MOLECULAR AND CELLULAR BIOLOGY, May 1994, p. 3364–3375
0270-7306/94/$04.00+0

Vol. 14, No. 5

Specific DNA Recognition and Intersite Spacing Are Critical for Action of the Bicoid Morphogen†

STEVEN D. HANES,[1]‡ GUY RIDDIHOUGH,[2] DAVID ISH-HOROWICZ,[2] AND ROGER BRENT[1]*

Department of Molecular Biology, Massachusetts General Hospital, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02114,[1] and Imperial Cancer Research Fund and Developmental Biology Unit, Department of Zoology, Oxford OX1 3PS, United Kingdom[2]

Received 18 August 1993/Returned for modification 28 October 1993/Accepted 27 January 1994

We examined DNA site recognition by Bicoid and its importance for pattern formation in developing *Drosophila* embryos. Using altered DNA specificity Bicoid mutants and appropriate reporter genes, we show that Bicoid distinguishes among related DNA-binding sites in vivo by a specific contact between amino acid 9 of its recognition α-helix (lysine 50 of the homeodomain) and bp 7 of the site. This result is consistent with our earlier results using *Saccharomyces cerevisiae* but differs from that predicted by crystallographic analysis of another homeodomain-DNA interaction. Our results also demonstrate that Bicoid binds directly to those genes whose transcription it regulates and that the amino acid 9 contact is necessary for Bicoid to direct anterior pattern formation. In both *Drosophila* embryos and yeast cells, Bicoid requires multiple binding sites to activate transcription of target genes. We find that the distance between binding sites is critical for Bicoid activation but that, unexpectedly, this critical distance differs between *Drosophila* and *S. cerevisiae*. This result suggests that Bicoid activation in *Drosophila* might require an ancillary protein(s) not present in *S. cerevisiae*.

The homeodomain is a conserved DNA-binding domain present in many transcription regulatory proteins (reviewed in reference 64). Homeodomain proteins control a variety of cell fate decisions in many organisms, including mating type determination in yeasts, cell lineage specification in nematodes, and establishment of body axis and cell identity in developing insect and vertebrate embryos (5, 7, 21, 22, 50).

In *Drosophila melanogaster*, proteins with similar homeodomains exhibit different biological functions, in part because of differences in DNA recognition (reviewed in references 32, 42, and 46). For example, the Bicoid homeodomain protein recognizes DNA sites that differ from those recognized by most other homeodomain proteins (13, 27). This site specificity is presumed to be important for its function. Bicoid directs the formation of anterior pattern in the early *Drosophila* embryo (20, 52). The *bicoid* gene (*bcd*) is transcribed maternally, and the mRNA is deposited in the anterior of the developing oocyte. Upon fertilization, the mRNA is translated and Bicoid protein diffuses to form an anterior-to-posterior concentration gradient (3, 11, 20). This gradient is important for Bicoid's action as a morphogen; cells that form in regions of higher Bicoid concentrations adopt more anterior cell fates (12). It is thought that Bicoid acts by stimulating spatially restricted transcription of the zygotic gap gene *hunchback* (11–14, 72) and other potential targets, including the head-specific genes *orthodenticle* (17), *empty spiracles*, and *buttonhead* (6) and the gap gene *Krüppel* (33).

We have studied Bicoid-DNA interaction by using genetic assays with *Saccharomyces cerevisiae* (27, 28). Results of these studies are consistent with a model in which Bicoid contains a recognition α-helix in its homeodomain but, unlike prokaryotic helix-turn-helix proteins (56, 57), Bicoid inserts the carboxy-terminal (rather than amino-terminal) end of this helix into the major groove of DNA. Whereas mutations at positions 1, 2, and 5 of the recognition helix do not affect Bicoid's DNA specificity, mutations at positions 9, 10, and 12 abolish its ability to recognize Bicoid sites. Most dramatically, a single Lys→Gln substitution at position 9 in the recognition helix (residue 50 of the homeodomain) changes Bicoid's DNA specificity so that it now recognizes sites bound by *Antennapedia* (Antp)-class homeodomain proteins (27). Mutations in the binding site that suppress amino acid 9 substitutions show that Lys-9 in the Bicoid recognition helix contacts bp 7 of the Bicoid site (TCTAATCCC), while Gln substituted at this position (as in Antp-class proteins) contacts bp 7 of Antp-class sites (e.g., TTTAATTGA [9]) (28). These experiments also indicate that the recognition helix is aligned with its carboxy terminus tilted toward the 5′ end of the binding site.

These genetic results are in good agreement with those of structural and biochemical studies of the interaction of other homeodomain proteins with DNA (reviewed in references 23, 30, 57, and 74). X-ray crystallography and nuclear magnetic resonance (NMR) analysis showed that the homeodomain adopts a three-α-helix structure (55, 61) in which helices 1 and 2 lie across the DNA backbone, helix 3 (the recognition helix) inserts into the major groove, and an arm amino terminal to helix 1 reaches into the minor groove (39, 54, 80). The unit formed by helix 2 and helix 3 is similar to the helix-turn-helix motif of prokaryotic transcriptional repressors (43, 55, 61). However, as first indicated by our work (27, 28), the geometry of this unit with respect to the DNA is quite different (39, 54, 80).

The results of our genetic studies differed in important details from those of two structural studies: an X-ray structure of an Engrailed homeodomain-DNA cocrystal shows that recognition helix residue 9 (Gln-50 in the Engrailed homeodomain) contacts bp 8 (TGTAATTAC [39]) instead of bp 7 as in our model (28), while an NMR study of an Antennapedia-DNA complex suggests contacts between residue 9 (Gln-50 in the Antennapedia homeodomain) and both bp 7 and bp 8 (TCTAATGGC [54]). These seemingly minor differences in

* Corresponding author.

† S.D.H. dedicates this paper to his first mentor, the late Edgar C. Henshaw of the University of Rochester Cancer Center.

‡ Present address: Axelrod Institute, Wadsworth Center for Laboratories and Research, NY State Department of Health, Albany, NY 12201-2002.

protein-DNA contact might have major implications for site recognition in vivo.

Here, we address these discrepancies and test explicitly the importance of Bicoid's DNA specificity in living *Drosophila* embryos. The results show that in *Drosophila* embryos, as in yeast cells, Bicoid distinguishes among related binding sites by a base-specific contact between recognition helix residue 9 and bp 7. We also show that Bicoid's DNA site specificity is necessary for its action as a morphogen. Finally, we show that Bicoid's ability to activate gene expression depends on the distance between its binding sites. In *Drosophila* embryos, Bicoid does not activate genes that contain Bicoid sites separated by one turn of the DNA helix, whereas Bicoid does activate genes in which Bicoid sites are separated by two and a half turns of the DNA helix. Unexpectedly, the opposite is true in *S. cerevisiae*. These results lead to the hypothesis that Bicoid activation requires ancillary proteins that differ between *S. cerevisiae* and *D. melanogaster*.

MATERIALS AND METHODS

P-element expression constructs. P elements that direct the synthesis of wild-type Bicoid, Bicoid-Q_9, or Bicoid-A_9 in *D. melanogaster* are derivatives of pUChsneo (68) into which were inserted modified versions of an 8.7-kb *Eco*RI genomic *bicoid* fragment (3, 19). The 8.7-kb fragment was previously shown to display normal Bicoid-regulated expression in vivo and to rescue *bcd*$^-$ embryos (3). The constructs were made in several steps. First, the 8.7-kb Bicoid *Eco*RI fragment was subcloned into the *Eco*RI site of pUChsneo, resulting in pUChsneo-8.7. Second, a 3.2-kb *Xba*I genomic Bicoid DNA fragment that includes the homeodomain was inserted into the same site of a pUC19 derivative, pUCΔAccIΔPstI, resulting in plasmid pX3.2. Third, the unique *Pst*I-*Sal*I fragment (401 bp) which contained most of the homeodomain (including the helix-turn-helix motif) was replaced with one of three different PCR products amplified from Bicoid cDNA yeast expression vectors (27) that encoded a wild-type, Q_9, or A_9 version of Bicoid. Replacement of the *Pst*I-*Sal*I fragment of pX3.2 resulted in pX3.2wt, pX3.2Q_9, and pX3.2A_9, respectively. The amplified regions were verified by DNA sequencing. Finally, the 3.2-kb *Xba*I inserts from these derivatives were used to replace the same *Xba*I fragment in partially digested, gel-purified pUChsneo-8.7, resulting in P-element vectors PE7wt2-5, PE7$Q_9$18-27, and PE7$A_9$11-20, respectively. Expression of wild-type Bicoid, Bicoid-Q_9, or Bicoid-A_9 is driven by the normal Bicoid promoter, and resulting transcripts contain normal Bicoid 3′ untranslated sequences.

P-element reporter constructs. Reporter genes were constructed with pβn27.1 (62), which carries a *hsp27-lacZ* fusion gene whose upstream regulatory sequences had been deleted and which encodes a nuclear localized form of β-galactosidase (62). Bicoid binding sites were inserted at the *Bam*HI site of the polylinker upstream of the *hsp27* TATA box.

Two sets of reporter constructs were made. One set carried six direct repeats of the minimal Bicoid binding site (and mutant versions thereof) identical to those used previously in *S. cerevisiae* (28). The center-to-center distance between sites was 11 bp. These constructs were made by insertion of complementary 70-bp oligonucleotides into the *Bam*HI site of pβn27.1. The general form for the upper strands was 5′-GATCA(TCTAATNNNTA)$_5$TCTAATTNNNT, where NNN was CCC for wild-type sites and either TGA or TAG for mutant sites. Another set of reporters carried four direct repeats of the 25-bp A3 Bicoid binding site (13) and were constructed with pairs of overlapping oligonucleotides. The center-to-center distance between sites was 25 bp. For each construct, the four oligonucleotides (two 45-mers and two 57-mers) were annealed to form a 106-bp fragment with *Bam*HI-compatible ends and were inserted into the same site of pβn27.1. The general form of the annealed fragment was GAT[CTGCCCA(TCTAATNNN)TTGACGCTGC]$_4$ for the upper strand. The 9-bp consensus Bicoid site is shown within parentheses. As before, NNN was CCC for wild-type sites and either TGA or TAG for the mutant sites. The scrambled spacer site construct was made with annealed oligonucleotides of the form GATC[AGTAAAC(TCTAATCCC)GGTCAT AGT]$_4$. Base pairs flanking the 9-bp Bicoid site were "scrambled" (by transversions G→T, A→C, T→G, and C→A) to alter the pattern of functional groups displayed in the major groove of DNA.

