

Functional Analysis of Mouse *Hoxa-7* in *Saccharomyces cerevisiae*: Sequences Outside the Homeodomain Base Contact Zone Influence Binding and Activation

MICHAEL K. GROSS AND PETER GRUSS*

*Abteilung für Molekulare Zellbiologie, Max-Planck-Institut für Biophysikalische Chemie,
37077 Göttingen, Germany*

Received 21 May 1993/Returned for modification 26 July 1993/Accepted 30 September 1993

The murine developmental control gene product, *Hoxa-7*, was shown to function as a DNA-binding transactivator in *Saccharomyces cerevisiae*. The importance of the ATTA core, the preference for antp class flanking nucleotides, the importance of Asn-51 of the homeodomain (HD), and the synergism of multiple binding sites all reflect properties that have previously been described for HOM or Hox proteins in tissue culture systems. A comparison of contact positions among genes of paralog groups and classes of mammalian HDs points to a lack of diversity in positions that make base contact, suggesting that besides the combination of HD amino acid-base pair contacts, another means of recognizing differences between targets must exist if Hox genes select different targets. The HD of antennapedia is identical to the *Hoxa-7* HD. The interaction of *Hoxa-7* with the exact sequence used in the nuclear magnetic resonance three-dimensional structural analysis on the antennapedia HD was studied. *Hoxa-7* binding and transactivation was influenced by sequences outside of the known base contact zone of this site. We conclude that *Hoxa-7* protein has a second means to interact with DNA or/and that the sequences flanking the base contact zone influence HD interactions by distorting DNA within the contact zone (base or backbone). This result is discussed in terms of DNA flexure and two modes of transcription used in *S. cerevisiae*.

Vertebrate *Hox* genes occur in related clustered arrays on four chromosomes. The expression domains in ectodermal and mesodermal structures along body and limb axes are characterized as overlapping, with the order of anterior expression boundaries along an axis generally correlating with the positions of the genes in a cluster. Experiments involving perturbed expression, overexpression, or targeted inactivation of particular *Hox* genes in a variety of organisms, with subsequent morphologic analyses, reveal evidence consistent with the idea that *Hox* genes specify axial positions in the head, trunk, and limb axes of the embryo (for reviews, see references 55 and 75). An analogous function is ascribed to the homeotic selector (*HOM*) genes of *Drosophila melanogaster* (reviewed in references 55 and 75). Homeoproteins bind DNA and can function as transcription factors (reviewed in references 32 and 46). Each Hox protein is thought to have target genes through which it exerts a segment identification role. Although no such experiments are yet reported in vertebrates, it is possible, by using domain swap experiments in the developing fly, to map target recognition specificity (for auto- or cross-regulatory targets) of *Drosophila* homeoproteins to within or near the homeodomain (HD) (18, 25, 45, 53). Moreover, several mammalian Hox proteins can mimic in vivo site selection or specification functions of *Drosophila* HOM proteins with corresponding HDs (52, 54), suggesting that enough information for some target site selectivity is contained within the conserved region near or within the HD of Hox proteins as well.

The three-dimensional structure of the HD-DNA complex is available for the domains of antennapedia (antp) (62),

engrailed (en) (42), and Mat $\alpha 2$ (89). The *Hoxa-7* HD is exactly identical in primary sequence to the antp HD throughout the N-terminal arm (N-arm) and the three helices. The *Drosophila*-derived en domain is only 27% homologous to the yeast-derived Mat $\alpha 2$ domain. In addition, the Mat $\alpha 2$ domain contains a three-amino-acid (aa) insertion between helices I and II. Despite these drastic differences in sequence, the backbone traces of the two domains overlap very closely (within 1-Å [0.1-nm] root mean square difference) in the region between the N-arm and the end of the recognition helix (RH) (89). In addition, the docking of these two strongly diverged HDs onto the DNA is extremely similar, an observation that provides a useful simplifying principle when recognition is considered (63). In addition, comparison of the bound and unbound structures of the antp (*Hoxa-7*) HD revealed no major structural change upon DNA binding in solution (5, 68, 69).

There is a central role for the absolutely conserved N-51 in recognizing an A in the major groove by a double H bond. This A is located in each core motif (TAAT for en; ACA for Mat $\alpha 2$). There are also evident differences between the two HD structures in base contacts. The similarly positioned recognition helices display different side chains (Mat $\alpha 2$, S-50, N-51, R-53; en, I-47, Q-50, N-51) to the major groove and thus make optimal contacts to the different combinations of base pairs displayed by the respective sites. The N-arm also displays several differences in base contacts in the minor groove (Mat $\alpha 2$, R7; en, R3, R5). Site selectivity of the Ultrabithorax (Ubx) and Deformed (Dfd) HDs have been mapped to the flexible N-arm by subdomain swapping experiments for DNA binding site selection in vitro (20) and for target site selection in vivo (49). These two domains are of different classes, have identical recognition helices (except residue 56 [Fig. 1]), and correspond to domains along a cluster. If one looks within classes of mammalian HDs,

* Corresponding author. Mailing address: Max Planck Institute of Biophysical Chemistry, Department of Molecular Cell Biology, P.O. Box 2841, 37018 Göttingen, Germany.

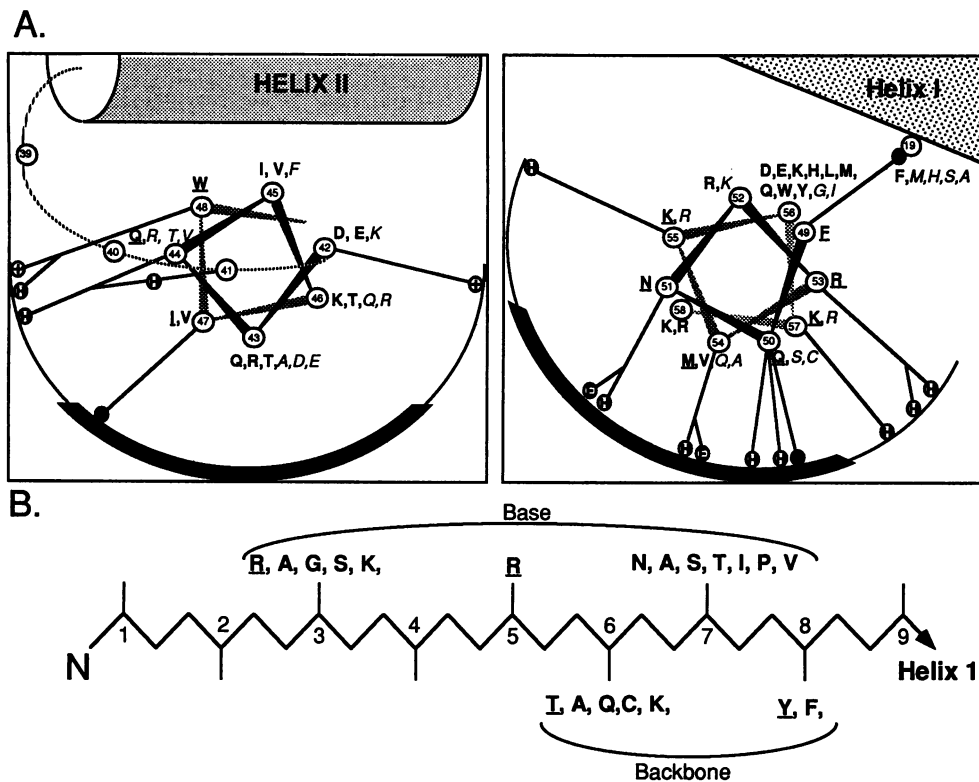


FIG. 1. Summary of possible RH and N-arm residues for all Hox HDs (bold) or other mammalian classes of HDs (italic). (A) Helical wheel diagrams show a view through the front (left; positions 42 to 48) and back (right; positions 49 to 58) sections of the RH as it lies in the major groove (semicircle) and makes contacts (lines) with the bases (thick semicircle) or sugar-phosphate (thin semicircle) moieties. The type of contact that was demonstrated at a given position for a particular residue (underlined) is indicated by H for hydrogen bond donor, E for hydrogen bond acceptor, + for electrostatic interaction, and ● for hydrophobic interaction. Branched lines indicate double contacts made by a single residue. Multiple lines emanating from one number indicate different types of contacts made by different residues at that position. No underlined residue at a position indicates it was only found in the Mat $\alpha 2$ structure. The very similar docking and backbone traces of en, Mat $\alpha 2$, and antp HDs on the DNA allow strong inferences to be made regarding the residues available for DNA contact at these positions in other HDs. All Hox HDs (bold) includes possible residues of all known mouse and human Hox domains from all four clusters and all 13 paralog groups (total of 36 Hox proteins per species). Symbols for other mammalian classes of HDs (italic) summarize possible residues in the HDs of mouse Evx1, and Evx2, human Evx2, and mouse En1, En2, Oct2, Oct4, Oct6, Pit1, Pax3, Pax6, Pax7, and Cdx1. Despite the large number of domains represented in this figure, the types of residues available at the contact positions are minimal and variability is typically between, not within, classes. The contacts summarized were obtained by a detailed reading of the three publications which present three-dimensional structures of HD-DNA complexes (see text). (B) The N-arm of the HD is shown schematically. Possible residues are shown for those positions at which a residue is known to contact DNA in any of the three determined structures. Position 7 makes contact only in the Mat $\alpha 2$ structure, and position 5 is conserved in all mammalian HDs examined (as described above). Thus, only position 3 can contribute to base sequence recognition if mammalian HDs wrap their N-arms into the minor groove similarly. However, unlike the RH, the N-arm is a flexible domain and could attach to DNA in a variety of ways and involving other positions. Helix 1 and helix 2 of the HD have only one residue that contacts DNA (Y-25), which is absolutely conserved in mammalian HDs. Thus, DNA binding differences among members of a class, if they exist, are likely due to differences in the N-arm or in domains outside the HD.

residues known to make base contact, in either minor or major grooves, are strongly conserved (Fig. 1). If the target specificity of mammalian Hox proteins lies within or near the HD, as has been demonstrated for several examples of the developmentally analogous HOM proteins of *D. melanogaster*, how can very closely related mammalian HDs (within a class) distinguish specific sequences? Alternatively, one may ask whether these proteins have different sets of targets in vivo or whether they merely act differentially on a common set of targets (based on adjacently bound proteins and/or on altered DNA topology, for example).

To establish a system that allows detailed and rapid analysis of the sequence specificity of the closely related HDs within the Hox cluster, we examined the function of Hoxa-7 in a sensitive effector/reporter system in yeast cells. Previous reports show that *Drosophila* HD proteins (20, 22,

71), the related *Drosophila* Dorsal and vertebrate c-Rel factors (37), *Drosophila* and vertebrate helix-loop-helix proteins (3, 10, 92), vertebrate hormone-dependent transcription factors (56, 58, 66, 73, 86), and vertebrate leucine zipper proteins (47, 80) can be functionally analyzed in yeast cells. DNA-binding specificities of different HD classes are readily resolved in such a system (31).

In the process of defining a transactivation and DNA binding function for Hoxa-7 from known consensus sites derived from the *Drosophila* or mouse genome, it became apparent that sequences outside of the base contact zone of the HD influence binding and activation. Mutational analysis of *cis*-acting elements demonstrated that both the previously defined core elements and a noncore, purine/pyrimidine (Pu/Py) asymmetric, AT-rich element were required for optimal transactivation and binding. The effect of site num-

ber and orientation was also examined. A point mutation in the HD reveals that it is required for all of the observed Hox-dependent effects. The particular flanking sequence that optimizes binding and activation has been demonstrated to induce intrinsic bends when it occurs in phased repeats (reviewed in reference 16). Either or both of two mechanisms, DNA distortion at the binding site or a second DNA-interacting domain in the Hoxa-7 protein, must be invoked to explain the data. The implications of each model are discussed in terms of DNA flexure and different modes of transactivation.

MATERIALS AND METHODS

Yeast strains and methods. *Saccharomyces cerevisiae* GFU8D (a *ura3-52 leu2 trp1*), a kind gift of H. D. Schmidt (Göttingen, Germany), was used in all experiments shown. Cells were grown in full medium until they were transformed with effector plasmids; thereafter all steps were done in minimal medium supplemented with the appropriate amino acids. Transformation was done by the standard lithium acetate method. Effector yeast strains, produced by transforming the host strain with effector plasmids and isolating clones, were used in reporter transformations. Three or more independent clones containing both effector and reporter (double transformants) were individually assayed for β -galactosidase (β -Gal) activity. The standard deviation was typically 10 to 20% of the average activity. Typically, triplicate measurements were made on several independent transformations, with the same relative activities of the reporters being compared but with different absolute activities between experiments. Differences in absolute levels measured in different experiments were due to differences in growth state (mid-log versus late log) or time after galactose induction. Hoxa-7-dependent transactivation is strongly dependent on these two parameters (data not shown). Thus, comparison of reporter construct activity was done only with double transformants grown and measured in parallel.

