Development 125, 4185-4193 (1998)
Printed in Great Britain © The Company of Biologists Limited 1998
DEV8517

Targeting gene expression to the head: the *Drosophila orthodenticle* gene is a direct target of the Bicoid morphogen

Qian Gao and Robert Finkelstein*

University of Pennsylvania School of Medicine, Department of Neuroscience, 215 Stemmler Hall, Philadelphia PA 19104, USA
*Author for correspondence (e-mail: finkelsr@mail.med.upenn.edu)

Accepted 26 August ; published on WWW 30 September1998

SUMMARY

The Bicoid (Bcd) morphogen establishes the head and thorax of the *Drosophila* embryo. Bcd activates the transcription of identified target genes in the thoracic segments, but its mechanism of action in the head remains poorly understood. It has been proposed that Bcd directly activates the cephalic gap genes, which are the first zygotic genes to be expressed in the head primordium. It has also been suggested that the affinity of Bcd-binding sites in the promoters of Bcd target genes determines the posterior extent of their expression (the Gene X model). However, both these hypotheses remain untested. Here, we show that a small regulatory region upstream of the cephalic gap gene *orthodenticle (otd)* is sufficient to recapitulate early *otd* expression in the head primordium. This region contains two control elements, each capable of driving *otd*-like expression. The first element has consensus Bcd target sites that bind Bcd in vitro and are necessary for head-specific expression. As predicted by the Gene X model, this element has a relatively low affinity for Bcd. Surprisingly, the second regulatory element has no Bcd sites. Instead, it contains a repeated sequence motif similar to a regulatory element found in the promoters of *otd*-related genes in vertebrates. Our study is the first demonstration that a cephalic gap gene is directly regulated by Bcd. However, it also shows that zygotic gene expression can be targeted to the head primordium without direct Bcd regulation.

Key words: *orthodenticle*, *bicoid*, *Drosophila melanogaster*, head development, pattern formation

INTRODUCTION

The molecular hierarchy that subdivides the trunk of the *Drosophila* embryo along its anteroposterior axis has been studied in great detail (reviewed by St. Johnston and Nusslein-Volhard, 1992; Pankratz and Jackle, 1993). In the first step of this cascade, gradients of maternal proteins activate the expression of the zygotic gap genes. The gap genes, which specify broad regions of the trunk, in turn activate the pair-rule genes, which establish the early parasegmental subdivisions. A similar molecular hierarchy subdivides the posterior (gnathal) region of the head (reviewed by Jurgens and Hartenstein, 1993).

Segmentation of the anterior (cephalic) region of the head is less well understood. The cephalic gap genes, which include *orthodenticle (otd)*, *empty spiracles (ems)*, *buttonhead (btd)*, and *sloppy paired (slp)*, are the first zygotic genes to act in this region. In the blastoderm embryo, these genes are expressed in broad stripes which include the primordia of multiple head segments (Dalton et al., 1989; Finkelstein and Perrimon, 1990; Wimmer et al., 1993; Grossniklaus et al., 1994). Since the pair-rule genes do not appear to function in cephalic segmentation, it has been proposed that anterior head segments are established directly by the cephalic gap genes (Cohen and Jurgens, 1990). Homologues of *otd* and *ems* have been identified in vertebrates (the *Otx* and *Emx* genes), where they are also required for anterior regionalization (reviewed by Finkelstein and Boncinelli, 1994).

In the *Drosophila* embryo, the maternal morphogen Bcd is required for both head and thorax development. *bcd* RNA is localized to the anterior pole of the oocyte, and Bcd protein diffuses to form an anteroposterior concentration gradient. In embryos lacking maternal *bcd* activity, the head and thorax are replaced by duplicated tail structures (Frohnhofer and Nusslein-Volhard, 1986). The *bcd* gene encodes a homeodomain protein (Frigerio et al., 1986; Berleth et al., 1988) that is capable of activating transcription in vitro (Struhl et al., 1989; Driever and Nusslein-Volhard, 1989).

According to an early model, the affinity of Bcd-binding sites in the promoters of target genes determines the posterior extent of their expression (Struhl et al., 1989; Driever et al., 1989). This hypothesis, referred to as the Gene X model, predicts that genes expressed specifically in the head primordium will contain low affinity Bcd sites, so that high levels of Bcd protein are required for their activation. Higher affinity Bcd sites would permit gene expression extending into the thoracic primordium. Other parameters, such as the spacing between Bcd sites (Hanes et al., 1994) and cooperative binding (Ma et al., 1996) have also been proposed to affect *bcd* target gene regulation. However, the importance of all these factors in the regulation of actual *bcd* target genes has not been determined.

To understand how *bcd* generates head-specific expression, it is critical to determine which zygotic genes expressed in the head are directly regulated by *bcd*. The gap gene *hunchback (hb)*, required for thoracic and gnathal development (Lehmann and Nusslein-Volhard, 1987), is a direct target of *bcd* (Struhl et al., 1989; Driever and Nusslein-Volhard, 1989). A second gap gene, *tailless (tll)*, required for the formation of the unsegmented termini of the embryo (Strecker et al., 1988), also contains functional Bcd sites in its promoter region and is probably directly regulated by *bcd* (Liaw and Lengyel, 1992). Finally, the gap gene *giant* is a candidate *bcd* target gene (Kraut and Levine, 1991). The targets of *bcd* regulation in the cephalic segments, however, remain undetermined.

The anterior domains of expression of *otd, ems, btd*, and *slp* appear very early in embryogenesis, and are absolutely dependent on maternal *bcd* function (Dalton et al., 1989; Finkelstein and Perrimon, 1990; Wimmer et al., 1993; Grossniklaus et al., 1994). In addition, increases in *bcd* dosage cause a posterior expansion of the expression domain of each gene. Although *bcd* is the only gene product absolutely required for cephalic gap gene expression, at least two additional genetic inputs are necessary for the correct specification of the boundaries of expression. The first is the terminal maternal system, required for the formation of the anterior and posterior termini of the embryo (reviewed by Sprenger and Nusslein-Volhard, 1993). The terminal cascade prevents *otd* expression near the anterior pole of the embryo, and also contributes to the activation of specific cephalic gap genes within their normal expression domains (Finkelstein and Perrimon, 1990; Grossniklaus et al., 1994; Gao et al., 1996). It has been suggested that the terminal system affects Bcd target gene expression indirectly, by phosphorylating Bcd and decreasing or increasing its potency as a transcription factor (Ronchi et al., 1993; Grossniklaus et al., 1994). The second genetic requirement for correct cephalic gap gene activation is the gap gene *hb*. In embryos lacking maternal and zygotic *hb* activity, cephalic gap gene expression shifts anteriorly (Simpson-Brose et al., 1994; Gao et al., 1996).

Despite these genetic studies, the molecular mechanism of cephalic gap gene activation remains undetermined. To understand how gene expression is targeted to the anterior head, we are focusing on the regulation of the *otd* gene. Our first goal is to determine whether *bcd* activates *otd* directly, and to understand why this activation is limited to the head primordium. Our second goal is to understand the contribution of other genes in establishing head-specific expression. Since no *bcd* homologue exists in vertebrates, this second objective is particularly important.