Germ line transformation and diagnostic PCR. Flies were grown on yeast extract-corn meal-molasses-malt extract-agar medium or yeast extract-corn meal-sucrose-agar medium at 18 or 25°C. P-element constructs were coinjected with pp25.7^{wc} helper DNA (37) into *bw;st* embryos. The P elements were all derivatives of pUChsneo (68) that contained an *hsp70-neomycin* fusion gene for selection of larvae on G418 medium. Survivors were outcrossed to the wild type, and transformants were selected with medium containing 1.2 to 1.6 mg of G418 (Geneticin; GIBCO) per ml as described by Steller and Pirrotta (68). Insertions were mapped to a specific chromosome by backcrossing to *bw;st* and reselection on G418 medium. Insertions were balanced with *FM7*, *CyO*, or *TM3* as appropriate. When possible, stocks were maintained as homozygous *bw* or *st* lines.

The identities of transformed stocks were checked with a PCR assay to distinguish among various P-element insertions in fly genomic DNA. For lines that carried the *bcd*$^+$, *bcd-Q*$_9$, or *bcd-A*$_9$ transgene, the assay detected base pair differences in the first position of the codon corresponding to amino acid 9 in the recognition helix. Diagnostic PCR primers differed only at the 3′-most base: 5′-GCCCAGGTGAAGATATGGTTTN, where N is A, C, or G, corresponding to a wild-type, Q_9, or A_9 derivative, respectively (see reference 27 for details of the mutations). For lines containing the four-site reporters, diagnostic primers for Bicoid, Antp 1, Antp 2, and the scrambled-context Bicoid sites were, respectively, CCGGGGATCTGC CCATCTAATCC, CCGGGGATCTGCCCATCTAATTGA, CCGGGGATCTGCCCATCTAATTAG, and CTCGGTAC CCGGGGATCAGTAAAC.

Genomic fly DNA was prepared by the method of Gloor and Engels (24). Briefly, two female flies were collected in microcentrifuge tubes, frozen on dry ice, and mashed in 100 μl of buffer containing 10 mM Tris (pH 8), 1 mM EDTA, 25 mM NaCl, and 200 μg of proteinase K per ml. The mixture was incubated at 37°C for 20 min and then at 95°C for 5 min to inactivate the proteinase K. After debris was pelleted, 2.5 μl was used directly for PCR or first extracted with phenol (no precipitation with ethyl alcohol). PCRs were done in standard buffer (1.5 mM $MgCl_2$) for 30 cycles of 1 min at 94°C, 1 min at 60°C, and 1.5 min at 72°C.

Rescue experiments. We tested whether *bcd*$^+$, *bcd-Q*$_9$, and *bcd-A*$_9$ transgenes rescued embryos from *bcd*^{E1}/*bcd*^{E1} mothers as follows: *Q*$_9$/*CyO* females were crossed with *bcd*^{E1}/*bcd*^{E1} males. The male progeny (*bcd*^{E1}/+;*Q*$_9$/+) were then crossed with *bcd*^{E1}/*TM3* females. Larvae were selected on G418 food, and *bcd*^{E1}/*bcd*^{E1};*Q*$_9$/+ females were collected and crossed with *bcd*^{E1}/*bcd*^{E1} males. No viable progeny were obtained from *bcd*^{E1}/*bcd*^{E1};*Q*$_9$/+ or *bcd*^{E1}/*bcd*^{E1};*A*$_9$/+ mothers. Viable progeny were obtained from *bcd*^{E1}/*bcd*^{E1};*bcd*$^+$/+ mothers, and all contained visible *bcd*^{E1}-linked markers *p*^{p} *e*^{s} *th*, confirming that

the embryos were derived from mothers deficient for endogenous *bcd*. One line that contained the bcd^+ transgene and two lines each that contained the $bcd\text{-}Q_9$ or $bcd\text{-}A_9$ transgene were tested. All contained single insertions on the second chromosome. A similar strategy was used to test rescue of anterior defects in embryos derived from homozygous bcd^{E1} *nos* mothers. bcd^- nos^- embryos fail to form either anterior structures or posterior structures (52). The double-mutant genotype was used because our bcd^- nos^- stocks are more fertile than our bcd^- stocks. Larval cuticles were prepared as previously described (78) except that vitelline membranes were not removed. Larvae from bcd^- nos^- mothers that carried either $bcd\text{-}Q_9$ or $bcd\text{-}A_9$ transgenes failed to form anterior structures; no head skeleton or thoracic denticle bands (T1, T2, or T3) were visible. In addition, duplication of posterior filzkörper material in the anterior and ectopic expression of denticles in the central portion of the embryo indicated lack of Bicoid activity. Larvae from mothers with bcd^+ transgenes formed normal anterior structures and did not duplicate filzkörper material or show ectopic denticle expression.

Embryo staining. Embryos were collected on apple juice plates, dechorionated, fixed (3.7% formaldehyde–phosphate-buffered saline), and devitellinized by standard procedures (78). *lacZ* expression was detected by immunohistochemical staining with mouse monoclonal antibody (1:500) to β-galactosidase (Promega) and alkaline phosphatase-conjugated goat anti-mouse immunoglobulin (1:1,000; Promega). Polyclonal anti-*eve* antiserum (48) was used at a dilution of 1:200. Staining in situ for *lacZ* RNA was done essentially as by Tautz and Pfeifle (73), except that RNA probes were used and the hybridization was carried out at 70°C and pH 4.5 (40). Stained embryos were washed briefly in ethanol and mounted with a methacrylate embedding kit (JB4; Polysciences). Embryos were photographed under Nomarski optics with a Zeiss Axiophot or Axioplan. For each of the experiments whose results are shown in Fig. 2, 3, 4, and 7, at least two independent transformant lines were tested and shown to give qualitatively similar results. Intensities of staining varied about twofold among individual $bcd\text{-}Q_9$ and $bcd\text{-}A_9$ lines, whereas the reporter lines were more uniform.

Yeast expression plasmids. Our previous studies with *S. cerevisiae* used LexA-Bicoid fusion protein derivatives. Here, we constructed vectors that express native Bicoid derivatives that lacked the LexA moiety. For wild-type Bicoid, plasmid pSH11-1 (27) was cut at the LexA-Bicoid junction (Bicoid amino acid 2) with *Bam*HI, the ends were made flush with mung bean nuclease, and an *Eco*RI blunt-end adaptor was attached. The adaptor sequence was 5′-AATTCATACAATG GCG (upper strand) and 5′-CGCCATTGTATG (lower strand), where the initiator ATG is underlined. The Bicoid cDNA was then released by digestion with *Eco*RI and inserted into the same site of pUC19. The *Eco*RI Bicoid-containing fragment was then inserted into the *Eco*RI site of yeast expression vector pJG4-1ΔE (26). The resulting plasmid, pSH-nBcd, encodes native Bicoid driven by the yeast ADH promoter/terminator and carries a yeast 2 μm replicator and *TRP1* selectable marker. Constructs pSH-nBcdQ_9 and pSH-nBcdA_9 express native Bicoid-Q_9 and Bicoid-A_9, respectively. They were made in the same way as pSH-nBcd, except the starting plasmids were pLexA-BcdQ_9 and pLexA-BcdA_9, respectively (27).

Yeast reporter plasmids. All reporter plasmids were derivatives of pLR1Δ1 (76) and contain the yeast *URA3* gene, a 2 μm replicator, and a *GAL1-lacZ* reporter gene from which the UAS_G was removed. Bicoid binding sites were inserted into the unique *Xho*I site 167 bp upstream of the *GAL1-lacZ* transcription start site.

Reporter genes that carry closely spaced binding sites have been described (28). These constructs contain six direct repeats of the Bicoid consensus binding site TCTAATCCC or the two mutant sites TCTAATTGA and TCTAATTAG, flanked by single-base-pair spacers: A(site)T. The mutant sites differ at three positions from the consensus (underlined). The center-to-center distance between sites is 11 bp.

The reporter genes that carry widely spaced Bicoid binding sites were constructed as follows. Each of the four different P-element vectors that contained widely spaced Bicoid binding sites (see above) were used as substrates for PCRs to amplify an ~106-bp fragment that carried the Bicoid binding sites and added *Sal*I ends. The PCR fragments were digested with *Sal*I, gel purified, and inserted into the *Xho*I site of pLR1Δ1. The resulting plasmids carried four direct repeats of the Bicoid consensus binding site TCTAATCCC or the Antp 1 site, TCTAATTGA, or Antp 2 site, TCTAATTAG. The binding sites were contained within a repeat unit with the configuration: CTGCCCA(site)TTGACGCTG or AGTAAAC(site)G GTCATAGT for the scrambled-spacer reporter. The center-to-center distance between sites is 25 bp.

Yeast methods. Yeast strain MGLD4-4a (**a** *ura3-52 leu2 his3 trp1 lys2 cyr*^{r}) was used throughout. Bicoid expression constructs and reporter *GAL1-lacZ* constructs were cotransformed into MGLD4-4a. Culture methods, DNA transformations, and β-galactosidase assays have been described (27). Units of β-galactosidase activity are expressed as $(1{,}000)A_{420}$/(minutes of reaction)(cell volume)(optical density at 600 nm) and varied by no more than 20% between individual transformants.

RESULTS

We introduced two sets of transgenes into *D. melanogaster*: transgenes that encode altered-DNA-specificity Bicoid proteins and reporter transgenes that carry the cognate binding sites (Fig. 1). Transgenes $bcd\text{-}Q_9$ and $bcd\text{-}A_9$ encode proteins in which Lys-9 of the Bicoid recognition helix is replaced by Gln or Ala, respectively. Both mutations abolish the ability of Bicoid to recognize Bicoid sites in *S. cerevisiae*, and the Lys→Gln substitution results in a mutant protein (Bicoid-Q_9) that instead recognizes Antp-class sites (27). Expression of $bcd\text{-}Q_9$, $bcd\text{-}A_9$, and a control bcd^+ transgene is driven by the normal *bcd* promoter from a genomic DNA fragment that also includes *bcd* 3′ untranslated sequences to ensure proper spatial and temporal expression (3, 49) of the mutant proteins.

Bicoid site reporter genes were made by replacing the upstream regulatory region of an *hsp27-lacZ* fusion gene (62) with four copies of a 25-bp sequence derived from the *hunchback* promoter. Each 25-bp repeat contains a 9-bp Bicoid site (TCTAATCCC [13]). Similarly designed reporter genes show Bicoid-dependent anterior expression in embryos (14, 70). Antp-class reporter genes were made with the same 25-bp repeat, except that critical 3′ base pairs in each Bicoid site were mutated, converting them into Antp-class sites (Antp 1, TCTAATTGA, and Antp 2, TCTAATTAG). These sites are recognized by Bicoid-Q_9 and Antp-class homeodomain proteins in *S. cerevisiae* (28).