DNA molecules. Reporter constructs, except CEN (centromere) and *cyc* (cytochrome *c*) constructs, were all derived from the standard basal reporter LR1 Δ 1 (85) (provided by R. Brent), either by inserting synthetic oligonucleotides directly into the *Xho*I (X) site or by inserting oligonucleotides into the *Bgl*III (Bg2) and *Not*I sites of a polylinker of unique sites that had been previously recombined into the X site (see Fig. 2 and 3 for sequences and junction details). All inserts were confirmed by sequencing (except K' and 2K' in Fig. 1, which were created by inserting *Sal*I fragments from 5K'TATACAT [29] into the X site of LR1 Δ 1). pHB-9 was provided by R. Brent (30). CEN 6antp was derived by inserting the *Sma*I-*Nco*I (filled-in) fragment of LR1 Δ 1-6antp, containing the upstream and *lacZ* regions, into the *Pvu*II sites of the vector pRS316 (74). 2.5antp-*cyc* was created by inserting the 3antp/X (Fig. 3 and 4) oligonucleotide into the X site of 2 μ m *cyc*-Z, which had been created by replacing the *Bam*HI (B) fragment of LR1 Δ 1, containing the minimal *GAL* promoter, with a Bg2-B fragment containing the *cyc* minimal promoter (Bg2 in a polylinker attached at the X site at -247 from ATG and B as a linker at +7 from ATG; provided by W. Kibbe). Sequencing subsequently revealed that recombination had removed the GAGAAA of one of the antp sites.

Effectors were derived by inserting a 700-bp fragment containing the protein coding region of Hoxa-7 (provided by M. Kessel) into a variety of yeast expression vectors. The insert was derived from the BB21 cDNA (41) by a series of steps which resulted in a B site joined to the natural *Alu*I site

31 bp upstream of the ATG and a *Hind*III (H3) site attached at the *Hae*III site 7 bp downstream of the stop codon and was inserted into Bluescript (Stratagene). ADH-Hox and Rev-Hox were derived by attaching the H3 linker (CCA AGCTTG) to the B site and using the resulting H3 fragment to replace the *lexA-GAL4* insert of pRB1027 (7) (provided by R. Brent). UHC63 and UHC were created by inserting the H3 fragment into YEP52 (8) (provided by H. D. Schmidt) in the sense and antisense orientations, respectively. pGHox1.1 was derived by insertion of the B-*Sal*I fragment from the Bluescript clone into pG-1 (provided by K. Yamamoto), a derivative of pGPD-2 (73) obtained by exchange of the bacterial replication origin with that of pUC18. UHA88 was derived a series of recombinations which result in the upstream region of LR Δ 20B (85) (provided by R. Brent), from B in *GAL10* to *Mae*II in the *GAL1* 5' nontranslated region, and the Hoxa-7 coding region fragment B to H3, replacing the *ADH* promoter and *lexA-GAL4* coding region of pRB1027. The junction between the *Mae*II site of the inducible *GAL* promoter and the B site at -31 of the Hoxa-7 coding region reads . . . TTAACGATACCGTCGGGGGATCC. . . (nucleotides of *Mae*II and B sites are underlined).

The point mutant iYEPHox1.1HTH70/71 and control non-mutant iYEPHox1.1HTH68/69 (Fig. 4) were derived from the helix-turn-helix (HTH) deletion mutant iYEPHox1.1 Δ HTH by insertion of synthetic oligonucleotides encoding the mutated and intact HTH regions, respectively, of the Hoxa-7 HD into the *Mlu*I and *Nde*I sites. The resultant two plasmids are identical to UHC63 except that they have a *Mlu*I and an inactivated *Nde*I site within the HD sequence, which were selected so that they do not alter the protein coding information. iYEPHox1.1 Δ HTH was created by inserting an appropriate synthetic oligonucleotide between the *Eco*RI and *Pst*I sites so that a precise deletion of the HTH is created, with adjacent *Mlu*I and *Nde*I sites at the deletion point. The sense sequence of the oligonucleotide used to create the HTH deletion mutant is AATTCATTTCAACCGCTACCTGACGCGTCATATGAAGTGGAAAGAAAGAGCATAAAGATGAGA GCCAGGCTCCCCTGCA. The antisense oligonucleotide is such that *Eco*RI and *Pst*I sticky ends are created at the 5' and 3' ends, respectively. The sequence of the normal HTH replacement oligonucleotide is (*Mlu*I)CGCGTCCGCGCCGCATCGAGATCGCTCACGCGCTCTGCCTCACTGAGCGCCAGATCAAGATCTGGTTCAG[AAT]GGCGT(*Nde*I). The mutant replacement oligonucleotide is identical except in the bracketed region, where it reads GCT. The antisense strands have sequences such that *Mlu*I and *Nde*I sticky ends are created as indicated.

Plasmids encoding Hox/ β -Gal fusions were used to obtain proteins used in binding studies, and plasmids encoding Hoxa-7/glutathione *S*-transferase fusions were used to obtain fusion protein for injecting rabbits and creating immunopurification columns. Hoxa-7/ β -Gal fusions will be described in detail elsewhere. Briefly, BAL 31 deletions from the 3' end of the Hoxa-7 protein coding region, with attached linkers containing H3 sites, were inserted into the H3 site of pMLB1113 (27). This created fusion of codons 1 to 214 (full length) or codons 1 to 85 (N terminus) of Hoxa-7 to the N terminus (at codon 8) of the *lacZ* coding region. The linker sequences between these two coding regions encode the peptide GTQYCR. Gex34 (a gift from M. Kessel), the expression vector used to produce antigen, was created by recombining the B-*Sma*I (N-88) fragment of the coding fragment described above into the fusion protein vector pGex3x (76) at the B-*Sma*I sites. The resulting fusion protein is composed of, from N to C terminus, 234 aa of glutathione

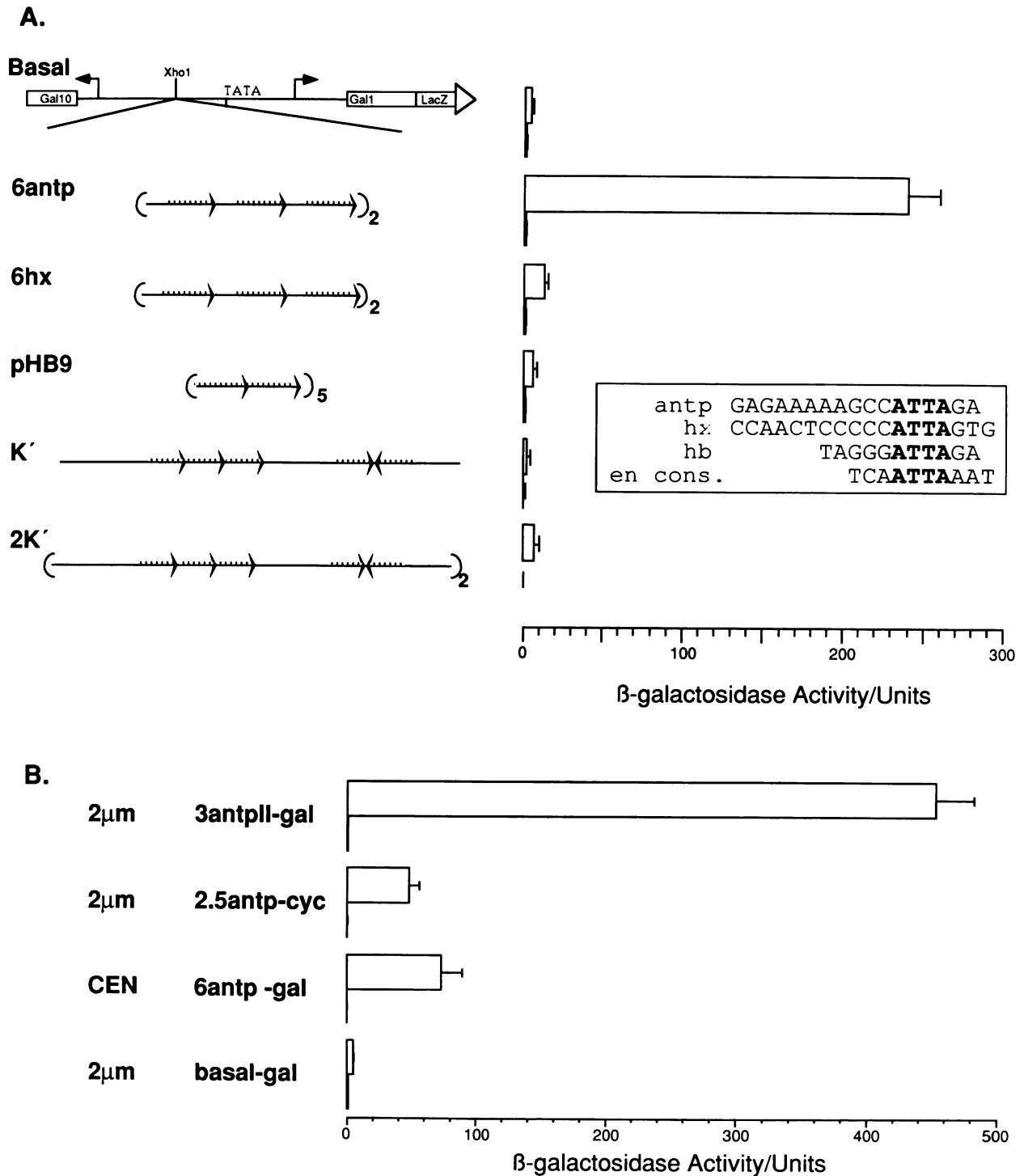


FIG. 2. *Hoxa-7* transactivation from natural HD binding sites. Each set of two bars has the indicated reporter. White bars indicate transformants containing functional *Hoxa-7* effector (+Hox); black bars indicate transformants containing a control effector plasmid (-Hox). (A) White bars indicate ADH-Hox effector; black bars indicate Rev-Hox controls. Control values, from top to bottom, were 0.4 ± 0.04 , 0.9 ± 0.4 , 0.5 ± 0.07 , 0.4 ± 0.2 , 0.2 ± 0.2 , and 0.0 ± 0.0 . Cells were grown in glucose selection medium. cons, consensus. (B) White bars indicate UHA88 effector; black bars indicate YEP52 controls. Control values were, from top to bottom, 0.5 ± 0.1 , 0.3 ± 0.1 , 0.1 ± 0.1 , and 1.0 ± 0.1 . Cells were grown in galactose selection medium. Error bars indicate the standard deviation of three parallel measurements.