In this study, we show that a 900 bp fragment upstream of the *otd* gene is sufficient to generate *otd*-like head expression. This fragment drives only early head expression, indicating that additional regulatory elements are required for later *otd* expression in the embryonic brain. Dissection of the 900 bp fragment shows that it contains two smaller elements, each sufficient to drive head expression. The sequence of the first element reveals three candidate Bcd-binding sites. We demonstrate that this element binds Bcd in vitro and that the three Bcd sites are essential for the function of this element in vivo. Consistent with the Gene X model, this regulatory element has a lower affinity for Bcd than an equivalent enhancer from the *hb* promoter. The second element does not bind Bcd. Instead it contains a sequence repeat resembling a motif required to target the expression of vertebrate *Otx* genes to cephalic mesenchyme (Kimura et al., 1997). We discuss the implications of our results for understanding head formation in *Drosophila* and higher animals.

MATERIALS AND METHODS

Fly stocks and P-element-mediated germ line transformation

Transgenes were introduced into *yw* homozygous flies. The dominant allele tor^{4021} causes constitutive *tor* activation and expansion of the embryonic termini (Strecker et al., 1989). The multiple copy *bcd* stock used was *BB bw/CyO*, in which the *BB bw* chromosome contains two copies of the *bcd* gene (Namba and Minden, 1997). For additional information regarding mutant alleles, see Lindsley and Zimm (1992). Embryos were injected as previously described (Spradling and Rubin, 1982). For each construct tested, at least three independent transformant strains were generated and tested.

Plasmid construction, in situ hybridization, and X-gal staining

Constructs including the *otd* transcription start site were made using the P element vector pCaSpeR-AUG-bgal (Thummel et al., 1988). Constructs lacking this region were generated using the vector pCaSpeR hs43 lacZ, which includes an *hsp70* minimal promoter (Thummel and Pirrotta, 1992). Regulatory fragments were inserted into the polylinker located 5′ of the *lacZ* gene in either vector. Larger deletions were generated using the restriction sites described in the figure legends. All other deletions were constructed by PCR using appropriate primers and verified by sequence analysis.

Whole-mount in situ hybridization was performed using digoxigenin-UTP-labeled *lacZ* or *otd* antisense RNA probes as previously described (Gao et al., 1996). X-gal staining was performed according to the method of Brand and Perrimon (1993).

Sequence analysis and mutagenesis

Sequence analysis was performed using automated Taq DyeDeoxy^{Tm} Terminater cycle sequencing reagents and an ABI PRISM 377 DNA Sequencer. Candidate binding sites for specific regulatory proteins were determined using the TFSEARCH program (can be obtained from http://pdap1.trc.rwcp.or.jp/research/db/TFSEARCH.html).

Inactivation of the three candidate Bcd binding sites in the 186 bp fragment was performed by PCR using appropriate oligonucleotides. Two base pairs in each site were changed as follows: B1 from CTTAATCT to CTTAGTGTC, B2 from TTTAATCGC to TTTAGTGGC, and B3 from GTTAATCCT to GTTAGTGCT. The altered base pairs are required for Bcd binding in vitro (Hanes and Brent, 1991). The oligonucleotides used were 26mers each containing the two base changes. All the mutagenized constructs were verified by sequence analysis.

Gel retardation assays

Gel retardation assays (Figs 6A and 7A-C) were performed in 20 μ1 of 1× BB buffer (prepared according to Ma et al., 1996) containing 0.1 mg/ml dI:dC (Pharmacia) for 20 minutes at room temperature. Longer reaction times yielded similar results, suggesting that binding assays had reached equilibrium by 20 minutes (also see Ma et al., 1996). 0.1 mg/ml bovine serum albumin (BSA) was also added to the reactions described in Fig. 6A. This full length recombinant Bcd protein, generously provided by X. Ma, was produced in insect Sf-9 cells using the baculovirus expression vector pJV NheI. To prevent nonspecific binding, 5 μg/lane of single-stranded DNA was added. The concentration of ^{32}P-labeled probes was 4.5×10^{-11} M. 3 μ1 of 20% Ficoll was added to each reaction before loading onto a native polyacrylamide gel in 0.5× TBE buffer. Unlabeled Bcd site competitor DNA was generated by annealing two 23 mer

oligonucleotides, Bcd1A (CTAGACGAGTCTAATCCCATGAT) and Bcd1B (CTAGATCATGGGATTAGACTCGT) (gifts from D. Wilson and C. Desplan). These oligonucleotides include a 19 bp region of overlap containing the 9 bp Bcd consensus binding site (TCTAATCCC).

Because of difficulties in obtaining sufficient quantities of full length Bcd, we used a truncated Bcd protein (amino acids 89-154, including the homeodomain) purified from bacteria for the experiments described in Fig. 7. Previous studies showed that this protein has similar binding properties to full length Bcd (Rivera-Pomar et al., 1995). We also performed smaller scale experiments with the full length protein that confirmed the results shown in the figure (data not shown). In the experiments shown, labeled and unlabeled fragments were added simultaneously and reactions performed for 20 minutes at room temperature.

RESULTS

Embryonic regulatory elements lie upstream of the *otd* gene

To localize the control elements required for embryonic head expression, we constructed a series of *lacZ* reporter fusions spanning the *otd* genomic region. Each of these constructs was introduced into flies by P-element-mediated transformation, and *lacZ* expression monitored during embryogenesis. We found that a 7.6 kb fragment extending upstream of the *otd* transcriptional start site was sufficient to recapitulate the endogenous pattern of *otd* head expression. We also tested additional regions upstream of the 7.6 kb fragment and downstream of the *otd* gene, and found that they did not drive head-specific expression (data not shown).

We then compared the pattern of endogenous *otd* expression to that driven by the 7.6 kb regulatory fragment (Fig. 1B-I). As described previously (Finkelstein and Perrimon, 1990; Gao et al., 1996), *otd* is expressed initially at relatively low levels in an anterior cap-like region of the syncytial blastoderm embryo (Fig. 1B). The posterior boundary of this early expression domain is not sharp, but is graded in intensity. Expression quickly fades from the anterior terminus, leaving a stripe extending from 75-92% egg length (EL) in the cellular blastoderm embryo (Fig. 1C). During this period, ventral expression also disappears (not shown; see Gao et al., 1996). By this stage, the anterior and posterior boundaries of *otd* expression are sharply defined. During germ band extension, *otd* expression becomes more complex, appearing at the ventral midline (Fig. 1D) and in other regions of the embryo. In the germ band-retracted embryo, expression can be seen in the anterior brain and in midline CNS cells (Fig. 1E). This expression persists through embryogenesis.

In the blastoderm embryo, *lacZ* expression driven by the 7.6 kb fragment is indistinguishable from endogenous *otd* expression (compare Fig. 1B and 1F; 1C and 1G). Later in embryogenesis, *lacZ* expression in the anterior head and in midline cells is similar, but not identical to *otd* expression at equivalent developmental stages (compare Fig. 1D and 1H; 1E and 1I). Expression of the transgene is less localized within the head primordium, and significantly weaker in midline cells. This suggests that additional regulatory elements are required for correct late expression.

Identification of the regulatory region that controls early *otd* head expression

The results described above indicate that the 7.6 kb fragment contains the regulatory elements that control *otd* expression in the blastoderm head primordium. To map these elements more precisely, we constructed a series of 5′ and 3′ deletions of this fragment (Fig. 2).