Bicoid-Q_9, but not wild-type Bicoid, activates Antp-class reporter genes. We crossed wild-type (bcd^+) females to males carrying different reporter genes and stained the progeny embryos for *lacZ* expression. Blastoderm-stage embryos (stage 5, 2 to 3 h after fertilization) that contain Bicoid site reporters show strong anterior-specific staining (Fig. 2), whereas no

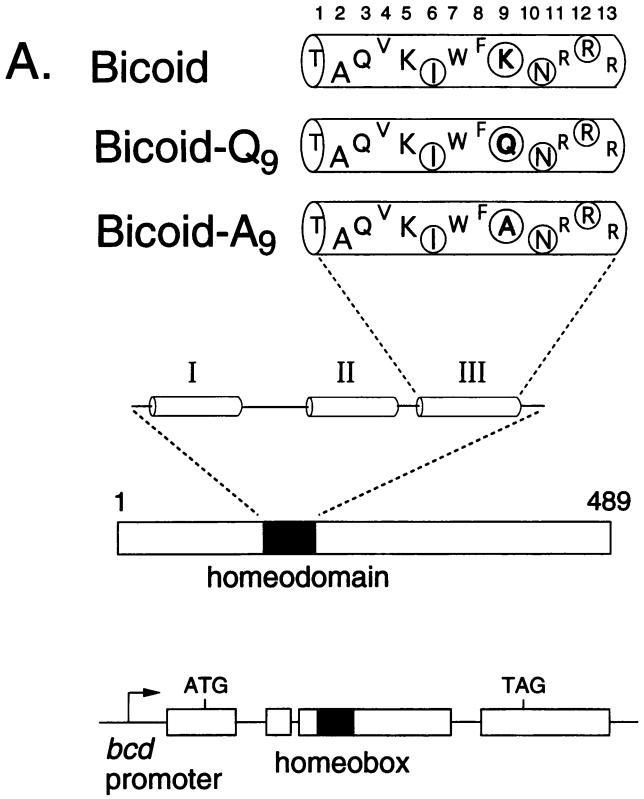

FIG. 1. Experimental design. (A) Schematic of Bicoid derivatives expressed in *D. melanogaster*, detailing homeodomain α-helix III (the recognition helix). Transgenes encoding Bicoid, Bicoid-Q_9, and Bicoid-A_9 were introduced into *D. melanogaster* by P-element-mediated transformation. Transgenes were constructed with *bcd* genomic DNA; the exon-intron structure shown is adapted from that of Berleth et al. (3). Transgene expression is driven by the normal *bcd* promoter as indicated. The spliced mRNA encodes a 489-amino-acid protein. Predicted α-helices in the Bicoid homeodomain are designated I, II, and III. Wild-type Bicoid contains a Lys (K) at position 9 in the recognition helix, Bicoid-Q_9 contains Gln (Q), and Bicoid-A_9 contains Ala (A) at this position. The circled residues at positions 6, 9, 10, and 12 have been shown to make important DNA contacts in Bicoid, Engrailed, and α2 (27, 28, 39, 80). (B) Schematic of reporter genes introduced into *D. melanogaster* by P-element transformation. Multiple copies of the indicated binding sites (Bicoid, Antp 1, and Antp 2) were installed upstream of an *hsp27-lacZ* fusion gene from which the heat shock response elements had been removed (62). Multiple sites were configured with an intersite spacing of either 11 bp (about one turn of the DNA helix) or 25 bp (about two and a half turns of the DNA helix). For details, see Materials and Methods. The Bicoid site is derived from the upstream regulatory region of *hunchback* (13). The Antp 1 and Antp 2 sites are mutated versions of the Bicoid site that bind Bicoid-Q_9 and Antp-class homeodomain proteins in *S. cerevisiae* (27, 28).

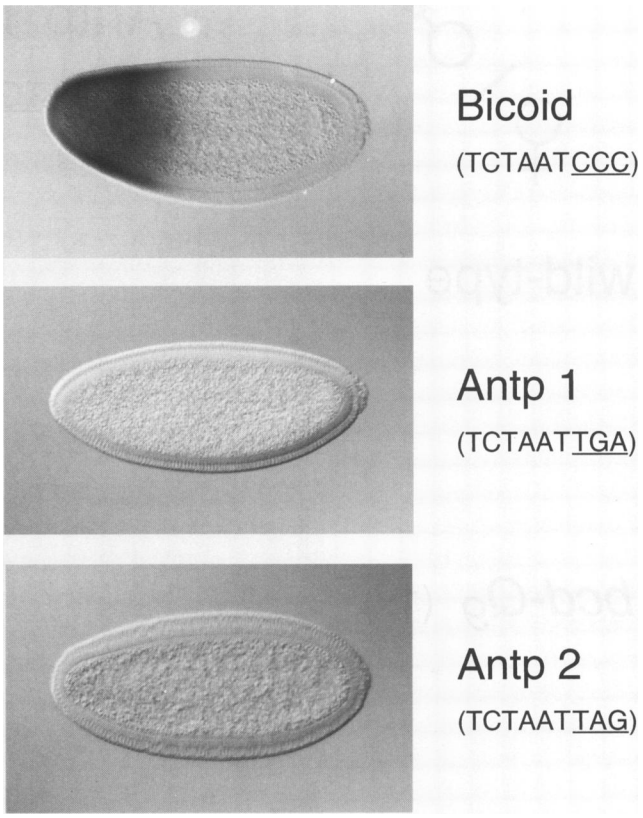

FIG. 2. Anterior-specific activation of Bicoid site target genes but not Antp site target genes. Whole-mount preparations of embryos were stained with anti-β-galactosidase antibodies and visualized under Nomarski illumination, with the anterior on the left and the dorsal facing up. Embryos were derived from wild-type (bcd^+) females that carried reporter *hsp27-lacZ* transgenes with the indicated binding sites. Staining patterns indicate that wild-type Bicoid recognizes Bicoid binding sites but not Antp-class binding sites in blastoderm-stage embryos.

anterior staining is observed in embryos that contain Antp 1 or Antp 2 reporters. Thus, in *D. melanogaster*, as in *S. cerevisiae*, wild-type Bicoid recognizes Bicoid sites but not Antp-class sites. The Antp site reporters are functional, as evidenced by their activation in neural cells at later stages (5 to 9 h [data not shown]) and their activation by Bicoid-Q_9 (see below). They are not activated in blastoderm-stage embryos by other homeodomain proteins that contain Gln at position 9 in their recognition helices (e.g., Even-skipped, Fushi tarazu, and Engrailed); this result is consistent with earlier studies with Antp site reporter constructs (75).

In contrast, Fig. 3 shows that both Antp site reporters are activated in the anterior of embryos from females homozygous for *bcd-Q_9* (i.e., carrying two copies of the *bcd-Q_9* transgene and two endogenous bcd^+ genes). Thus, appropriate 3′ base pair changes in the binding site suppress the Lys→Gln mutation in the Bicoid recognition helix. Base pairs in Antp 1 and