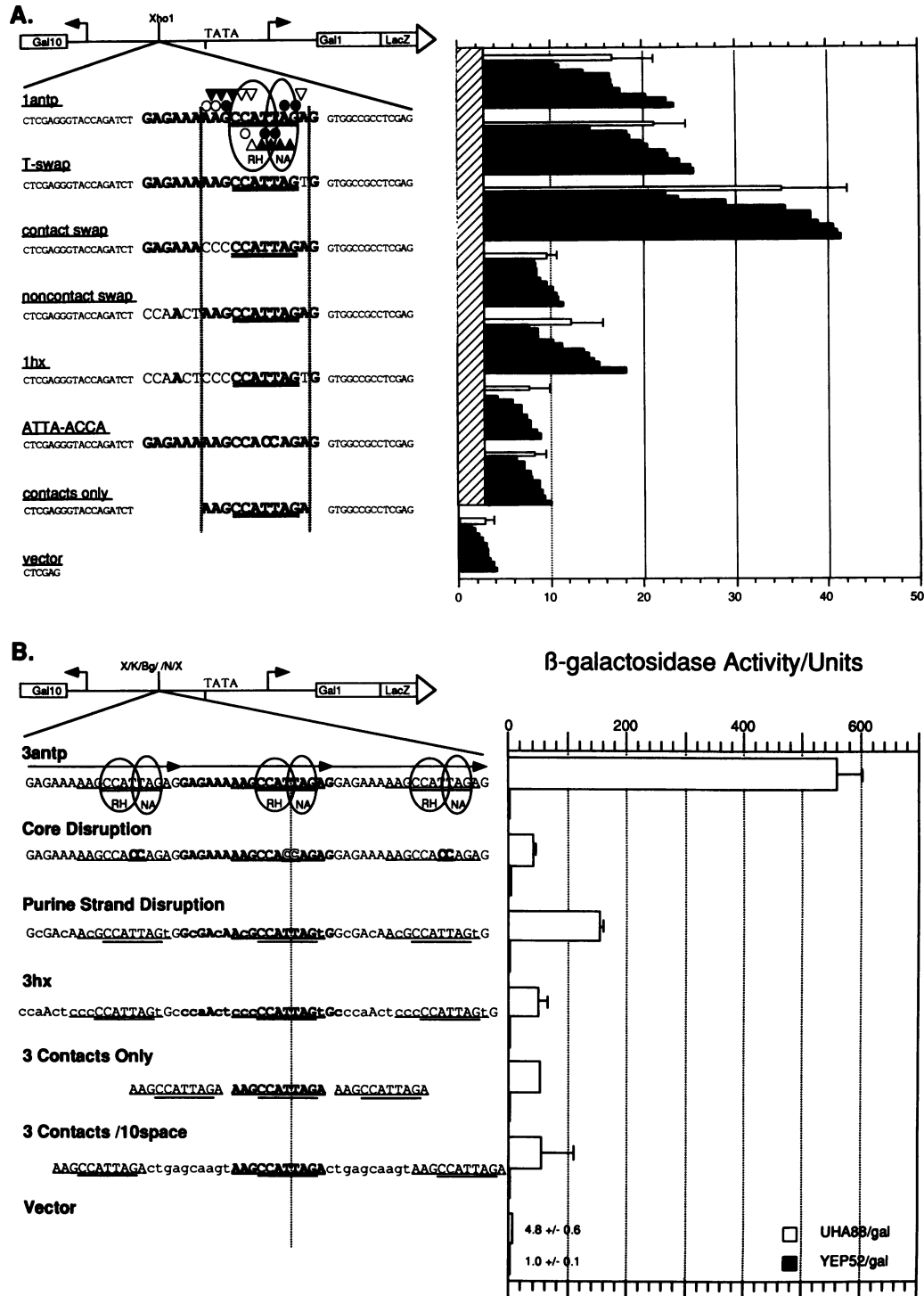


FIG. 3. Hoxa-7 transactivation from context and core mutants of the antp and hx sites. Small letters indicate residual polylinker sequences after insertion of oligonucleotides with the normal or mutant sites (large letters) into the minimal *GAL10-GAL1* promoter context (scale drawn; see Fig. 8). The base contact zone is doubly underlined, and the orientations of RH and N-arm (NA) contact zones are schematically indicated by circles. TATA indicates the canonical TATAAA box, and the arrow indicates the major transcription initiation site of the wild-type promoter. Dimethyl sulfide (circles) and ethylation interference (triangles) patterns determined for the antp (Hoxa-7) HD on this site (1) are indicated. Solid and open symbols indicate strong and weak interference, respectively. (A) Black bars show individual values (arranged by magnitude) of 10 parallel measurements for cells grown in galactose selection medium. These values were used to compute an average and standard deviation (white bars with error bars). Values shown are for transactivation with the strongest effector, UHA88. Controls for each reporter with use of the construct Rev-Hox or UHC (–Hox plasmids) all have values of 1 U or less (not shown). The vertical lines demarcate the interference zone in the antp site. Within the site, bold and plain letters indicate base pairs found at that position in the antp and hx sites, respectively. (B) Measurements for the UHA88 (+Hox) and YEP52 (–Hox) effectors were performed on three clones grown in galactose medium. The bar height and error bars indicates the average and standard deviation, respectively. Values for the vector without an insert are

S transferase, 9 aa of factor X cleavage site, 13 aa encoded by the B linker and Hoxa-7 5' nontranslated region, aa 1 to 88 of Hoxa-7 coding region, and 4 aa of the vector. β -Gal NT, used to create the immunoselection column (below), was the result of recombining a B-SmaI fragment encoding the N terminus of Hoxa-7 into the B-H3 (filled-in) sites of pUR292 (70). The resulting plasmid encodes (N to C terminus) β -Gal, 6 aa from linker, and aa 1 to 88 of Hoxa-7.

Protein methods. Hoxa-7/ β -Gal fusion proteins were isolated essentially as described elsewhere (83), using *Escherichia coli* DH5 Δ lac (27). Protein eluted from the substrate affinity column was concentrated by lyophilization, cleared by centrifugation at $10^4 \times g$ for 10 min, dialyzed against storage buffer (10 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid [HEPES; pH 7.6], 5 mM NaCl, 1 mM dithiothreitol [DTT], 0.1 mM EDTA, 10% glycerol), and frozen as aliquots at -70°C . Protein concentration was determined by the Bradford assay (full length at 0.12 $\mu\text{g}/\mu\text{l}$, N terminus at 0.26 $\mu\text{g}/\mu\text{l}$), and purity was examined by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) followed by silver staining. Fusion proteins were electrophoretically pure, with only one noticeable minor band that seemed due to degradation (data not shown).

Antigen was prepared from Gex34 transformed into *E. coli* DH5 α . Mid-log-phase cultures (3 liters) were induced for 7 h with 1 mM isopropylthiogalactopyranoside (IPTG), harvested at $4,000 \times g$, resuspended in 1/15 volume of lysis buffer (50 mM Tris [pH 8], 25% sucrose, 1 mM EDTA), and frozen. Quickly thawed suspensions were adjusted to 1 mM phenylmethylsulfonyl fluoride (PMSF) and treated with 20 strokes of a loose pestle in a Dounce homogenizer. Extract was made 1 mg/ml in lysozyme, kept 15 min on ice, sonicated three times for 20 s each time on ice at maximum impulse, and kept on ice 20 min longer. Three volumes of detergent buffer (0.2 M NaCl, 2 mM EDTA, 20 mM Tris [pH 7.5], 1% Nonidet P-40, 1% sodium deoxycholate) was added, followed by 20 strokes with a Dounce homogenizer (tight pestle) and centrifugation at $10^4 \times g$ for 20 min. The pellet was resuspended in 0.5% Triton X-100-1 mM EDTA and centrifuged again as above. The pellet was resuspended in 8 M urea-20 mM NaCl-20 mM KPO₄ (pH 7.5) by rocking at 4°C for 24 h. Inclusion bodies thus solubilized were separated by large-scale SDS-PAGE, and the appropriate band was recovered by gel elution as described previously (27). Recovered protein was dialyzed against 50 mM (NH₄)₂CO₃, lyophilized to dryness, and resuspended in 70% formic acid at 5 to 10 mg/ml. CNBr was added at a 100:1 molar ratio with respect to the molarity of methionine residues in each Gex34 preparation; the mixture was covered with N₂ and allowed to incubate 48 h at room temperature in the dark. Ten volumes of water was added, and the sample was lyophilized to dryness. CNBr cleavage products were separated on large-scale double stacking gels (72), and the appropriate 92-aa peptide gel, isolated as described above, was dialyzed against phosphate-buffered saline (PBS). A 5-week-old New Zealand White rabbit was injected with 300 μg and boosted with 250 μg of antigen, and serum was harvested as described previously (27).

Immune sera collected at various times were pooled, (NH₄)₂SO₄ (0.29 g/ml) was added, and the sera were kept on ice 1 h and centrifuged at $10^4 \times g$ for 30 min. The pellet was resuspended in and dialyzed to equilibrium against PBS. After clarification, this immunoglobulin fraction was repeatedly passed over two preabsorption columns (until no protein was retained) and one immunoselection column as described previously (27). The preabsorption columns were prepared as follows. (i) IPTG-induced Gex3x-containing bacteria (as described above) were resuspended in 0.2 M sodium borate-0.2 M NaCl-1 mM PMSF, treated for 4 min with glass beads in a Bead-Beater, and clarified by centrifugation at $10^4 \times g$ for 15 min. This soluble extract was coupled to CNBr-activated Sepharose as described previously (27). (ii) Total soluble protein was prepared from a culture of *E. coli* DH5 Δ lac harboring the parental plasmid pMLB1113 as described previously (83), dialyzed to equilibrium against 100 mM Na₂CO₃ (pH 8.3)-0.5 M NaCl, concentrated to 20 mg/ml by ultrafiltration (Amicon; as specified), and coupled to Affi-Gel 10 (Bio-Rad; as specified). The immunoselection column was prepared by coupling 7 mg of purified (as described above) β -Gal/NT to Affi-Gel 10. The final eluate from the immunoselection column was dialyzed to equilibrium against PBS, made 10% in glycerol, and stored as aliquots at -70°C . This preparation was able to specifically detect the CNBr fragment, representing the N-terminal 88 aa of Hoxa-7, in Western blot (immunoblot) analyses (data not shown).

Yeast whole cell extracts (see Fig. 7) were prepared from frozen cell pellets derived from mid-log-phase yeast cells grown in galactose selection medium for 11 h. Pellets (0.3 g) were resuspended in 1 ml of extraction buffer (0.4 M [Fig. 7A] or 0.8 M [Fig. 7B] NaCl, 20 mM HEPES [pH 7.6], 1 mM EDTA, 10% glycerol, 1 mM DTT, bestatin [40 $\mu\text{g}/\text{ml}$], pepstatin [0.7 $\mu\text{g}/\text{ml}$], aprotinin [2 $\mu\text{g}/\text{ml}$], leupeptin [0.5 $\mu\text{g}/\text{ml}$], PMSF [35 $\mu\text{g}/\text{ml}$]), 2.5 g of glass beads was added, and the mixture was vortexed eight times for 30 s each time at maximum speed. The extract was pipetted away from the glass beads and preclarified by centrifugation in a microfuge at 12,000 rpm. The supernatant was further clarified by centrifugation at $10^5 \times g$ for 10 min in a Beckman U2 centrifuge. Protein concentration (36) was normalized by using extraction buffer, and samples were frozen at -70°C . All steps were performed at or below 4°C .

Transactivation assays. Measurements were performed on cells in mid- to late-log-phase growth in minimal medium (A_{600} between 0.5 and 1.2). Culture (200 μl) was mixed with 800 μl of Z buffer (100 mM NaPO₄, 10 mM KCl, 1 mM MgSO₄, 35 mM β -mercaptoethanol [pH 7.0]) containing 0.006% SDS, 40 μl of CHCl₃ was added, tubes were vortexed for 10 s, and 200 μl of 4-mg/ml *o*-nitrophenyl- β -D-galactoside (ONPG) was added at time zero. Tubes were prewarmed (before addition of ONPG) and incubated at 28°C until a medium yellow color was observed. The reaction was then terminated by addition of 500 μl of 1 M Na₂CO₃. Absorbance was measured at 420 nm for the reaction and at 600 nm for the culture. Units of β -Gal activity were determined as $A_{420}(1,000)/A_{600}(0.2)(\text{time in minutes})$. Generally,

written next to bars. Values for the YEP52 controls are, from top to bottom, 1.4 ± 0.07 , 2.9 ± 0.1 , 2.0 ± 0.03 , 1.8 ± 0.3 , 1.1 ± 0.04 , and 0.8 ± 0.4 . The interference zone is underlined, the bold section delineates the center repeat of the triple site, and outline font highlights the core disruptions. The flanking upstream (Bg2) and downstream (NotI) sequences are as indicated for the 1antp reporter in panel A. The vertical line aligns the sites along the central A base pair of the center repeat. Uppercase and lowercase letters indicate base pairs that are the same or differ from the antp site, respectively (aligned on the central A base pair for each repeat).

