MOLECULAR AND CELLULAR BIOLOGY, Nov. 1993, p. 6941–6956
0270-7306/93/116941-16$02.00/0

Vol. 13, No. 11

Cooperative Binding of an *Ultrabithorax* Homeodomain Protein to Nearby and Distant DNA Sites

PHILIP A. BEACHY,[1]* JACOB VARKEY,[2] KEITH E. YOUNG,[1] DORIS P. VON KESSLER,[1] BENJAMIN I. SUN,[1] AND STEPHEN C. EKKER[1]

Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205,[1] and Department of Biological Sciences, Savannah State College, Savannah, Georgia 31404[2]

Received 14 July 1993/Accepted 17 August 1993

Cooperativity in binding of regulatory proteins to multiple DNA sites can heighten the sensitivity and specificity of the transcriptional response. We report here the cooperative DNA-binding properties of a developmentally active regulatory protein encoded by the *Drosophila* homeotic gene *Ultrabithorax* (*Ubx*). We show that naturally occurring binding sites for the *Ubx*-encoded protein contain clusters of multiple individual binding site sequences. Such sites can form complexes containing a dozen or more *Ubx*-encoded protein molecules, with simultaneous cooperative interactions between adjacent and distant DNA sites. The distant mode of interaction involves a DNA looping mechanism; both modes appear to enhance transcriptional activation in a simple yeast assay system. We found that cooperative binding is dependent on sequences outside the homeodomain, and we have identified regions predicted to form coiled coils carboxy terminal to the homeodomains of the *Ubx*-encoded protein and several other homeotic proteins. On the basis of our findings, we propose a multisite integrative model of homeotic protein action in which functional regulatory elements can be built from a few high-affinity sites, from many lower-affinity sites, or from sites of some intermediate number and affinity. An important corollary of this model is that even small differences in binding of homeotic proteins to individual sites could be summed to yield large overall differences in binding to multiple sites. This model is consistent with reports that homeodomain protein targets contain multiple individual binding site sequences distributed throughout sizable DNA regions. Also consistent is a recent report that sequences carboxy terminal to the *Ubx* homeodomain can contribute to segmental specificity.

Cooperative DNA binding is a common property among proteins that function in the control of transcription. In prokaryotes, where such proteins have been best studied, binding typically involves a limited number of discrete DNA sites, and cooperativity ensues from the favorable energies of protein-protein interaction. In some cases, such interactions are sufficiently strong to mediate protein association in the absence of DNA (e.g., dimerization and tetramerization of the *lac* repressor; see reference 67), whereas in other cases, oligomerization depends on the presence of appropriately arranged DNA-binding sites (e.g., dimerization of the LexA repressor; see reference 42).

Cooperativity in eukaryotes is suggested by the greater-than-additive increase in transcriptional response as regulator-binding sites are added to a target promoter. Such a supralinear increase can involve activator proteins of the same type or of different types; current evidence suggests that some or all of the apparent cooperativity between distinct proteins may result from the ability of the basic transcription machinery to interact simultaneously with multiple activators (49, 50, 82; reviewed in reference 76). In addition, several eukaryotic regulatory proteins are intrinsically capable of cooperative binding to multiple DNA sites as homo- or heteromultimers in the absence of the basic transcription machinery (see, for example, references 80, 91, and 96).

Whether intrinsic or extrinsic, cooperativity augments the overall affinity of a particular regulatory protein for DNA. Cooperativity also tightens the protein concentration range within which a group of DNA sites makes the transition from the unoccupied to the occupied state. Multiple binding sites thus could equip a promoter for function as a biological switch with sensitivity to small changes in concentration of a particular regulatory protein, while sites for multiple distinct regulators might produce a flexible yet decisive response to multiple proteins (for more extensive discussions of cooperative binding, see reference 62 and references therein).

Cooperative binding is particularly attractive as the mechanistic basis for some of the regulatory phenomena observed in developmental regulation. In *Drosophila* embryogenesis, for example, establishment and refinement of spatial information result largely from interactions between transcription factors encoded in successive tiers of a genetic hierarchy (81). At several levels in this hierarchy, spatial information residing in smooth concentration gradients of regulators is interpreted to yield discrete spatial boundaries in expression of target genes. A well-known case is activation of the gap gene, *hunchback*, at a particular position along the anterior-posterior axis in a threshhold-dependent response to the concentration gradient of a regulatory protein encoded by *bicoid*. Appropriate activation at the *hunchback* promoter requires multiple DNA-binding sites for the *bicoid*-encoded protein (reviewed in reference 81), consistent with a cooperative binding mechanism. Multiple binding sites occur near the promoters of genes at many other levels in the early hierarchy and appear to be required for appropriate regulation (for examples, see reviews in references 60 and 79). In most cases, however, cooperative binding of regulators to these multiple sites has not been directly demonstrated. We report here the cooperative binding properties of a protein that functions within the early regulatory hierarchy; this

* Corresponding author.

regulatory protein is encoded by *Ultrabithorax* (*Ubx*), one of the eight homeotic genes that function to specify segmental differentiation pathways.

Ubx operates within the segmented framework established in the embryo by the segmentation genes; its function is to specify the distinguishing features of parasegments 5 and 6, which together constitute a contiguous region including the posterior thorax and a portion of the first abdominal segment (reviewed in references 5 and 19). The *Ubx* transcription unit produces a family of closely related nuclear proteins (6, 44, 58) which function in the control of transcription (38, 46, 69, 88). Each *Ubx* protein family member contains the homeodomain (reviewed in reference 75), a 61-amino-acid sequence that is present in all other homeotic proteins and in the protein products of some genes at other levels of the early regulatory hierarchy (the *bicoid* protein, for example, is a homeodomain protein). The proteins encoded by homeotic genes are believed to implement specific morphogenetic pathways by direct transcriptional regulation of downstream target genes.

We report here a series of biochemical studies of *Ubx* protein binding which demonstrate that naturally occurring binding sites for the *Ubx* protein contain multiple individual recognition sequences. We found, through binding studies to variants of these sites, that *Ubx* protein complexes are progressively stabilized by increasing numbers of individual recognition sequences. This cooperative stabilization extends to sites separated by as much as 200 bp, by a mechanism involving DNA looping. We further found that distant and nearby sites are capable of mediating synergistic activation of transcription by the *Ubx* protein in *Saccharomyces cerevisiae*. These binding properties provide a physical basis for understanding several recently published genetic studies indicating that, in *Drosophila melanogaster*, homeodomain protein target sites contain multiple individual sites dispersed throughout extensive DNA regions (see Discussion). We also found that cooperative binding of the *Ubx* protein to multiple DNA sites requires amino acid sequences outside the homeodomain, consistent with recent studies showing that such sequences contribute to the segmental specificity of embryonic function (14).

MATERIALS AND METHODS

Protein purification. UbxHD, a 72-amino-acid residue peptide containing the *Ubx* homeodomain, was expressed in *Escherichia coli* and purified as previously described (20). For full-length UBX Ib, a *Ubx* isoform containing all three of the optional elements, the simplicity and yield of the previously reported purification (7) were improved significantly by replacement of four conventional chromatography steps with sequence-specific DNA chromatography (40, 66) and Ni^{2+} interaction chromatography (61). Since the homeodomain is located at the carboxy terminus of the full-length *Ubx* protein and the amino terminus contains a relatively large number of histidine residues that appear to interact with Ni^{2+}, this combination of steps also provided effective separation of full-length protein from numerous abundant partially degraded forms. The improved protocol follows.

UBX Ib was expressed in *E. coli* with growth, induction, and lysis of the cells essentially as previously described (7), except that the lysis buffer was supplemented with 2 μg of pepstatin A (Sigma) per ml. Beginning with ~100 g of cells (from a 20-liter culture) and following removal of DNA by precipitation with polyethyleneimine (also described previously), ~450 ml of extract was obtained at a protein concentration of ~15 mg/ml. These and all subsequent steps were carried out at 4°C. Ammonium sulfate was added to ~25% saturation (13.4 g/100 ml of lysate). The pellet was washed with buffer Y (final concentrations of 0.2 M NaCl, 50 mM Tris-HCl, pH 7.4 [measured at 22°C], 1 mM EDTA, 1 mM dithiothreitol, and 5% glycerol) that also contained 2 μg of pepstatin A per ml, 0.25 mM phenylmethylsulfonyl fluoride, and ammonium sulfate to 17.5% saturation (from a solution saturated at 4°C). The remaining pellet was retained and solubilized in a minimum volume by repeated extraction with small aliquots of buffer Y containing pepstatin A and phenylmethylsulfonyl fluoride until the protein concentration in the extracts began to fall. The solubilized material was pooled and passed through a column (1 by 12 cm; 10 ml) of single-stranded DNA-Sepharose (see below) to remove the *E. coli* single-stranded DNA-binding protein present in the extract. This column contained 580 μg of sonicated, denatured Bluescript plasmid DNA per ml of gel; it was washed in 0.1 N NaOH prior to equilibration with buffer Y. Protein not retarded by the first column was applied to a second DNA column (1 by 10 cm; ~8 ml) pre-equilibrated in buffer Y; this column contained 2 mg of the annealed oligonucleotides 5′-AGACG$(CCATTA)_6$-3′ and 5′-$(TAATG G)_6$-3′ per ml of gel. Both the single-stranded and the tailed double-stranded DNA matrices were made by conjugating DNA to cyanogen bromide-activated Sepharose 4B (Pharmacia; see reference 95).

Following a wash in 16 ml of buffer Y, the second DNA column was developed with an 80-ml gradient from 0.3 to 0.75 M NaCl; a peak centered at ~0.53 M contained almost exclusively UBX Ib and its breakdown products. This material was pooled and dialyzed rapidly against IMAC buffer (0.8 M NaCl, 25 mM $NaPO_4$ [pH 7.5], 5% glycerol). The protein was then applied to a column (1.5 by 10 cm; ~18 ml) of Ni^{2+}-charged metal-chelating Sepharose FF (Pharmacia) that had been equilibrated with IMAC buffer. The column was underlaid with a 4-ml bed of uncharged resin to trap any released Ni^{2+}. After washing with 22 ml of IMAC buffer, a 140 ml gradient from 0 to 250 mM imidazole hydrochloride (pH 7.5) in IMAC buffer was applied to the column, with predominantly full-length protein eluting between 110 and 140 mM imidazole. These fractions were pooled and dialyzed against IMAC buffer to remove imidazole, and the chromatography was repeated in a scaled-down form with a smaller column (1 by 12 cm; 10 ml). The final pool was dialyzed against IMAC buffer containing 0.5 mM EDTA and 1 mM dithiothreitol, aliquotted, and stored at −80°C (final protein concentration, 140 μg/ml or 3.3 μM). Full-length UBX Ib accounted for more than 90% of the protein in the final pool, and the total yield was ~2 mg of protein. UBX Ib protein was monitored throughout purification by Coomassie staining of denaturing gels (47) and by Western blotting (immunoblotting) (11) with monoclonal antibody FP3.38 (93).

Oligonucleotides and DNA fragments for binding assays. For Fig. 1A, oligonucleotides were synthesized and annealed to yield duplex DNA with the sequence 5′-TTAT TATCCACATTATCAGCGGCATTATTGTTATTATT-3′ (the underlined sequences are numbered 1 to 4 from left to right) with *Eco*RI (5′-AATT-3′) and *Not*I (5′-GGCC-3′) sticky ends at the 5′ ends of the top and bottom strands, respectively. The duplex region corresponds to the U-A binding site from +47 to +85 with respect to the *Ubx* transcription start site (7). For sub 1,4 and sub 1,3,4 (Fig. 1B and 1C, respectively), oligonucleotides were synthesized with the indicated ATTA core sequences replaced by 5′-

CGGC-3'. These annealed oligonucleotides were gel purified, 5' end labelled with [γ-^{32}P]ATP, end filled with unlabelled nucleotides, and further purified with a NICK column (Pharmacia). For Fig. 1D, a 108-bp genomic DNA fragment containing U-B and extending from +222 to +329 with respect to the start site of *Ubx* transcription was obtained by gel purification from a *Dde*I-*Sty*I digestion of pPB260 (see below). This fragment was further purified with a NACS column (Bethesda Research Laboratories), 5' end labelled with [γ-^{32}P]ATP, end filled with unlabelled nucleotides, and purified again with a NICK column.