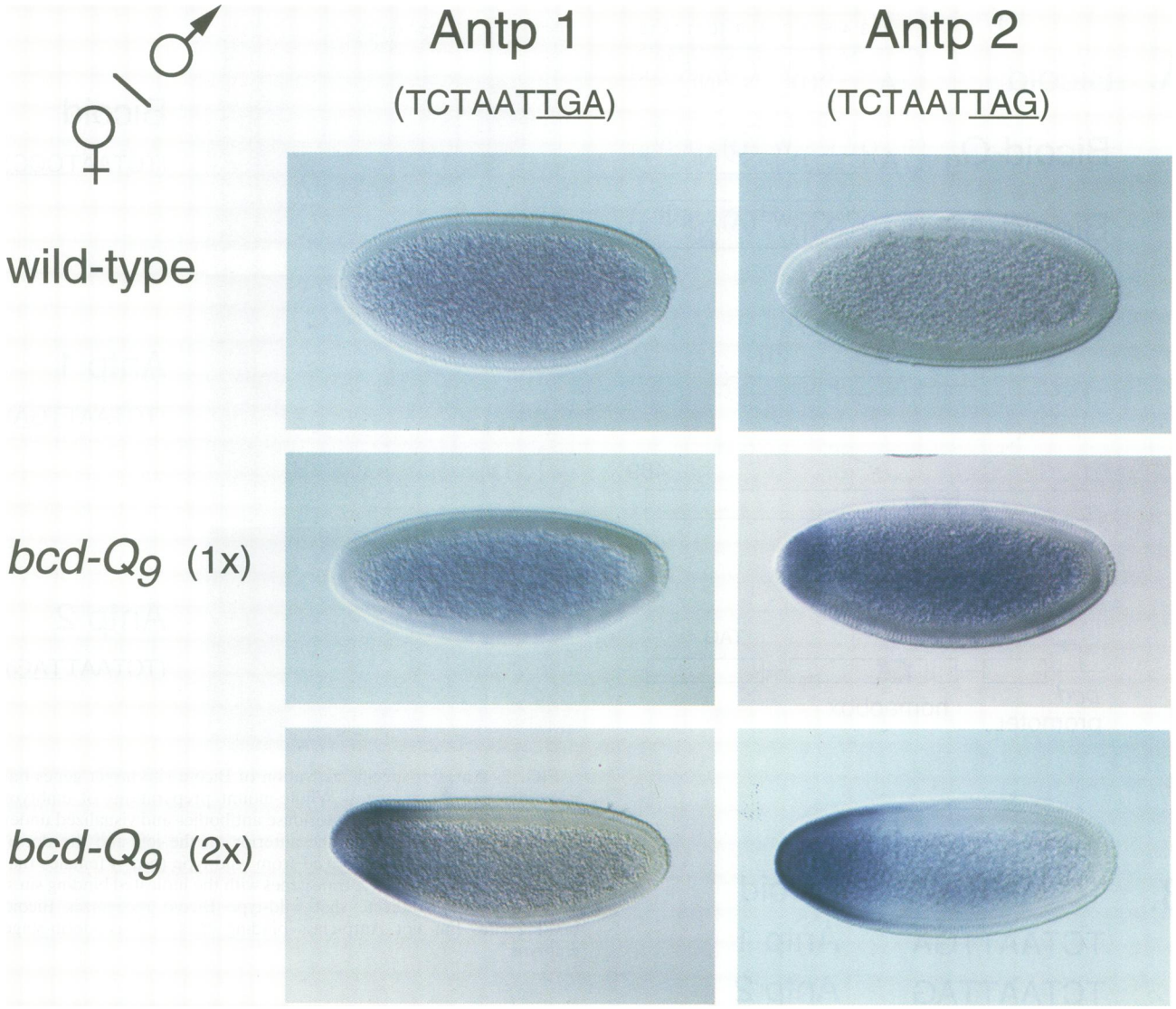

FIG. 3. A single amino acid substitution at position 9 in the homeodomain recognition helix (helix III) causes Bicoid to have an altered target gene specificity in vivo. Embryos from wild-type (bcd^+/bcd^+) females or from wild-type females that also contained one (1x) or two (2x) copies of $bcd\text{-}Q_9$ were mated to males that carried either the Antp 1 (column 1) or the Antp 2 (column 2) reporter genes and were stained with anti-β-galactosidase antibodies. The $bcd\text{-}Q_9$ embryos show dose-dependent anterior activation of Antp 1 and Antp 2 reporter genes. Anterior-specific staining is weaker than that seen for wild-type Bicoid (Fig. 2, top panel), probably because $bcd\text{-}Q_9$ embryos in this figure carried only a single copy of the indicated reporter gene, whereas embryos in Fig. 2 (top panel) carried two copies of the Bicoid site reporter. Control crosses (not shown) in which $bcd\text{-}Q_9$ transgenes were paternally derived showed no anterior staining. Both Antp 1 and Antp 2 sites contain the same base pair (T · A) at position 7 but contain different base pairs at position 8 (G · C versus A · T). Note that the Antp 1 reporter is recognized by Bicoid-Q_9 despite the fact it does not contain thymine at position 8, which in the X-ray structure (39) is seen to be contacted by residue 9 (Gln-50).

Antp 2 that differ from the Bicoid site are found at positions 7, 8, and 9. The only change common to both Antp sites is a C→T transition at position 7, indicating that any base-specific contact by Gln must be to this position (see Discussion). Base pairs 8 and 9 also contribute to overall DNA-binding affinity, as evidenced by the stronger activation of Antp 2 reporters compared with that of Antp 1 reporters (Fig. 3). The fact that different nucleotides are present at positions 8 and 9, however, indicates that any contacts to these positions cannot be strictly base specific.

Bicoid-Q_9 activation of Antp site reporters is dose dependent and shows a maternal effect. Embryos from females with one copy of $bcd\text{-}Q_9$ stain only weakly, and the domain of expression is more anteriorly restricted (Fig. 3). Embryos from control crosses in which $bcd\text{-}Q_9$ is paternally inherited do not stain (not shown). These results are as expected if anterior staining is dependent upon activation by Bicoid-Q_9.

Bicoid-A_9 is a new altered-specificity protein. The above results are consistent with a previously proposed model in which Gln-9 in the recognition helix of Bicoid-Q_9 forms two hydrogen bonds with adenine of T · A at position 7 (28, 65). Alanine cannot form hydrogen bonds, and in *S. cerevisiae* Bicoid-A_9 does not recognize either Bicoid sites or the Antp 1 site (28). We introduced Bicoid-A_9 into *D. melanogaster* as a

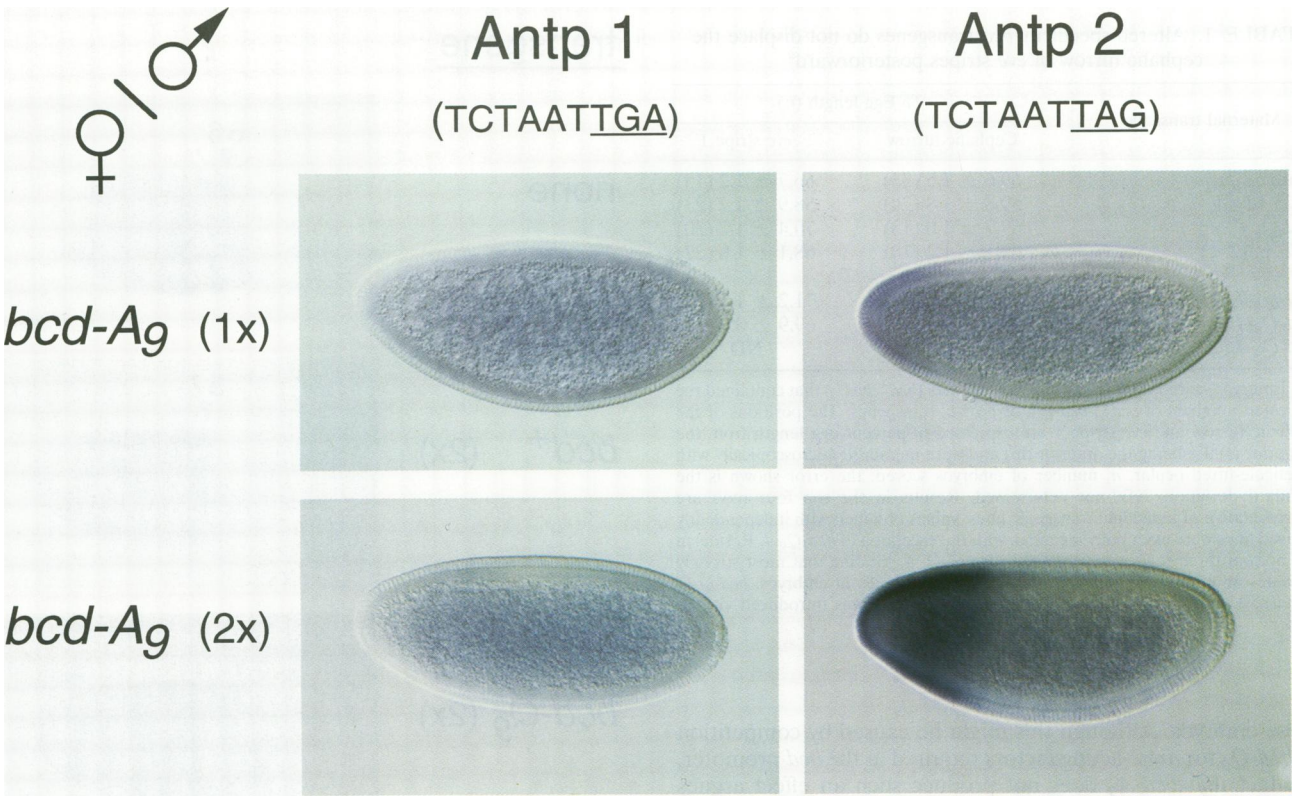

FIG. 4. Bicoid-A_9 is a new altered-specificity mutant. Embryos from wild-type (bcd^+/bcd^+) females that also contained one (1x) or two (2x) copies of the *bcd-A*$_9$ transgene were stained with anti-β-galactosidase antibodies. Females had been crossed to males that carried either the Antp 1 or Antp 2 site reporter genes. The *bcd-A*$_9$ embryos show dose-dependent anterior activation of the Antp 2 reporter but not of the Antp 1 reporter. Control crosses (not shown) in which *bcd-A*$_9$ transgenes were paternally derived show no anterior staining.

negative control for DNA recognition, and as expected Bicoid-A_9 does not recognize Bicoid sites (described below) or Antp 1 sites (Fig. 4) in embryos. However, to our surprise, Bicoid-A_9 does recognize Antp 2 sites (Fig. 4). Activation of Antp 2 reporters is dependent on maternal *bcd-A*$_9$ gene dosage (Fig. 4) and does not occur if *bcd-A*$_9$ is paternally inherited (not shown).

Transcription assays with *S. cerevisiae* confirm that Bicoid-A_9 recognizes Antp 2 but not Bicoid sites or Antp 1 (see below). Thus, Bicoid-A_9 is a new altered-specificity protein that recognizes a subset of Antp-class sites, suggesting that alanine at position 9 in the recognition helix might make a specific DNA contact (see Discussion).

Bicoid-Q_9 and Bicoid-A_9 cannot direct formation of anterior structures. Like other transcriptional activators, Bicoid is modular; its DNA-binding activity and gene activation functions are separable (10, 27). Bicoid's activation domain is necessary for anterior pattern formation and can be replaced by heterologous activation domains (10). To show that Bicoid's DNA recognition function is also necessary for anterior pattern formation, we used Bicoid-Q_9 and Bicoid-A_9, which are altered in proper site recognition but are otherwise indistinguishable from wild-type Bicoid. Like the wild-type protein, they are maternal effect, dose-dependent activators of anterior transcription.

We performed crosses to introduce bcd^+, *bcd-Q*$_9$, and *bcd-A*$_9$ transgenes into females homozygous for bcd^{E1}, a frameshift mutation that produces no detectable Bicoid in embryos (11, 70). In contrast to the control bcd^+ transgene, neither *bcd-Q*$_9$ nor *bcd-A*$_9$ rescues bcd^{E1} embryos to produce viable progeny (data not shown). We also tested whether bcd^+, *bcd-Q*$_9$, and *bcd-A*$_9$ transgenes could restore anterior structures to bcd^- nos^- double-mutant embryos. Neither *bcd-Q*$_9$ nor *bcd-A*$_9$ corrects anterior defects in these embryos, whereas the control bcd^+ transgene corrects anterior defects to yield larvae indistinguishable from nos^- single-mutant embryos (data not shown).

To obtain an independent measure of Bicoid-Q_9 and Bicoid-A_9 activity in embryos, we measured the position of the cephalic furrow, an invagination of anterior cells that demarcates the division between head and thorax. Increases in maternal bcd^+ gene dosage displace this furrow towards the posterior (12). Two extra copies of bcd^+ displace the cephalic furrow posteriorly by about 10% egg length (Table 1). Two maternal copies of either *bcd-Q*$_9$ or *bcd-A*$_9$ cause no such shift. We confirmed this result by visualizing the striped pattern of the *even-skipped* (*eve*) pair-rule protein which is also known to be displaced posteriorly in response to increasing bcd^+ gene dosage (12). In *bcd-Q*$_9$ and *bcd-A*$_9$ embryos, no posterior shift of the *eve* stripes is observed (Table 1; Fig. 5), confirming that altered-specificity mutants lack wild-type Bicoid activity.