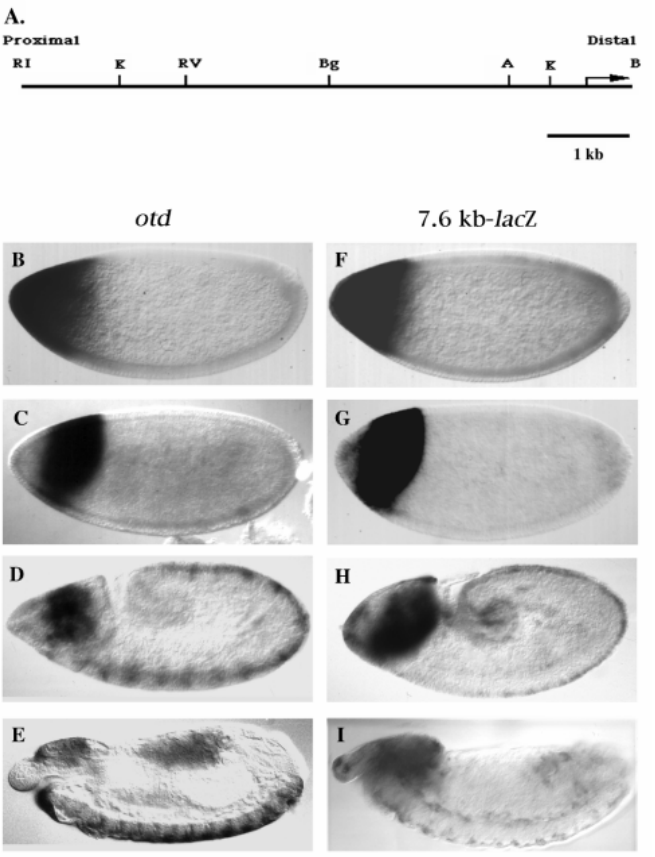

Fig. 1. *otd* embryonic regulatory elements are contained within a 7.6 kb upstream region. (A) Map of the *otd* upstream regulatory region. Shown is a restriction map of the 7.6 kb upstream regulatory fragment. This fragment includes 513 bp of the *otd* 5¢ untranslated region (UTR). The transcriptional initiation site (indicated by an arrow) was defined by primer extension experiments (V. Iyer, Q. Gao, and R. Finkelstein, unpublished results) and adds 35 bp to the 5′ UTR compared to the previously published sequence (Finkelstein et al., 1990). Abbreviations: RI, *Eco*RI; K, *Kpn*I; RV, *Eco*RV; Bg, *Bgl*II; A, *Ase*I; B, *Bam*HI. (B-I) The 7.6 kb-*lacZ* fusion gene mimics embryonic *otd* expression. B-E show the pattern of *otd* RNA expression during embryogenesis, as detected by in situ hybridization. F-I show the expression pattern of the 7.6 kb-*lacZ* fusion gene, detected by in situ hybridization (F-H) or X-gal staining (I). Both *otd* and the 7.6 kb-*lacZ* fusion gene are initially expressed in a cap covering the anterior end of the syncytial blastoderm embryo (B,F). At this stage, the posterior border of expression is graded. As cellularization proceeds, expression disappears from the anterior terminus and the posterior border sharpens (C,G). In the germ band-extended embryo, both *otd* and the 7.6 kb-*lacZ* fusion gene RNA are expressed in the procephalic region and in midline precursor cells (D,H). The patterns of *otd* and *lacZ* expression are no longer identical by this stage, with the fusion gene expressed in a larger region of the head, and at lower levels in midline cells. In germ band-retracted embryos, expression is seen in the brain and ventral nerve cord (E,I). Again the expression patterns are no longer identical, with the fusion gene expressed in a larger region of the brain, as well as in two spots in the clypeolabral region. In all panels, anterior is to the left and dorsal is up.

Deleting 2.1 kb from the 5′ end of the 7.6 kb fragment had no discernible effect on early *lacZ* expression (data not shown). However, further deletion of a 1.8 kb *Eco*RV-*Bgl*II fragment completely eliminated head expression, indicating that essential control elements reside in this region. When fused to a minimal heterologous promoter, the 1.8 kb fragment drove strong *lacZ* expression in the blastoderm head primordium (Fig. 2C). This expression was indistinguishable, in intensity and position, from that produced by the 7.6 kb fragment. Expression driven by this fragment disappears by approximately 5 hours after egg laying (data not shown), demonstrating that a different control region is necessary to initiate or maintain *otd* expression in the head and brain.

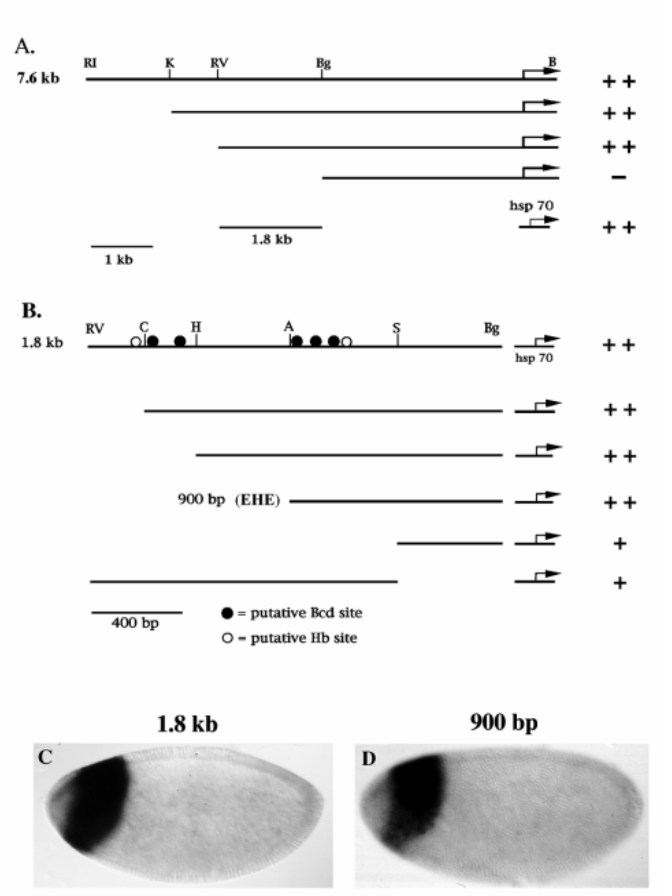

Fig. 2. Localizing the control region required for early *otd* head expression. (A) 5′ deletion analysis reveals that a 1.8 kb *Eco*RV-*Bgl*II fragment contains the regulatory elements that control early head expression. Progressive deletions of the 7.6 kb fragment described in Fig. 1 are shown. To the right of each construct, the intensity of *lacZ* expression is indicated. Two pluses indicate blastoderm expression approximately equivalent to that driven by the 7.6 kb fragment (see Fig. 1C). The first two deletions, to the *Kpn*I and *Eco*RV sites respectively, affect neither the intensity nor the spatial extent of *lacZ* expression. Deletion to the *Bgl*II site completely eliminates expression, indicating that important regulatory elements lie within the 1.8 kb *Eco*RV-*Bgl*II fragment. This fragment, fused to a minimal *hsp70* promoter, produces a blastoderm head stripe indistinguishable from that generated by the 7.6 kb regulatory region (compare Figs 1C and 2C). (B) Further deletion analysis of the 1.8 kb fragment shows that the Early Head Enhancer resides primarily within a 900 bp fragment. Deletion to the *Cla*I site, which eliminates a putative Hb site, does not affect *lacZ* expression. Further deletion to a *Hin*cII site, which eliminates two Bcd sites, very slightly reduces the intensity of expression. Deletion to the *Avr*II site has no further effect on *lacZ* expression. A larger deletion (to the *Spe*I site) significantly reduces expression, indicating that important regulatory sequences lie between the *Avr*II and *Spe*I sites. Since this construct still drives *otd*-like expression, however, a head-specific control element must lie between the *Spe*I and *Bgl*II sites. This was confirmed by a 3′ deletion to the *Spe*I site. (C,D) *lacZ* expression (detected by in situ hybridization) driven by the 1.8 kb (*Eco*RV-*Bgl*II) and 900 bp (*Avr*II-*Bgl*II) fragments. Each fragment drives strong *otd*-like expression in the head primordium. C, *Cla*I; H, *Hin*cII; S, *Spe*I.; other abreviations as in Fig. 1A.