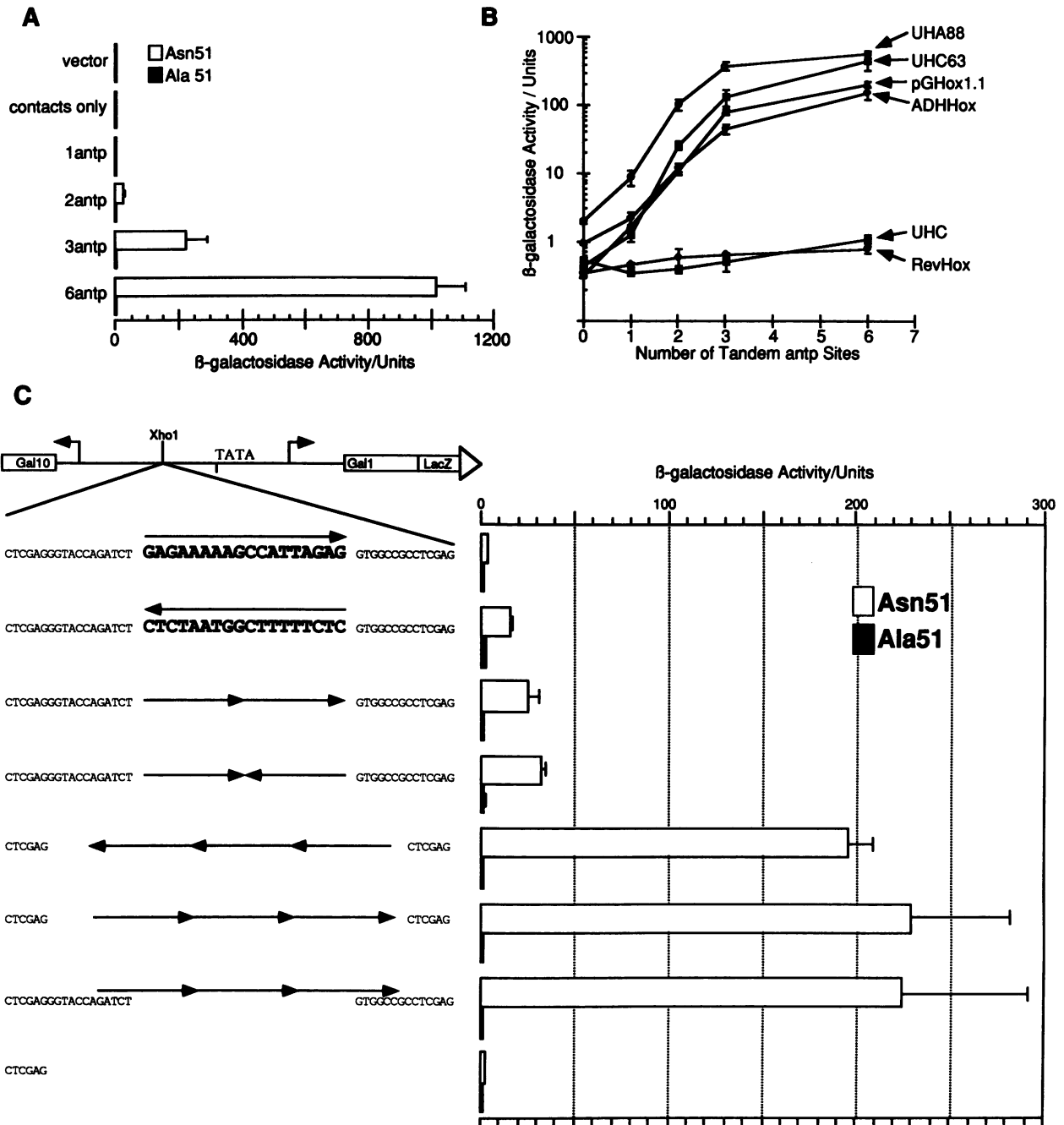


FIG. 4. Effect of point mutant, effector type, site number, orientation, and edges on transactivation. (A) Reporters with direct site repeats (ATTA in the sense orientation) were assayed in three cotransformant clones grown in galactose selection medium. For each, the activity of the control nonmutant (Asn-51; iYEPHox1.1HTH68/69) and the point mutant (Ala-51; iYEPHox1.1HTH70/71) are compared. Values for the point mutant are, from top to bottom, 0.8 ± 0.06 , 0.9 ± 0.2 , 1.2 ± 0.2 , 1.1 ± 0.2 , 1.0 ± 0.1 , and 1.3 ± 0.4 . (B) Logarithmic plot of activities and standard errors for all combinations of five different reporters with six different effectors. Effectors contain the Hoxa-7 coding region under the control four different yeast promoters (see Materials and Methods). Three independent clones were measured for each value. Site numbers of 0, 1, 2, 3, and 6 correspond to the reporters LR1 Δ 1 and its derivatives 1antp, 2antp, 3antp, and 6antp, respectively. All are direct repeats, and only the basal plasmid and 6antp has XhoI edges rather than Bg2-NotI edges (Fig. 3). (C) Diagrams are drawn as described in the legend to Fig. 3. Each bar represents the average and standard deviation of measurements on three independent cotransformant clones. White bars (+Hox) and black bars (-Hox) are for the control nonmutant and point mutant, respectively (as in panel A). Values for the low point mutant signals are, from top to bottom, 1.2 ± 0.2 , 1.7 ± 0.4 , 1.1 ± 0.2 , 1.1 ± 0.4 , 0.7 ± 0.03 , 1.0 ± 0.2 , 1.0 ± 0.1 , and 0.8 ± 0.06 . Differing edges are shown as text. Arrows indicate sequences as in the top two constructs.

reactions showing similar activities (Fig. 3A) were initiated and terminated at the same time.

Band shift analyses. In all experiments, binding reaction mixtures were mixed in the following order to give a total volume of 20 μ l: 1 μ l of competitor DNA [poly(dI-dC) or calf thymus], 4 μ l of 5 \times SBB (125 mM HEPES [pH 7.6], 50% glycerol, 10 mM DTT, 250 mM KCl, 0.25 μ g of bovine serum albumin per ml), water as required, 10⁴ cpm of oligonucleotide probes labelled by Klenow fragment fill-in with [α -³²P]dCTP, proteins (premixed). Reaction mixtures were kept on ice during mixing and for 2 h thereafter. Complexes were separated on 4% polyacrylamide-1 \times Tris-borate-EDTA-10% glycerol gels at 4°C in an electric field of 5 to 15 V/cm. No loading dye was added.

RESULTS

Hoxa-7 transactivates a reporter gene in *S. cerevisiae*. To utilize the simple eukaryote *S. cerevisiae* in the analysis of Hox proteins, a Hox-specific effect needed to be demonstrated (Fig. 2A). Known binding sites for either the antp HD (BS2 [57]; designated antp here), the Hoxa-5 protein (footprint b [60]; designated hx here), the en, paired, zerknüllt, Ftz, oct2, and even skipped proteins (K' fragment, which contains sequences from -949 to -708 of the en promoter including the five binding sites from which the original en consensus was defined [17, 29, 34, 43]), and the bicoid protein (30; designated the hb site here) were tested in a silent reporter vector (LR1 Δ 1). Sites were multimerized to maximize transactivation effects (29). Figure 2 shows diagrams of reporter genes with their respective insert sequences. Hoxa-7 was clearly able to transactivate reporter genes containing HD binding sites in *S. cerevisiae*. The transactivation was dependent on the inserted sites and on the presence of an effector plasmid containing the Hoxa-7 coding region in the sense orientation.

Effector and reporter genes were on plasmids containing replication functions of the 2 μ m circle. Typically, these plasmids are present at 50 to 100 copies per cell. Odenwald et al. (60) show that the related Hoxa-5 binds sequences within replication origins, suggesting that Hox proteins could influence DNA replication of the reporter plasmids containing Hox binding sites. In addition, it was possible that Hox expression caused selection of clones with low effector plasmid copy numbers (overexpression of transcription factors is often detrimental). A compensatory increase in reporter copy number would then result in a higher signal in Hox-expressing cotransformants. To test whether the Hox-specific reporter signal was only a copy number effect, the reporter transcription unit with the strongest transactivation signal was recombined into a CEN-based plasmid which is maintained at one to two copies per cell. Figure 1B demonstrates transactivation by Hoxa-7 of the CEN reporter. A clear difference in reporter gene activity is observed between Hoxa-7-expressing and control strains. Thus, Hoxa-7 did not exert its effect by directly influencing replication or abundance of the reporter plasmid.

To test whether the Hoxa-7-specific transactivation could also be achieved in a different promoter context, the best transactivation site was inserted into a reporter gene containing the cytochrome *c* minimal promoter. Clearly, Hoxa-7-specific transactivation is observed in the alternative promoter context (Fig. 2B). Thus, Hoxa-7 specific transactivation is not absolutely dependent on specific factors that bind only to the residual *GAL1-GAL10* promoter context (see Fig. 8).

In our attempt to use *S. cerevisiae* as a Hox-free system in which the function of each Hox protein can be measured independently (and thus comparatively) and the interactions of HD proteins with each other can be excluded (not possible in mammalian cell systems), it was important to demonstrate that the Hoxa-7-specific transactivation was not influenced by interactions with the yeast HD proteins α 1, α 1, and α 2. To test this possibility, isogenic strains of the two haploid mating types (YPH499 and YPH500) and diploids (YPH500) (74) were transformed with high-, middle-, or low-activity reporters. No qualitative differences in transactivation were observed between these three strains for any effector-reporter combination tried (data not shown). Thus, the Hoxa-7-specific transactivation is identical with various combinations of α 1, α 1 and α 2 present. One other HD protein, the PHO2 (BAS2) protein, has been discovered in *S. cerevisiae*. We have not yet tested whether this influences Hoxa-7-specific transactivation.

Point mutation in the Hoxa-7 HD abolishes transactivation.

The transactivation effects described above were clearly Hoxa-7 and *cis*-element specific. However, they could conceivably have been the result of a bridging mechanism or of a secondary activation mechanism. The former implies that a yeast factor bound to the inserted *cis*-acting element of the reporter gene but was unable to function until Hoxa-7 was present to form a bridge to the general transcription apparatus. The latter implies that Hoxa-7 functions to alter the expression of a yeast factor which then acts directly on the *cis*-acting element inserted into the reporter gene.

To discern whether Hoxa-7 functions via a bridging or binding mechanism, a point mutation was engineered into the Hoxa-7 HD. Asn-51 of the HD engages in two hydrogen bonds with the central adenine (TAATNN) in the major groove. Furthermore, Asn-51 is conserved in all HDs except that of pem (46). Removal of the amide functional group from Asn-51, as ain Ala-51, results in a complete ablation of Hoxa-7-specific transactivation for all reporter constructs described above or below. Figure 4 clearly demonstrates this for several reporters which produce increasing amounts of transactivation signal with the intact Hoxa-7 coding region. The three-dimensional structure surrounding this amino acid and the hydrogen bonding interactions are well understood for this and other HDs. It is unlikely that the alteration of the Asn-51 side chain significantly distorts the remainder of the Hoxa-7 protein structure. If transactivation were due to a bridging mechanism, it would be extremely unlikely that the amide functional group, absolutely essential for DNA binding in the major groove, is also essential for bridging. Therefore, DNA binding by the Hoxa-7 HD is required for transactivation and the bridging mechanism can be excluded.

It is considerably more difficult to exclude, *in vivo*, the possibility that Hoxa-7 is transactivating the reporter by a secondary activation mechanism. Typically, a direct effect is assumed if the *cis*-acting element required for transactivation *in vivo* can be shown to specifically bind to the protein *in vitro*. The specific binding of Hoxa-7, in yeast extracts or as a highly purified fusion protein, to the antp and hx binding sites is demonstrated below (Fig. 5 to 7).

Mutational analysis of *cis*-acting elements by transactivation. Although the most striking, Hoxa-7-specific, transactivation was from reporter genes containing either antp or hx sites, the absolute levels of transactivation of these two constructs differed by 1 order of magnitude. The base sequence specificity of the HD is determined by major groove-RH base contacts 3 bp upstream and by minor

groove-N-arm base contacts 3 bp downstream of the central T in the ATTA core sequence. The HD of antp (*Hoxa-7*) is identical to that of *Hoxa-5* in all residues known to make base contact in the antp or en structure. The identical RHs, and the lack of differences at positions which contact DNA (except residue 6, a backbone contact), of these HDs simplistically leads to the assumption that all three proteins should recognize identical base sequences. The antp site and hx site were isolated independently on the basis of their interaction with the purified, bacterially produced antp HD (57) and the baculovirus-produced *Hoxa-5* protein (60), respectively. As expected, the best sites identified in both of these studies have identical base contact zones (CCATTAG). The surprisingly drastic difference in function of *Hoxa-7* from these two sites suggested that these sites differ in flanking sequences critical to either binding or activation.

Comparison of hx and antp sites (Fig. 3A) suggested a mutation strategy to identify the nucleotides flanking the base contact zone which were critical to transactivation. The region of the antp site known to be in proximity to the *Hoxa-7* (antp) HD (interference zone) is indicated by triangles and circles (1). Only the base pairs on the edges of this zone do not have inferred contacts (base or backbone). The base contact zone of the *Hoxa-7* (antp) HD is underlined. Outside the base contact zone, sequences of the two sites differ in backbone contact and noncontacted regions. Transactivation via the antp site was to be reduced to the level of the hx site by swapping the differing nucleotides either upstream (contact swap) or downstream (T-swap) of the identical base contact zones or by swapping the differing nucleotides outside the interference zone (noncontact swap). The experiment was initially performed with single sites to simplify interpretation. Transactivation observed with single site reporters was far lower than would be expected for a linear dependence on binding site number, indicating that cooperative effects were contributing to the high signals seen with multiple sites (see below). However, use of a strong effector and numerous measurements allowed comparison of these single-site mutants. Figure 3A shows that transactivation from the 1antp site was only slightly greater (1.6-fold, with vector background subtracted) than that from 1hx. This contrasted greatly with the >10-fold difference between 6-mer or 3-mer sites (Fig. 2 and 3B) and suggested that the affinity of *Hoxa-7* for the antp site increased with site number, whereas the very similar hx site seemed to lack this property. Contrary to expectations, the contact swap mutant increased transactivation somewhat, suggesting that CCC is more favorable to *Hoxa-7* interaction *in vivo* than is the AAG used in the characterized *Hoxa-7* (antp) HD-DNA complex (1). The noncontact swap mutant reduced transactivation to levels comparable to that from 1hx as did the contacts-only mutant, which essentially replaced the GAGAA of 1 antp with flanking sequences and changed the position of the site by a half turn with respect to the promoter context. The T-swap did not influence *Hoxa-7* transactivation. Thus, sequences outside the interference zone and those immediately flanking the base contact zone influenced transactivation *in vivo*.

The importance of the core ATTA sequence to transactivation was tested with a site (ATTA-ACCA) containing mutations in the core which eliminate binding of *Hoxa-7* (see below). Although the transactivation of the 1antp site was impaired by this mutation, the ACCA mutant retained the ability to transactivate specifically at just barely below the level of 1hx. This was not due to a spacing effect because reporters with other inserts at this site are transactivated at

the level of the basal reporter (data not shown; see Discussion). Thus, regions outside the *Hoxa-7* (antp) HD contact area, notably the Pu/Py asymmetric region (GAGAAA AAG), were as important to transactivation as the ATTA core was. The inability to abolish transactivation by mutating either of these components alone suggests that both were required for optimal transactivation.