The *bcd-Q*$_9$ transgenic line (*bcd-Q*$_9$*-4*) that most strongly activates anterior expression of Antp site reporters causes a small but reproducible shift of the cephalic furrow and the *eve* >^{9}stripes towards the anterior (Table 1; Fig. 5). The extent of this anterior shift is dose dependent and shows a maternal effect (Table 1, and data not shown). The results suggest that Bicoid-Q_9 reduces the amount of wild-type Bicoid activity in

TABLE 1. Altered-specificity *bcd* transgenes do not displace the cephalic furrow or *eve* stripes posteriorward[a]

Maternal transgene	% Egg length (*n*)	
	Cephalic furrow	*eve* stripe 1
None	69.6 ± 1.5 (33)	66.7 ± 1.2 (15)
bcd^+ (2×)	59.8 ± 1.8 (24)	58.9 ± 1.2 (23)
$bcd\text{-}Q_9$ (2×)	69.7 ± 1.0 (18)	70.4 ± 1.2 (20)
$bcd\text{-}A_9$ (2×)	69.1 ± 1.2 (18)	68.1 ± 1.0 (22)
$bcd\text{-}Q_9$-*4* (2×)	72.1 ± 1.4 (49)	71.2 ± 1.3 (27)
$bcd\text{-}Q_9$-*4* (1×)	71.4 ± 1.0 (20)	69.9 ± 0.9 (14)
$bcd\text{-}Q_9$-*4* (2×, paternal)	68.2 ± 0.9 (25)	ND

[a] Embryos scored were from wild-type females (bcd^+/bcd^+) that contained the indicated numbers of bcd^+, $bcd\text{-}Q_9$, or $bcd\text{-}A_9$ transgenes. The positions of the cephalic furrow and *eve* stripe 1 are expressed in percent egg length from the posterior (100% being the anterior tip) and were measured microscopically with a reticule-fitted ocular. *n*, number of embryos scored; the error shown is the standard deviation. ND, not determined. Results in the top four rows are representative of individual transgenic lines; values obtained with independently derived lines were essentially identical, with the exception of $bcd\text{-}Q_9$-*4*. Results in the bottom three rows were obtained for $bcd\text{-}Q_9$-*4*, the line that most strongly activates anterior expression of Antp-class target genes in embryos. $bcd\text{-}Q_9$-*4* (paternal) indicates a control cross in which $bcd\text{-}Q_9$-*4* was introduced via the father.

these embryos. Although this might be caused by competition of $bcd\text{-}Q_9$ for *trans*-acting factors required at the *bcd* promoter, the fact that $bcd\text{-}A_9$ does not produce such an effect argues against this possibility. Instead, we favor the idea that monomers of Bicoid-Q_9 and wild-type Bicoid associate to form mixed oligomers that bind to Bicoid sites with reduced affinity or that are sequestered to other sites in the genome, mechanisms analogous to those described for *lac* i^{-d} and λ CP dominant negative repressor mutants (1, 53). Lack of a dominant negative effect by Bicoid-A_9 might be explained by preferential formation of homooligomers as seen for the homeodomain protein HNF1 (which also has an alanine at position 9 of its recognition helix [51]) or by the possibility that mixed oligomers retain adequate affinity for Bicoid sites.

Bicoid requires a different intersite spacing in flies and yeasts. Bicoid requires multiple binding sites to activate gene expression in *S. cerevisiae* and *D. melanogaster* (13, 27, 70). The reporters described above carry four copies of the Bicoid or *Antp*-class binding sites spaced with a center-to-center distance of 25 nucleotides, about two and a half turns of the helix in B-form DNA. This arrangement places adjacent binding sites on opposite faces of the DNA helix. These reporters are strongly activated in embryos by the appropriate Bicoid derivatives (Fig. 2, 3, and 4). In contrast, otherwise equivalent reporter genes whose sites (six copies) have a center-to-center distance of 11 nucleotides, about one turn of the DNA helix, are not activated in embryos. This arrangement places all sites on (approximately) the same side of the DNA helix. We do not detect activation of these reporters by using either antibodies against β-galactosidase protein as in Fig. 2, 3, and 4 (data not shown) or the more sensitive method of in situ hybridization with antisense *lacZ* RNA probes (Fig. 6; data not shown). At later stages of development when Bicoid is no longer present (5 to 9 h), these constructs are expressed, indicating that they are functional (data not shown).

Activation of reporters with widely spaced (25 bp) sites might be due to a sequence-specific activator that binds nucleotides that flank the Bicoid sites but which are absent in reporters with closely spaced (11 bp) sites. We ruled out this possibility by testing activation of a scrambled-spacer reporter

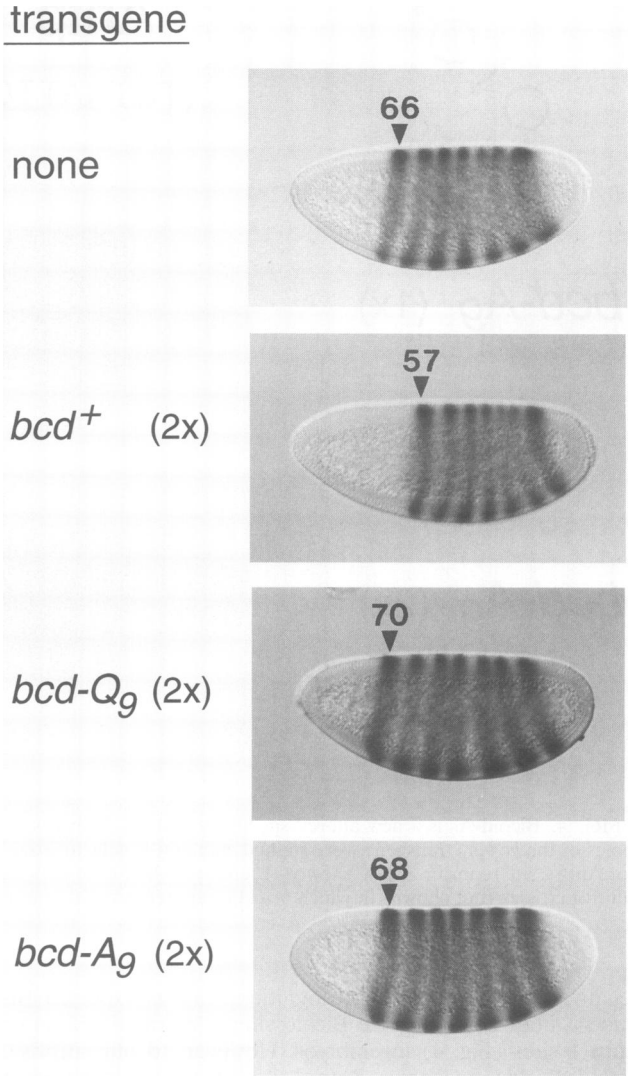

FIG. 5. Altered-specificity Bicoid proteins do not alter the embryonic fate map, as does wild-type Bicoid. Embryos were collected from wild-type (bcd^+/bcd^+) females that carried two copies of the bcd^+, $bcd\text{-}Q_9$, or $bcd\text{-}A_9$ transgenes and were stained with antibodies to the *even-skipped* protein. The number indicates the position of *eve* stripe one (percent egg length from the posterior). The *eve* stripes are displaced posteriorly by about 10% egg length by two copies of the bcd^+ transgene but not by two copies of either the $bcd\text{-}Q_9$ or the $bcd\text{-}A_9$ transgene. $bcd\text{-}Q_9$ causes a small (ca. 3%) but reproducible shift of the stripes anteriorward. Embryos shown are representative of results given in Table 1.

in which the Bicoid sites and their intersite distances are maintained but each flanking nucleotide is altered. Figure 6 shows that this reporter is activated, indicating that flanking nucleotides are not a target for a sequence-specific activator. Indeed, this reporter gene is more strongly activated, and expression extends more posteriorly, than any Bicoid-responsive gene reported to date (two independent *D. melanogaster* lines gave identical results).

Results of Fig. 6 demonstrate that Bicoid activation in embryos is highly dependent on the distance between binding sites and perhaps on their relative positions about the DNA axis. We find that Bicoid activation in yeasts is also highly dependent on the distance between binding sites. However, as

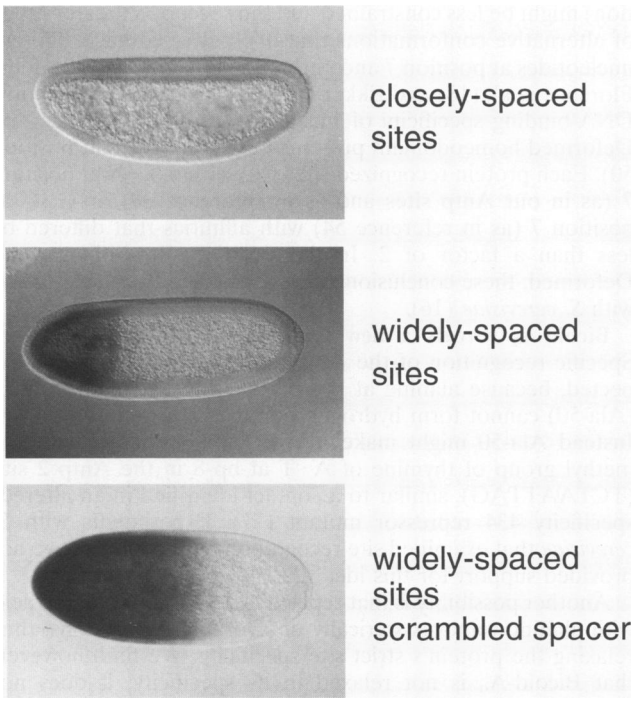

FIG. 6. Intersite distance is critical for Bicoid-dependent gene activation in embryos. *lacZ* expression was detected by in situ hybridization with an antisense RNA probe. Embryos were from wild-type (*bcd*$^+$) females that contained reporter genes with multiple Bicoid binding sites arranged in one of three different ways (see Materials and Methods). Closely spaced (top panel), six copies of the Bicoid site (TCTAATCCC) spaced with a center-to-center distance of 11 bp. Widely spaced (middle panel), four copies of the Bicoid site spaced with a center-to-center distance of 25 bp. Widely spaced sites, scrambled spacer (bottom panel), are identical to the widely spaced construct except that each base pair flanking the Bicoid sites was changed, but the distance between sites was maintained. The widely spaced sites are similar to the A3 site of the *hunchback* promoter (14). The faint anterior and posterior stripes in embryos containing closely spaced sites are not Bicoid specific; equivalent constructs carrying Antp-class sites give identical patterns (not shown).