We then determined the effects of additional deletions within the 1.8 kb fragment (Fig. 2B). The sequence of this fragment revealed five candidate Bcd sites and two possible Hb sites (see below). 5′ deletions, which removed two Bcd and one Hb site, very slightly decreased the level of *lacZ* expression but had no effect on its position (data not shown). Additional 5′ and 3′ deletions revealed however that a 900 bp fragment is essential for strong, head-specific *lacZ* expression. Fused to a heterologous promoter, this fragment generated *otd*-like head expression approximately equivalent in intensity to that driven by the 1.8 kb fragment (Fig. 2B,C). Since the 900 bp fragment is the smallest contiguous regulatory region capable of driving strong head expression, we will refer to it as the Early Head Enhancer (EHE).

The Early Head Enhancer responds to *bcd* and *tor* regulation

We tested next whether expression driven by the EHE responds to maternal cues in a similar fashion to *otd* expression. As described above, early *otd* head expression depends on *bcd* and, to a lesser extent, *tor* activity. Decreasing or increasing maternal *bcd* dosage causes endogenous *otd* expression to shift anteriorly or posteriorly (Finkelstein and Perrimon, 1990; Fig. 3A,B). Constitutive *tor* activation results in a posterior expansion of the *otd* domain (Gao et al., 1996; Fig. 3C).

We introduced the 900 bp-*lacZ* reporter construct into genetic backgrounds in which maternal *bcd* or *tor* activity is altered. The boundaries of *lacZ* expression shifted in a similar manner to those of the endogenous *otd* domain (Fig. 3D-F). This indicates that the EHE contains regulatory elements that control the response of *otd* to these two maternal patterning systems.

The Early Head Enhancer contains two independent control elements

To understand how the EHE functions, we mapped it at higher resolution (Fig. 4A). Progressive 5′ deletions showed that a 186 bp element at its 5′ end is critical for maintaining the intensity of early head expression. This region contains the three putative Bcd sites in the EHE. Deletion of this element significantly decreased the intensity of *lacZ* expression, without significantly affecting its spatial extent (compare Fig. 4B and 4C). Further 5′ deletions, which removed a putative Hb site, had no obvious effect on the level or position of *lacZ* expression.

3′ deletions revealed a second important control element at the opposite end of the EHE. Removal of 173 bp from the 3′ end of the 900 bp fragment also reduced the intensity of *lacZ* expression (Fig. 4D). Again, the spatial extent of early head expression was not significantly altered.

Deletion of the region between the 186 bp and 173 bp

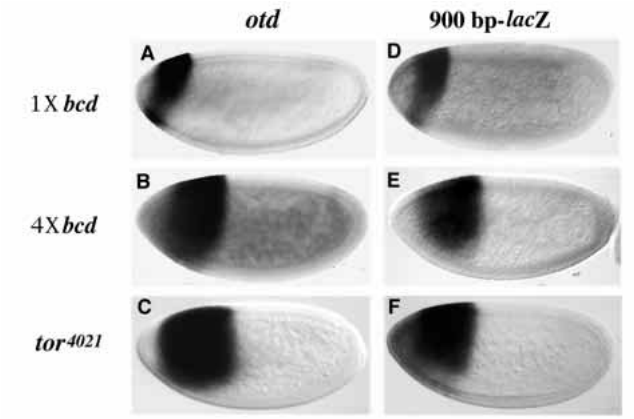

Fig. 3. The 900 bp fragment contains *bcd* and *tor* response elements. (A-C) *otd* expression and (D-F) *lacZ* expression, each detected by in situ hybridization. The embryos in A,D and B,E are from females carrying one copy or four copies of a wild-type *bcd* gene respectively. The embryos in C,F are from mothers carrying the gain-of-function allele tor^{4021}. (A,D) Decreasing maternal *bcd* dosage causes the *otd* blastoderm stripe to shift anteriorly (compare to Fig. 1C). A similar shift is seen in *lacZ* expression driven by the 900 bp fragment. (B,E) Increasing *bcd* dosage causes a posterior shift in *otd* and *lacZ* expression. (C,F) Constitutive *tor* activity causes *otd* expression to expand posteriorly. *lacZ* expression also shifts posteriorly, although to a lesser extent.

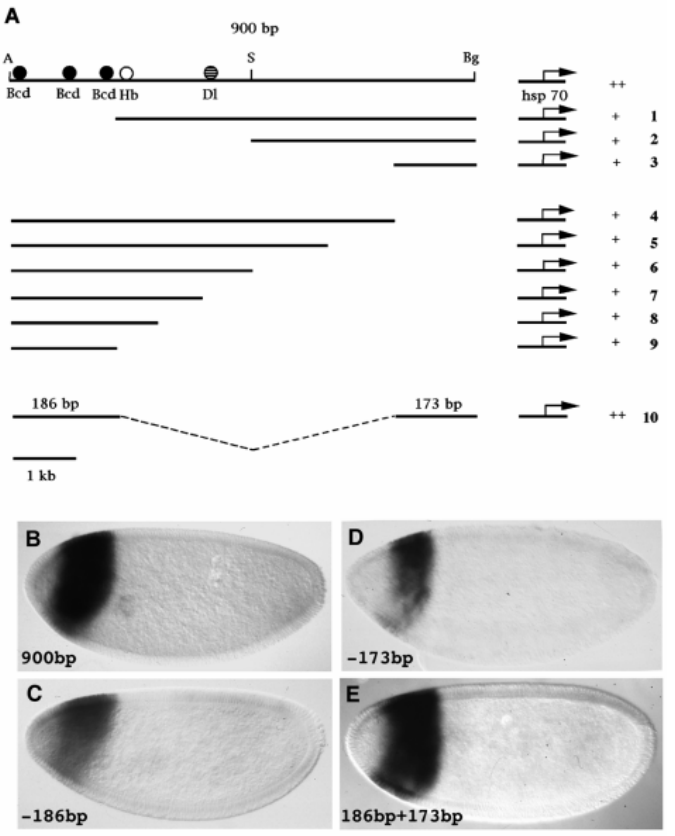

Fig. 4. The 900 bp fragment contains two independent regulatory elements. (A) The top line is a schematic representation of the 900 bp fragment showing putative binding sites for Bcd, Hb, and Dl. Below are a series of deletion constructs (1-12), all of which give *otd*-like head expression. These include three 5′ deletions (1-3), six 3′ deletions (4-9), and an internal deletion (10). Pluses indicate the approximate level of *lacZ* expression. (1-3) A 186 bp 5′ deletion, which removes all three putative Bcd sites (construct 1), significantly reduces the intensity of *lacZ* expression but does not affect its spatial position (compare B and C below) Further deletion of putative Hb and Dl sites (constructs 2 and 3) has no further effect on expression levels (data not shown). (4-9) Deletion of 173 bp at the 3′ end of the 900 bp fragment (construct 4) also significantly reduces *lacZ* expression without altering its position (D). Larger 3′ deletions (constructs 5-9) have no further effect on expression. (10) Fusion of the 186 bp and 173 bp fragments regenerates expression levels equivalent to those produced by the 900 bp fragment (E). (B-E) *lacZ* expression driven by the 900 bp fragment (B) and deletions there of (C-E). (B) The 900 bp fragment drives strong *lacZ* expression in the blastoderm head primordium. (C) The 186 bp 5′ deletion (construct 1), which removes three putative Bcd sites, significantly reduces the intensity of *lacZ* expression, but does not alter its position. (D) A 173 bp 3′ deletion (construct 4) affects *lacZ* expression approximately as in C. (E) An internal deletion that fuses the 186 bp and 173 bp fragments (construct 10) generates *lacZ* expression equivalent to that driven by the 900 bp fragment. This deletion, which removes a putative Dl-binding site, reduces the ventral retraction of *lacZ* expression.