For the experiments of Fig. 2, the synthetic U-A site and its variants (see above) were inserted into *Eco*RI-*Not*I-digested Bluescript KS Minus (pBS; from Stratagene). Fragments containing U-A and its variants (125 bp) were obtained from this series of plasmids by digestion with *Eco*O109I and *Hin*fI and labelled by filling the *Eco*O109I ends with [α-^{32}P]dCTP and dGTP.

The experiments in Fig. 3A utilized fragments from a pBS derivative (pBND) designed to test for DNA bending. pBND contains the 352-bp *Cla*I-*Bam*HI fragment of pBR322 inserted in the *Bam*HI-*Cla*I sites of pBS (CBI) and a second copy of this pBR322 fragment (CB2) inserted in the same orientation as CB1 into the *Kpn*I site of pBS via attachment of 8-bp *Kpn*I linkers to the filled ends of the fragment. The 365-bp fragment utilized for the left half of Fig. 3A extended from the *Eco*RV site (labelled by replacement synthesis with T4 polymerase) in CB2 to the Cfr10I site in CB1. The fortuitously occurring weak site in Fig. 3A derived from one end of CB1. A derivative of pBND carrying a *Ubx* genomic fragment inserted at the *Hin*cII site between CB1 and CB2 was the source for the 324-bp fragment in the right half of Fig. 3A. The genomic insert was the end-filled *Hin*dIII-*Mlu*I fragment from pPB260 (see below; −46 to +119 with respect to the start of transcription) and contains the U-A-binding site. The fragment for the right half of Fig. 3A extended from an *Xho*I site between CB2 and the genomic fragment (labelled by replacement synthesis with T4 polymerase) to the *Cfr*10I site in CB1; the fragment thus contained primarily the U-A genomic region near the labelled end and the same portion of CB1 as in the left half of Fig. 3A but none of CB2.

The experiments in Fig. 3B and Fig. 4A, B, and D utilized a *Hin*dIII-*Eco*RI insert present in pBS derivative pPB260. This insert, kindly supplied by M. Biggin (10), extended from a *Hin*dIII site artificially introduced upstream of −44 to a naturally occurring *Eco*RI site beginning at +357 (coordinates are given with respect to the start site of *Ubx* transcription in the genomic sequence; see reference 44). Other derivatives of this insert (Fig. 3B; see Fig. 6B) were made by replacing the U-A and U-B regions (+46 to +88 and +224 to +310, respectively) with the sequences described by Biggin and Tjian (10). One further derivative (Fig. 4C) was made by replacing the spacer sequences between U-A and U-B (+120 to +221; removed by using the *Mlu*I and *Dde*I sites) with the 102-bp *Hae*III fragment from pBS. For binding studies of pPB260 and its derivatives, 3'-end labelling was accomplished by filling with α-^{32}P-labeled deoxynucleoside triphosphates at the upstream *Hin*dIII site (Fig. 4B) or at the downstream *Eco*RI site (Fig. 4A, C, and D) or by replacement synthesis with T4 polymerase at the *Bam*HI downstream flanking vector site (Fig. 3B). Constructs using pPB260 and its derivatives for yeast transcriptional assays are described below.

Binding assays and quantitative analysis. Equilibrium electrophoretic mobility shift assays in Fig. 1A to D were done as previously described (21), with DNA fragments at concentrations of approximately 100, 50, 75, and 75 pM, respectively. The basic DNase I footprinting protocol, as used for UBX Ib and UbxHD in Fig. 4, is given in reference 21; DNA concentrations were approximately 6, 60, 3, and 3 pM in panels A to D, respectively. Footprinting reactions for dissociation rate measurements (Fig. 2 and 3) utilized this basic protocol for binding of protein to end-labelled DNA but with addition of specific unlabelled competitor DNA for various periods of time prior to DNase I treatment (20). The reaction mixtures contained 8 nM protein and 75 pM labelled DNA for Fig. 2, 5 nM protein and 20 pM DNA for Fig. 3A, and 10 nM protein and 8 pM DNA for Fig. 3B; reaction mixtures in Fig. 3B also contained bovine serum albumin (crystallized, fraction V; Sigma) at 20 μg/ml. The competitor DNA (sequence in reference 20) contained a single nearly optimal *Ubx* binding site and was added to a concentration of 50 nM (5- to 10-fold excess relative to protein) after incubation of protein with labelled DNA.

The electrophoretic mobility shift assay (Fig. 1B and E) was quantified by excision of bands corresponding to free and bound DNAs from dried gels and direct counting in liquid scintillation cocktail. Autoradiograms from footprinting gels were quantified with a Computing Densitometer (Molecular Dynamics). After background subtraction to remove optical density contributed by the film, the digitized images were used to sum the total optical density corresponding to individual bands or blocs of bands. For each panel of Fig. 2, a single band within site 2 was quantified; a single band located within the portion of the weak site denoted by the filled rectangle was quantified for Fig. 3A. The larger DNA fragment size and resulting poorer resolution in Fig. 3B and D required use of a bloc of bands located in the mid-lower portion of the U-A site. Also, because the extremely long incubation times in Fig. 3B and D were consistently associated with a decrease in recoverable label, normalization to account for the total number of counts per track was as follows: intensities of the bloc of bands monitored in U-A were divided by the intensities of a bloc of bands located above U-A whose intensities were not affected by protein. Normalization was not necessary to generate linear plots for the other dissociation rate experiments.

The fraction of DNA bound (O_t) was calculated from the formula $O_t = 1 - (I_t/I_{max})$, where I_t is the band intensity at time t and I_{max} is the band intensity in the absence of protein. Dissociation rates (K_d) were determined from the plots in Fig. 2D and 3C and D by using the formula ln(fraction of DNA bound) = $-K_d t$ (t = time), and half-lives were calculated from the formula $t_{1/2} = -\ln(0.5)/K_d$.

Electron microscopy. Complexes were prepared for electron microscopy by incubating 60 ng of UBX Ib and 50 ng of a purified DNA fragment in 50 μl of buffer M (50 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 0.1 mM disodium EDTA, 0.1 mM dithiothreitol, 10 mM Tris-HCl, pH 7.4). These concentrations of protein and, especially, DNA (~4 and ~30 nM, respectively) were somewhat higher than those used in the binding assays, and this facilitated the scanning process used to identify complexes on the grids. Following incubation for 10 min at 22°C, the binding reaction mixture was transferred to ice and 6 μl of 1% glutaraldehyde (EM Sciences) was added. Fixation continued on ice for 15 min and was followed by transfer of 5-μl samples to polylysine-coated thin carbon on a copper grid. After 2 min of adsorption, excess sample was removed by the wicking action of filter paper and the grid was washed with 2 drops of deionized H_2O. Specimens were stained sequentially for 1 min each in 0.01% uranyl acetate and 0.01% uranyl formate. Excess

stain was removed, and the specimens were rinsed with deionized H_2O. The samples were air dried and photographed in a Philips 420 electron microscope with a super twin lens operating in the dark-field mode at 40 kV.

Yeast transcriptional activation. The multimeric sites assayed (see Fig. 6A) were prepared by ligating kinase-treated and annealed oligonucleotides PB122 (5′-GAAGCCATTAA 3′) and PB123 (5′-TCTTAATGGCT-3′) at a 20-fold molar excess into the separately prepared arms of pBS extending from *Alw*NI to a partially filled *Bam*HI site and from *Alw*NI to a partially filled *Sal*I site. Candidates for multimerized binding sites were identified by polymerase chain reaction and confirmed by sequence analysis. Fragments containing no, one, and multiple tandemly repeated sites were excised by using the flanking *Xba*I and *Xho*I sites from pBS and inserted into target plasmid pSEΔ1′ (20) digested with *Xba*I and *Xho*I. The targets assayed (see Fig. 6B) were prepared by transferring fragments excised with *Xba*I and *Xho*I from the derivatives of pPB260 (described above) into pSEΔ1′. Use of the *Xho*I site in pSEΔ1′ for the adjacent-site series and the distant-site series placed the insertion 174 bp upstream of the transcription start site. The transcriptional response of the adjacent- and distant-site series to induction of UBX Ib in *S. cerevisiae* was measured by assaying for β-galactosidase activity as previously described (20).

RESULTS

Multiple recognition sequences within genomic binding sites. Although some progress has been made in the identification of downstream candidates for regulation by homeotic proteins (26, 28, 29, 35, 39, 64, 65, 77), only recently have systematic studies on the number and arrangement of binding sites required for this regulation begun. We therefore began our characterization of DNA binding by a *Ubx* protein with U-A and U-B, the first two naturally occurring binding sites to be biochemically identified (7) (see also Discussion). These genomic sites and others located elsewhere are associated with large DNase I footprints (often greater than 40 bp). The sequences at many of these naturally occurring sites consist primarily of simple repeated elements such as the triplet TAA or hexanucleotides such as TAATCG and TAATGG (7, 21). The large size of these sites is not essential for binding, since the sequence $(TAA)_4$ is also bound (7) and the optimal recognition sequence for the *Ubx* homeodomain (5′-TTAATGGCC-3′ [20]) is considerably smaller than the footprints at these genomic sites. Because the most important positions in the homeodomain optimal site correspond to the TAAT core, it seemed likely that the genomic sites, which contain multiple repeats of this motif, were composites of clustered individual sites; this finding, in turn, suggested the possibility of cooperative binding.

To test these ideas, we began by defining the number of individual recognition sequences in the U-A and U-B binding sites. These two sites are associated, respectively, with 44- and 87-bp DNase I protections and are centered at positions +64 and +262 with respect to the *Ubx* transcription start site (7). Inspection of the U-A sequence revealed the presence of four TAAT motifs spaced at intervals of about 10 bp, suggesting that U-A comprises four individual binding site sequences. From our previous work with the *Ubx* homeodomain, we expected all four of these sequences to be bound, albeit at lower affinity than the optimal site. We used the purified *Ubx* homeodomain (UbxHD) for these studies because it displays a fundamental DNA sequence preference nearly identical to that of the intact UBX Ib protein (20) and readily produces discrete complexes in electrophoretic mobility shift assays, unlike intact protein (data not shown). Indeed, electrophoretic mobility shift assays with a radiolabelled DNA fragment containing the U-A site and increasing concentrations of the purified *Ubx* homeodomain (21) demonstrated that four types of complexes can be resolved, corresponding to single, double, triple, and quadruple occupancy of the four TAAT motifs present (Fig. 1A). As shown in Fig. 1B, replacement of the two outer TAAT motifs with GCCG sequences resulted in the formation of only two distinct complexes, even at the highest protein concentrations, while replacement of sites 1, 3, and 4 permitted formation of only a single complex (Fig. 1C). This result confirms that the number of resolvable complexes indeed corresponds to the number of binding sites.

Within the U-B site, a central cluster of three overlapping TAAT motifs is flanked by a pair of overlapping TAAT motifs at a distance of about three helix turns upstream and another pair of overlapping TAAT motifs two to three turns downstream. In addition to these complete motifs, several tetranucleotides in the regions between the overlapping clusters differ by a single base from the TAAT sequence (see reference 7 for the sequence of the U-B site). The complexity of the sequence at U-B makes it difficult to anticipate the number of binding sites, since immediately adjacent or overlapping TAAT sequences cannot be bound simultaneously by two homeodomains (23, 43), and the relative affinities of sites with incomplete TAAT cores are known for only a few sequences (21, 23). Because the DNase I footprint at U-B is about twice the size of the U-A footprint, twice the number of sites might be expected if individual sites are packed at the same density. Indeed, eight strong complexes and a weaker ninth complex were observed with the U-B fragment (Fig. 1D). It is unlikely that any of the observed complexes at U-B are due to nonspecific binding, since the altered U-A fragments in panels B and C formed only one or two complexes under the same conditions, even though they are identical in size to the fragment in panel A, which accommodates four homeodomains. Other experiments with DNA fragments containing various numbers and arrangements of naturally occurring or synthetic binding sites indicate that these conditions are permissive of complex formation only with individual sites expected to show some degree of affinity above that for nonspecific DNA (data not shown). We thus conclude that the naturally occurring U-A and U-B sites contain four and eight or nine individual recognition sequences, respectively, packed at a density corresponding approximately to the number of helical turns within the footprinted region.