A second set of experiments was designed subsequently to determine whether the results obtained by analysis of single sites could explain the much larger differences observed with multimerized sites. Figure 3B shows transactivation obtained with 3antp versus 3hx sites differed by at least 10-fold. Experiments with weaker effectors gave larger fold differences (data not shown). Again, disruption of the core ATTA sequence in all three binding sites reduced but did not abolish transactivation (see Discussion). Comparison of the 3antp and 3hx sites led to the observation that the base contact zones were connected by uninterrupted, A-rich, purine strands in 3antp and by interrupted, C-rich, pyrimidine strands in 3hx. This led to speculation that the AT richness and/or the Pu/Py asymmetry of the connecting DNA influenced the topology of the core binding motif *in vivo* to give different levels of transactivation. The purine strand disruption (PSD) mutant was designed to interrupt the contiguous Pu/Py asymmetric strand at every third residue without changing the spacing or base contact zones. The PSD mutant also changed the spacer DNA from AT rich to GC rich. As expected, the mutant did not destroy transactivation but ameliorated it to a level between those of 3antp and 3hx. Thus, the orientation of the Pu/Py asymmetry with respect to the core and/or the AT richness of the spacer was crucial to transactivation.

Removal of the spacing sequences between interference zones (three contacts only) or insertion of 10 bp from a nonfootprinted portion of the immunoglobulin heavy-chain promoter (three contacts/10-bp space) resulted in the same activation levels as for 3hx. Thus, at three different spacings of the base contact zones (4, 12, and 14 bp), similar transactivation levels were seen, indicating that transactivation is not exquisitely sensitive to spacing. These results also suggest that the reduced transactivation observed in PSD and 3hx was not the result of a repressor interacting with the spacer sequence. Although the effects are more striking, results with 3-mer sites correlate well with those obtained for single sites. Not only the core motif but also the flanking Pu/Py strand (also AT rich) was important for optimal transactivation.

Effect of site multimerization on transactivation. The dramatic differences in transactivation levels achieved with single and 3-mer sites prompted an investigation of the site number dependence. Figure 4 shows that *Hoxa-7*-specific transactivation increased exponentially with site number at least up to 3 antp sites and thereafter increased linearly. The increase in activity with site number was dependent on the effector plasmid in use (Fig. 4B). The *Hoxa-7* coding region was placed under the control of the *GAL1* (UHA88), *GAL10* (UHC63), *GPD* (pGHox1.1), or *ADH* (ADH-Hox) promoters. Two constructs containing the reverse orientation coding region behind the *GAL10* (UHC) and *ADH* (Rev-Hox) promoters were also tested. The two strongest effector plasmids (UHA88 and UHC63) clearly showed nonlinearity, hence cooperativity, of transactivation with site number. Intermediate-strength effectors (ADH-Hox and pGHox1.1) also gave significant nonlinearity, but it was not as apparent at higher site number. Presumably the differences observed with various effector plasmids are due to different levels of

Hoxa-7 being produced (see above). Thus, the various strengths of synergism observed with different effector plasmids likely reflect the ability of the particular effector to fill the available sites with Hoxa-7. Perhaps an optimum complex size exists which maximizes transactivation per bound Hoxa-7. The size of this complex would be between three and six sites. The clustering of HD binding sites in promoter sequences from *Drosophila* and mouse genes suggests that such synergism may be a general feature of natural HOM/Hox targets. Future investigations will be required to determine the optimum number of sites, spacing, and spacer composition. The yeast system will provide a simple, quantitative means to do this with individual homeoproteins and with defined homeoprotein mixtures. Synergism may be mediated by cooperativity in DNA binding, cooperativity in transactivation, or cooperativity in chromatin structuring.

Orientation effects on transactivation. To test whether the exponentially higher signals seen with multimer binding sites are due to a stronger binding site being formed at the junction point of direct antp repeats, direct versus inverted repeats of antp binding sites were compared (Fig. 4C). The two reporters were equally effective in transactivation assays, indicating that no extra protein binding site or advantageous structure is formed at the junction. In addition, this experiment suggests that the synergistic effect observed with site multimerization does not depend on the orientation of the bound Hoxa-7 proteins with respect to each other.

Orientation-dependent function of transcription factors is relatively rare and usually ascribed only to the TATA-binding protein or TF2D complex. A threefold effect of orientation on transactivation was observed with a single antp site. The better-functioning orientation places the center of the core motif 8 bp further upstream and has a Py-rich region downstream on the sense strand (Fig. 4C). The orientation effect may be due to edge effects.

The orientation dependence of a 3antp site was measured to determine whether the orientation dependence of a single site was amplified upon multimerization. The 3antp site functioned in an orientation- and edge-independent manner. Thus, the orientation-dependent function to Hoxa-7 measured for a single site was obscured by its homosynergistic properties.

Purified fusion protein binding. Transactivation by Hoxa-7 in *S. cerevisiae* is likely to occur by a direct binding rather than a bridging mechanism. However, it is difficult to exclude secondary activation, the activation of some factor X by Hoxa-7, which in turn activates the reporter. Typically this cannot be demonstrated in tissue culture systems, although it is more likely to occur there than in yeast cells, as a result of the large number of similar *Hox* genes in a cluster controlled in part by auto- and cross-regulatory interactions. DNA binding of Hoxa-7 to the antp site was examined in vitro to determine by an independent method whether a specific interaction exists between Hoxa-7 and the antp site. If Hoxa-7 binds to the antp site in vitro, it is likely that this also occurs in vivo, given these data and the evidence presented above.

A protein (N214/ β -Gal) containing 214 aa of the Hoxa-7 protein, including the HD and lacking only 15 C-terminal residues, fused to the N terminus of β -Gal was isolated by substrate affinity chromatography to electrophoretic homogeneity. A similar fusion protein (N85/ β -Gal) containing only the N-terminal 85 aa of Hoxa-7 was similarly prepared as a control. Gel shift analysis with the labelled 3antp oligonucleotide and the purified fusion proteins revealed specific bands with the N214/ β -Gal fusion protein that were absent

with equal or greater amounts of N85/ β -Gal fusion protein (Fig. 5A). To identify the shifted band by a second means, an affinity-purified antibody preparation (Ab1) was used to supershift the band. Ab1 was derived from a rabbit injected with a glutathione *S*-transferase/Hoxa-7 fusion protein derivative. Addition of Ab1 to the binding reaction resulted in a dose-dependent appearance of bands of lower mobility but did not lead to a clear reduction in the intensity of the original band. This is most likely due to recruitment of more protein into complexes as was seen with extracts (below). Ab1 did not result in a shift by itself or in combination with N85/ β -Gal. These results were consistent with the idea that the Hoxa-7 protein interacts with the 3antp site via its HD.

Binding to mutant sites. To determine whether the in vivo transactivation results could be explained by a simple correlation with in vitro binding affinities, the 3-mer oligonucleotides tested in reporter constructs were gel isolated and labelled by Klenow filling of sticky ends, and bands were isolated again. The individually labelled oligonucleotides, of comparable specific activity, were incubated with identical amounts of N214/ β -Gal fusion protein and separated on native gels (Fig. 5B). Strongest binding was seen for the 3antp oligonucleotide, regardless of ends. The core mutation abolished binding, and the other mutations reduced binding to lower levels. Thus, binding affinities of the oligonucleotides to the fusion protein correlated directly with transactivation levels. Only the three contacts/no space probe bound better than expected from transactivation results, perhaps because the closer spacing allows the β -Gal moiety of the fusion protein to dimerize, an effect not possible in the transactivation system. Also, the absolute elimination of binding with the core mutation does not correlate with the merely reduced transactivation for this mutation (see Discussion). Nevertheless, binding experiments with highly purified Hoxa-7/ β -Gal fusion demonstrated that both the core and the spacer sequences were important for binding, a recapitulation of the transactivation results.

Multimerization enhances binding. The homosynergistic transactivation described above (Fig. 3 and 4) for multiple binding sites could be a result of either cooperative transactivation or cooperative binding. To distinguish between these models, the 3antp/N214/ β -Gal mobility shift band was competed for with unlabelled specific competitors containing different numbers and orientations of tandem antp binding sites. The competitors were of different lengths but had identical ends. Figure 6 shows that the unlabelled specific oligonucleotides containing more than one site all competed for the protein with similar efficiencies. Only the single antp site was unable to compete, even at much higher concentrations of sites. The inability of 1antp to compete was not due to an intrinsic inability of the fusion protein to bind a short piece of DNA because the fusion protein could bind to it when it was the labelled probe (data not shown). Thus, a single adjacent site in either orientation led to increased binding. However, the transactivational cooperativity seen in vivo had its maximum inflection between two and three sites, whereas the inflection for in vitro binding seems to occur between one and two sites. The reason for this difference may be the native Hoxa-7 protein, its concentration relative to that of the DNA, the supercoiled DNA topology, or the presence of other factors in the in vivo yeast system. These results are consistent with the hypothesis that Hoxa-7 homosynergism functions by cooperative binding.

Binding with yeast extracts. Standard attempts at achieving Hoxa-7-specific gel shifts with whole cell and nuclear extracts derived from yeast strains which gave maximal trans-

activation remained unsuccessful. However, addition of an affinity-purified antibody, directed against a glutathione *S*-transferase/Hoxa-7 N-88 fusion protein, allowed a Hoxa-7-specific band to be observed (Fig. 7A). Yeast whole cell extracts prepared from Hoxa-7-transformed and control strains were mixed with increasing amounts of the antibody preparation prior to addition of the labelled 3antp probe. Electrophoretic analysis revealed no difference in pattern between test and control extracts in the absence of antibody. However, in the presence of antibody, bands of retarded mobility were present only in Hox-containing lanes. The antibody without extract did not induce these shifts. This finding is consistent with the idea that Hoxa-7 is expressed from the effector plasmid and that the bands are the result of Hoxa-7 binding to the 3antp probe. The three bands likely represent different site filling of the 3-mer probe. Two explanations are consistent with the appearance of Hoxa-7-specific bands upon antibody addition. First, the antibody, directed against the N-terminal 88 aa of Hoxa-7, competes with a yeast protein (or the N terminus) that occludes the HD binding surface. Second, the antibody may serve as a bridge between two Hoxa-7 proteins, thereby promoting synergistic binding to the tandem sites. In the fusion protein binding assays this bridge may be the dimerization domains of the *lacZ* moiety. At any rate, this technique presents a means to visualize the Hox protein produced in yeast cells.

The binding activities of 3-mer mutant oligonucleotide binding sites were compared in yeast extracts (Fig. 7B). Extracts prepared from yeast cells transformed with either Hoxa-7 or control effector plasmids were mixed with equal amounts of antibody and then with each of the seven individually labelled oligonucleotides of comparable specific activity. Figure 7B shows that a Hoxa-7-specific band was clearly evident with the 3antp probes (regardless of ends). A faint Hoxa-7-specific signal is present with the PSD probe. Longer exposures allowed weaker yet Hox-specific bands to be detected for all 3-mers except for the core mutation (Fig. 7B, bottom). Thus, the core mutant abolished binding, 3antp had the strongest binding, and the other oligonucleotides showed signals of intermediate strength that correlate with transactivation data. Thus, binding in extracts was also consistent with the model that the degree of transactivation from a 3-mer site is closely related to the affinity of Hoxa-7 for the site.

DISCUSSION

The results above describe the first evidence that a vertebrate HD protein can function as a DNA-binding transactivator in yeast cells. In such a system, several functional properties expected by analogy to *Drosophila* HOM proteins can be recapitulated and sensitively measured. Hoxa-7-mediated transactivation in *S. cerevisiae* (i) displays a preference for consensus sequences with the flanking CC dinucleotide (as in CCATTA) which are known to be preferred by antp class recognition helices (46) (crystal and nuclear magnetic resonance structures) (sites containing the flanking dinucleotides for bicoid [GG] or en [CA] type only give marginal and spurious Hoxa-7-specific transactivation that is barely above background [Fig. 1] [26a]); (ii) is absolutely dependent on the Asn-51 amide group, which makes two H bonds to the central A; (iii) requires the central TT dinucleotide in the ATTA core for binding; and (iv) displays homosynergism with multimerized binding sites (29). Thus, the data are consistent with a mode of transactivation similar to that described for several HOM proteins (13, 44, 87) and

