt,S.J., Miller,J.R., Powell,D.J. and Duboule,D. (1986) *Nature*, **324**, 662–664.
Gaunt,S.J., Sharpe,P.T. and Duboule,D. (1988) *Development*, **104** (suppl.), 169–181.
Gaunt,S.J., Krumlauf,R. and Duboule,D. (1989) *Development*, **107**, 131–142.
Gehring,W.J. (1987) *Science*, **236**, 1245–1252.
Gibson,T.J., Postma,J.P.M., Brown,R.S. and Argos,P. (1988) *Protein Eng.*, **2**, 209–218.
Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) *Mol. Cell. Biol.*, **2**, 1044–1051.
Graham,A., Papalopulu,N. and Krumlauf,R. (1989) *Cell*, **57**, 367–378.
Herbomel,P., Bousachot,B. and Yanív,M. (1984) *Cell*, **39**, 653–662.
Hope,I.A. and Struhl,K. (1986) *Cell*, **46**, 885–894.
Ingham,P.W. (1988) *Nature*, **335**, 25–34.
Jones,K.A., Yamamoto,K.R. and Tjian,R. (1985) *Cell*, **42**, 559–572.
Joseph,L.J., Le Beau,M.M,. Jamieson,G.A., Acharya,S., Shows,T.B., Rowley,J.D. and Sukhatme,V.P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 7164–7168.
Kadonaga,J.T., Jones,K.A. and Tjian,R. (1986) *Trends Biochem. Sci.*, **11**, 20–23.
Kadonaga,J., Carner,K.R., Maslarz,F.R. and Tjian,R. (1987) *Cell*, **51**, 1079–1090.
Kadonaga,J.T., Cousey,A.J., Ladika,J. and Tjian,R. (1988) *Science*, **242**, 1566–1570.
Köster,M., Pieler,T., Pöting,A. and Knöchel,W. (1988) *EMBO J.*, **7**, 1735–1741.
Lee,M.S., Gippert,G.P., Soman,K.V., Case,D.A. and Wright,P.E. (1989) *Science*, **245**, 635–637.
Lemaire,P., Relevant,O., Bravo,R. and Charnay,P. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 4691–4695.
Levine,M. and Hoey,T. (1988) *Cell*, **55**, 537–540.
Ma,J. and Ptashne,M. (1987a) *Cell*, **51**, 113–119.
Ma,J. and Ptashne,M. (1987b) *Cell*, **48**, 847–853.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499–560.
McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J. (1984) *Cell*, **7**, 403–408.
Meisfeld,R., Godowski,P.J., Maler,B.A. and Yamamoto,K.R. (1987) *Science*, **236**, 423–427.
Mermod,N., O'Neil,E.A., Kelly,T. and Tjian,R. (1989) *Cell*, **58**, 741–753.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20–78.
Milbrandt,J. (1987) *Science*, **238**, 797–799.
Miller,J., McLachlan,A.D. and Klug,A. (1985) *EMBO J.*, **4**, 1609–1614.
Mitchell,P.J. and Tjian,R. (1989) *Science*, **245**, 371–378.
Morishita,K., Parker,D.S., Mucenski,M.L., Jenkins,N.A., Copeland,N.G. and Ihle,J.N. (1988) *Cell*, **54**, 831–840.
Murphy,P., Davidson,D.R. and Hill,R.E. (1989) *Nature*, **341**, 156–159.
Raymondjean,M., Cereghini,S. and Yaniv,M. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 757–761.
Rosenberg,U.B., Schröder,C., Preiss,A., Keinlin,A., Coté,S., Riede,I. and Jäckle,H. (1986) *Nature*, **319**, 336–339.
Rosenberg,A.H., Lade,B.N., Chui,D.S., Lin,S.W., Dunn,J.J. and Studier,F.W. (1987) *Gene*, **56**, 125–135.
Rubin,M.R., Toth,L.E., Patel,M.D., D'Eustachio,P. and Nguyen-Huu,M.C. (1986) *Science*, **233**, 663–667.
Ruppert,J.M., Kinzler,K.W., Wong,A.J., Bigner,S.H., Kao,F.-T., Law,M.L., Seuanez,H.N., O'Brien,S.J. and Vogelstein,B. (1988) *Mol. Cell. Biol.*, **8**, 3104–3113.
Schneider,I. (1972) *J. Embryol. Exp. Morphol.*, **27**, 253–265.
Strauss,F. and Varshavsky,A. (1984) *Cell*, **37**, 889–901.
Struhl,K. (1987) *Cell*, **50**, 841–846.
Studier,F.W. and Moffatt,B.A. (1986) *J. Mol. Biol.*, **189**, 113–130.
Sukhatme,V.P., Cao,X., Chang,L.C., Tsai-Morris,C.-H., Stamenkovich,D., Ferreira,P.C.P., Cohen,D.R., Edwards,S.A., Shows,T.B., Curran,T., Le Beau,M.M. and Adamson,E.D. (1988) *Cell*, **53**, 37–43.
Tautz,D., Lehmann,R., Schnürch,H., Schuh,R., Seifert,E., Keinlin,A,. Jones,J. and Jäckle,H. (1987) *Nature*, **327**, 383–389.
Wildeman,A., Sassone-Corsi,P., Grünström,T., Zenke,M. and Chambon,P. (1984) *EMBO J.*, **3**, 3129–3133.
Wilkinson,D., Bhatt,S., Chavrier,P., Bravo,R. and Charnay,P. (1989a) *Nature*, **337**, 461–464.
Wilkinson,D.G., Bhatt,S., Cook,M., Boncinelli,E. and Krumlauf,R. (1989b) *Nature*, **341**, 405–409.
Wolgemuth,D.J., Engelmyer,E., Duggal,R.N., Gizang-Ginsberg,E., Mutter,G.L., Ponzetto,C., Viviano,C. and Zakeri,Z.F. (1986) *EMBO J.*, **5**, 1229–1235.
Wolgemuth,D.J., Viviano,C.M., Gizang-Ginsberg,E., Frohman,M.A., Joyner,A.L. and Martin,G.R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5813–5817.

Received on December 11, 1989; revised on January 15, 1990