Cooperative binding to adjacent sites. To assay for interactions between proteins bound at adjacent sites, a DNase I footprinting method was used to measure dissociation rates for complexes between the intact *Ubx* protein and the mutated derivatives of U-A described in the previous section. Equilibrium measurements were unsuitable as a standard for comparison, since at protein concentrations low enough to reveal a cooperative effect, significant activity loss occurred during the time required for protein to reach equilibrium with DNA fragments containing multiple sites (data not shown; see also reference 96). Dissociation rate measurements have the added advantage that neither protein concentration nor actual binding activity is important provided sufficient protein is present to produce complete site occupancy prior to addition of a competitor. The DNase I footprinting method used in these experiments permitted monitoring of a particular site or sites in a variety of distinct

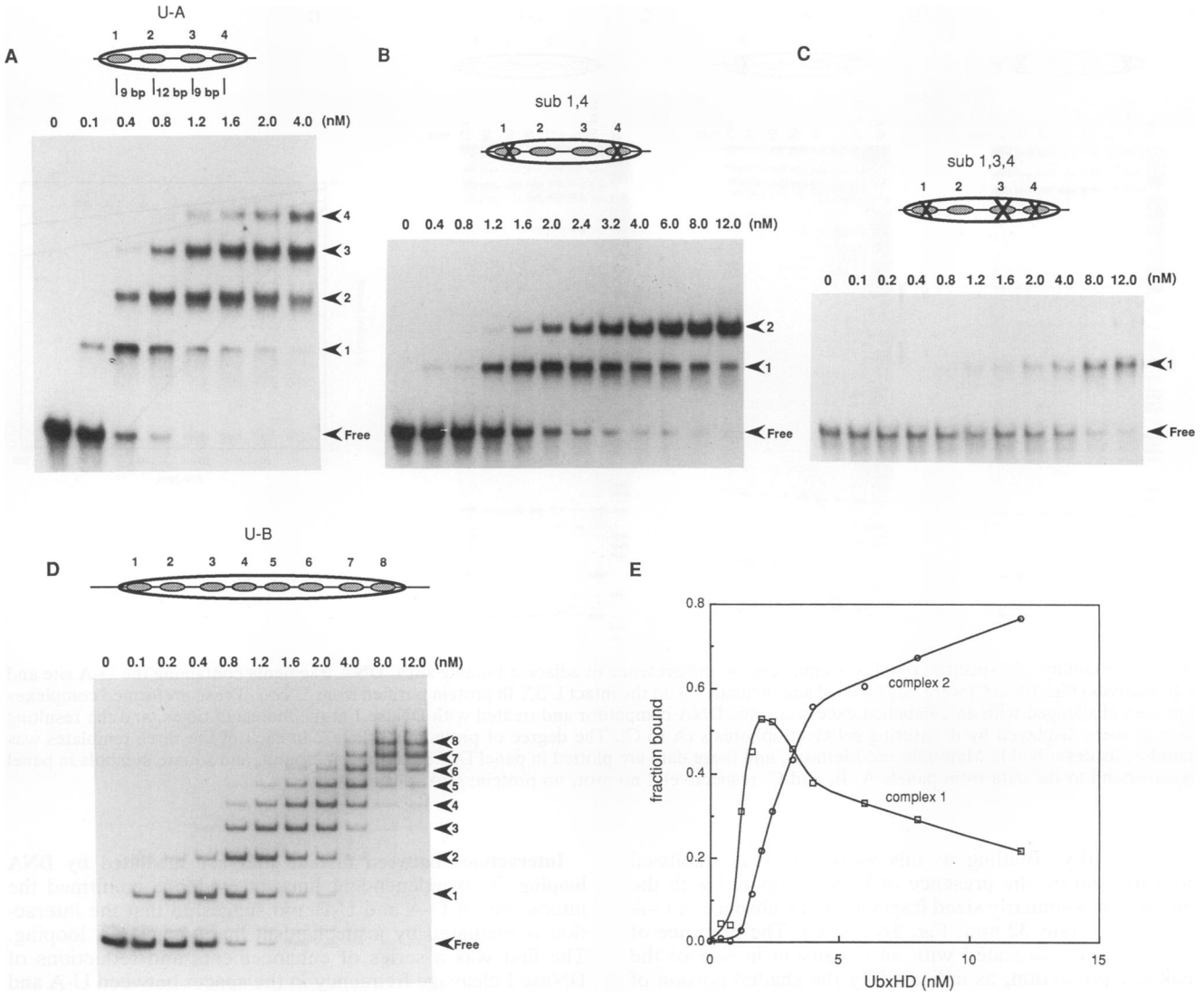

FIG. 1. Genomic DNA-binding sites for *Ubx* homeodomain occur in clusters. (A to D) Binding of the purified *Ubx* homeodomain (UbxHD) was assayed by electrophoretic mobility shift of DNA fragments containing the U-A and U-B genomic binding sites (7). Four TAAT core sequences present in the wild-type U-A sequence (A) are represented by shaded ovals and are located at center-to-center distances of 9, 12, and 9 bp as indicated above panel A. The DNA fragments used for panels B and C were the same as those used for panel A, except that two and three of these core TAAT sequences were replaced by the sequence GCCG. Note that the number of distinct complexes for each fragment corresponds to the number of specific binding sites remaining. Eight or nine distinct complexes were resolved with a DNA fragment containing U-B (D). UbxHD concentrations are specified in nanomolar units throughout. (E) The fraction of labelled DNA (the two-site template in panel B) present as the singly or doubly occupied complex was plotted as a function of UbxHD concentration. The singly occupied complex reached a maximum level of ~50% of the input DNA at 2.0 and 2.4 nM UbxHD.

sequence contexts. The protein used for these experiments was the UBX Ib isoform containing all three of the optional elements whose presence depends upon the pathway selected for splicing of the primary transcript. This protein was purified nearly to homogeneity, from the *E. coli* source reported previously (7), by using a simpler and more efficient procedure (see Materials and Methods).

The complex formed by UBX Ib with the single site present in the mutated U-A fragment in Fig. 2A was relatively unstable, with a half-life of less than 30 min. The half-life of binding to this isolated site (site 2) was extended to ~160 min by addition of a second site (site 3 [Fig. 2B]) by using another mutated U-A fragment. As shown in Fig. 2C, binding to site 2 was prolonged further to give a half-life greater than 4 h when sites 1 and 4 were reintroduced; indeed, the entire U-A region reconstituted in this fragment was stably bound throughout this time period. The stability of a *Ubx* protein complex with a single site can thus be extended by nearly 10-fold through addition of sites at adjacent positions.

Cooperative binding to nonadjacent sites. In the process of assaying genomic sites in various sequence contexts, we found that composite sites like U-A and U-B are capable of stabilizing binding to sites located some distance away. The first such situation we studied involved the U-A sequence positioned about seven turns away from a weak site fortuitously present in a cloning vector (Fig. 3A). The vector site contained a single TAAT motif with several partial motifs

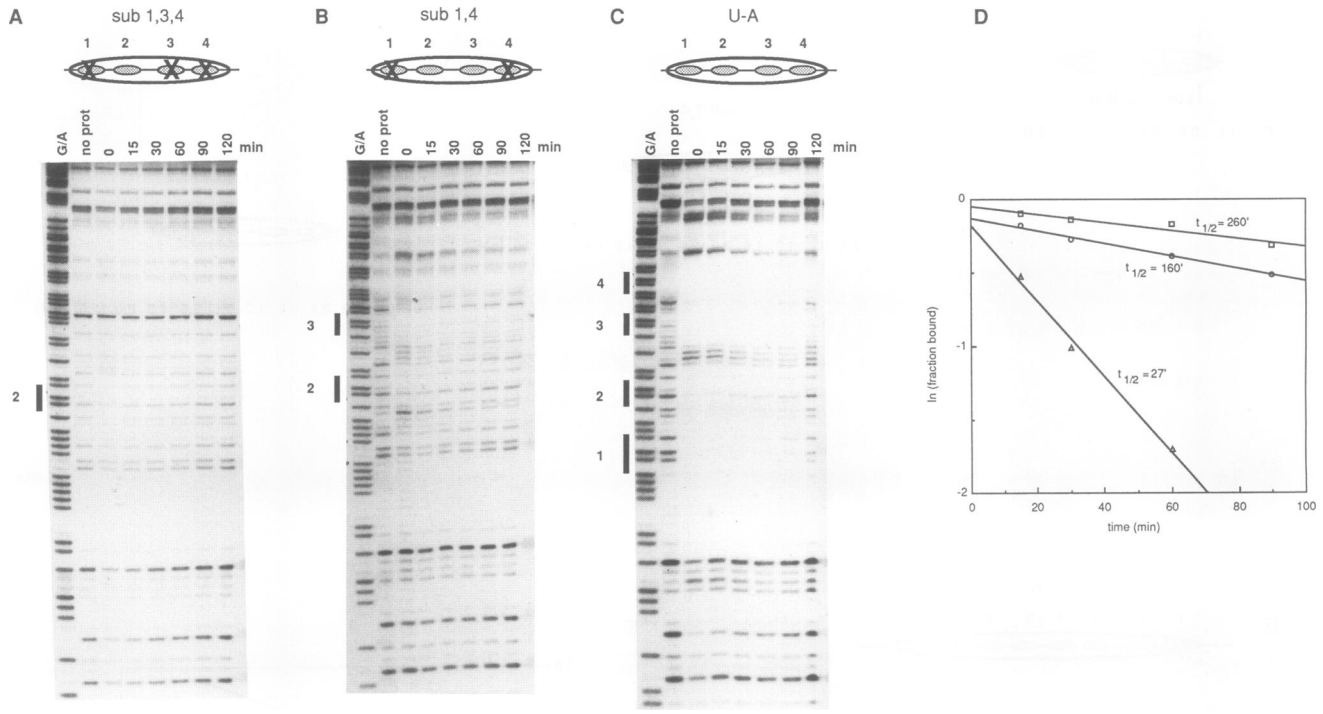

FIG. 2. Stability of a specific complex is enhanced by the presence of adjacent binding sites. DNA fragments containing the U-A site and its derivatives (Fig. 1A to C) were end labelled and incubated with the intact UBX Ib protein purified from *E. coli*. These preformed complexes were then challenged with an unlabelled excess specific DNA competitor and treated with DNase I at the indicated times, and the resulting products were displayed by denaturing gel electrophoresis (A to C). The degree of protection at site 2 in each of the three templates was quantified as described in Materials and Methods, and these data are plotted in panel D. The triangular, round, and square symbols in panel D correspond to the data from panels A, B, and C, respectively. no prot, no protein; G/A, purine positions.

located nearby. Binding to this weak site was stabilized about fivefold by the presence of U-A, compared with the same site on a similarly sized fragment in the absence of U-A ($t_{1/2}$ = 151 versus 32 min; Fig. 3A and C). The presence of U-A was also associated with an expansion in size of the weak-site protection, as indicated by the shaded portion of the rectangle in the right half of Fig. 3A. The half-life of the weak site was extended to approximately that of the U-A site, suggesting that the side-by-side interactions that lend stability to U-A (see above) can occur simultaneously with the interaction between U-A and the more distant weak vector site.

In the second situation we studied, U-A was present in its natural context, located at a center-to-center distance of 198 bp from the U-B site. Higher salt concentrations were used to facilitate the measurements, since under standard conditions (150 mM KOAc), the stability of complexes when both sites are present is considerably greater than 12 h (data not shown). Two independent dissociation rate experiments (Fig. 3B and D) each showed that binding to the U-A site is stabilized by the presence of U-B. Figure 3B shows the footprint gel from one of these experiments performed at 250 mM KOAc, and Fig. 3D shows the quantification of the second experiment, performed at 200 mM KOAc. From such data, we estimated that the presence of U-B at 200 mM KOAc stabilized binding by UBX Ib to U-A almost threefold (from 4.8 to 12.5 h; Fig. 3D). Twelve hours represents a considerable fraction of the time required to complete embryogenesis. Such stable complexes may therefore play a role in establishing and maintaining a determined state within the nuclei of differentiating cells.

Interactions between distant sites are mediated by DNA looping. Two independent lines of evidence confirmed the interaction of U-A and U-B and suggested that the interaction is mediated by a mechanism involving DNA looping. The first was a series of enhancements and reductions of DNase I cleavage frequency in the spacer between U-A and U-B. These changes appeared in parallel with the protection of U-A and U-B as the UBX Ib protein concentration increased (Fig. 4A and B). The average spacing between adjacent enhancements and between adjacent reductions was between 10 and 11 bases (data not shown), corresponding to the DNA helical repeat and suggestive of DNA curvature (17). As seen in Fig. 4C, replacement of the normal spacer sequences with a fragment from pBR322 also produced protein-dependent periodic alterations in DNase I cleavage. Although the exact pattern of altered cleavages was not identical, presumably because of some degree of sequence dependence by DNase I, the occurrence and regularity of the alterations suggest that changes in cleavage of the normal spacer are not the result of specific interactions with UBX Ib.

