

Targeted Chemical Wedges Reveal the Role of Allosteric DNA Modulation in Protein–DNA Assembly

Rocco Moretti^{†,¶}, Leslie J. Donato^{†,¶}, Mary L. Brezinski^{†,¶}, Ryan L. Stafford[‡], Helena Hoff[‡], Jon S. Thorson[§], Peter B. Dervan[‡], and Aseem Z. Ansari^{†,*}

[†]Department of Biochemistry and The Genome Center, University of Wisconsin-Madison, Madison, Wisconsin 53706, [‡]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, and [§]School of Pharmacy, University of Wisconsin-Madison, Madison, Wisconsin 53705, [¶]These authors contributed equally to this study.

ABSTRACT The cooperative assembly of multiprotein complexes results from allosteric modulations of DNA structure as well as direct intermolecular contacts between proteins. Such cooperative binding plays a critical role in imparting exquisite sequence specificity on the homeobox transcription factor (Hox) family of developmental transcription factors. A well-characterized example includes the interaction of Hox proteins with extradenticle (Exd), a highly conserved DNA binding transcription factor. Although direct interactions are important, the contribution of indirect interactions toward cooperative assembly of Hox and Exd remains unresolved. Here we use minor groove binding polyamides as structural wedges to induce perturbations at specific base steps within the Exd binding site. We find that allosteric modulation of DNA structure contributes nearly 1.5 kcal/mol to the binding of Exd to DNA, even in the absence of direct Hox contacts. In contrast to previous studies, the sequence-targeted chemical wedges reveal the role of DNA geometry in cooperative assembly of Hox–Exd complexes. Programmable polyamides may well serve as general probes to investigate the role of DNA modulation in the cooperative and highly specific assembly of other protein–DNA complexes.

*Corresponding author,
ansari@biochem.wisc.edu.

Received for review December 18, 2007
and accepted February 17, 2008.

Published online April 18, 2008

10.1021/cb700258r CCC: \$40.75

© 2008 American Chemical Society

The physical basis of DNA sequence recognition by its ligands (proteins or small molecules) has been investigated in great detail over the years, and several underlying principles have been defined (1–3). The main contributors to the specificity of DNA sequence recognition are interactions between protein side chains and the edges of the base pairs in the cognate DNA site. Efforts to rationally redesign these direct interactions have yielded some exciting successes (4–8). A major limitation of focusing on re-engineering direct interactions, however, is that recognition of sequence-dependent structural properties of DNA and other modes of *indirect readout* are also critical specificity determinants (9–15). Often a combination of direct and indirect interactions governs exquisite DNA sequence specificity displayed by its ligands (9, 11, 14–22). Specificity in molecular recognition is also achieved by cooperative binding of multiple proteins to closely appositioned DNA sequences (23). Binding of one protein to its DNA cognate site in the enhancer can alter DNA structure and conformational flexibility of a juxtaposed site, thus allosterically improving the energetics for the binding of the next protein (11, 15, 18, 20). However, because of the concomitant protein–protein interactions between binding proteins and the difficulty in truly dissecting the contribution of the DNA modulation to cooperative assembly, the latter effect is often overlooked.

Here we explore the role of DNA in homeobox transcription factor (Hox)–extradenticle (Exd) function. Extradenticle (Exd) is a transcription factor that plays a vital role in animal development by guiding the Hox family

of developmental master regulators to unique DNA binding sites (24, 25). The apparent similarity of consensus sequences bound by different members of the Hox family *in vitro* has led to intensive examination of the source of the exquisite regulatory specificity of the Hox proteins *in vivo* (24, 26, 27). Several lines of evidence suggest that DNA sequence selectivity of different Hox proteins is influenced by cooperative binding with Exd (24, 28–35). Most of this cooperative interaction is the result of a conserved N-terminal YPWM tetrapeptide on the Hox protein binding into a pocket on the Exd/Pbx1 homeodomain (Figure 1, panel a). In addition to this direct interface between Hox and Exd/Pbx1, crystal structures of four Hox–Exd/Pbx1–DNA complexes reveal a discernible perturbation of DNA geometry in the ternary complex (32–35). Importantly, there is compelling evidence that Hox–Exd–DNA complex may not require interaction between the Hox N-terminal docking peptide and Exd (28, 36). In this ternary complex, it is possible that DNA structural modulation, along with the coupled decrease in conformational entropy of the binding site due to binding of one partner, enhances DNA binding by the other partner protein. However, separating the allosteric DNA contributions from the direct protein–protein contributions toward cooperative binding has been difficult.