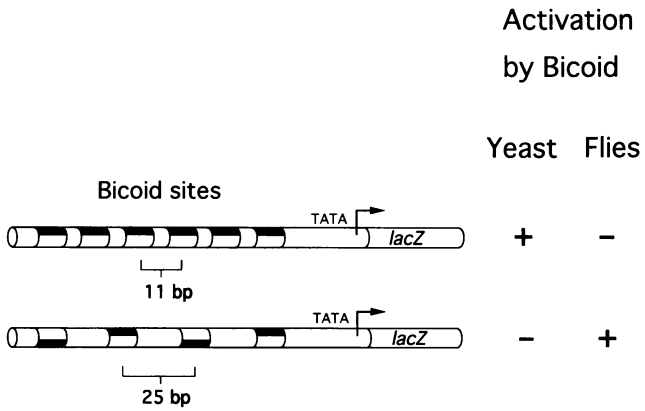

FIG. 7. The distance between binding sites required for Bicoid to activate transcription is different in *D. melanogaster* and *S. cerevisiae*. The schematic represents reporter genes that carry multiple Bicoid binding sites with a center-to-center distance of 11 or 25 bp; this arrangement places adjacent sites on (approximately) the same or opposite sites of the DNA helix, respectively. In the figure, binding sites are represented by darkened segments. Only half of each site is visible, the remainder being on the back side of the DNA helix. Results are summarized from experiments with *D. melanogaster* (Fig. 2, 3, and 5, and data not shown) and *S. cerevisiae* (Table 2).

shown in Table 2, the critical distance is different from that in *D. melanogaster*. In *S. cerevisiae*, the widely spaced site reporters are activated only weakly, if at all, whereas the closely spaced site reporters are activated strongly. This difference between Bicoid activation in the two organisms (summarized in Fig. 7) may reflect the need for species-specific ancillary proteins to help Bicoid cooperatively activate gene expression (see Discussion).

DISCUSSION

Amino acid 9 of the Bicoid recognition helix contacts bp 7. The X-ray structure of an Engrailed homeodomain-DNA cocrystal shows a direct interaction between recognition helix residue 9 (Gln-50) and the thymine methyl group of A · T at position 8 (TGTAATTAC [39]). Our data, however, suggest that this contact is not critical for Bicoid-DNA recognition in

TABLE 2. Sequence and intersite distance requirements for Bicoid to activate transcription in *D. melanogaster* and *S. cerevisiae*

Binding site and intersite spacing[a]	*D. melanogaster*[b]			*S. cerevisiae*[c]		
	Bicoid	Bicoid-Q_9	Bicoid-A_9	Bicoid	Bicoid-Q_9	Bicoid-A_9
Bicoid (TCTAATCCC)						
A, close	–	–	–	161	<0.1	<0.1
B, wide	++++	ND	ND	39	<0.1	<0.1
C, wide	++++	ND	ND	23	<0.1	<0.1
Antp 1 (TCTAATTGA)						
A, close	–	–	–	<0.1	97	<0.1
B, wide	–	++	–	<0.1	<0.1	<0.1
Antp 2 (TCTAATTAG)						
A, close	–	–	–	<0.1	141	117
B, wide	–	+++	+++	<0.1	<0.1	<0.1

[a] A, [A(site)T]$_6$; B, [CTGCCCA(site)TTGACGCTG]$_4$; C, [AGTAAAC(site)GGTCATAGT]$_4$.
[b] *Drosophila* results are based on in situ staining for β-galactosidase protein or *lacZ* mRNA in embryos (Fig. 2, 3, and 4; also data not shown). + and – values were assigned on an arbitrary scale as estimated by anterior staining intensity of blastoderm-stage embryos. ND, not determined.
[c] Yeast results are given in units of β-galactosidase activity assayed as previously described (27). Yeast cells were cotransformed with plasmids that encoded Bicoid derivatives and plasmids containing the indicated binding sites positioned upstream of a *GAL1-lacZ* reporter gene.

Drosophila embryos; Bicoid-Q_9 recognizes both the Antp 1 site (TCTAATTGA) and the Antp 2 site (TCTAATTAG) despite the fact that the former does not contain a thymine methyl group at position 8. Instead, the *Drosophila* results are consistent with a model in which Bicoid distinguishes among related homeodomain binding sites by a specific contact between recognition helix residue 9 and bp 7 (28). This model is based on recognition of a series of mutant sites by Bicoid and Bicoid-Q_9 in *S. cerevisiae*; any changes at position 7 abolished recognition, whereas changes at position 8 (and 9) were tolerated. We note, however, that a single-base-pair change at position 7 is not sufficient to completely switch site recognition by a homeodomain protein (28, 58); rather, the identity of bp 8 (and 9) influences the ability of Bicoid to discriminate base pairs at position 7 (28).

In the present study, three base pairs are changed in the Bicoid site to create Antp 1 and Antp 2, but only the C→T change at position 7 is common to both sites; Antp 1 and Antp 2 differ at positions 8 and 9, yet both are recognized. Therefore, any base-specific contact by Gln-50 of Bicoid-Q_9 must be to position 7. A contact to bp 7 (by our numbering system) is also seen in an X-ray crystal structure of the yeast homeodomain protein α2 complexed with DNA (80). In this case, residue 50 is serine, and the interaction occurs over a relatively long distance (0.39 nm). An NMR model of the Antennapedia homeodomain-DNA complex indicates that Gln-50 contacts both bp 7 and bp 8 (54), perhaps via water-mediated hydrogen bonds (81).

A possible explanation for the difference between our results and those of Kissinger et al. (39) and Otting et al. (54) is that conditions used for X-ray crystallography and NMR, such as the extremely high protein and DNA concentrations (10^{-3} to 10^{-4} M), might stabilize interactions that do not occur in vivo. In the X-ray structure of Engrailed (39), it is also possible that forces generated by packing of protein and DNA into a cocrystal alter the nucleotide contacted by Gln-50.

Another explanation for these differences is that each study used a different binding site. The site used for NMR contained GGC at bp 7, 8, and 9 (TCTAATGGC). Although this site is bound with high affinity by the isolated Antennapedia homeodomain in vitro (2), it is not recognized in *S. cerevisiae* by either Bicoid-Q_9 or the *fushi tarazu* homeodomain protein (28, 29), suggesting that contacts to the GC base pairs described in the NMR model are not preferred by Gln-50 in vivo. The site used in the X-ray structure is similar to our Antp 2 site but differs at positions 2 and 9 (TGTAATTAC). While it is conceivable that nucleotide differences at these positions might change the Gln-50 contact from bp 7 to bp 8 by altering local DNA structure, we do not favor this hypothesis. Such sequence-specific alterations are rare (57), and the X-ray structure of Kissinger et al. (39) does not show distortions in the B-form DNA consistent with such a mechanism.

A final explanation for the difference between our results and those of the structural studies is that different homeodomains were used. Amino acids that are not conserved among the homeodomains of Bicoid, Engrailed, and Antennapedia might affect the positioning of Gln-50 within the major groove. For example, residue 54, which differs among the three proteins (Arg, Ala, and Met, respectively), protrudes into the major groove in proximity to residue 50 (54, 80) and may affect its position. Our results with *S. cerevisiae* clearly demonstrated that for wild-type Bicoid, the side chain of Lys-50 is positioned so that it can tolerate only C · G at position 7 and that of Gln-50 in Bicoid-Q_9 can tolerate only T · A at position 7 (28). The side chain of Gln-50 in Engrailed and Antennapedia (and other homeodomains that normally contain Gln at this position) might be less constrained and thus occupy a greater range of alternative conformations that allow it to contact different nucleotides at position 7 and perhaps 8. In support of this idea, Florence et al. (18) and Ekker et al. (16) examined the in vitro DNA-binding specificity of Fushi Tarazu, Ultrabithorax, and Deformed homeodomain proteins (all contain Gln at position 50). Each protein recognized sites with either T · A at position 7 (as in our Antp sites and as in reference 39) or G · C at position 7 (as in reference 54) with affinities that differed by less than a factor of 2. In the case of Ultrabithorax and Deformed, these conclusions were supported by in vivo assays with *S. cerevisiae* (16).

Bicoid-A_9 reveals a new homeodomain-DNA interaction. Specific recognition of the Antp 2 site by Bicoid-A_9 is unexpected, because alanine at position 9 in its recognition helix (Ala-50) cannot form hydrogen bonds to base pairs in DNA. Instead Ala-50 might make a hydrophobic contact with the methyl group of thymine of A · T at bp 8 in the Antp 2 site (TCTAATTAG), similar to a contact identified in an altered-specificity 434 repressor mutant (77). Experiments with *S. cerevisiae* that examined site recognition by Bicoid-A_9 have not provided support for this idea (29).

Another possibility is that replacement of a bulky amino acid with alanine removes sterically unfavorable interactions, thus relaxing the protein's strict site specificity. We find, however, that Bicoid-A_9 is not relaxed in its specificity; it does not recognize Bicoid or Antp 1 sites in *D. melanogaster* or *S. cerevisiae* (Fig. 4; Table 2) or nine closely related binding sites in *S. cerevisiae* (29).

Finally, it is possible that specific recognition is the combined result of the lack of steric hindrance by Ala-50 and the twofold rotational symmetry of Antp 2 over bp 2 to 9 (TCTAATTAG). In this model, Bicoid-A_9 monomers could bind Antp 2 in either of two orientations, resulting in a twofold increase in the affinity of each monomer for its site. Cooperative interactions between adjacent Bicoid-A_9 monomers might amplify this effect, resulting in a binding affinity significantly higher than that for nonspecific DNA. Such a model would not require a specific DNA contact by Ala-50. Whatever model pertains, determining the nature of Bicoid-A_9 DNA recognition will be useful in understanding site specificity by homeodomain proteins that normally contain Ala at position 9 in their recognition helices (51).

Bicoid's DNA site specificity is necessary for anterior pattern formation. Bicoid is modular; its DNA binding and gene activation functions are separable. Driever et al. (10) demonstrated that its activation function is necessary for it to direct pattern formation in vivo. Our results now formally demonstrate that Bicoid's DNA recognition function is also necessary for it to direct pattern formation in vivo. While capable of activating anterior-specific transcription of appropriate reporter genes at the proper time during development (blastoderm stage), altered-specificity Bicoid proteins cannot rescue anterior pattern defects in *bcd*$^-$ embryos. In addition, neither *bcd-Q_9* nor *bcd-A_9* produces the normal *bcd*$^+$ dose-dependent posterior shift of the cephalic furrow or *eve* stripes. Indeed, some *bcd-Q_9* lines exhibit a dose-dependent shift of the cephalic furrow and *eve* stripes towards the anterior, mimicking the effect of reduced *bcd* activity.