Similar protein-induced periodic changes in the frequency of DNase I cleavage were noted previously (31) upon looping-mediated cooperative binding by the lambda repressor to operator sequences spaced by five or six turns of DNA. These periodic changes are attributable to bending-induced structural distortions of the minor groove of DNA, which is the site of DNase I cleavage. Because the DNA is preferentially bent away from DNase I at the site of cleavage (85), these distortions would be expected to enhance cleavage along the outer surface of the curve and suppress it along the

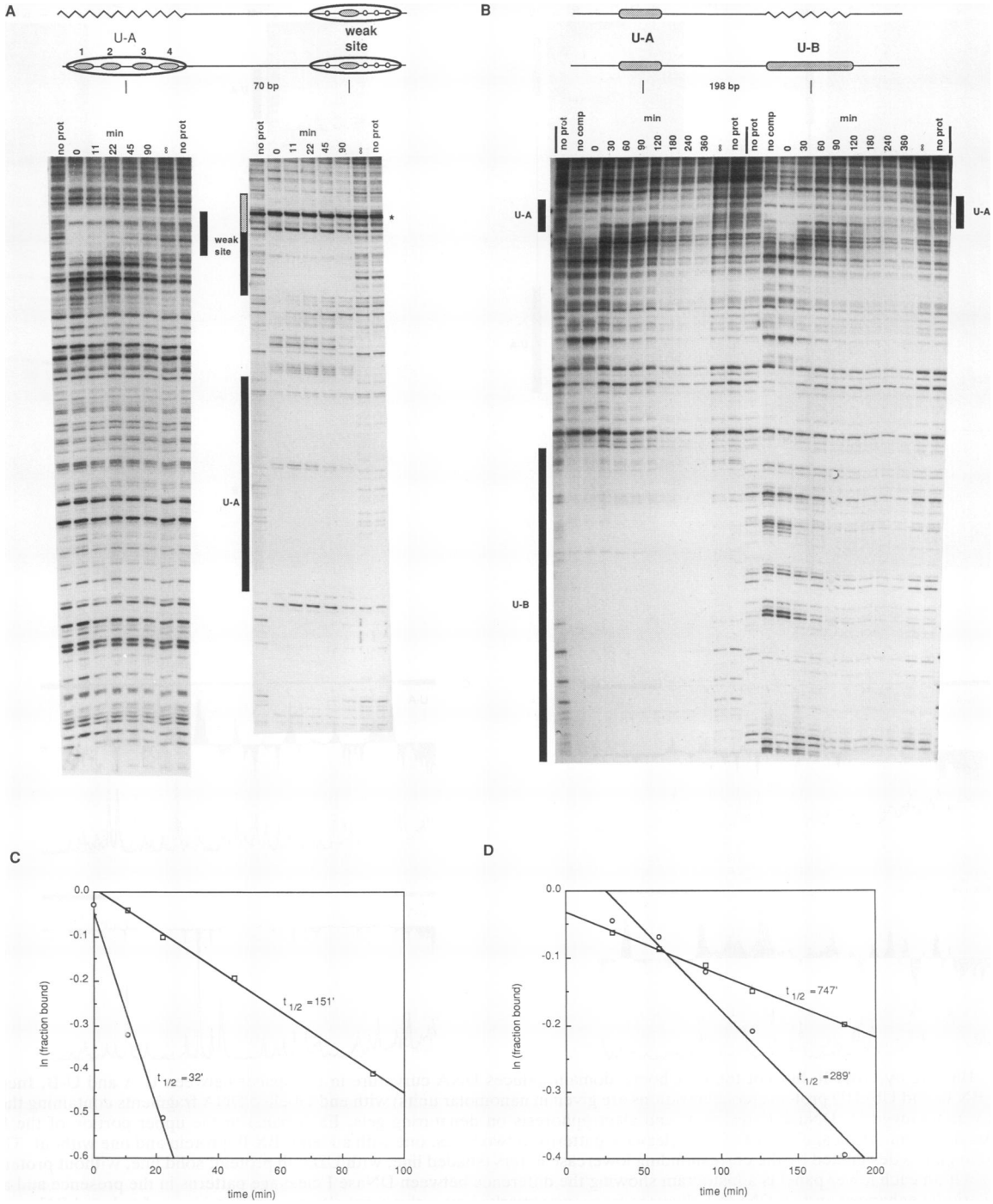

FIG. 3. Stability of specific complexes is enhanced by the presence of distant binding sites. Stabilities of specific complexes were measured as a function of the presence or absence of distant sites in a dissociation rate assay like that whose results are shown in Fig. 2 (see Materials and Methods for details). (A and C) Binding to a weak site fortuitously present in vector sequences (left half of panel A; round symbols in panel C) was stabilized about fivefold by the added presence of the U-A site (right half of panel A; square symbols in panel C). The weak site contains a single TAAT core sequence (shaded oval) and several three-of-four matches to the TAAT core (small circles); the center of the TAAT core is located 70 bp from the center of the U-A site. The presence of U-A induces expansion of the weak-site footprint to include adjacent sequences with partial core homology (shaded portion of the rectangle in the right panel). A contaminating DNA fragment proved difficult to remove from the labelled fragment used in the right half of panel A; this fragment comprised <5% of the total counts and did not contribute significantly to the DNase I cleavage pattern (aside from a band, marked by the asterisk, corresponding to the uncleaved fragment). (B and D) U-A and U-B are present at their naturally occurring positions in the left portion of panel B (center-to-center distance of 198 bp); in the right portion, the U-B site has been replaced with inert sequences. Qualitatively, in panel B, binding to the U-A site was stabilized severalfold by the presence of the U-B site (compare 240′ on the left side with 60′ or 90′ on the right). A second experiment was analyzed quantitatively, and the data are plotted in panel D (see text and Materials and Methods). no prot, no protein; no comp, no competition.

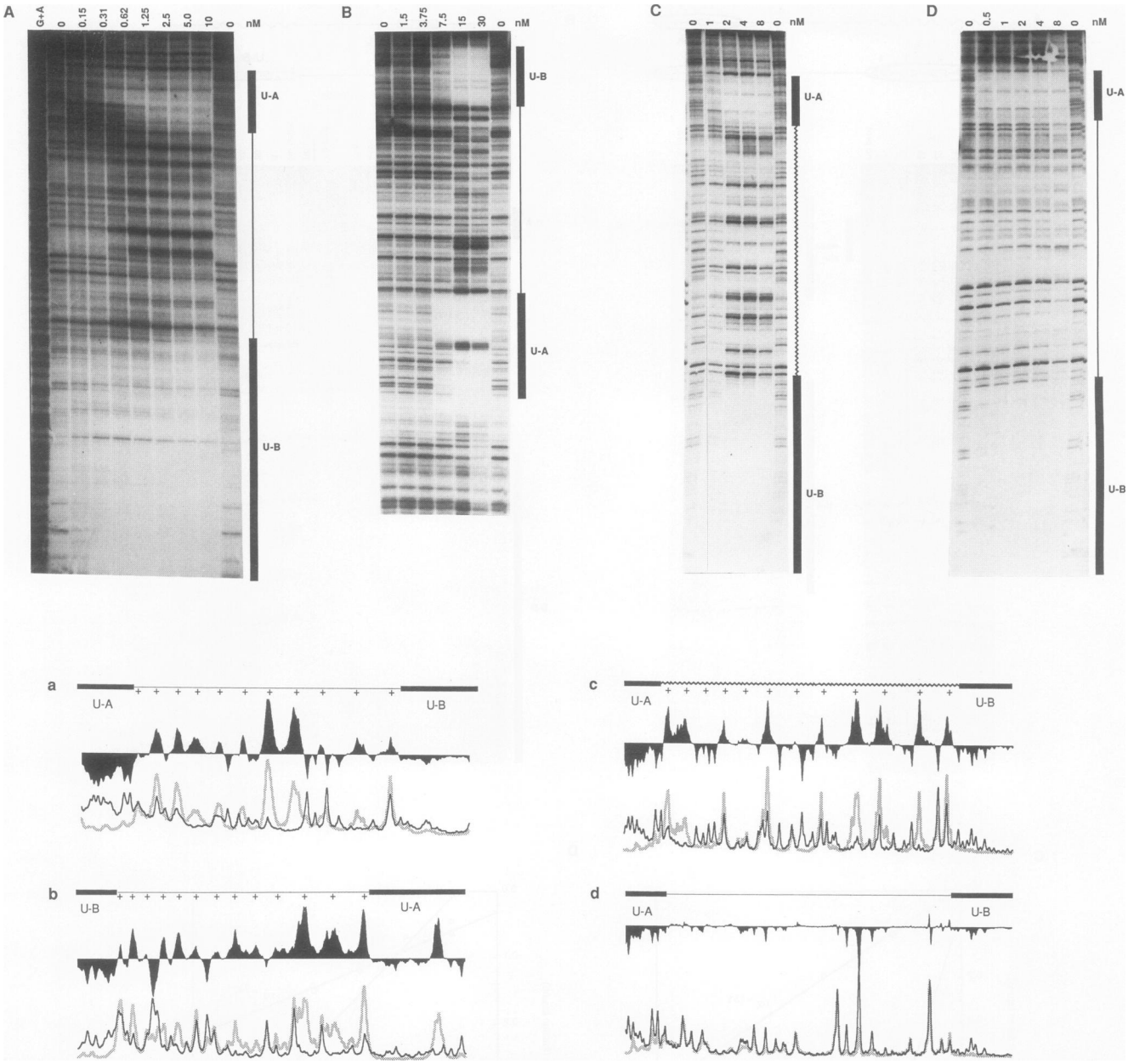

FIG. 4. Binding by UBX Ib but not the *Ubx* homeodomain induces DNA curvature in the spacer between U-A and U-B. Incubation of purified UBX Ib and UbxHD proteins (concentrations are given in nanomolar units) with end-labelled DNA fragments containing the U-A and U-B sites was followed by DNase I treatment and electrophoresis on denaturing gels. Each panel in the upper portion of the figure was analyzed by densitometric scanning of DNase I cleavage patterns in two lanes, one with added UBX Ib protein and one without. These scans are plotted in panels designated by the corresponding lowercase letters (shaded line, with UBX Ib protein; solid line, without protein). Above these plots within each lower panel is a histogram showing the difference between DNase I cleavage patterns in the presence and absence of UBX Ib. Periodic enhancements of DNase I cleavage frequency (marked by a plus sign) alternate with regions of reduced DNase I cleavage frequency (unmarked). These changes in DNase I cleavage are observed upon labelling of either strand (A and B) and also when the naturally occurring spacer is replaced by a fragment identical in length from pBR322 (C). Apparent differences in the protein concentration required for protection in panel B are due to the use of a higher DNA concentration and a UBX Ib preparation with activity lower than that used for panels A and C (see Materials and Methods). Regular alterations in DNase I cleavage of the spacer are not produced by binding of UbxHD, although the U-A and U-B sites are protected (D).

inner surface. DNA curvature is additionally supported by the average offset of ~2 nucleotides in the locations of enhancements on the two strands of the DNA (data not shown): as a result of double-helix geometry, the positions most affected by minor groove distortion lie several nucleotides apart on complementary strands of the DNA (17). Periodic enhancements and reductions of DNase I cleavage frequency were also apparent in the spacer between U-A and the weak vector site (Fig. 3A), suggesting that DNA looping underlies the stabilization of weak-site binding in that case as well.

Changes in DNase I cleavage of the spacer induced by UBX Ib were observed with several different protein preparations at subnanomolar concentrations and under a variety of conditions: these conditions included the presence of Na^+ ranging in concentration from 50 to 250 mM, the presence of nonionic detergent (Nonidet P-40 at 0.025 and 0.1%), and the presence of bovine serum albumin (10, 25, 50, and 100 μg/ml). The altered cleavage pattern also persisted for many hours after addition of unlabelled specific competitor DNA added in considerable excess (50 nM) relative to protein and labelled DNA (10 nM and 20 pM, respectively). These findings support our view that the changes in spacer cleavage are the result of specific protein-protein contacts and not a consequence of nonspecific aggregation.