To probe the contribution of DNA conformation to DNA binding by protein complexes, we used minor groove binding small molecules as sequence-specific structural wedges. Polyamides subtly perturb DNA structure, expanding the minor groove width and decreasing the conformational flexure of the bound DNA site (37). Direct steric occlusion of minor groove binding proteins by polyamides is not surprising; however, allosteric inhibition of proteins that bind to the obverse major groove has also been demonstrated (ref 37 and references therein). Hairpin polyamides are typically composed of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) heterocycles connected by amide bonds. They have a rigid crescent shape that fits into the minor groove and tracks the curvature and rise of DNA for nearly a half turn (38). When presented in a stacked hairpin format, these molecules display affinities that rival natural DNA binding proteins, have high DNA sequence selectivity, and by using specific combinations of heterocycles can be engineered to target any of the four base pairs (38).

Here, we examine the contribution of polyamide-induced DNA structural perturbations to the binding of Exd to a naturally occurring Exd cognate site. In applying these engineered molecules as sequence-specific structural wedges, we find that DNA modulation contributes ~1.5 kcal/mol to the binding of Exd to its cognate site. The data also suggest that polyamides can be used as a general structural tool to probe the contribution of DNA structure toward cooperative assembly of other protein–DNA complexes. The dissection of both direct and allosteric forces that govern sequence specificity of DNA binding molecules will greatly aid in understanding the principles of molecular recognition. Eventually, these principles will enable the creation of synthetic molecules that can target desired sites in the genome with exquisite precision.

RESULTS AND DISCUSSION

Cooperative Binding of Hox and Exd to a Cognate

DNA Site. To examine the role of molecular interactions that mediate cooperative DNA binding we focused on a *Drosophila melanogaster* Hox–Exd complex. Ultrathorax (Ubx) is one of the best-characterized members of the Hox family of transcription factors, and it plays a pivotal role in fly development (39–41). The crystal structure of the Ubx–Exd–DNA complex reveals that the Trp residue of the Ubx/Hox docking peptide is buried in Exd (Figure 1, panel a) (32). Substitution of this Trp residue greatly compromises cooperative assembly of Ubx/Hox–Exd complexes on target DNA sites (30, 31, 42). We therefore generated a Ubx mutant lacking the critical Trp residue as well as three additional flanking residues to probe the extent to which the remaining direct or indirect interactions contribute to cooperative binding with Exd (Figure 1, panel b). The resulting protein retains its ability to bind DNA efficiently on its own (see Supplementary Figure 1). The DNA binding modules of the wild-type and mutant proteins were purified, and their affinity, individually or in cooperative pairs with Exd, for the cognate site was determined by electrophoretic mobility shift assays (EMSAs) (Figure 1, panels c–e).

Both wild-type and mutant Ubx homeodomains bind DNA with similar affinities; however, the mutant polypeptide was compromised in its ability to form a stable cooperative complex with Exd (Figure 1, panel d). The affinity of Exd for its cognate 5′TGAT3′ site increases by 2–3 orders of magnitude in the presence of Ubx bearing an intact YPWM docking peptide (compare panel c