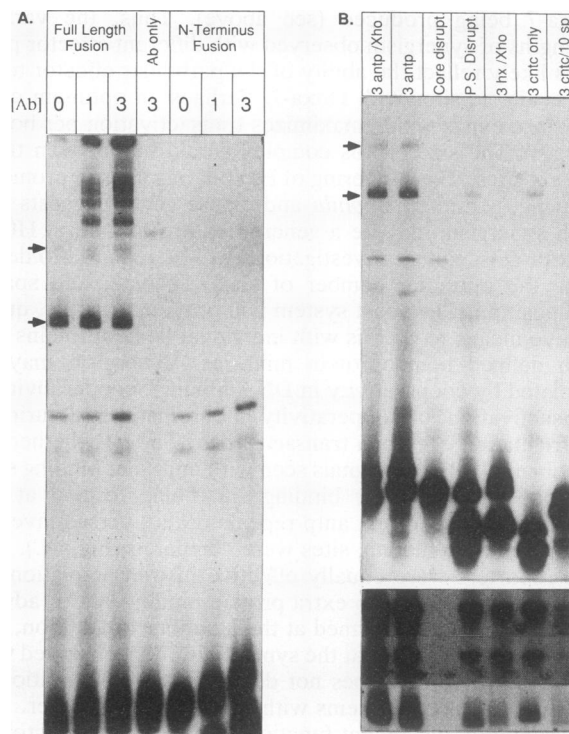


FIG. 5. Binding studies with purified Hoxa-7/ β -Gal fusion proteins. (A) Binding reactions using labelled 3 antp (Bg2-*Nor1* ends) oligonucleotide and 1 μ g of poly(dI-dC) were assembled with full-length fusion protein (0.36 μ g), control N-terminus fusion protein (0.78 μ g), or water (Ab only). Either 0 μ l (lanes 0), 1.5 μ l (lanes 1), or 4.5 μ l (lanes 3) of affinity-purified antibody preparation (see Materials and Methods) was premixed with the fusion proteins before addition to the binding reaction. Arrows indicate HD-specific bands seen only with the HD-containing fusion protein. Other lower-mobility complexes (or aggregation in the well) are seen only when both antibody and HD are present, suggesting that the antibody is specifically recognizing the Hoxa-7 moiety of the fusion protein. One faint nonspecific band is seen with the antibody alone. Several higher-mobility complexes are seen with both N-terminus and full-length fusions and are absent in the Ab only lane. These bands are consistent with a second binding function of the Hoxa-7 N-terminal 88 aa. (B) Band-isolated probes of comparable specific activity were mixed with full-length fusion protein (0.36 μ g) of the same source as in panel A. The gel and the top strip show 24-h and 4-week exposures, respectively, of binding using 0.3 μ g of poly(dI-dC) nonspecific competitor. The middle and bottom strips show relative band intensities with 0.3 μ g of calf thymus DNA and 1 μ g of poly(dI-dC), as nonspecific competitors. Strips show the band corresponding to the bottom arrow. A higher-mobility band corresponds to those seen in panel A for N-terminus binding. Note that this band is retained where the HD-specific bands (arrows) are lost in the core disruption mutant. Note also that this band is abolished in all sites which do not retain spacers from native binding sites (PSD [P.S. Disrupt.], three contacts only [3 cntc only], and three contacts/10-bp space [3 cntc/10 sp]). This is also consistent with a second binding domain (see Discussion).

recently for several Hox proteins (4, 65, 91). These properties are expected because the HD of Hoxa-7 is identical to that of antp in the N-arm and helices, and has an RH 3 identical to those of Dfd, Ubx, and Ftz HDs (except residue 56).

We have also shown that Hoxa-7 recognizes known consensus binding sequences *in vitro*, using purified fusion protein of bacterial origin. In addition, we compare binding

with whole eukaryotic cell extracts used as the source. The reason that this has only rarely (65) been demonstrated may be related to the reason why an antibody to the N terminus is required to visualize Hoxa-7-specific bands in our system. Either the N terminus occludes the HD binding surface by attaching to a second protein (or until a second protein binds it) or a second protein is normally required to link individual Hoxa-7 molecules on tandem binding sites. In our binding assays, the role of this second protein would be mimicked by the affinity purified antibody against the N-terminal 85 residues.

We have also demonstrated that Hoxa-7 senses sequences outside the base contact zone. Results from transactivation and in vitro DNA binding studies correlate in this conclusion and suggest the direct interpretation that Hoxa-7 recognizes these sequences by a second DNA binding domain (see below). However, the peculiar nature of the flanking sequence which it prefers for transactivation and binding, allows other plausible interpretations, involving intrinsic curvature and/or a relation to the dAdT elements described by Struhl (77).

Spacer DNA-induced distortion or a second binding domain. The sequence GAGAAAAG resides adjacent to the base contact zone in the antp site which is preferred in binding and activation to the various mutant sites (Fig. 3) that have identical base contact zones but differ in the flanking region. We have discussed the elimination of bridging models by use of a point mutation in the HD (Fig. 4; Results). A second model for the preferred activation of the antp versus hx sites (particularly as multimers) invokes a second protein which interacts with the GAGAAAAG spacer of antp site but not with the spacer of hx. This model seems particularly attractive in view of the reports of constitutive, cobinding coactivators such as MCM and serum response factor which increase the specificity of site recognition for yeast HD protein MAT α 2 and the mammalian HD proteins Phox-1 and SAP1 (28, 40, 82, 84).

However, several lines of evidence suggest that a second factor is not involved in the current situation. (i) The >98% pure Hoxa-7/ β -Gal fusion protein of bacterial origin, which can be assumed to lack the putative cobinding protein, also binds preferentially to the 3-mer sites with the GAGAAAAG spacer. (ii) In gel shifts with yeast extracts, the core disruption mutant shows no additional bands that are not also found in 3antp. (iii) In gel shifts with yeast extracts, the Hox-specific bands observed for 3hx (+3 bp) and PSD (=size) oligonucleotides are of the same size as the Hox-specific complexes observed for 3antp oligonucleotide. If a second protein was preferentially intercalating at the spacer, one would expect to see a slower mobility in the 3antp, Hox-specific complex. (iv) If a yeast protein were binding specifically to the junction of the tandem sites, one would see a difference in direct and inverted repeats. This was not observed (Fig. 4C). We cannot exclude the possibility that a second protein interacts with the spacer in the transactivation assay, that the β -Gal moiety leads to a preference for the 3antp site in vitro, or that a yeast protein specifically displaces a Hoxa-7 molecule from the 3antp site to give a complex that is fortuitously of the same mobility as the 3hx complex. These explanations require an individual mechanism to account for each observation. However, taken together, the results suggest that the preferential use of the antp site in vivo is not likely due to a yeast factor that specifically recognizes the antp spacer sequence. Thus, we must explain the preference for the Pu/Py asymmetric spacer in vitro and in vivo by one of these two models: (i) the spacer

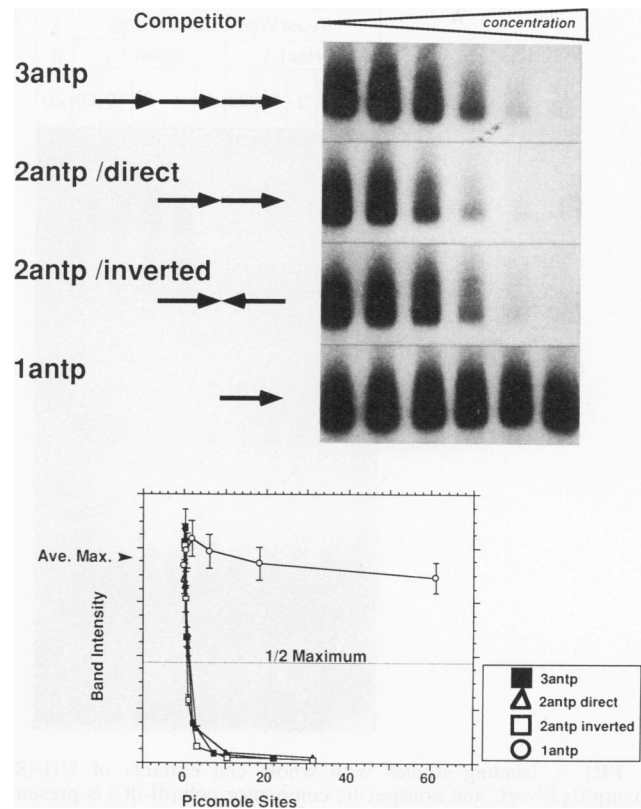


FIG. 6. Adjoining site in either orientation enhances fusion protein binding. Binding reactions were as in Fig. 5 except that increasing amounts of specific competitors were included. Strips correspond to bottom arrows in Fig. 5A and B; identical results were obtained for bands at the top arrows. Values plotted are derived by densitometry. "Picomole sites" refers to individual antp sites, regardless of whether they are tandem or single. Competitor amounts in slots for each strip differ slightly (from left to right): 3antp, 3, 10, 30, 100, and 300 ng; 2antp, 4.5, 15, 45, 150, and 450 ng; 1 antp, 10, 30, 100, 300, and 1,000 ng.

DNAs influence the HD-DNA interaction by distorting the DNA within the contact zone (base or backbone), and (ii) a second DNA binding domain exists in the Hoxa-7 protein. Note that these two models are not mutually exclusive. We will discuss each model below in view of available literature concerning the AT-rich Pu/Py tract.

Nature of the active spacer. The spacer sequences of the 3antp site resemble the dA-dT and DED48 elements which have been intensively studied in *S. cerevisiae* by the groups of K. Struhl and R. Kornberg. No specific binding proteins have been reported, although such elements may interact with proteins by general recognition (14, 26, 88; see below). These Pu/Py asymmetric AT-rich elements have unique physical and biochemical properties, occur frequently in yeast and vertebrate promoters (77), and are involved in DNA conformation-dependent constitutive transcription which functions by a fundamentally different mechanism than acid activation (12, 64, 78). They also prefer locations at the edges of nucleosomes in vitro (50, 67), and they synergistically aid the transactivation of a variety of DNA-binding factors in vivo (9).

The spacer sequence also contains an A tract. Crystal structures of A tracts have shown that they are essentially straight in the region of AA (or TT) base steps (59, 90), but

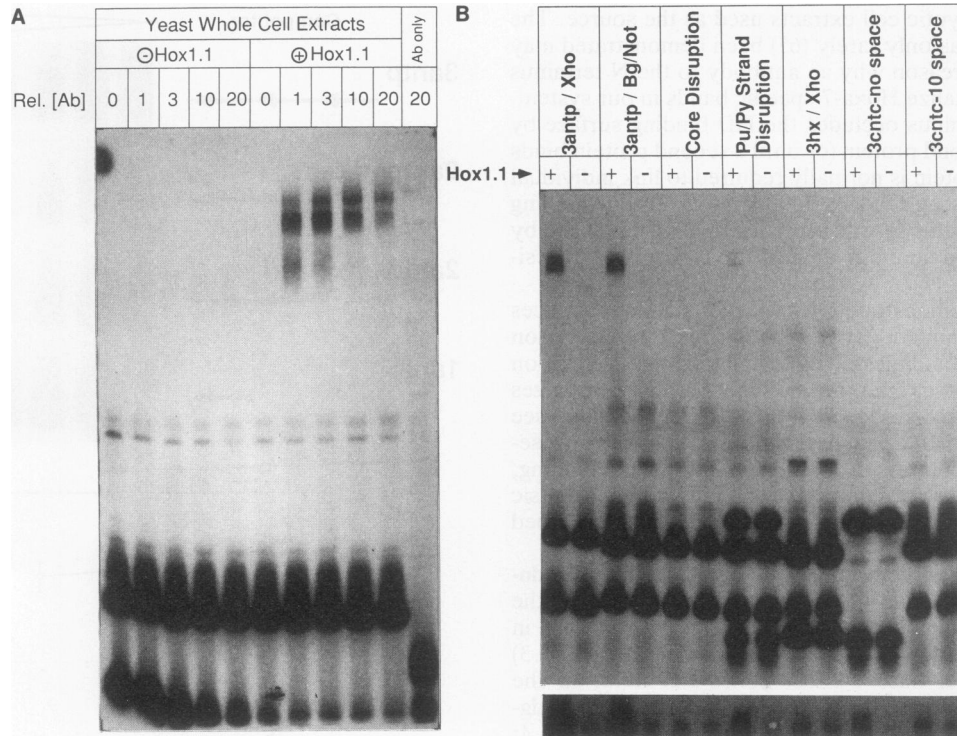


FIG. 7. Binding studies with whole cell extracts of UHA88 (+Hox) and UHC (-Hox) effector strains. (A) Labeled probe is 3antp/Bg2/NotI, and nonspecific competitor poly(dI-dC) is present at 1 μ g. Binding reactions were assembled (see Materials and Methods) by using 400 mM NaCl whole cell extracts (18 μ g of total protein) premixed with 0, 0.3, 1, 3, or 6 μ l of affinity-purified antibody preparation (same source as in Fig. 5A). (B) Labeled, band-isolated probes of comparable specific activity were used to assemble binding reactions with whole cell extracts (with or without Hox; 800 mM NaCl), 3 μ l of affinity-purified antibody (in all lanes), and 0.3 μ g of poly(dI-dC). The gel was autoradiographed for 24 h; the strip at the bottom is a 5-week exposure corresponding to the prominent Hox-specific band. Although the intensity of the Hox-specific band was reduced for the PSD and 3hx probes, the mobility is identical to that of 3antp. Heterogeneity of band-isolated probes was also observed in the absence of protein (data not shown). Designations are as for Fig. 5.

bends may occur at junctions with mixed sequence DNA because of differences in propeller twist, base pair inclination, and base stacking on the two sides of the junction. In addition, A tracts are thought to be particularly stiff because DNA has only one, rather than two, preferred roll angles (11) at the TT (or AA) base step. Both nucleosome exclusion by long dA-dT homopolymers (67) or positioning by shorter dA-dT homopolymers (50), and intrinsic curvature of DNA with phased A tracts (for a review, see reference 16) is explained by these properties (see references cited above and discussion below).