As with other studies that make use of allele-specific suppression to suggest direct contacts between macromolecules (31), our results with altered-specificity Bicoid proteins and reporter genes with the cognate binding sites strongly suggest that Bicoid works by binding directly to the promoter regions of genes whose transcription it regulates, rather than by regulating the transcription of an intermediary protein(s). A

similar approach was used by Schier and Gehring (63) to suggest that *fushi tarazu* protein interacts directly with its upstream autoregulatory element.

Although Bicoid is sufficient to direct anterior-specific transcription, the posterior boundary of expression of our reporters is often more diffuse than that of endogenous Bicoid targets. This fact suggests that other proteins may contribute to establishment of the posterior border in normal embryos. Such proteins could be expressed throughout the embryo or could be regionalized, as is the *hb* protein, which appears to collaborate with *bcd* in stimulating transcription of *Kr* (35), *even-skipped* stripe 2 (66), and possibly *hb* itself (8).

Ancillary proteins may mediate cooperative gene activation by Bicoid in *D. melanogaster*. Bicoid is a weak activator and requires multiple binding sites to activate gene expression. This activation is cooperative; in yeast cells, tissue culture cells, and *Drosophila* embryos, we and others have observed a more than additive increase in transcription from reporter genes that carry increasing numbers of Bicoid sites (13, 27, 29, 70). In principle, this cooperativity could be due to interactions between Bicoid monomers that facilitate high-affinity DNA binding or to cooperative interactions of DNA-bound monomers with proteins of the general transcription machinery (i.e., synergy [4]). At least some component of cooperative gene activation by Bicoid is due to DNA-binding cooperativity (29). This idea is consistent with the observation that Bicoid-Q_9 antagonizes Bicoid activity in embryos, perhaps by the formation of mixed oligomers that bind Bicoid sites with reduced affinities.

Gene activation by Bicoid exhibits an unexpected constraint. We find that the distance between binding sites is critical for Bicoid-dependent activation but that this critical distance differs in *D. melanogaster* and *S. cerevisiae* (summarized in Fig. 7). In *D. melanogaster*, Bicoid activates reporter genes in which sites are widely spaced (25 bp, center to center), consistent with the fact that Bicoid sites in upstream regions of known or suspected Bicoid-regulated *Drosophila* genes are typically separated by relatively large distances (20 to 150 bp) (13, 33, 66). The wide spacing in our reporters places adjacent sites on opposite sides of the DNA helix about two and a half turns apart. In *S. cerevisiae*, Bicoid activates reporter genes in which sites are closely spaced (11 bp, center to center); this places adjacent sites on the same side of the DNA helix about one turn apart. In prokaryotic model systems, side-of-helix-specific presentation of binding sites has been shown to be important for cooperative interactions between DNA-bound proteins (15, 34). Our results suggest a species-specific side-of-helix dependence in eukaryotes that has not been previously observed.

How can one protein (Bicoid) prefer different site spacings in different organisms? For a possible explanation, we look to the yeast homeodomain protein α2 (the *MATα2* product). α2 forms a complex with the MCM1 protein that promotes cooperative DNA binding by the α2 dimer to its operator (38). Interestingly, MCM1 imposes a site spacing requirement on α2 so that the MCM1-α2 complex only recognizes half-sites spaced two and a half turns apart (67). α2 also forms a complex with the yeast **a**1 protein. The **a**1-α2 complex binds to operators of similar sequence but with half-sites spaced only 1.3 turns apart (25). Like α2, we suggest that Bicoid might require interaction with other proteins in order to cooperatively bind DNA. These ancillary proteins may act differently on Bicoid in *D. melanogaster* and *S. cerevisiae* such that they impose different site spacing requirements.

It is also possible that ancillary proteins might affect Bicoid's activation function rather than its DNA binding, similar to the effect that interaction of the herpes simplex virus protein VP16 has on activation by the Oct1 homeoprotein (41, 47, 60, 69). Ancillary proteins might facilitate cooperative interaction of Bicoid with the transcription machinery if and only if Bicoid monomers are positioned (spaced) appropriately on the DNA.

We do not know what these ancillary proteins might be. Perhaps they are proteins that interact specifically with Bicoid. Or, they may be general factors that affect many transcription activators, for example, chromatin components or global positive regulators such as those encoded by the yeast *SWI/SNF* genes (45, 59, 82; reviewed in reference 79) and the *Drosophila trithorax* gene (36). In *S. cerevisiae*, activation by Bicoid requires the *SWI2/SNF2* product (44), and it is possible that its *Drosophila* homolog, the product of *brahma* (71), is required for Bicoid activation in *D. melanogaster*. These proteins might impose different site spacing requirements in each organism.

We hope to identify genes encoding the putative ancillary proteins by using interaction trap assays (26) or by direct selection for *Drosophila* proteins that help Bicoid activate the widely spaced reporters in *S. cerevisiae*. Isolation of such proteins will help us understand how Bicoid works as a concentration-dependent activator of genes involved in early pattern formation and may illuminate mechanisms by which other important developmental regulators cooperate to generate region- and tissue-specific gene transcription.

ACKNOWLEDGMENTS

We are grateful to Mark Wainwright for help with germ line transformations, Catherine Bramwell for help in maintaining fly stocks, Wolfgang Driever and Daniel St. Johnston for *bcd* stocks, and Liz Gavis and Ruth Lehmann for help with photomicroscopy. S.D.H. thanks Chao-Ting Wu for the introduction to *Drosophila* techniques and the members of the ICRF Developmental Biology unit in Oxford for tolerating repeated visits and providing a stimulating atmosphere. We thank Russ Finley, Helen Francis-Lang, Erica Golemis, Domingos Henrique, Joanne Kamens, Ze'ev Paroush, and Robin Wharton for helpful discussions and/or comments on the manuscript. We also thank the reviewers for some useful suggestions.

This work was supported by the Human Frontiers Science Program, the Imperial Cancer Research Fund, and a Company of Biologists Ltd. Development Travel Award (S.D.H.). D.I.-H. is an International Research Scholar of the Howard Hughes Medical Institute. R.B. was supported by the Pew Scholars Program.

REFERENCES

1. **Adler, K., K. Beyreuther, E. Fanning, N. Geisler, B. Gronenborn, A. Klemm, B. Müller-Hill, M. Pfahl, and A. Schmitz.** 1972. How the *lac* repressor binds DNA. Nature (London) **237**:322–327.
2. **Affolter, M., A. Percival-Smith, M. Müller, W. Leupin, and W. J. Gehring.** 1990. DNA binding properties of the purified Antennapedia homeodomain. Proc. Natl. Acad. Sci. USA **87**:4093–4097.
3. **Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein, G. Frigerio, M. Noll, and C. Nüsslein-Volhard.** 1988. The role of localization of *bicoid* RNA in organizing the anterior pattern of the *Drosophila* embryo. EMBO J. **7**:1749–1756.
4. **Carey, M., Y. S. Lin, M. Green, and M. Ptashne.** 1990. A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature (London) **345**:361–364.
5. **Chalfie, M.** 1993. Homeobox genes in Caenorhabditis elegans. Curr. Opin. Genet. Dev. **3**:275–277.
6. **Cohen, S. M., and G. Jürgens.** 1990. Mediation of *Drosophila* head development by gap-like segmentation genes. Nature (London) **346**:482–485.
7. **De Robertis, E. M., E. A. Morita, and K. W. Y. Chow.** 1991. Gradient fields and homeobox genes. Development **112**:669–678.
8. **Desplan, C. (The Rockefeller University).** 1992. Personal communication.
9. **Desplan, C., J. Theis, and P. H. O'Farrell.** 1988. The sequence specificity of homeodomain-DNA interaction. Cell **54**:1081–1090.