fragments had no obvious effect on *lacZ* expression in the head primordium. When these two fragments were fused and juxtaposed to a minimal promoter, they produced head expression similar in intensity and extent to that driven by the intact EHE (Fig. 4E). This construct also mimicked the response of the EHE to alterations in maternal *bcd* and *tor* expression (data not shown).

These experiments revealed that the activity of the EHE resides primarily within two small regulatory elements, each sufficient to drive *otd*-like expression in the head primordium. As mentioned above, the 186 bp element contains three candidate Bcd binding sites. Each of these sites contains 6 of the 9 nucleotides defined as a high affinity Bcd site in the *hb* promoter (Fig. 5A; Driever and Nusslein-Volhard, 1989a). In particular, each site contains the TAATC core critical for the recognition of purified Bcd protein in vitro (Hanes and Brent, 1991; D. Wilson and C. Desplan, personal communication). The presence of these sequences suggested that Bcd binds directly to the 186 bp fragment.

The 173 bp element does not contain candidate Bcd sites or consensus recognition sites for any of the proteins involved in early head development. However, it does contain a 6 bp motif (G G/C ATCT) tandemly repeated six times (Fig. 5B). A deletion which specifically removes all six copies of this repeat eliminated head-specific expression (data not shown).

As described, the removal of the single putative Hb site did not obviously affect the function of the EHE. This is consistent with our previous observation that *hb* plays a relatively minor role in *otd* activation (Gao et al., 1996). In contrast, the loss of a putative *dl* site that lies between the 186 bp and 173 bp fragments prevented the ventral retraction of *lacZ* expression (Fig. 4E and Q. Gao, unpublished results). This is consistent with our previous finding that *dl* is required for this retraction (Gao et al., 1996). Finally, the EHE also contains possible binding sites for the product of the terminal gap gene *huckebein* (data not shown), which is involved in repressing *otd* expression at the anterior terminus of the embryo (Gao et al.,

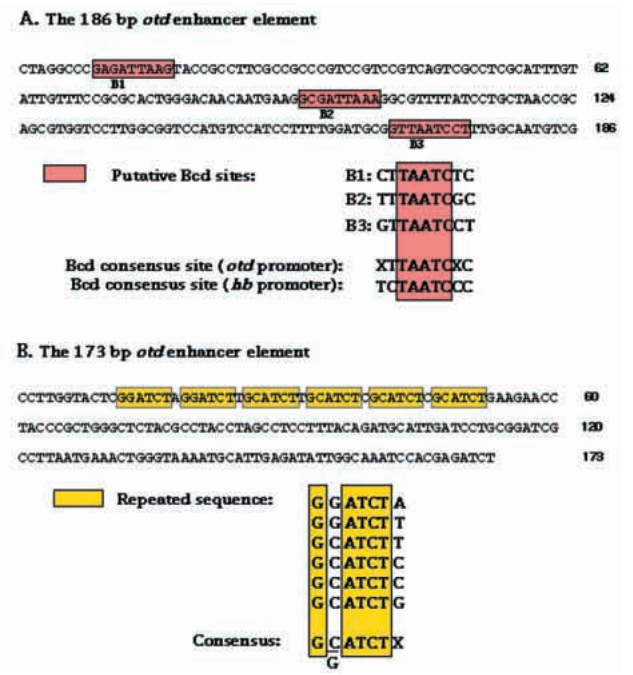

Fig. 5. Sequence motifs in the 186 and 173 bp elements. (A) The 186 bp element contains three candidate Bcd-binding sites (B1-B3). A comparison of the sequences of these sites to that of the 9 bp Bcd consensus site defined in the *hb* promoter is shown (Driever and Nusslein-Volhard, 1989). (B) The 173 bp element contains no candidate Bcd sites. It does contain a 6 bp motif repeated six times.

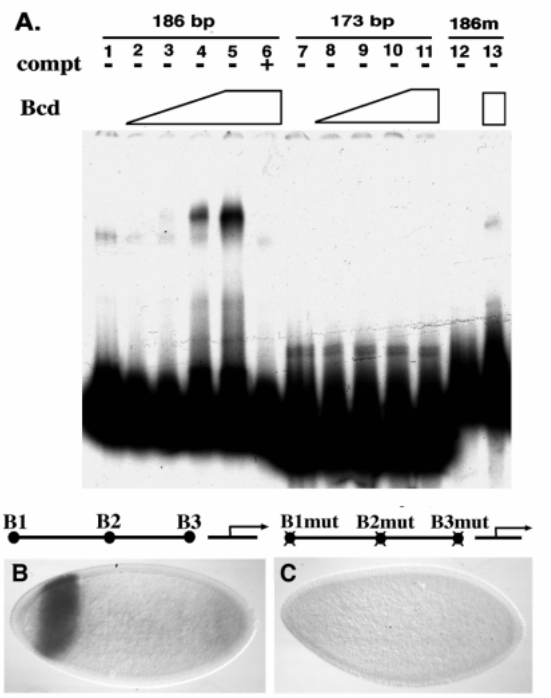

Fig. 6. Bcd binds to the 186 bp element and is required for its activity in vivo. (A) Gel retardation assays performed with Bcd protein purified from insect cells (see Materials and Methods). The DNA fragments tested were the 186 bp element (lanes 1-6), the 173 bp element (lanes 7-11), and the mutagenized 186 bp element lacking consensus Bcd sites (lanes 12,13). The volumes of Bcd protein added to the reactions were as follows; lanes1 and 7, 0 ml; lanes 2 and 8, 0.0625 ml; lanes 3 and 9, 0.25 m1; lanes 4 and 10, 1.0 ml; and lanes 5 and 11, 4.0 ml. 4 ml protein was added to the reactions in lanes 6 and 13, and no protein was added top the reaction in lane 12. To the reaction in lane 6, an unlabeled oligonucleotide containing a high affinity Bcd-binding site was added as competitor (see Materials and Methods). (B) *lacZ* expression in the blastoderm head primordium driven by the wild-type 186 bp element. (C) Inactivation of the three Bcd sites in the 186 bp element eliminates *lacZ* expression.

1996). We did not analyze the function of these putative sites further. However, *lacZ* expression driven by the *Eco*RV-*Avr*II fragment, which lacks these sites, retracts only partially from the anterior pole (Fig. 7C).

The Early Head Enhancer binds Bcd with lower affinity than a *hb* enhancer

One of the goals of this study was to determine whether Bcd directly activates *otd* in the head primordium. We therefore tested whether purified Bcd protein binds to either of the two control elements within the EHE. Consistent with the sequence of the 186 bp element, gel retardation assays showed that it binds Bcd in vitro (Fig. 6A). No binding to the 173 bp element could be detected.