Direct visualization of looped and linear complexes. A second line of evidence that confirms our analysis of the unusual DNase I cleavages resulted from visualization of protein-DNA complexes by dark-field electron microscopy (Fig. 5). The DNA fragment in these complexes contained the U-A and U-B binding sites, as illustrated schematically at the top of Fig. 5g. Figure 5a shows a field containing several examples each of linear and looped molecules, and individual examples are shown at higher magnification in Fig. 5b to f (linear molecules) and h to q (loops). Bound UBX Ib in these micrographs appears as a thickening or increased density (see arrows; compare with free DNA in Fig. 5u and v). The distances between binding sites and the ends of the fragment (Fig. 5g) were calculated from footprinting experiments, and they agreed well with the measurements from micrographs of actual complexes (Fig. 5 legend).

Looped forms were observed in ~40% of those complexes in which the U-A and U-B sites of an isolated DNA molecule were occupied. This is a lower frequency of loop formation than in the footprinting experiments, in which cleavage by DNase I at certain sites was completely abolished, suggesting looping of nearly all DNA molecules; the discrepancy may be due to the lower protein-to-DNA ratio used in the binding reactions for spreading. The length of the loop contour consistently measured a distance corresponding to 235 ± 8.6 bp. This dimension is consistent with contact between proteins bound at the two sites and suggests that they are aligned not center to center but in a manner that maximizes loop circumference and minimizes torsional stress in the DNA. The inner diameter of this loop should average ~234 Å (1 Å = 0.1 nm), insufficiently small to completely block access by a molecule with the dimensions of DNase I (45 by 40 by 35 Å [85]), although perhaps sufficiently small to hinder it and thus aid in reducing cleavage along the inner surface of the DNA loop (see above).

Complexes involving more than one DNA molecule were also observed, and several examples are shown in panels R to T. The points of DNA contact between molecules in these intermolecular complexes were usually associated with protein density at positions corresponding to U-A and U-B within the fragments. Panel R shows such a *trans* association involving both binding sites on each molecule, and panels S and T show the association of a smaller U-B-containing fragment with sites U-B and U-A, respectively, on the larger DNA. Such interactions in *trans* have also been reported for the *lac* repressor, a molecule also capable of DNA loop-mediated cooperativity between distant sites (45).

The formation of protein bridges between unlinked DNA molecules may be relevant to transvection, a phenomenon that has long been known to occur at *Ubx* and other loci in *D. melanogaster* and may have its counterpart in other eukaryotes (reviewed in reference 86). Transvection may be thought of as the pairing- or proximity-dependent provision of regulatory information by one chromosome to its homolog. If the association of proteins bound in *cis* to distant sites is important for regulation, it is possible that the association of proteins bound in *trans* to an unlinked DNA molecule could also contribute to regulation. The molecular genetics of the *Ubx* locus suggest that for transvection to occur, the promoter must be associated or in some way communicate with regulatory sequences from the ~40-kb *bxd/pbx* region and ~25-kb *abx/bx* region located, respectively, upstream and downstream of the *Ubx* transcription start site. It is not known whether these regions contain *Ubx* binding sites which might associate with the sites near the promoter or whether autoregulation plays a role in transvection. It should be noted that the product of the *zeste* gene, which is required for transvection, binds specifically to sites near the *Ubx* promoter, as well as to other sites in the regulatory regions just mentioned (8), and could play a facilitating role in mediating contacts between distant sites.

Multiple nearby and distant sites act synergistically in transcriptional activation. We tested the significance of binding cooperativity in vivo by using β-galactosidase as a reporter gene to monitor the regulatory potency of a variety of binding site combinations. We chose a yeast system, described more fully by Ekker et al. (20), to facilitate rapid testing of multiple sequences and to reduce the potential for interfering activities from homeodomain proteins present in homologous systems such as *Drosophila* embryos or cultured cells. This yeast system offers the additional features of inducible UBX Ib expression and control over target copy number, neither one easily achieved by the standard transfection procedures for assays in cultured cells. The ability to control target copy number is particularly important in attempting to measure the interactions of linked sites in *cis* when the protein is also capable of interacting with unlinked sites in *trans* (Fig. 5r to t).

We began by assaying the DNA fragments utilized for Fig. 1A to C, but because of an endogenous yeast activity that appeared to activate U-A site variants in which GCCG was present as a replacement for TAAT, we were unable to determine the effect of binding site number upon transcription. As an alternative, we tested a series of DNA fragments containing zero to eight nearly optimal *Ubx* binding sites, with multiple sites arranged in a directly repeated orientation at a spacing of 11 bp. As seen in Fig. 6A, the increase in transcriptional response is supralinear at lower site numbers, with a plateau reached at four or five sites and above; this behavior gives a sigmoidal appearance to the response curve and is suggestive of cooperativity.

To test the effects of interactions between distant sites, we compared the activation produced by a fragment containing U-A and U-B with the activation seen when U-A, U-B, or both sites were replaced by inert sequences; these substitutions maintained the size, spacing, and orientation of the test fragment with respect to the basal promoter. Transcriptional activation with one or both sites missing was reduced to 15% or less of that obtained with both sites present (Fig. 6B), consistent with a role for the cooperativity observed in vitro for binding to distant sites.

The *Ubx* homeodomain peptide does not bind cooperatively. Our previous work with a *Ubx* homeodomain peptide gave no indication of cooperativity in binding to DNA (21). To test this possibility explicitly, we carried out a variety of

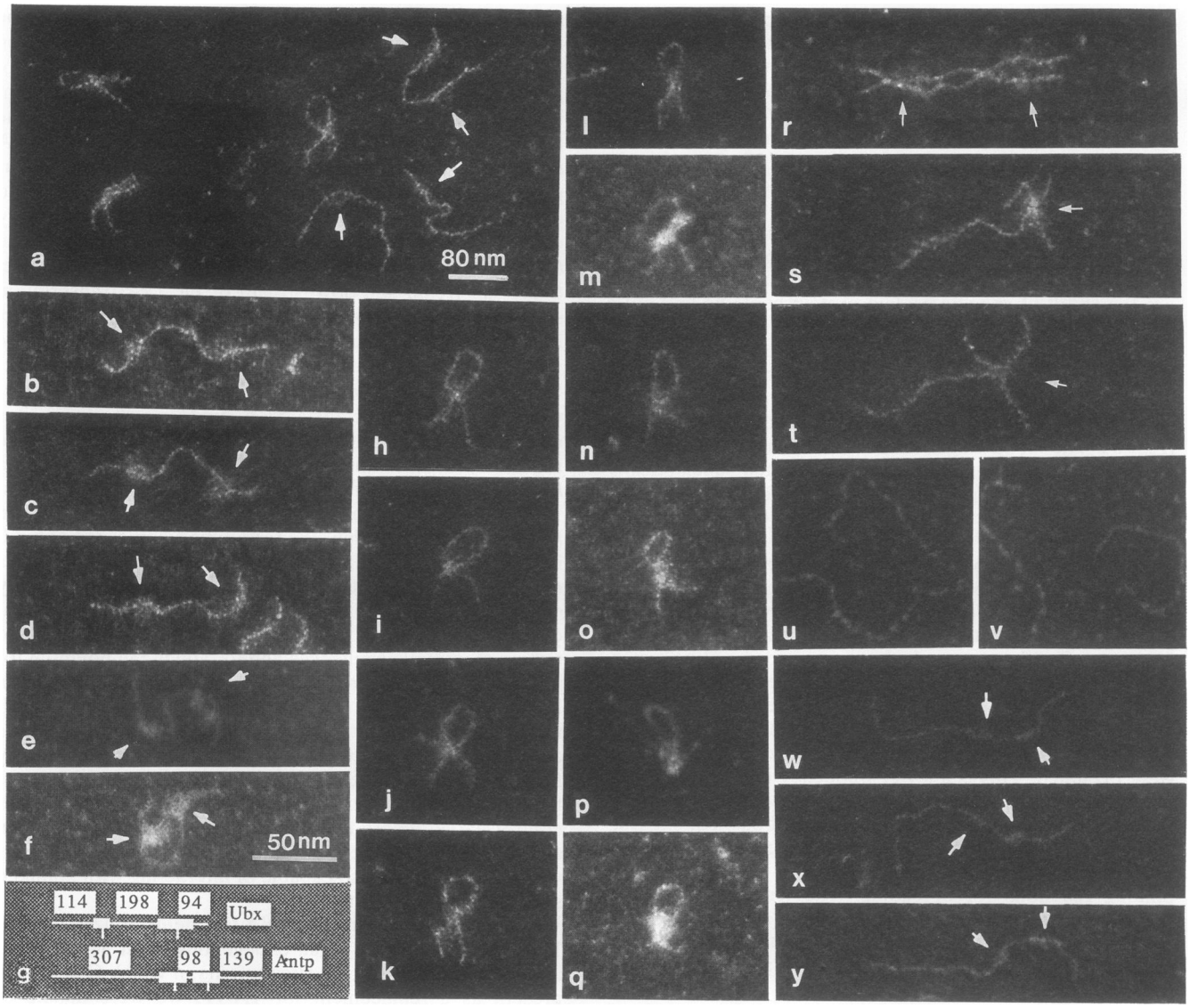

FIG. 5. Visualization of DNA complexes with the UBX Ib protein. UBX Ib complexes with specifically bound DNA fragments were prepared for viewing by dark-field electron microscopy as detailed in Materials and Methods. (a to q) Panel a shows a field containing linear and looped forms of UBX Ib complexes with a 406-bp genomic DNA fragment containing U-A and U-B (the insert from pPB260, also used in Fig. 3B and 4A, B, and D; see the schematic at the top of panel g). Individual complexes are shown in panels b to f (linear forms) and h to q (looped forms). Looped forms were observed in ~40% of complexes in which protein was associated with both sites (16 of 35 in one experiment and 76 of 196 in another). The end-to-center and center-to-center distances measured from micrographs for eight linear complexes (scaled in base pairs) correspond closely to the distances expected on the basis of footprinting (7) (see the top of panel g), with normalized means (± the standard deviations) of 102 ± 10, 213 ± 13, and 91 ± 8.2 bp. The mean loop circumference measured for 13 looped complexes was 235 ± 8.6 bp. (r to t) Several examples of protein-mediated intermolecular associations were also observed, involving either both binding sites on each molecule (r) or only one binding site from each molecule (as in panels s and t, in which the shorter DNA is a 242-bp fragment containing U-B). (u and v) Appearance of DNA spread in the absence of protein. (w to y) Linear complexes with a 544-bp genomic DNA fragment that extends from +13 to +556 with respect to the *Antp* P1 transcription start site. DNase I footprinting of this fragment identified A-1 and A-2, two binding sites located at end-to-center and center-to-center distances of 307, 98, and 139 bp (reference 7; see the bottom of panel g). These locations agree well with the normalized means of 310 ± 11, 103 ± 7.4, and 131 ± 11 bp measured from eight micrographs. The A-1 and A-2 binding sites were not further characterized with respect to the number of individual binding sites present nor with regard to cooperativity in binding by UBX Ib. The arrows indicate bound proteins.

experiments with the templates described above. Figure 4D shows that although the U-A and U-B sites were well protected by the homeodomain in a DNase I footprinting experiment, the spacer between these sites did not show the changes in cleavage pattern characteristic of looping induced by the intact protein. We also failed to observe stabilization of homeodomain-DNA complexes by addition of distant or nearby sites in footprinting experiments with the DNA fragments in Fig. 3 and 4 (data not shown). Finally, as shown in Fig. 1E, we ruled out cooperativity between two adjacently bound homeodomains by quantitative analysis of the equilibrium binding experiment shown in Fig. 1B. In a