A tracts have been implicated in chromatin opening by their interaction with nuclear scaffolds and histone H1 (15, 38, 39). Interestingly, a recent review focuses on potential HD binding cores also found in elements that are involved in chromatin structuring (6).

The antp site-derived spacer (GAGAAAAG) resembles sequences which occur between the neighboring bacterial HTH binding cores in natural phage 434 operator sites (GAAAAA in O_R3 and others in [2]) and in a Hin recombinase site (GAGAAAAG [51]). A similar sequence occurs between the HD binding cores in a haploid-specific operator (AGAAAAG [19]) which binds an $\alpha 1$ - $\alpha 2$ heterodimer.

We will discuss the two working models for Hoxa-7 preference for the GAGAAAAG spacer in light of some of the unique physical/biochemical and biological properties outlined above.

Intrinsically bent DNA. Many base sequences can impart

systematic curvature to DNA, but most such effects are small compared with the special effect produced by runs of homopolymeric dA-dT base pairs (A tracts) (16). The binding of Hoxa-7 to repeated base contact cores in vitro, when they are separated by A tracts (3antp) and less well to sequences that are separated by disrupted A tracts or other spacers, suggests that binding may be sensitive to local three-dimensional curvature. Such curvature could be best maintained in vitro by A tracts but could also be generated by other sequences if they are phased appropriately. Recall that these tracts are out of reach of the HD base contact zone in the antp site. In model 1, the curvature effected by the phased A tracts results in rearrangement of the DNA structure in the HD contact zones (base or backbone) so that new amino acid side chains are brought into proximity of the DNA (candidates would be residues 43, 46, and 58 of the RH or any residue of the N-arm). In model 2, the HD contact zones remain conformationally unaltered, but a second DNA interaction domain in Hoxa-7 engages in general recognition with the A tract. General recognition typically involves recognition of the minor groove of AT-rich DNA by a variety of β -turn or extended conformations of proline-containing peptides as short as seven residues (reviewed in reference 14). Candidate sequences containing prolines flanked by hydrophilic amino acids are common in the N termini of Hox proteins.

If intrinsic curvature is sensed in vitro in the experiments described above, it does not necessarily mean that intrinsi-

cally bent sequences are required *in vivo*. Rather, DNA is typically bent around the nucleosome with certain rotational and translational preferences, based on nonrandom phasing of particular base steps which, in the minor groove, prefer to be on the inside (for example, AAA/TTT) or outside (for example, GGG/CCC) of the curve (reviewed in reference 81). The *in vitro* binding preference of Hoxa-7 to sites with flanking, stiff A tracts (and thus curvature when multimerized) may only fortuitously recapitulate the normal recognition situation. Alternatively, it may reflect the relative ability of the spacer DNA to wrap around (catabolite activator gene protein [24]) or bend between (λ repressor [33] or phage 434 repressor [35]) the monomers bound at the core motifs.

Molecular interpretation in two modes of transactivation. Two fundamentally different modes of transcriptional activation have been delineated in *S. cerevisiae* (79), one by domain activation the other by modulation of chromatin/DNA structure (topology mode). Figure 8A shows the inter-

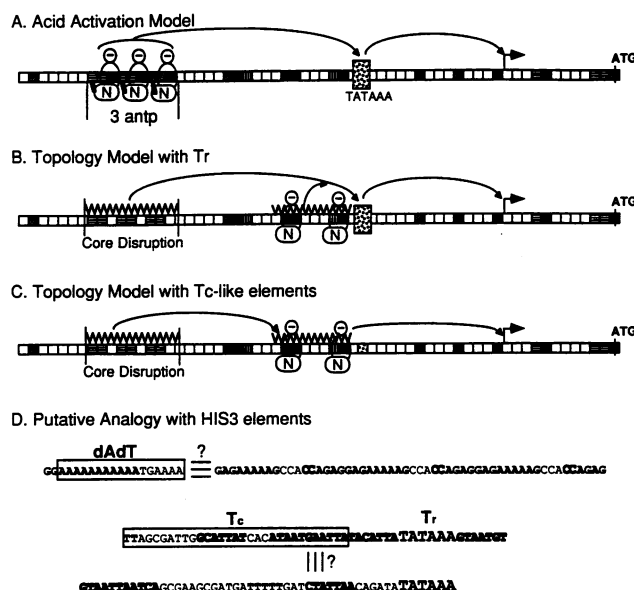


FIG. 8. Working models of transactivation. Scale diagrams show the minimal *GAL1* promoter context in which transactivation of *cis* elements was performed (Fig. 2 to 4) with 3antp (A) and triple core disruption (B and C) inserts. Each box represents a half turn of the DNA helix (four boxes = 21 bp) and is shaded according to the type of sequence element that it contains. The minimal promoter in the reporter contains 15 potential HD binding sites (core motifs) between the start codons of *GAL1* and *GAL10* (11 on the *GAL1* side are shown as stippled or black boxes). The base contact zones surrounding these cores were compared with those of binding sites obtained by screening random oligonucleotides for Ubx, Dfd (20, 21), and Ftz (23) HD binding. Four of the potential binding sites have base contact zones identical to those of high-affinity binding sites isolated in these screens (black boxes), strongly suggesting that Hoxa-7 (same class HD as Ubx, Ftz, and Dfd) could interact with these sites (models B and C). Horizontally and vertically hatched boxes represent A tracts or T tracts of 5 bp or longer. The canonical TATA element and the developing preinitiation complex are patched. (A) Transactivation depends on the affinity of Hoxa-7 for the binding sites inserted into the minimal context. In this model, the Pu/Py asymmetric spacers in 3antp potentiate binding synergy and are absent in other 3-mer sites (Fig. 3, 5, and 7). Transactivation domains interact directly or by adapters with the growing preinitiation complex to aid the initiation of transcription (arrows). A potential activation domain of Hoxa-7 is the C terminus, consisting

action of Hoxa-7 expected in the 3antp reporter context in a simple domain activation model. The amount of Hox protein bound determines the amount of transcription. The orientation and edge insensitivity (Fig. 4C) (6), the dependence on site number (Fig. 4A and B), the dependence on Hox dose (Fig. 4B), and the function in a second minimal context (Fig. 2B) are evidence that is consistent with the idea that Hoxa-7 is functioning as a classic enhancer-binding protein that mediates its orientation- and (presumably) distance-independent effect, presumably by an activation domain contacting, directly or by bridging factors, the growing preinitiation complex. The action of Hoxa-7 in this model requires a functional HD (Fig. 4) and core motif (Fig. 3).

However, several observations are inconsistent with this simple interpretation. First, mutation of the core motifs in a manner known to abolish Hoxa-7 (antp) HD binding (1) in the single or 3-mer studies (Fig. 3A and B) did not return the activity to the baseline that was measured for the vector without an insert. Rather, an intermediate level was measured in both cases. This result demonstrates that another feature of the antp site influences transactivation. Moreover, this second feature also results in Hoxa-7 responsiveness (compare YEP52 [-Hox] and UHA88 [+Hox] bars in Fig. 3B). Core disruption reporters (1- and 3-mer) gave same activity as the vector when cotransformed Ala-51 mutant effector, while the corresponding Asn-51 effector recapitulated the Hox responsiveness (data not shown). Thus, the responsiveness was absolutely dependent on the amide group of Asn-51. Second, the threefold orientation dependence of the 1antp site also does not fit into the classic enhancer description. These discrepancies could be explained by two models. (i) A second DNA binding domain in Hoxa-7 specifically binds to the spacer. This property was not recapitulated *in vitro* with whole cell extracts. (ii) The spacer sequences synergize with residual Hoxa-7 binding sites (Fig. 8B). In fact, we have selected dA-dT-like elements at the active spacer position of this minimal promoter context based on Hox responsiveness (unpublished data).

What implications do these models have for the role of

of 15 glutamic acid residues. A domain of Hoxa-7 likely to interact with the Pu/Py asymmetric AT-rich spacer is the proline-rich N terminus which contains potential β -turn motifs for general recognition. (B and C) These models of transactivation to explain how the core disruption mutants mediate a Hoxa-7-dependent transactivation (Fig. 3) even though the *cis*-acting element inserted into the silent LR1 Δ 1 reporter does not to bind Hoxa-7. These two models are based on analogy with the models for basal transcription of the *HIS3* promoter described by Struhl (79). (B) When produced, Hoxa-7 is bound at optimal base contact zones of the minimal promoter context. The residual active spacer, functioning as a dA-dT-like element, acts synergistically with the bound Hoxa-7 on the preinitiation complex forming on the canonical TATAAA box and thereby enhances transcription. Either of the two elements alone is not sufficient to give appreciable activation. (C) The activation of transcription by the dA-dT-like active spacer can be implemented only if Hoxa-7 binds to the Tc-like elements (the optimal HD sites) in minimal promoter. This model is analogous to a model in which a Tc-binding protein is thought to mediate *HIS3* basal transcription from the dA-dT element utilizing Tc-like elements. (D) The topologically functioning Tc and dA-dT elements of the *HIS3* promoter would correspond to the residual, optimal HD base contact zones and the Pu/Py asymmetric AT-rich active spacer, respectively. The zigzag lines in models B and C suggest that these elements may function by an analogous topological mechanism as the elements in the *HIS3* promoter.

Hoxa-7 in the vertebrate nuclear milieu in which it normally must exert its function and in which promoters are not compact? If Hoxa-7 functioned in the domain activation mode, it would be able to act from a distance by looping out the intervening sequences to the preinitiation complex. If Hoxa-7 only influenced or sensed topology, it could do so (i) near the initiation site, (ii) near the binding sites of domain activators in distantly located enhancers, (iii) near splice sites, or (iv) near termination sites. This much wider range of effects would be consistent with its role as a developmental control factor and would allow control at the commitment step of transcriptional initiation (48, 61) as well as at the level of basal or induced expression.

ACKNOWLEDGMENTS

We thank M. Kessel for valuable guidance in setting up the project, Hans Peter Geithe for oligonucleotide synthesis, H. D. Schmitt and the D. Gallwitz group for yeast strains and methods, the K. Weber group for help in protein chemistry, Ralf Altschäffel for photography, and Anja Faisst for help in construction of deletion mutants. We thank R. Brent, K. Yamamoto, R. Sikorski, and G. F. Merrill for sending plasmids, yeast, and bacteria.

This work was supported by postdoctoral fellowships from the Alexander von Humboldt Stiftung, the NATO/NSF, and the Max-Planck Gesellschaft to M.K.G.