We also tested whether the three Bcd consensus sites in the 186 bp fragment are required for its activity in vivo. Using site-directed mutagenesis, we inactivated all three sites and verified that they no longer bind Bcd (Fig. 6A, lane 13). The mutated 186 bp fragment was reintroduced into flies, and *lacZ* expression monitored in the resulting embryos. Removal of the three Bcd consensus sites completely abolished head expression (compare Fig. 6B and 6C). This indicates that these three sites are essential for the head-specific regulatory activity of this fragment.

In its original form, the Gene X model predicted that a gene expressed specifically in the head primordium would have lower affinity Bcd sites than genes expressed more posteriorly. As mentioned earlier, in addition to the affinity of isolated Bcd sites, subsequent studies showed that intersite spacing, the number of sites, and cooperative binding effects all contribute to the affinity of regulatory regions for Bcd. We therefore compared the overall affinity of the 186 bp element for Bcd to that of a 250 bp enhancer from the *hb* promoter. The *hb* enhancer drives *lacZ* expression across both the head and thoracic primordia and binds Bcd with high affinity (Driever and Nusslein-Volhard, 1989). We found that significantly higher Bcd levels were required in gel retardation assays to shift the labeled 186 bp fragment than the labeled *hb* regulatory element (Fig. 7A,B). In a second assay, we compared the abilities of the two fragments to compete away Bcd protein from binding to Bcd sites. Consistent with the previous result, higher levels of the 186 bp fragment were necessary to inhibit Bcd binding (Fig. 7C). These experiments demonstrate that, consistent with the Gene X model, the overall affinity of the *otd* regulatory element for Bcd is lower than that of the *hb* enhancer.

Correct *otd* regulation requires the repression of posterior expression

The 1.8 kb regulatory fragment contains two candidate Bcd and

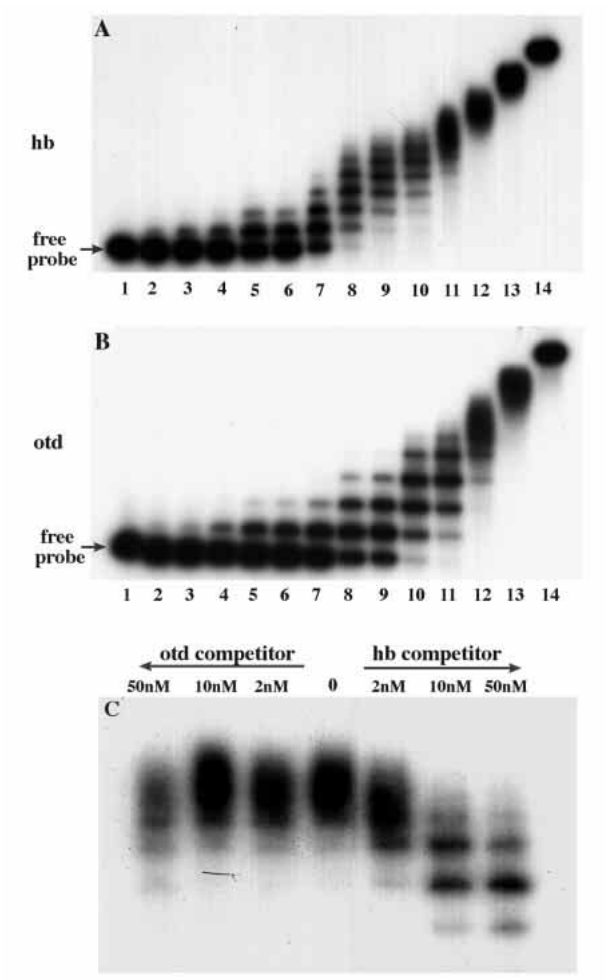

Fig. 7. The 186 bp element binds Bcd with lower affinity than a regulatory element from the *hb* promoter. (A,B) Gel retardation assays performed with a truncated Bcd protein purified from bacteria (see Materials and Methods). Similar results were obtained with full length Bcd (data not shown). The ^{32}P-labeled fragments are the 186 bp element (A) and a 250 bp enhancer from the *hb* gene, which contains at least 6 Bcd sites (B). Arrows indicate position of free probes. The concentrations of Bcd protein in the reactions shown in lanes 1-14 were 0 nM, 39 nM, 78 nM, 156 nM, 312 nM, 468 nM, 624 nM, 1.248 mM, 1.872 mM, 2.496 mM, 3.744 mM, 4.992 mM, 9.984 mM, and 19.968 mM respectively. (A) As Bcd concentration increases, progressive shifting of the *hb* regulatory fragment occurs. Multiple bands are generated by Bcd binding to each of the sites in the fragment, and possibly by the formation of super complexes of Bcd (Ma et al., 1996). Note that a Bcd concentration of 1.872 mM almost completely eliminates free probe. (B) Shifting of the 186 bp element by increasing concentrations of Bcd. A Bcd concentration of 3.744 mM, twice that required for the *hb* enhancer, is necessary to eliminate most of the free probe. (C) Competition experiment. ^{32}P-labeled 186 bp fragment, in the presence of 5.0 mM truncated Bcd protein (approximately equivalent to the concentration used in tjhe reactions in lane 12 of panels A and B), was incubated with increasing amounts of unlabeled competitor fragments. These competitor fragments, the 186 bp *otd* fragment or the *hb* enhancer fragment, were added at the same time as the labeled fragment. It can be seen that the *hb* fragment competes away Bcd protein more effectively than the 186 bp fragment. Equivalent results were obtained using the ^{32}P-labeled *hb* probe with the same two cold competitors (data not shown).

one Hb site that lie upstream of the EHE. As mentioned earlier, deletion of the region containing these sites caused a slight decrease in the intensity of expression. To determine whether this region is sufficient to drive early head expression, we generated more fusion constructs and tested their functions in vivo.

Unexpectedly, we found that a 526 bp *Eco*RV-*Hin*cII fragment containing these sites drove both anterior and posterior *lacZ* expression (Fig. 8D,E). This expression resembled that of the terminal gap gene *tll* (Pignoni et al., 1990, 1992), suggesting that this fragment contains terminal system