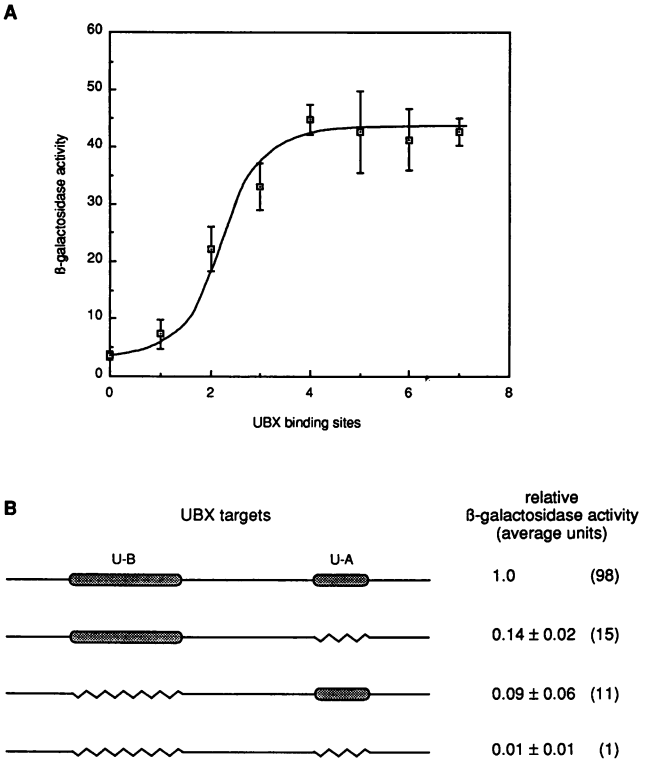

FIG. 6. Role of cooperative binding in transcriptional activation by UBX Ib. Transcriptional activation was assayed in a simple yeast system as a function of the number and arrangement of *Ubx* protein binding sites. UBX Ib protein expression was under control of the inducible GAL1 promoter, and binding sites to be tested were inserted upstream of a basal promoter driving transcription of the β-galactosidase reporter gene (described more fully in reference 20; see text). (A) β-Galactosidase activity was measured as a function of the number of adjacent binding sites inserted upstream of the basal promoter. Each point was plotted as the mean of four independent determinations plus or minus the standard deviation. Multiple sites were spaced by a center-to-center distance of 11 bp. (B) Several variants of the 406-bp genomic DNA fragment containing U-A and U-B (Fig. 3 to 5; see Materials and Methods) were tested upstream of the basal promoter in the β-galactosidase reporter construct. Normalized values from each of three independent experiments (the construct containing U-A and U-B equals 1.0) and are given plus or minus the standard deviation; average β-galactosidase activity is shown in parentheses.

situation involving cooperative binding to two sites similar in affinity, prior association of a protein molecule with the first site would enhance binding to the second; the level of singly occupied complexes during a protein titration should therefore reach a maximum level somewhat below the 50% level of total DNA expected if binding were independent (12, 91). The homeodomain peptide, consistent with the results from footprinting reactions, produced a maximum level slightly greater than 50% and therefore appears not to bind cooperatively.

The lack of cooperativity in *Ubx* homeodomain binding to multiple sites is consistent with the lack of dyad symmetry found in the optimal binding site defined previously (20). It is also consistent with the behavior of the *Ubx* homeodomain as a monomer in gel filtration (data not shown) and with the work of several other groups indicating that the homeodomain binds as a monomer (2, 23, 43, 57, 63). Given the behavior of the intact protein, these data suggest that cooperative binding of UBX Ib is mediated by amino acid sequences lying outside the homeodomain.

DISCUSSION

A working model for cooperative DNA binding by UBX Ib. Our studies of an *Ultrabithorax* homeotic gene product have revealed some properties not previously associated with this class of *Drosophila* DNA-binding proteins. The fundamental findings are summarized schematically in Fig. 7. UBX Ib molecules appear to bind cooperatively such that the stability of a complex is a function of not only individual site affinity (A) but also the number of individual sites within a cluster (B) and the presence of multiple separate clusters within a DNA region (C). Cooperativity in binding is depicted in Fig. 7 by contact between bound molecules. Because the *Ubx* homeodomain peptide fails to bind cooperatively, UBX Ib is drawn as a modular protein with its binding domain separate from its interactive surfaces. Interaction surfaces for adjacent sites are drawn as distinct donors and acceptors because of the geometry suggested by direct orientation of the repeated multiple sites within U-A and U-B (Fig. 7B). Since nearby cooperativity and distant cooperativity appear to be simultaneously manifested, distinct surfaces presumably are used for these two modes of interaction (Fig. 7C). The location and nature of these interaction surfaces remain to be determined, and the drawing is presented as a working model. This model has been useful in suggesting an explanation for the expansion of the weak-site footprint upon interaction with U-A (Fig. 3A): in interactions between sites of disparate sizes, binding of additional proteins to low-affinity DNA sites (lightly shaded in Fig. 7C) may be promoted by the favorable energies available through both the adjacent and distant modes of interaction.

The results of transcriptional activation assays in a simple yeast system were consistent with our findings of cooperative binding in vitro. We found that increasing numbers of adjacent sites led to a sigmoidal transcriptional response curve. In addition, we found previously, by using the same assay system, that transcriptional response can vary dramatically when the number of sites is held constant at four but the strengths of individual sites are varied (20). UBX Ib also appears to be able to mediate a cooperative transcriptional response to binding sites located some distance apart, as indicated by the synergy resulting from the combined presence of U-A and U-B in this assay system. Cooperative binding thus appears to function as a regulatory mechanism by which UBX Ib can integrate not only the number and quality of adjacent sites in a cluster but also the number of site clusters within a DNA region.

Intrinsic versus extrinsic cooperativity in vivo. One mechanism for cooperative integration in vivo is the formation of multiple contacts between the basal transcription apparatus and bound transcription factors (extrinsic cooperativity; see reference 62). This mechanism becomes particularly important when a factor is present at concentrations that are saturating for binding. UBX Ib is not present at saturating concentrations in our yeast assay system, since the strengths of individual sites can dramatically influence the response when the number of sites is held constant at four (20). Indeed, the difference in transcriptional response in vivo for the well-studied four-site clusters b and k parallels very closely the difference in complex stability measured in vitro (47-fold in vivo versus 43-fold in vitro; see Fig. 5 of reference 20). Thus, although we cannot rule out some role for

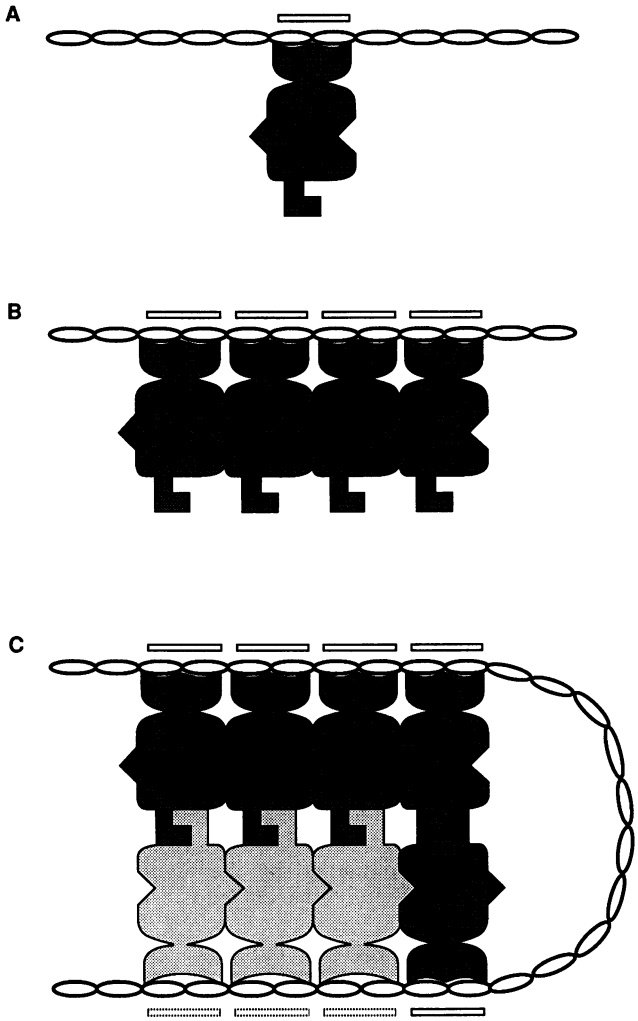

FIG. 7. A model for integration of multiple nearby and distant sites by UBX Ib binding. UBX Ib protein is depicted schematically as it interacts with a single site (A), several adjacent sites (B), and in a DNA loop anchored by interactions with sites at two distant locations (C). See the text for further discussion.

extrinsic cooperativity in the yeast assays, our in vitro results suggest that intrinsic cooperativity plays a major role.

Intrinsic cooperativity in binding to adjacent, as well as distant, DNA sites is a property of a number of prokaryotic regulators (reviewed in references 1, 30, 52, and 73). Cooperative binding is somewhat less well established but has also been demonstrated among eukaryotic transcriptional regulators. For example, the dimeric chicken progesterone receptor has been shown to bind cooperatively to two adjacent sites (91) and the *Drosophila* segmentation protein *even-skipped* appears to bind cooperatively (87). In addition, the stabilities of complexes formed by trimers of the heat shock transcription factor (96), by the human Hox 2.1 homeodomain protein (25), and by the Oct-2 POU homeodomain protein (48) increase dramatically with an increase in the number of adjacent binding sites. Electron microscopic evidence has also been obtained for homomeric interaction of eukaryotic transcriptional activators when bound to distant sites (33, 84, 89), although cooperativity in these cases has not been measured. The ability of UBX Ib to bind multiple adjacent and distant sites cooperatively thus is not unique among DNA-binding proteins, although it is the first example among the *Drosophila* homeotic proteins. Intrinsic binding cooperativity has been suggested for several other homeodomain-containing *Drosophila* patterning proteins (15, 16, 32), but the use of complex extracts or of fusions to oligomerizing proteins for these in vitro studies did not eliminate alternative conclusions.

An additional form of cooperativity may derive from homeodomain protein interactions with other transcription factors, as shown for yeast proteins α2 and MCM1 (80) and as suggested for the *fushi tarazu* protein (22, 24, 72). However, this possibility cannot be addressed biochemically until DNA-binding partners for *Drosophila* homeodomain proteins have been identified.

Implications of cooperativity for the design of regulatory elements. The ability of UBX Ib to integrate the properties of multiple binding sites organized in distinct clusters suggests that regulatory information at target sites may be distributed at several locations throughout large DNA regions. Binding cooperativity of the type we observed also suggests that functionally equivalent regulatory elements could be built from a few high-affinity sites, from many lower-affinity sites, or from sites of some intermediate number and affinity. How well do these expectations match the characteristics of regulatory elements active in an embryo?

The U-A and U-B sites near the *Ubx* promoter fit these expectations, since multiple clustered sites are found distributed throughout an extensive DNA region, but the role of these sites in the embryo is unclear. U-A and U-B are unusually well conserved in evolution (94), and *Ubx* appears to be autoregulated, not only in the epidermis and its imaginal derivatives (36) but also in the visceral mesoderm (9). The U-A and U-B sites were present in the reporter constructs used to study autoregulation; their analysis in ectodermal derivatives nevertheless has been slowed by the complexity and dispersion through large DNA regions of other essential regulatory sequences (37, 55). The U-A site has actually been implicated in *Ubx* autoregulation in the visceral mesoderm (56), but this analysis was complicated by the fact that at least part of the autoregulation is indirect, requiring expression of the *decapentaplegic* gene (34, 59, 90). Thus, more in vivo study will be required before U-A and U-B can be considered bona fide examples of regulatory elements for a homeodomain protein.

A better established *Ubx* target element is a recently described sequence that mediates repression by *Ubx* and *abd-A* proteins of the *Antp* P2 promoter in tracheal cells (4). Repression appears to involve a total of 30 binding sites distributed throughout a 2.3-kb region. On the basis of systematic mutagenesis studies, the repression appears to require a minimum number (14 to 19) of these 30 sites, without particular preference as to location or identity. The requirement for a large number of binding sites dispersed in an extensive DNA region provides an in vivo illustration of the principles derived from our in vitro work. Nearly all of the binding sites are suboptimal in sequence, suggesting that this regulatory element belongs to the class described above which contains many low-affinity sites.

Several other regulatory elements for homeodomain proteins have recently been identified. The best established case is for autoregulation of *fushi tarazu* (*ftz*), a segmentation gene located within the *Antennapedia* homeotic gene complex that contains a homeodomain very similar to those of the homeotic genes. By mutating binding sites to match the specificity of a mutated *ftz*-encoded protein, several functional sites for *ftz* autoregulation were identified (71, 72). As

these researchers recognized, their results implicate factors in addition to simple binding of DNA by the *ftz* protein, yet their work clearly identified a physiologically active autoregulatory element that contains multiple individual binding sites for the *ftz* protein distributed throughout an ~400-bp DNA region.

Other recently identified candidate elements for regulation by homeotic proteins appear to have in common the characteristic that multiple binding sites are distributed throughout a sizable DNA region (27, 39, 64, 92). In addition, several examples of regulatory elements for the more distantly related homeodomain protein encoded by *bicoid* have been identified (18, 78, 83), and these elements also appear to share the characteristics of multiple binding sites dispersed within a region of several hundred base pairs. Although high-affinity binding sites occur in most of these regulatory elements, lower-affinity sites are present and apparently play an important role in target elements for the *Dfd*, *ftz*, and *bcd* proteins. The characteristics of these regulatory elements suggest that other homeodomain proteins also have integrative properties based on cooperative binding to multiple dispersed sites.