10. **Driever, W., J. Ma, C. Nüsslein-Volhard, and M. Ptashne.** 1989. Rescue of *bicoid* mutant *Drosophila* embryos by Bicoid fusion proteins containing heterologous activating sequences. Nature (London) **342:**149–153.
11. **Driever, W., and C. Nüsslein-Volhard.** 1988. A gradient of *bicoid* protein in *Drosophila* embryos. Cell **54:**83–93.
12. **Driever, W., and C. Nüsslein-Volhard.** 1988. The *bicoid* protein determines position in the *Drosophila* embryo in a concentration-dependent manner. Cell **54:**95–104.
13. **Driever, W., and C. Nüsslein-Volhard.** 1989. The bicoid protein is a positive regulator of *hunchback* transcription in the *Drosophila* embryo. Nature (London) **337:**138–143.
14. **Driever, W., G. Thoma, and C. Nüsslein-Volhard.** 1989. Determination of spatial domains of zygotic gene expression in the *Drosophila* embryo by the affinity of binding sites for the Bicoid morphogen. Nature (London) **340:**363–367.
15. **Dunn, T. M., S. Hahn, S. Ogden, and R. F. Schleif.** 1984. An operator at −280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc. Natl. Acad. Sci. USA **81:**5017–5020.
16. **Ekker, S. C., D. P. von Kessler, and P. A. Beachy.** 1992. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. EMBO J. **11:**4059–4072.
17. **Finkelstein, R., and N. Perrimon.** 1990. The orthodenticle gene is regulated by *bicoid* and *torso* and specifies *Drosophila* head development. Nature (London) **346:**485–488.
18. **Florence, B., R. Handrow, and A. Laughon.** 1991. DNA-binding specificity of the *fushi tarazu* homeodomain. Mol. Cell. Biol. **11:**3613–3623.
19. **Frigerio, G., M. Burrie, D. Bopp, S. Baumgartner, and M. Noll.** 1986. Structure of the segmentation gene *paired* and the *Drosophila* PRD gene set as part of a gene network. Cell **47:**735–746.
20. **Frohnhöffer, H. G., and C. Nüsslein-Volhard.** 1986. Organization of anterior pattern in the *Drosophila* embryo by the maternal gene *bicoid*. Nature (London) **324:**120–125.
21. **Gehring, W. J.** 1985. Homeotic genes, the homeo box, and the genetic control of development. Cold Spring Harbor Symp. Quant. Biol. **50:**243–251.
22. **Gehring, W. J.** 1987. Homeoboxes in the study of development. Science **236:**1245–1252.
23. **Gehring, W. J., M. Müller, M. Affolter, A. Percival-Smith, M. Billeter, Y.-Q. Qian, G. Otting, and K. Wüthrich.** 1990. The structure of the homeodomain and its functional implications. Trends Genet. **6:**323–329.
24. **Gloor, G., and W. Engels.** 1992. Single-fly DNA preps for PCR. Drosophila Information Service **71:**148–149.
25. **Goutte, C., and A. D. Johnson.** 1988. a1 alters the DNA binding specificity of α2 repressor. Cell **52:**875–882.
26. **Gyuris, J., E. Golemis, H. Chertkov, and R. Brent.** 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. Cell **75:**791–803.
27. **Hanes, S. D., and R. Brent.** 1989. DNA specificity of the Bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell **57:**1275–1283.
28. **Hanes, S. D., and R. Brent.** 1991. A genetic model for interaction of the homeodomain recognition helix with DNA. Science **251:** 426–430.
29. **Hanes, S. D., and R. Brent.** Unpublished data.
30. **Harrison, S.** 1991. A structural taxonomy of DNA binding proteins. Nature (London) **353:**715–719.
31. **Hartman, P. E., and J. R. Roth.** 1973. Mechanisms of suppression. Adv. Genet. **17:**1–105.
32. **Hayashi, S., and M. P. Scott.** 1990. What determines the specificity of action of homeodomain proteins? Cell **63:**883–894.
33. **Hoch, M., E. Seifer, and H. Jäckle.** 1991. Gene expression mediated by cis-acting sequences of the *Krüppel* gene in response to the *Drosophila* morphogens *bicoid* and *hunchback*. EMBO J. **10:**2267–2278.
34. **Hochschild, A., and M. Ptashne.** 1986. Cooperative binding of λ repressors to sites separated by integral turns of the DNA helix. Cell **44:**681–697.
35. **Hülskamp, M., C. Pfeifle, and D. Tautz.** 1990. A morphogenetic gradient of *hunchback* protein organizes the expression of the gap genes *Krüppel* and *knirps* in the early *Drosophila* embryo. Nature (London) **346:**577–580.
36. **Ingham, P.** 1985. Genetic control of the spatial pattern of selector gene expression in *Drosophila*. Cold Spring Harbor Symp. Quant. Biol. **50:**201–208.
37. **Karess, R. E., and G. M. Rubin.** 1984. Analysis of P transposable element functions in *Drosophila*. Cell **38:**135–146.
38. **Keleher, C. A., C. Goutte, and A. D. Johnson.** 1988. The yeast cell-type-specific repressor a2 acts cooperatively with a non-cell-type-specific protein. Cell **53:**927–936.
39. **Kissinger, C. R., B. Liu, E. Martin-Blanco, T. B. Kornberg, and C. O. Pabo.** 1990. Crystal structure of an engrailed homeodomain-DNA complex a 2.8 A resolution: a framework for understanding homeodomain-DNA interactions. Cell **53:**927–936.
40. **Klinger, M., and P. Gergen (State University of New York at Stony Brook).** 1991. Personal communication.
41. **Lai, J.-S., M. A. Cleary, and W. Herr.** 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domain. Genes Dev. **6:**2058–2065.
42. **Laughon, A.** 1991. DNA binding specificity of homeodomains. Biochemistry **30:**11357–11368.
43. **Laughon, A., and M. P. Scott.** 1984. Sequence of a *Drosophila* segmentation gene: protein structure homology with DNA binding proteins. Nature (London) **310:**25–31.
44. **Laurent, B. C., and M. Carlson.** 1992. Yeast SNF2/SWI2, SNF5, and SNF6 proteins function coordinately with the gene-specific activators GAL4 and Bicoid. Genes Dev. **6:**1707–1715.
45. **Laurent, B. C., M. A. Treitel, and M. Carlson.** 1991. Functional interdependence of the yeast *SNF2*, *SNF5*, and *SNF6* proteins in transcriptional activation. Proc. Natl. Acad. Sci. USA **88:**2687–2691.
46. **Levine, M., and T. Hoey.** 1988. Homeobox proteins as sequence-specific transcription factors. Cell **55:**537–540.
47. **Luo, Y., H. Fujii, T. Gerster, and R. G. Roeder.** 1992. A novel B cell-derived coactivator potentiates the activation of immunoglobulin promoters by octamer-binding transcription factors. Cell **71:**231–241.
48. **Macdonald, P. M., P. W. Ingham, and G. Struhl.** 1986. Isolation, structure and expression of *even-skipped*: a second pair-rule gene of Drosophila containing a homeobox. Cell **47:**721–734.
49. **Macdonald, P. M., and G. Struhl.** 1988. cis-acting sequences responsible for anterior localization of *bicoid* mRNA in *Drosophila* embryos. Nature (London) **336:**595–598.
50. **McGinnis, W., and R. Krumlauf.** 1992. Homeobox genes and axial patterning. Cell **68:**283–2302.
51. **Mendel, D. B., and G. R. Crabtree.** 1991. HNF-1, a member of a novel class of dimerizing homeodomain proteins. J. Biol. Chem. **266:**667–680.
52. **Nüsslein-Volhard, C., H. G. Frohnhöffer, and R. Lehman.** 1987. Determination of anteroposterior polarity in *Drosophila*. Science **238:**1675–1681.
53. **Oppenheim, A. B., and D. Noff.** 1975. Deletion mapping of *trans* dominant mutations in the λ repressor gene. Virology **63:**553–556.
54. **Otting, G., Y.-Q. Qian, M. Billeter, M. Müller, M. Affolter, W. Gehring, and K. Wüthrich.** 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. EMBO J. **9:**3085–3092.
55. **Otting, G., Y.-Q. Qian, M. Müller, M. Affolter, W. Gehring, and K. Wüthrich.** 1988. Secondary structure determination for the Antennapedia homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix. EMBO J. **7:**4305–4309.
56. **Pabo, C., and R. T. Sauer.** 1984. Protein-DNA recognition. Annu. Rev. Biochem. **53:**293–321.
57. **Pabo, C., and R. T. Sauer.** 1992. Transcription factors: structural families and principles of DNA recognition. Annu. Rev. Biochem. **61:**1053–1095.
58. **Percival-Smith, A., M. Müller, M. Affolter, W. J. Gehring.** 1990. The interaction with DNA of wild-type and mutant fushi tarazu homeo domains. EMBO J. **9:**3967–3974.
59. **Peterson, C. L., and I. Herskowitz.** 1991. Characterization of the yeast *SWI1*, *SWI2*, and *SWI3* genes, which encode a global

activator of transcription. Cell **68**:573–583.
60. **Pomerantz, J. L., T. M. Kristie, and P. A. Sharp.** 1992. Recognition of the surface of a homeo domain protein. Genes Dev. **6**:2047–2057.
61. **Qian, Y. Q., M. Billeter, G. Otting, M. Müller, W. J. Gehring, and K. Wüthrich.** 1989. The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell **59**:573–580.
62. **Riddihough, G., and D. Ish-Horowicz.** 1991. Individual stripe regulatory elements in the *Drosophila hairy* promoter respond to maternal, gap and pair-rule genes. Genes Dev. **5**:840–854.
63. **Schier, A. F., and W. Gehring.** 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. Nature (London) **356**:804–807.
64. **Scott, M. P., J. W. Tamkun, and G. W. Hartzell III.** 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta **989**:25–48.
65. **Seeman, N. C., J. M. Rosenberg, and A. Rich.** 1976. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. USA **73**:804–808.
66. **Small, S., R. Kraut, T. Hoey, R. Warrior, and M. Levine.** 1991. Transcriptional regulation of a pair-rule gene stripe in *Drosophila*. Genes Dev. **5**:827–839.
67. **Smith, D. L., and A. D. Johnson.** 1992. A molecular mechanism for combinatorial control in yeast: MCM1 protein sets the spacing and orientation of the homeodomains of an α2 dimer. Cell **68**:133–142.
68. **Steller, H., and V. Pirrotta.** 1985. A transposable P vector that confers G418 resistance to *Drosophila* larvae. EMBO J. **4**:167–171.
69. **Stern, S., M. Tanaka, and W. Herr.** 1989. The Oct-1 homeo domain directs formation of a multiprotein-DNA complex with the HSV *trans*-activator VP16. Nature (London) **341**:624–630.
70. **Struhl, G., K. Struhl, and P. M. Macdonald.** 1989. The gradient morphogen *bicoid* is a concentration-dependent transcriptional activator. Cell **57**:1259–1273.
71. **Tamkun, J. W., R. Deuring, M. P. Scott, M. Kissinger, A. M. Pattatucci, T. C. Kaufman, and J. A. Kennison.** 1992. brahma: a regulator of *Drosophila* homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. Cell **68**:561–572.
72. **Tautz, D., R. Lehmann, H. Schnurch, R. Schuh, A. K. Seifert, K. Jones, and H. Jäckle.** 1987. Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. Nature (London) **327**:383–389.
73. **Tautz, D., and C. Pfeifle.** 1989. A non-radioactive *in situ* hybridization method for localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. Chromosoma **98**:81–85.
74. **Treisman, J., E. Harris, D. Wilson, and C. Desplan.** 1992. The homeodomain: a new face for the helix-turn-helix? Bioessays **14**:145–150.
75. **Vincent, J.-P., J. A. Kassis, and P. H. O'Farrell.** 1990. A synthetic homeodomain binding site acts as a cell type specific, promoter specific enhancer in *Drosophila* embryos. EMBO J. **9**:2573–2578.
76. **West, R. W., R. R. Yocum, and M. Ptashne.** 1984. *Saccharomyces cerevisiae GAL1-GAL10* divergent promoter region: location and function of the upstream activating sequence UAS_G. Mol. Cell. Biol. **4**:2467–2478.
77. **Wharton, R. P., and M. Ptashne.** 1987. A new-specificity mutant of 434 repressor that defines an amino acid-base pair contact. Nature (London) **326**:888–891.
78. **Wieschaus, E., and C. Nüsslein-Volhard.** 1986. Looking at embryos, p. 199–226. *In* D. B. Roberts (ed.), *Drosophila*: a practical approach. IRL Press, Oxford.
79. **Winston, F., and M. Carlson.** 1992. Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. Trends Genet. **8**:387–391.
80. **Wolberger, C., A. K. Vershon, B. Liu, A. D. Johnson, and C. Pabo.** 1991. Crystal structure of MAT α2 homeodomain-operator complex suggests a general model for homeodomain-DNA interaction. Cell **67**:517–528.
81. **Wüthrich, K. (Eidgenössische Technische Hochschule, Zürich, Switzerland).** 1992. Personal communication.
82. **Yoshinaga, S. K., C. L. Peterson, I. Herskowitz, and K. R. Yamamoto.** 1992. Roles of SWI1, SWI2, and SWI3 proteins for transcriptional enhancement by steroid receptors. Science **258**: 1598–1604.