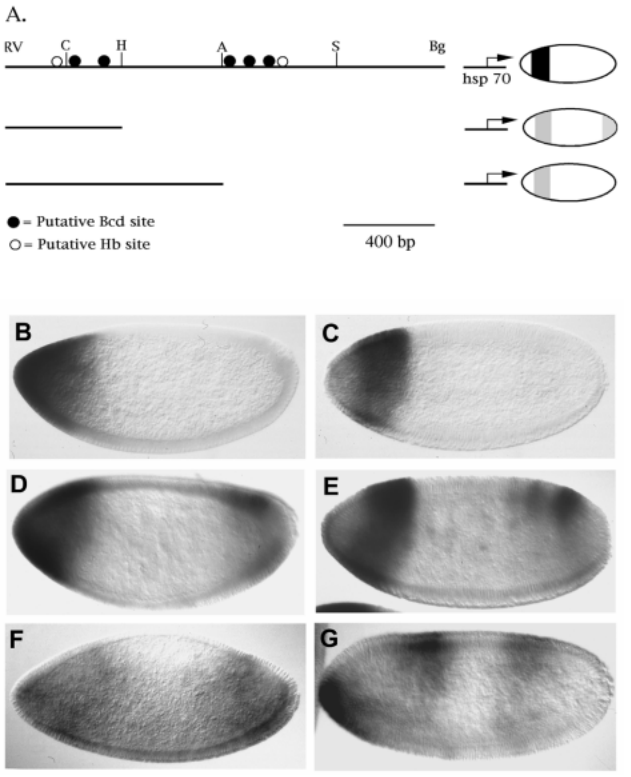

Fig. 8. Correct *otd* expression requires posterior repression. In A, the top line shows the 1.8 kb regulatory fragment, which controls early *otd* expression in the head primordium (abbreviations as in Fig. 3A). Below are shown two 3′ deletion constructs. *lacZ* expression is schematized on the right. B-G show the results of in situ hybridization of various transgenic embryos with a digoxigenin-labeled *lacZ* probe. (A) As described earlier, the *Eco*RV-*Bgl*II 1.8 kb fragment generates *otd*-like anterior expression. A smaller *Eco*RV-*Hin*cII subfragment generates not only anterior, but also ectopic posterior expression. Restoration of the region between the *Hin*cII and *Avr*II sites eliminates posterior expression, suggesting the presence of a negative regulatory region in that fragment. (B,C) The anterior, *otd*-like expression driven by the *Eco*RV-*Avr*II subfragment in syncytial (B) and cellular (C) blastoderm embryos. Note that expression levels are significantly reduced compared to those driven by the entire 1.8 kb fragment (see Fig. 3B). (D,E) The ectopic posterior expression conferred by the *Eco*RV-*Hin*cII subfragment in syncytial (D) and cellular (E) blastoderm embryos. The posterior domain in D becomes replaced by discrete stripes of expression in E. (F,G) The *Eco*RV-*Hin*cII subfragment responds to terminal system regulation. Anterior and posterior expression both expand in syncytial (F) and cellular (G) blastoderm embryos derived from *tor*^{4021} females.

response elements. Consistent with this idea, the anterior and posterior expression domains specified by this fragment both expanded in embryos derived from tor^{D} females (Fig. 8F,G).

Since *otd* is not expressed at the posterior pole, we hypothesized that additional regulatory elements exist that prevent posterior expression. To test this idea, we examined expression driven by a larger regulatory fragment extending to the 5′ end of the EHE. This fragment drove expression only in the head primordium, indicating that it contains a negative regulatory element that represses posterior expression (Fig. 8B,C).

DISCUSSION

Identification of a regulatory region that specifically drives early head expression

The cephalic gap genes are initially expressed in broad stripes within the blastoderm head primordium. Later, each gene is transcribed in a more restricted region of the anterior head and brain. An important issue in anterior patterning is whether the early and late phases of cephalic gap gene expression are independently controlled.

Here, we showed that a 7.6 kb *otd* regulatory fragment accurately recapitulates early expression in the head primordium. This fragment continues to drive expression in a smaller region of the anterior head and brain later in embryonic development. Dissection of the 7.6 kb fragment led to the identification of a 900 bp region (the EHE) that contains the major control elements responsible for early head expression. Shortly after the cellular blastoderm stage, *lacZ* expression produced by the EHE disappears entirely. This demonstrates that there are indeed distinct control elements required for the re-initiation or maintainance of *otd* expression later in embryogenesis.

bcd and the mechanism of *otd* activation

One of the goals of this study was to determine if *bcd* directly regulates early *otd* expression. Dissection of the EHE showed that it contains a 186 bp fragment sufficient to generate *otd*-like expression in the head primordium. This fragment contains Bcd consensus sites that are required for its function in vivo and bind Bcd in vitro. This demonstrates that Bcd, or a protein with similar binding specificity, is required to activate *otd* expression in the head primordium. Since no such protein has been identified that is expressed early enough to function in *otd* activation, this result strongly suggests that Bcd participates directly in the regulation of a cephalic gap gene.

We have also shown that, consistent with the Gene X model, the 186 bp fragment binds Bcd with lower affinity than an analogous regulatory element from the *hb* gene. In future studies, it will be important to analyze the binding properties of the 186 bp fragment in more detail. It will be interesting to evaluate the contributions of spacing between Bcd sites, cooperative binding effects, and other parameters in determining the affinity of this regulatory element for Bcd. In addition, our results do not rule out the possibility that setting the posterior boundary of *otd* expression also requires repressors that prevent expression in the thoracic primordium. In this regard, it will be important to determine whether the 186 bp element can suppress the more posterior expression driven by the *hb* enhancer.

The terminal system and *otd* regulation

We also demonstrated that the terminal maternal system participates directly in *otd* activation. A 526 bp fragment upstream of the EHE drives expression at both the anterior and posterior poles of the embryo. This fragment also responds to alterations in maternal *tor* activity. It is therefore likely that a transcription factor in the terminal system cascade is directly involved in *otd* activation. Since *otd* is not normally expressed at the posterior pole, control elements must exist that prevent posterior expression. We show that at least some of these elements lie immediately downstream of the 526 bp fragment.

Why should a cephalic gap gene be regulated by the maternal terminal system? *otd*-related genes have been identified in every invertebrate yet analyzed, but *bcd* homologues have only been found in dipterans (Schroder and Sander, 1993). This implies that different maternal gene products must have contributed to *otd* activation in other invertebrates. It has been proposed that *hb* was the primary activator of the cephalic gap genes in primitive invertebrates (Simpson-Brose et al., 1994). However, we have shown that the contribution of *hb* to *otd* expression in the fruitfly embryo is relatively minor. A second possibility is that prior to the appearance of *bcd*, the terminal system played a more important role in *otd* activation than it does in *Drosophila*. Analysis of *otd* in other invertebrates will be required to test this hypothesis.

An evolutionary conserved head-specific regulatory element?

Perhaps the most surprising result of our study is the existence of the 173 bp regulatory element. This element is sufficient to generate early *otd*-like expression, but does not contain consensus Bcd sites or bind Bcd in vitro. This indicates that the 173 bp fragment must contain binding sites for a different activator of early head expression. However, since *lacZ* expression driven by the 173 bp fragment is eliminated in embryos lacking *bcd* (data not shown), this activator must, at least in *Drosophila*, be *bcd*-dependent. The only clue regarding the functional specificity of this activator is the reiterated sequence motif required for the activity of this regulatory element.

The mechanism of *Otx* gene activation in vertebrate embryos is poorly understood. In a recent study, the *cis*-acting elements necessary for *Otx2* expression in the pufferfish and mouse were compared (Kimura et al., 1997). Two short sequence motifs, shared between mouse and fish, were shown to be required for expression in mesencephalic neural crest. One of these motifs (TAAATCTG) shows similarity to the repeat unit we identified in the 173 bp fragment. It will be important to determine whether this repeat unit is sufficient to drive early head expression in flies. It will also be interesting to determine whether the vertebrate control element can direct head-specific expression in the fruitfly embryo.

We are particularly grateful to J. Ma and R. Rivera-Pomar for providing purified full length and truncated Bcd proteins respectively. We thank D. Wilson and C. Desplan for the Bcd site oligos Bcd1A and Bcd 1B, N. Perrimon for the tor^{4021} stock, and J. Minden for the

multiple copy *bcd* stock. We also thank X. Ma, J. Ma, R. Rivera-Pomar, D. Wilson, C. Desplan, and A. Sehgal for helpful discussions and D. Burz and S. Hanes for advice regarding Bcd binding assays. Finally, we are grateful to the two reviewers, who aided significantly in improving the manuscript. This research was funded by National Institutes of Health grant 5R01 GM47985 to R. F.