Role of sequences outside the homeodomain in Ubx function. Although earlier binding studies demonstrated that the intact *Ubx* protein has the same individual site sequence preference as the homeodomain peptide (20), the studies reported here (Fig. 1B and E and 4D) suggest that sequences outside the homeodomain mediate cooperative binding to multiple sites. Chan and Mann (14) have recently shown that sequences carboxy terminal to the *Ubx* homeodomain can contribute to the specificity of homeodomain function. In ectopic expression experiments, for example, the presence or absence of the *Ubx* carboxy-terminal tail was a critical determinant for the segmental specificity of a chimeric protein containing the *Antp* homeodomain preceded by *Ubx* amino-terminal sequences.

Whether this in vivo function of the carboxy terminus correlates with the binding cooperativity we observed in vitro must be addressed by further biochemical and genetic studies. It is worth noting, however, that a recently developed sequence analysis algorithm (51) predicts a coiled-coil conformation for a 43-residue region at the carboxy terminus of UBX Ib with a probability greater than 0.99; no such structure is predicted for the *Antp* protein (Fig. 8A). In the work of Chan and Mann (14), two sequence changes appeared to preserve the in vivo activity of the *Ubx* carboxy-terminal tail. The first of these was a premature truncation removing 17 residues of the carboxy terminus; the second was an insertion of five alanine residues between the homeodomain and the carboxy-terminal tail. Despite the fundamental nature of these changes, coiled-coil structures of 30 and 39 residues, respectively, are predicted (data not shown). The preservation of a predicted coiled coil in each of these mutants supports the relevance of this structure for normal function of the *Ubx* protein.

The accuracy of the Lupas algorithm for structure prediction is not fully established, but it does suggest a mechanism for pairwise interaction between UBX Ib molecules. The possibility of a *Ubx* coiled-coil structure is particularly intriguing in the light of recent findings that association of *lac* repressor dimers, the mechanistic basis for DNA looping by the *lac* repressor, involves a coiled-coil motif near the carboxy terminus of the protein (3, 13). Although earlier gel filtration studies demonstrated that the intact *Ubx* protein has a large Stokes' radius suggestive of a dimer (7), a more recent analysis that included sedimentation measurements

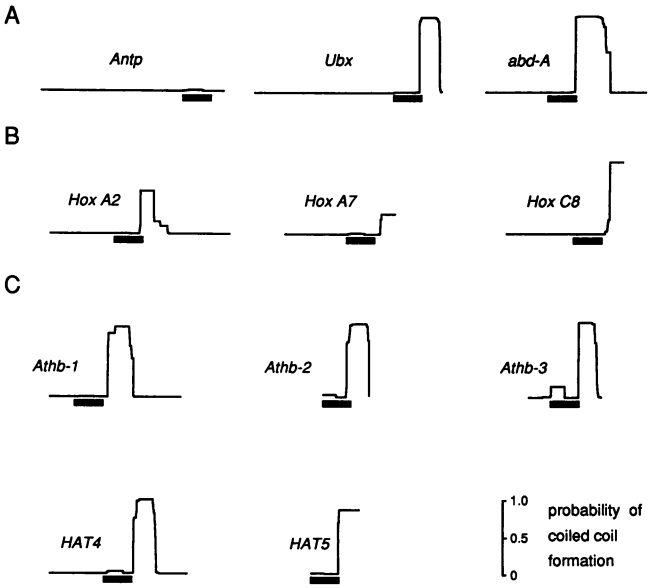

FIG. 8. Homeodomain proteins containing coiled coils. The algorithm of Lupas et al. was applied to the amino acid sequences of *Drosophila* (A), mammalian (B), and plant (C) homeodomain proteins. The location of the 60-residue homeodomain within each sequence is indicated by the solid bar, and the predicted probability of coiled-coil formation is indicated by the thin line (scale in panel C). All sequences were obtained from release 4.0 of the *Entrez* data base (National Center for Biotechnology Information); references may be found in the text and in reviews by Scott et al. (75) and McGinnis and Krumlauf (54). (A) The sequences of all eight *Drosophila* homeotic proteins were tested, and only *Ubx* and *abd-A* showed a significant probability of coiled-coil formation. The *Antp* sequence is included for comparison (see text). (B) Of the mammalian sequences available, the three shown display the highest probabilities of coiled-coil formation. The nomeclature used is that of Scott (74). (C) Most reported plant homeodomain proteins contain a leucine zipper–coiled-coil motif immediately following the homeodomain (see text).

indicates that the protein is a monomer in solution and that its large Stokes' radius is due to its unusual shape (10a). Under the conditions of these studies, the UBX Ib protein thus appears to exist in solution as a monomer. The association of UBX Ib monomers to form dimers or higher-order oligomers nevertheless could occur in the presence of DNA (see, for example, reference 42), and the energy of this favorable interaction could then account for the cooperative binding behavior we observed.

We have applied the sequence analysis algorithm of Lupas et al. (51) to all eight *Drosophila* homeotic proteins and found that in addition to *Ubx*, the *abdominal-A* protein (41) contains a predicted 71-residue coiled coil that begins immediately adjacent to the homeodomain in the same position as the *Ubx* coiled coil (Fig. 8A). In addition, several of the available protein sequences encoded by *Hox* genes show some probability of coiled-coil formation at the analogous position. Three of these, *Hox C8*, *Hox A2*, and *Hox A7* (nomenclature in accordance with reference 74), are shown in Fig. 8B. The amino acid sequences encoded by several recently characterized plant homeodomain genes reveal the presence of a leucine zipper motif immediately to the carboxy-terminal side of the homeodomain (53, 68, 70); application of the coiled-coil algorithm to these sequences indicates a high probability of coiled-coil structure formation,

which is true for only a subset of all heptad leucine repeats (51). Whether these predicted structures play a role in homomeric interactions between homeodomain proteins or in heteromeric interactions with other proteins will have to be addressed by further biochemical studies. The similarities between these sequences nevertheless suggest a common molecular mechanism for producing cooperative binding by homeodomain proteins.

ACKNOWLEDGMENTS

P.A.B. is grateful for help provided by colleagues at the Carnegie Institution Department of Embryology, where this work was initiated, and to M. Krasnow for communication of unpublished results. We thank A. Collector and C. Wendling for oligonucleotide synthesis, T. Tullius and M. Beer for providing equipment, and M. Biggin for the *Ubx* promoter deletion construct. A. Lupas supplied the computer code for the coiled-coil algorithm, and J. Meehan graciously assisted in adapting it for use with the Macintosh. We are also grateful to J. Corden, S. Desiderio, K. Matthews, and R. Schleif for helpful comments on the manuscript.

During the early stages of this work, P.A.B. was a DuPont Fellow of the Life Sciences Research Foundation and a Neuroscience Fellow of the Alfred P. Sloan Foundation.

REFERENCES

1. **Adhya, S.** 1989. Multipartite genetic control elements: communication by DNA loop. Annu. Rev. Genet. **23:**227–250.
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. **Alberti, S., S. Oehler, B. von Wilcken-Bergmann, H. Krämer, and B. Müller-Hill.** 1991. Dimer-to-tetramer assembly of lac repressor involves a leucine heptad repeat. New Biol. **3:**57–62.
4. **Appel, B., and S. Sakonju.** 1993. Cell-type specific mechanisms of transcriptional repression by the homeotic gene products UBX and ABD-A in *Drosophila* embryos. EMBO J. **12:**1099–1109.
5. **Beachy, P. A.** 1990. A molecular view of the *Ultrabithorax* homeotic gene of *Drosophila*. Trends Genet. **6:**46–51.
6. **Beachy, P. A., S. L. Helfand, and D. S. Hogness.** 1985. Segmental distribution of bithorax complex proteins during *Drosophila* development. Nature (London) **313:**545–551.
7. **Beachy, P. A., M. A. Krasnow, E. R. Gavis, and D. S. Hogness.** 1988. An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. Cell **55:**1069–1081.
8. **Benson, M., and V. Pirrotta.** 1988. The *Drosophila zeste* protein binds cooperatively to sites in many gene regulatory regions: implications for transvection and gene regulation. EMBO J. **7:**3907–3915.
9. **Bienz, M., and G. Tremml.** 1988. Domain of *Ultrabithorax* expression in *Drosophila* visceral mesoderm from autoregulation and exclusion. Nature (London) **333:**576–578.
10. **Biggin, M. D., and R. Tjian.** 1988. Transcription factors that activate the *Ultrabithorax* promoter in developmentally staged extracts. Cell **53:**699–711.

10a. **Baffinger, N., B. Johnson, and M. Krasnow.** Personal communication.

11. **Burnette, W. N.** 1981. Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radio-iodinated protein A. Anal. Biochem. **112:**195–203.
12. **Cann, J. R.** 1989. Phenomenological theory of gel electrophoresis of protein-nucleic acid complexes. J. Biol. Chem. **264:**17032–17040.
13. **Chakerian, A. E., V. M. Tesmer, S. P. Manly, J. K. Brackett, M. J. Lynch, J. T. Hoh, and K. S. Matthews.** 1991. Evidence for leucine zipper motif in lactose repressor protein. J. Biol. Chem. **266:**1371–1374.
14. **Chan, S.-K., and R. S. Mann.** 1993. The segment identity functions of *Ultrabithorax* are contained within its homeodomain and carboxy-terminal sequences. Genes Dev. **7:**796–811.
15. **Dearolf, C. R., J. Topol, and C. S. Parker.** 1989. The *caudal* gene product is a direct activator of *fushi tarazu* transcription during *Drosophila* embryogenesis. Nature (London) **341:**340–343.
16. **Desplan, C., J. Theis, and P. H. O'Farrell.** 1988. The sequence specificity of homeodomain-DNA interaction. Cell **54:**1081–1090.
17. **Drew, H. R., and A. T. Travers.** 1985. DNA bending and its relation to nucleosome positioning. J. Mol. Biol. **186:**773–790.
18. **Driever, W., G. Thoma, and C. Nusslein-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.
19. **Duncan, I. M.** 1987. The bithorax complex. Annu. Rev. Genet. **21:**285–319.
20. **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.
21. **Ekker, S. C., K. E. Young, D. P. von Kessler, and P. A. Beachy.** 1991. Optimal DNA sequence recognition by the *Ultrabithorax* homeodomain of *Drosophila*. EMBO J. **10:**1179–1186.
22. **Fitzpatrick, V. D., A. Percival-Smith, C. J. Ingles, and H. M. Krause.** 1992. Homeodomain-independent activity of the *fushi tarazu* polypeptide in *Drosophila* embryos. Nature (London) **356:**610–612.
23. **Florence, B., R. Handrow, and A. Laughon.** 1991. DNA binding specificity of the *fushi tarazu* homeodomain. Mol. Cell. Biol. **11:**3613–3623.
24. **Furukubo-Tokunaga, K., M. Müller, M. Affolter, L. Pick, U. Kloter, and W. J. Gehring.** 1992. In vivo analysis of the helix-turn-helix motif of the *fushi tarazu* homeodomain of *Drosophila melanogaster*. Genes Dev. **6:**1082–1096.
25. **Galang, C. K., and C. A. Hauser.** 1992. Cooperative DNA binding of the highly conserved human Hox 2.1 homeodomain gene product. New Biol. **4:**558–568.
26. **Gould, A. P., J. J. Brookman, D. I. Strutt, and R. A. White.** 1990. Targets of homeotic gene control in *Drosophila*. Nature (London) **348:**308–312.
27. **Gould, A. P., and R. A. H. White.** 1992. *Connectin*, a target of homeotic gene control in *Drosophila*. Development **116:**1163–1174.
28. **Graba, Y., D. Aragnol, P. Laurenti, V. Garzino, D. Charmot, H. Berenger, and J. Pradel.** 1992. Homeotic control in *Drosophila*; the *scabrous* gene is an in vivo target of *Ultrabithorax* proteins. EMBO J. **11:**3375–3384.
29. **Hinz, U., A. Wolk, and R. Renkawitz-Pohl.** 1992. *Ultrabithorax* is a regulator of β3 tubulin expression in the *Drosophila* visceral mesoderm. Development **116:**543–554.
30. **Hochschild, A.** 1990. Protein-protein interactions and DNA loop formation, p. 107–138. *In* N. R. Cozzarelli and J. C. Wang (ed.), DNA topology and its biological effects. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
31. **Hochschild, A., and M. Ptashne.** 1986. Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. Cell **44:**681–687.
32. **Hoey, T., and M. Levine.** 1988. Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. Nature (London) **332:**858–861.
33. **Hofmann, J. F.-X., T. Laroche, A. H. Brand, and S. M. Gasser.** 1989. RAP-1 factor is necessary for DNA loop formation in vitro at the silent mating type locus *HML*. Cell **57:**725–737.
34. **Hursh, D., R. W. Padgett, and W. M. Gelbart.** 1993. Cross regulation of *decapentaplegic* and *Ultrabithorax* transcription in the embryonic visceral mesoderm of *Drosophila*. Development **117:**1211–1222.
35. **Immergluck, K., P. A. Lawrence, and M. Bienz.** 1990. Induction across germ layers in *Drosophila* mediated by a genetic cascade. Cell **62:**261–268.
36. **Irvine, K. D., J. Botas, S. Jha, R. S. Mann, and D. S. Hogness.** 1993. Negative autoregulation by *Ultrabithorax* controls the