REFERENCES

- 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.
- Aggarwal, A. K., D. W. Rodgers, M. Drottar, M. Ptashne, and S. C. Harrison. 1988. Recognition of a DNA operator by the repressor of phage 434: a view at high resolution. *Science* **242**:899-907.
- Amati, B., S. Dalton, M. W. Brooks, T. D. Littlewood, G. I. Evan, and H. Land. 1992. Transcriptional activation by the human c-myc oncoprotein in yeast requires interaction with Max. *Nature (London)* **359**:423-426.
- Arcioni, L., A. Simeone, S. Guazzi, V. Zappavigna, E. Boncinelli, and F. Mavilio. 1992. The upstream region of the human homeobox gene HOX3D is a target for regulation by retinoic acid and HOX homeoproteins. *EMBO J.* **11**:265-277.
- Billeter, M., Y.-Q. Quian, G. Otting, M. Müller, W. J. Gehring, and K. Wüthrich. 1990. Determination of the three-dimensional structure of the *Antennapedia* homeodomain from *Drosophila* in solution by ¹H nuclear magnetic resonance spectroscopy. *J. Mol. Biol.* **214**:183-197.
- Boulikas, T. 1992. Homeotic protein binding sites, origins of replication and nuclear matrix anchorage sites share the ATTA and ATTTA motifs. *J. Cell. Biochem.* **50**:111-123.
- Brent, R., and M. Ptashne. 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* **43**:729-736.
- Broach, J. R., Y.-Y. Li, L.-C. C. Wu, and M. Jayaram. 1983. Vectors for high-level, inducible expression of cloned genes in yeast, p. 83-117. *In* M. Inouye (ed.), *Experimental manipulation of gene expression*. Academic Press, Inc., New York.
- Buchman, A. R., and R. D. Kornberg. 1990. A yeast ARS-binding protein activates transcription synergistically in combination with other weak activating factors. *Mol. Cell. Biol.* **10**:887-897.
- Cabrera, C. V., and M. C. Alonso. 1991. Transcriptional activation by the heterodimers of the *achaete-scute* and *daughterless* gene products of *Drosophila*. *EMBO J.* **10**:2965-2973.
- Calladine, C. R., and H. R. Drew. 1986. Principles of sequence-dependent flexure of DNA. *J. Mol. Biol.* **192**:907-918.
- Chen, W., S. Tabor, and K. Struhl. 1987. Distinguishing between mechanisms of eukaryotic transcriptional activation with bacteriophage T7 RNA polymerase. *Cell* **50**:1047-1055.
- Chouinard, S., and T. C. Kaufman. 1991. Control of expression of the homeotic labial (lab) locus of *Drosophila melanogaster*: evidence for both positive and negative autogenous regulation. *Development* **113**:1267-1280.
- Churchill, M. A. E., and A. A. Travers. 1991. Protein motifs that recognize structural features of DNA. *Trends Biol. Sci.* **16**:92-97.
- Coll, M., C. A. Frederick, A. H.-J. Wang, and A. Rich. 1987. A bifurcated hydrogen-bonded conformation in the d(A.T) base pairs of the dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* **87**:2289-2293.
- Crothers, D. M. 1990. Intrinsically bent DNA. *J. Biol. Chem.* **265**:7093-7096.
- Desplan, C., J. Theis, and P. H. O'Farrell. 1988. The sequence specificity of homeodomain-DNA interaction. *Cell* **54**:1081-1090.
- Dessain, S., C. T. Gross, M. A. Kuziora, and W. M. McGinnis. 1992. Antp-type homeodomains have distinct DNA binding specificities that correlate with their different regulatory functions in embryos. *EMBO J.* **11**:991-1002.
- Dranginis, A. M. 1992. Interaction of the MATA1/MATA1alpha2 heterodimer with the operator of a haploid-specific gene. *Yeast* **8**:S65.
- 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.
- 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.
- Fitzpatrick, V. D., and C. J. Ingles. 1989. The *Drosophila* fushi tarazu polypeptide is a DNA-binding transcriptional activator in yeast cells. *Nature (London)* **337**:666-668.
- Florence, B., R. Handrow, and A. Laughon. 1991. DNA-binding specificity of the fushi tarazu homeodomain. *Mol. Cell. Biol.* **11**:3613-3623.
- Gartenberg, M. R., and D. M. Crothers. 1988. DNA sequence determinants of CAP-induced bending and protein binding affinity. *Nature (London)* **333**:824-829.
- Gibson, G., A. Schier, P. LeMotte, and W. J. Gehring. 1990. The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**:1087-1103.
- Gilmour, D. S., G. H. Thomas, and S. C. R. Elgin. 1989. *Drosophila* nuclear proteins bind to regions of alternating C and T residues in gene promoters. *Science* **245**:1487-1490.
- Gross, M. K. Unpublished data.
- Gross, M. K., and G. F. Merrill. 1988. Regulation of thymidine kinase protein levels during myogenic withdrawal from the cell cycle is independent of mRNA regulation. *Nucleic Acids Res.* **16**:11625-11643.
- Grueneberg, D. A., S. Natesan, C. Alexandre, and M. Z. Gilman. 1992. Human and *Drosophila* homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science* **257**:1089-1095.
- Han, K., M. S. Levine, and J. L. Manley. 1989. Synergistic activation and repression of transcription by *Drosophila* homeobox proteins. *Cell* **56**:573-583.
- 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.
- Hanes, S. D., and R. Brent. 1991. A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**:426-430.
- Hayashi, S., and M. P. Scott. 1990. What determines the specificity of action of *Drosophila* homeodomain proteins? *Cell* **63**:883-894.
- 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.
- Hoey, T., and M. Levine. 1988. Divergent homeobox proteins recognize similar DNA sequences in *Drosophila*. *Nature (London)* **332**:858-861.
- Hogan, M. E., and R. H. Austin. 1987. Importance of DNA

- stiffness in protein-DNA binding specificity. *Nature (London)* **329**:263-266.
36. Kalb, V. F., Jr., and R. W. Bernlohr. 1977. A new spectrophotometric assay for protein in cell extracts. *Anal. Biochem.* **82**:362-371.
 37. Kamens, J., and R. Brent. 1991. A yeast transcription assay defines distinct Rel and Dorsal DNA recognition sequences. *New Biol.* **3**:1005-1013.
 38. Käs, E., E. Izaurralde, and U. K. Laemmli. 1989. Specific inhibition of DNA binding to nuclear scaffolds and histone H1 by distamycin. *J. Mol. Biol.* **210**:587-599.
 39. Käs, E., L. Poljak, Y. Adachi, and U. K. Laemmli. 1993. A model for chromatin opening: stimulation of topoisomerase II and restriction enzyme cleavage of chromatin by distamycin. *EMBO J.* **12**:115-126.
 40. Keleher, C. A., S. Passmore, and A. D. Johnson. 1989. Yeast repressor $\alpha 2$ binds to its operator cooperatively with yeast protein Mcm1. *Mol. Cell. Biol.* **9**:5228-5230.
 41. Kessel, M., U. Drescher, and P. Gruss. 1987. The murine homeo domain protein Hox 1.1. *Ann. N.Y. Acad. Sci.* **511**:88-100.
 42. Kissinger, C. R., B. Liu, E. Martin-Bianco, T. B. Kornberg, and C. O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. *Cell* **63**:579-590.
 43. Ko, H. S., P. Fast, W. McBride, and L. M. Staudt. 1988. A human protein specific for the immunoglobulin octamer DNA motif contains a functional homeobox domain. *Cell* **55**:135-144.
 44. 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.
 45. Kuziora, M. A., and W. McGinnis. 1989. A homeodomain substitution changes the regulatory specificity of the deformed protein in *Drosophila* embryos. *Cell* **59**:563-571.
 46. Laughon, A. 1991. DNA binding specificity of homeodomains. *Biochemistry* **30**:11357-11367.
 47. Lech, K., K. Anderson, and R. Brent. 1988. DNA-bound Fos proteins activate transcription in yeast. *Cell* **52**:179-184.
 48. Lewin, B. 1990. Commitment and activation at PolII promoters: a tail of protein-protein interactions. *Cell* **61**:1161-1164.
 49. Lin, L., and W. McGinnis. 1992. Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev.* **6**:1071-1081.
 50. Losa, R., and F. Thoma. 1990. Poly(dA).poly(dT) rich sequences are not sufficient to exclude nucleosome formation in a constitutive yeast promoter. *Nucleic Acids Res.* **18**:3495-3502.
 51. Mack, D. P., J. P. Sluka, J. A. Shin, J. H. Griffin, M. I. Simon, and P. B. Dervan. 1990. Orientation of the putative recognition helix in the DNA-binding domain of Hin recombinase complexed with the Hix site. *Biochemistry* **29**:6561-6567.
 52. Malicki, J., K. Schughart, and W. McGinnis. 1990. Mouse *Hox-2.2* specifies thoracic segmental identity in *Drosophila* embryos and larvae. *Cell* **63**:961-967.
 53. Mann, R. S., and D. S. Hogness. 1990. Functional dissection of Ultrabithorax proteins in *Drosophila melanogaster*. *Cell* **60**:597-610.
 54. McGinnis, N., M. A. Kuziora, and W. McGinnis. 1990. Human HOX4.2 and *Drosophila deformed* encode similar regulatory specificities in *Drosophila* embryos and larvae. *Cell* **63**:969-976.
 55. McGinnis, W., and R. Krumlauf. 1992. Homeobox genes and axial patterning. *Cell* **68**:283-302.
 56. Metzger, D., J. H. White, and P. Chambon. 1988. The human oestrogen receptor functions in yeast. *Nature (London)* **334**:31-36.
 57. Müller, M., M. Affolter, W. Leupin, G. Otting, K. Wüthrich, and W. J. Gehring. 1988. Isolation and sequence-specific DNA binding of the Antennapedia homeodomain. *EMBO J.* **7**:4299-4304.
 58. Nawaz, Z., M.-J. Tsai, D. P. McDonnell, and B. W. O'Malley. 1992. Identification of novel steroid-response elements. *Gene Expression* **2**:39-46.
 59. Nelson, H. C. M., J. T. Finch, B. F. Luisi, and A. Klug. 1987. The structure of an oligo(dA).oligo(dT) tract and its biological implications. *Nature (London)* **330**:221-226.
 60. Odenwald, W. F., J. Garbern, H. Arnheiter, E. Tournier-Lasserre, and R. A. Lazzarini. 1989. The Hox-1.3 homeo box protein is a sequence-specific DNA-binding phosphoprotein. *Genes Dev.* **3**:158-172.
 61. Ohkuma, Y., M. Horikoshi, R. G. Roeder, and C. Desplan. 1990. Engrailed, a homeodomain protein, can repress in vitro transcription by competition with the TATA box-binding protein transcription factor IID. *Proc. Natl. Acad. Sci. USA* **87**:2289-2293.
 62. Otting, G., Y. Q. Qian, M. Billeter, M. Müller, M. Affolter, W. J. 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.
 63. Pabo, C. O., and R. T. Sauer. 1992. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**:1053-1095.
 64. Ponticelli, A. S., and K. Struhl. 1990. Analysis of *Saccharomyces cerevisiae his3* transcription in vitro: biochemical support for multiple mechanisms of transcription. *Mol. Cell. Biol.* **10**:2832-2839.
 65. Pöpperl, H., and M. S. Featherstone. 1992. An autoregulatory element of the murine Hox-4.2 gene. *EMBO J.* **11**:3673-3680.
 66. Privalsky, M. L., M. Sharif, and K. R. Yamamoto. 1990. The viral erbA oncogene protein, a constitutive repressor in animal cells, is a hormone-regulated activator in yeast. *Cell* **63**:1277-1286.
 67. Prunell, A. 1982. Nucleosome reconstitution on plasmid-inserted poly(dA).poly(dT). *EMBO J.* **1**:173-179.
 68. 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.
 69. Quian, Y. Q., G. Otting, K. Furukubu-Tokunaga, M. Affolter, W. J. Gehring, and K. Wüthrich. 1992. NMR structure determination reveals that the homeodomain is connected through a flexible linker to the main body in the *Drosophila* Antennapedia protein. *Proc. Natl. Acad. Sci. USA* **89**:10738-10742.
 70. Rütger, U., and B. Müller-Hill. 1983. Easy identification of cDNA clones. *EMBO J.* **2**:1791-1794.
 71. 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.
 72. Schägger, H., and G. von Jagow. 1987. Tricine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368-379.
 73. Schena, M., and K. R. Yamamoto. 1988. Mammalian glucocorticoid receptor derivatives enhance transcription in yeast. *Science* **241**:965-967.
 74. Sikorski, R. S., and P. Hieter. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**:19-27.
 75. Slack, J. M. W., and D. Tannahill. 1992. Mechanisms of anteroposterior axis specification in vertebrates. *Development* **114**:285-302.
 76. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase. *Gene* **67**:31-40.
 77. Struhl, K. 1985. Naturally occurring poly(dA-dT) sequences are upstream promoter elements for constitutive transcription in yeast. *Proc. Natl. Acad. Sci. USA* **82**:8419-8423.
 78. Struhl, K. 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. *Mol. Cell. Biol.* **6**:3847-3853.
 79. Struhl, K. 1987. Promoters, activator proteins, and the mechanism of transcriptional initiation in yeast. *Cell* **49**:295-297.
 80. Struhl, K. 1988. The *jun* oncoprotein, a vertebrate transcription factor, activates transcription in yeast. *Nature (London)* **332**:649-650.
 81. Travers, A. A. 1991. DNA bending and kinking—sequence dependence and function. *Curr. Opin. Struct. Biol.* **1**:114-122.
 82. Treisman, R., R. Marias, and J. Wynne. 1992. Spatial flexibility

- in ternary complexes between SRF and its accessory proteins. *EMBO J.* **11**:4631–4640.
83. Ullman, A. 1984. One-step purification of hybrid proteins which have β -galactosidase activity. *Gene* **29**:27–31.
84. Vershon, A. K., and A. D. Johnson. 1993. A short, disordered protein region mediates interactions between the homeodomain of the yeast α 2 protein and the MCM1 protein. *Cell* **72**:105–112.
85. 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 UASg. *Mol. Cell. Biol.* **4**:2467–2478.
86. Wilson, T. E., T. J. Fahrner, M. Johnston, and J. Milbrandt. 1991. Identification of the DNA binding site for NGFI-B by genetic selection in yeast. *Science* **252**:1296–1300.
87. Winslow, G. M., S. Hayashi, M. Krasnow, D. S. Hogness, and M. P. Scott. 1989. Transcriptional activation by the Antennapedia and fushi tarazu proteins in cultured *Drosophila* cells. *Cell* **57**:1017–1030.
88. Winter, E., and A. Varshavsky. 1989. A DNA binding protein that recognizes oligo(dA).oligo(dT) tracts. *EMBO J.* **8**:1867–1877.
89. Wolberger, C., A. K. Vershon, B. Liu, A. D. Johnson, and C. O. Pabo. 1991. Crystal structure of a MAT α 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* **67**:517–528.
90. Yoon, C., G. G. Prive, D. S. Goodsell, and R. E. Dickerson. 1988. Structure of an alternating-B DNA helix and its relationship to A-tract DNA. *Proc. Natl. Acad. Sci. USA* **85**:6332–6336.
91. Zappavigna, V., A. Renucci, J.-C. Izpisua-Belmonte, G. Urier, C. Peschle, and D. Duboule. 1991. Hox4 genes encode transcription factors with potential auto- and cross-regulatory capacities. *EMBO J.* **10**:4177–4187.
92. Zervos, A. S., J. Gyuris, and R. Brent. 1993. Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites. *Cell* **72**:223–232.