REFERENCES

Berleth, T., Burri, M., Thoma, G., Bopp, D., Richstein, S., Frigerio, G., Noll, M. and Nusslein-Volhard, C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO J.* **7**, 1749-1756.

Brand, A. and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401-415.

Cohen, S. M. and Jurgens, G. (1990). Gap-like segmentation genes that mediate *Drosophila* head development. *Nature* **346**, 482-485.

Dalton, D., Chadwick, R. and McGinnis, W. (1989). Expression and embryonic function of *empty spiracles*: a *Drosophila* homeobox gene with two patterning functions on the anterior-posterior axis of the embryo. *Genes Dev.* **3**, 1940-1956.

Driever, W. and Nusslein-Volhard, C. (1989). The bicoid protein is a positive regulator of *hunchback* transcription in the early *Drosophila* embryo. *Nature* **337**, 138-143.

Driever, W., Thoma, G. and Nusslein-Volhard, C. (1989). Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the *bicoid* morphogen. *Nature* **340**, 363-367.

Finkelstein, R. and Boncinelli, E. (1994). From fly head to mammalian forebrain: the story of *otd* and *Otx. Trends Genet.* **10**, 310-315.

Finkelstein, R. and Perrimon, N. (1990). The *orthodenticle* gene is regulated by *bicoid* and *torso* and specifies *Drosophila* head development. *Nature* **346**, 485-488.

Finkelstein, R., Smouse, D., Capaci, T. M., Spradling, A. C. and Perrimon, N. (1990). The *orthodenticle* gene encodes a novel homeodomain protein involved in the development of the *Drosophila* nervous system and ocellar visual structures. *Genes Dev.* **4**, 1516-1527.

Frigerio, G., Burri, M., Bopp, D., Baumgartner, S. and Noll, M. (1986). Structure of the segmentation gene *paired* and the Drosophila PRD gene set as part of a gene network. *Cell* **47**, 735-746.

Frohnhofer, H. G. and Nusslein-Volhard, C. (1986). Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. *Nature* **324**, 120-125.

Gao, Q., Wong, Y. and Finkelstein, R. (1996). *orthodenticle* regulation during embryonic head development in *Drosophila*. *Mech. Dev.* **56**, 3-15.

Grossniklaus, U., Cadigan, K. M. and Gehring, W. J. (1994). Three maternal coordinate systems cooperate in the patterning of the *Drosophila* head. *Development* **120**, 3155-3171.

Hanes, S. D. and Brent, R. (1991). A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**, 426-430.

Hanes, S. D., Riddihough, G., Ish-Horowicz, D. and Brent, R. (1994). Specific DNA recognition and intersite spacing are critical for action of the Bicoid morphogen in *Drosophila melanogaster*. *Mol. Cell. Biol.* **14**, 3364-3375.

Jurgens, G. and Hartenstein, V. (1993). The terminal regions of the body pattern. In *The Development of Drosophila melanogaster*. pp. 687-746. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Kimura, C., Takeda, N., Suzuki, M., Oshimura, M., Aizawa, S. and Matsuo, I. (1997). *Cis*-acting elements conserved between mouse and pufferfish *Otx2* genes govern the expression in mesencephalic neural crest cells. *Development* **124**, 3929-3941.

Kraut, R. and Levine, M. (1991). Spatial regulation of the gap gene *giant* during *Drosophila* development. *Development* **111**, 601-609.

Lehmann, R. and Nusslein-Volhard, C. (1987). *hunchback*, a gene required for segmentation of an anterior and posterior region of the *Drosophila* embryo. *Dev. Biol.* **119**, 402-417.

Liaw, G.-J. and Lengyel, J. A. (1992). Control of *tailless* expression by *bicoid*, *dorsal*, and synergistically interacting terminal system regulatory elements. *Mech. Dev.* **40**, 47-61.

Lindsley, D. L. and Zimm, G. G. (1992). *The genome of Drosophila melanogaster*. Ner york: Academic Press.

Ma, X., Yuan, D., Diepold, K., Scarborough, T. and Ma, J. (1996). The *Drosophila* morphogenetic protein Bicoid binds DNA cooperatively. *Development* **122**, 1195-1206.

Namba, R. and Minden, J. S. (1997). *Drosophila* embryonic pattern repair: how embryos respond to Bicoid dosage alteration. *Development* **124**, 1393-1403.

Pankratz, M. J. and Jackle, H. (1993). Blastoderm Segmentation. In *The Development of Drosophila melanogaster*. pp. 467-516. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Pignoni, F., Steingrimsson, E. and Lengyel, J. A. (1992). *bicoid* and the terminal system activate *tailless* expression in the early *Drosophila* embryo. *Development* **115**, 239-251.

Pignoni, R., Baldarelli, R. M., Steingrimsson, E., Diaz, R. J., Patapoutian, A., Merriam, J. R. and Lengyel, J. A. (1990). The *Drosophila* gene *tailless* is expressed at the embryonic termini and is a member of the steroid receptor superfamily. *Cell* **62**, 151-163.

Rivera-Pomar, R., Lu, X., Perrimon, N., Taubert, H. and Jackle, H. (1995). Activation of posterior gap gene expression in the *Drosophila* blastoderm. *Nature* **376**, 253-256.

Ronchi, E., Treisman, J., Dostatni, N., Struhl, G. and Desplan, C. (1993). Down-regulation of the Drosophila morphogen bicoid by the torso receptor-mediated signal transduction cascade. *Cell* **74**, 347-355.

Schroder, R. and Sander, K. (1993). A comparison of transplantable *bicoid* activity and partial *bicoid* homeobox sequences in several *Drosophila* and blowfly species (Calliphoridae). *Roux's Arch. Dev. Biol.* **203**, 34-43.

Simpson-Brose, M., Treisman, J. and Desplan, C. (1994). Synergy between two morphogens, Bicoid and Hunchback, is required for anterior patterning in *Drosophila*. *Cell* **78**, 855-955.

Spradling, A. C. and Rubin, G. M. (1982). Transposition of cloned P elements into Drosophila germ line chromosomes. *Science* **218**, 341-347.

Sprenger, F. and Nusslein-Volhard, C. (1993). The terminal system of axis determination in the *Drosophila* embryo. In *The Development of Drosophila melanogaster*. (M. Bate and A. Martinez-Ariass), pp. 365-386. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

St. Johnston, D. and Nusslein-Volhard, C. (1992). The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68**, 561-572.

Strecker, T. R., Halsell, S. R., Fischer, W. W. and Lipshitz, H. D. (1989). Reciprocal effect of hyper- and hypoactivity mutations in the *Drosophila* pattern gene *torso*. *Science* **243**, 1062-1066.

Strecker, T. R., Merriam, J. R. and Lengyel, J. A. (1988). Graded requirement for the zygotic terminal gene *tailless* in the brain and tail region of the *Drosophila* embryo. *Development* **102**, 721-734.

Struhl, G., Struhl, K. and MacDonald, P. M. (1989). The gradient morphogen *bicoid* is a concentration-dependent transcriptional activator. *Cell* **57**, 1259-1273.

Thummel, C. S., Boulet, A. M. and Lipshitz, H. D. (1988). Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74**, 445-456.

Thummel, C. S. and Pirrotta, V. (1992). New pCaSpeR P element vectors. *Drosophila Information Service* **71**, 150.

Wimmer, E. A., Jackle, H., Pfeifle, C. and Cohen, S. M. (1993). A *Drosophila* homologue of human Sp1 is a head-specific segmentation gene. *Nature* **366**, 690-694.