level and pattern of its expression. Development **117**:387–399.
37. **Irvine, K. D., S. L. Helfand, and D. S. Hogness.** 1991. The large upstream control region of the *Drosophila* homeotic gene *Ultrabithorax*. Development **111**:407–424.
38. **Johnson, F. B., and M. A. Krasnow.** 1990. Stimulation of transcription by an *Ultrabithorax* protein *in vitro*. Genes Dev. **4**:1044–1052.
39. **Jones, B., and W. McGinnis.** 1993. The regulation of *empty spiracles* by *Abdominal-B* mediates an abdominal segment identity function. Genes Dev. **7**:229–240.
40. **Kadonaga, J. T., and R. Tjian.** 1986. Affinity purification of sequence-specific DNA binding proteins. Proc. Natl. Acad. Sci. USA **83**:5889–5893.
41. **Karch, F., W. Bender, and B. Weiffenbach.** 1990. *abd-A* expression in Drosophila embryos. Genes Dev. **4**:1573–1587.
42. **Kim, B., and J. W. Little.** 1992. Dimerization of a specific DNA-binding protein on the DNA. Science **255**:203–206.
43. **Kissinger, C. R., B. Liu, E. Martin-Blanco, T. B. Kornberg, and C. O. Pabo.** 1990. Crystal structure of an engrailed homeodomain/DNA complex at 2.6 A resolution: a framework for understanding homeodomain/DNA interactions. Cell **63**:579–590.
44. **Kornfeld, K., R. B. Saint, P. A. Beachy, P. J. Harte, D. A. Peattie, and D. S. Hogness.** 1989. Structure and expression of a family of *Ultrabithorax* mRNAs generated by alternative splicing and polyadenylation in *Drosophila*. Genes Dev. **3**:243–258.
45. **Krämer, H., M. Niemöller, M. Amouyal, B. Revet, B. von Wilcken-Bergmann, and B. Müller-Hill.** 1987. *lac* repressor forms loops with linear DNA carrying two suitably spaced *lac* operators. EMBO J. **6**:1481–1491.
46. **Krasnow, M. A., E. E. Saffman, K. Kornfeld, and D. S. Hogness.** 1989. Transcriptional activation and repression by *Ultrabithorax* proteins in cultured *Drosophila* cells. Cell **57**:1031–1043.
47. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
48. **LeBowitz, J. H., R. G. Clerc, M. Brenowitz, and P. A. Sharp.** 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. Genes Dev. **3**:1625–1638.
49. **Lin, Y.-S., M. Carey, M. Ptashne, and M. R. Green.** 1990. How different eukaryotic transcriptional activators can cooperate promiscuously. Nature (London) **345**:359–361.
50. **Lin, Y.-S., and M. R. Green.** 1991. Mechanism of action of an acidic transcriptional activator in vitro. Cell **64**:971–981.
51. **Lupas, A., M. VanDyke, and J. Stock.** 1991. Predicting coiled coils from protein sequences. Science **252**:1162–1164.
52. **Matthews, K. S.** 1992. DNA looping. Microbiol. Rev. **56**:123–136.
53. **Mattsson, J., E. Soderman, M. Svenson, C. Borkird, and P. Engstrom.** 1992. A new homeobox-leucine zipper gene from Arabidopsis thaliana. Plant Mol. Biol. **18**:1019–1022.
54. **McGinnis, W., and R. Krumlauf.** 1992. Homeobox genes and axial patterning. Cell **68**:283–302.
55. **Müller, J., and M. Bienz.** 1991. Long range repression conferring boundaries of Ultrabithorax expression in the Drosophila embryo. EMBO J. **10**:3147–3155.
56. **Müller, J., F. Thuringer, M. Biggin, B. Zust, and M. Bienz.** 1989. Coordinate action of a homeoprotein binding site and a distal sequence confers the *Ultrabithorax* expression pattern in the visceral mesoderm. EMBO J. **8**:4143–4151.
57. **Müller, M., M. Affolter, W. Leupin, G. Otting, K. Wütrich, and W. J. Gehring.** 1988. Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain. EMBO J. **7**:4299–4304.
58. **O'Connor, M. B., R. Binari, L. A. Perkins, and W. Bender.** 1988. Alternative RNA products from the *Ultrabithorax* domain of the bithorax complex. EMBO J. **7**:435–445.
59. **Panganiban, G. E. F., R. Reuter, M. Scott, and F. M. Hoffmann.** 1990. A *Drosophila* growth factor homolog, *decapentaplegic*, regulates homeotic gene expression within and across germ layers during midgut morphogenesis. Development **110**:1041–1050.
60. **Pankratz, M. J., and H. Jaeckle.** 1990. Making stripes in the *Drosophila* embryo. Trends Genet. **6**:287–292.
61. **Porath, J., and B. Olin.** 1983. Immobilized metal ion affinity adsorption and immobilized metal ion affinity chromatography of biomaterials. Serum protein affinities for gel-immobilized iron and nickel ions. Biochemistry **22**:1621–1630.
62. **Ptashne, M.** 1992. A genetic switch. Blackwell and Cell Press, Cambridge.
63. **Qian, Y. Q., M. Billeter, M. Otting, M. Muller, W. J. Gehring, and K. Wuthrich.** 1989. The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell **59**:573–580.
64. **Regulski, M., S. Dessain, N. McGinnis, and W. McGinnis.** 1991. High affinity binding sites for the *Deformed* protein are required for the function of an autoregulatory enhancer of the *Deformed* gene. Genes Dev. **5**:278–286.
65. **Reuter, R., G. E. F. Panganiban, F. M. Hoffman, and M. P. Scott.** 1990. Homeotic genes regulate the spatial expression of putative growth factors in the visceral mesoderm of *Drosophila* embryos. Development **110**:1041–1050.
66. **Rosenfeld, P. J., and T. J. Kelly.** 1986. Purification of nuclear factor I by DNA recognition site affinity chromatography. J. Biol. Chem. **261**:1398–1408.
67. **Royer, C. A., A. E. Chakerian, and K. S. Matthews.** 1990. Macromolecular binding equilibria in the lac repressor system: studies using high-pressure fluorescence spectroscopy. Biochemistry **29**:4959–4966.
68. **Ruberti, I., G. Sessa, S. Lucchetti, and G. Morelli.** 1991. A novel class of plant proteins containing a homeodomain with a closely linked leucine zipper motif. EMBO J. **10**:1787–1791.
69. **Samson, M.-L., L. Jackson-Grusby, and R. Brent.** 1989. Gene activation and DNA binding by *Drosophila Ubx* and *abd-A* proteins. Cell **57**:1045–1052.
70. **Schena, M., and R. W. Davis.** 1992. HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. Proc. Natl. Acad. Sci. USA **89**:3894–3898.
71. **Schier, A. F., and W. J. Gehring.** 1992. Direct homeodomain-DNA interaction in the autoregulation of the *fushi tarazu* gene. Nature (London) **356**:804–807.
72. **Schier, A. F., and W. J. Gehring.** 1993. Functional specificity of the homeodomain protein *fushi tarazu*: the role of DNA-binding specificity *in vivo*. Proc. Natl. Acad. Sci. USA **90**:1450–1454.
73. **Schleif, R.** 1992. DNA looping. Annu. Rev. Biochem. **61**:199–223.
74. **Scott, M. P.** 1992. Vertebrate homeobox gene nomenclature. Cell **71**:551–553.
75. **Scott, M. P., J. W. Tamkun, and G. W. Hartzell.** 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta **989**:25–48.
76. **Sharp, P. A.** 1991. TFIIB or not TFIIB? Nature (London) **351**:16–18.
77. **Simcox, A. A., E. Hersperger, A. Shearn, J. R. S. Whittle, and S. M. Cohen.** 1991. Establishment of imaginal discs and histoblast nests in Drosophila. Mech. Dev. **34**:11–20.
78. **Small, S., A. Blair, and M. Levine.** 1992. Regulation of *even-skipped* stripe 2 in the *Drosophila* embryo. EMBO J. **11**:4047–4057.
79. **Small, S., and M. Levine.** 1991. The initiation of pair-rule stripes in the Drosophila blastoderm. Curr. Opin. Genet. Dev. **1**:255–260.
80. **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 $\alpha 2$ dimer. Cell **68**:133–142.
81. **St. Johnston, D., and C. Nüsslein-Volhard.** 1992. The origin of pattern and polarity in the Drosophila embryo. Cell **68**:201–219.
82. **Stringer, K. F., C. J. Ingles, and J. Greenblatt.** 1990. Direct and selective binding of an acidic transcriptional activator domain to the TATA-box factor TFIID. Nature (London) **345**:783–786.
83. **Struhl, G., K. Struhl, and P. M. Macdonald.** 1989. The gradient morphogen *bicoid* is a concentration-dependent transcriptional activator. Cell **57**:1269–1273.
84. **Su, W., S. Jackson, R. Tjian, and H. Echols.** 1991. DNA looping

between sites for transcriptional activation: self-association of DNA-bound Sp 1. Genes Dev. **5**:820–826.
85. **Suck, D., A. Lahm, and C. Oefner.** 1988. Structure refined to 2A of a nicked DNA octanucleotide complex with DNase I. Nature (London) **332**:464–468.
86. **Tartof, K. D., and S. Henikoff.** 1991. Trans-sensing effects from Drosophila to humans. Cell **65**:201–203.
87. **TenHarmsel, A., R. J. Austin, N. Savenelli, and M. D. Biggin.** 1993. Cooperative binding at a distance by *even-skipped* protein correlates with repression and suggests a mechanism of silencing. Mol. Cell. Biol. **13**:2742–2752.
88. **Thali, M., M. M. Müller, M. DeLorenzi, P. Mattias, and M. Bienz.** 1988. *Drosophila* homeotic genes encode transcriptional activators similar to mammalian OTF-2. Nature (London) **336**: 598–601.
89. **Théveny, B., A. Bailly, C. Rauch, M. Rauch, E. Delain, and E. Milgrom.** 1987. Association of DNA-bound progesterone receptors. Nature (London) **329**:79–81.
90. **Thüringer, F., and M. Bienz.** 1993. Indirect autoregulation of a homeotic *Drosophila* gene mediated by extracellular signaling. Proc. Natl. Acad. Sci. USA **90**:3899–3903.
91. **Tsai, S. Y., M.-J. Tsai, and B. W. O'Malley.** 1989. Cooperative binding of steroid hormone receptors contributes to transcriptional synergism at target enhancer elements. Cell **57**:443–448.
92. **Vachon, G., B. Cohen, C. Pfeifle, M. E. McGuffin, J. Botas, and S. M. Cohen.** 1992. Homeotic genes of the Bithorax complex repress limb development in the abdomen of the Drosophila embryo through the target gene Distal-less. Cell **71**:437–450.
93. **White, R. A. H., and M. Wilcox.** 1984. Protein products of the bithorax complex in *Drosophila*. Cell **39**:163–171.
94. **Wilde, C. D., and M. Akam.** 1987. Conserved sequence elements in the 5′ region of the *Ultrabithorax* transcription unit. EMBO J. **6**:1393–1401.
95. **Wu, C., S. Wilson, B. Walker, I. Dawid, T. Paisley, V. Zimarino, and H. Ueda.** 1987. Purification and properties of Drosophila heat shock activator protein. Science **238**:1247–1253.
96. **Xiao, H., O. Perisic, and J. T. Lis.** 1991. Cooperative binding of Drosophila heat shock factor to arrays of a conserved 5 bp unit. Cell **64**:585–593.