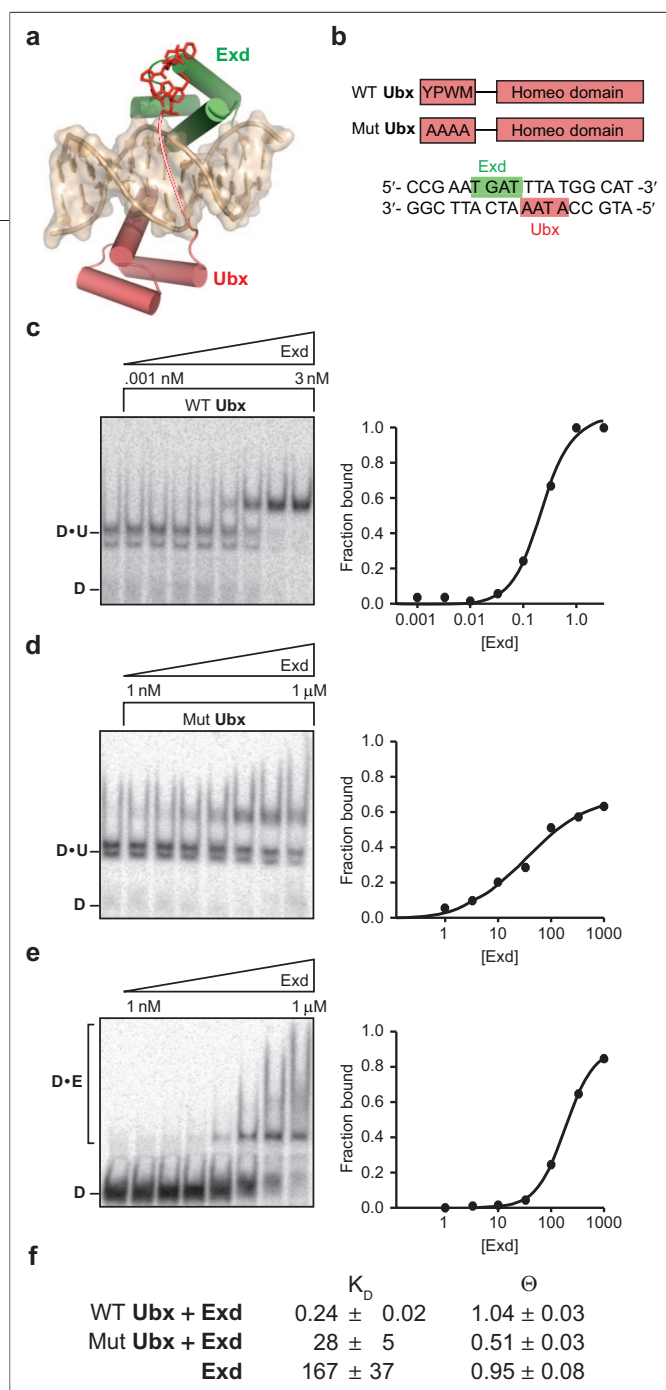


Figure 1. Contributions of direct and indirect interactions to Ubx–Exd ternary complex formation. **a)** The crystal structure of the Ubx–Exd–DNA ternary complex, showing the relative binding orientation of the two proteins. The N-terminal portion of Ubx reaches across the DNA, contacting Exd and stabilizing the ternary complex. **b)** Schematic of wild-type and mutant Ubx proteins and the DNA sequence used for Ubx–Exd–DNA interactions. Wild-type Ubx contains the YPWM interaction peptide, while the peptide is mutated to alanines in the mutant Ubx. **c–e)** The ability of Ubx to recruit Exd to DNA was evaluated *via* EMSA (in 0.5x TBE). Both wild-type Ubx (**c**) and the interaction peptide mutant (**d**) were able to improve binding of Exd over the Ubx free condition (**e**). Note the lowered Exd concentration range used for wild-type Ubx. **f)** The dissociation constant (K_D) and the fraction occupancy at projected saturation (Θ) of each condition was determined by fitting the fraction DNA bound values obtained by EMSA to the Hill binding equation. Wild-type Ubx improves Exd binding by 800-fold, 100-fold of which is due to the YPWM interaction peptide.

to panel e in Figure 1). The mutant version of Ubx lacking these four residues shows a 100-fold decrease in its ability to cooperatively stabilize Exd binding to its cognate site (Figure 1, panels c and d). Intriguingly, mutant Ubx was still capable of enhancing the affinity of Exd for its site by 7–10-fold (Figure 1, panels d and e). This residual enhancement of Exd–DNA interaction suggests a role for DNA modulations in recruitment of Exd to its cognate site. It should be noted that at high concentrations Exd appears to form nondistinct, lower mobility complexes, leading to “streaking” on the gel (Figure 1, panel d). These complexes disappear in the presence of Ubx and a distinct ternary complex is apparent.

DNA Groove Dimensions in Hox–Exd/Pbx1

Structures. As a first step toward understanding the role of DNA structure in Exd–DNA interactions, we examined the groove dimensions of DNA in four crystal structures of *Drosophila* and human Hox–Exd/Pbx1–DNA complexes (Figure 2) (32–35). We used CURVES, an algorithm designed to extract the measure of DNA groove dimensions and nucleobase stacking geometries (43–45). CURVES models the DNA backbone with a cubic spline interpolation of phosphate positions, allowing for continuous groove measurements throughout most of the structured DNA in the complex, albeit this treatment does not define groove dimensions at the DNA ends. We compared groove dimensions of the published structures with those of idealized B-form DNA (Figure 2, in gray). The analysis indicates that the major groove at the core GA base step of the 5' TGAT^{3'} Exd/Pbx1 binding site is wider than typical B-form DNA (Figure 2, panel a). In parallel, the depth of the major groove increases, likely to accommodate the recognition helix of Exd/Pbx1 (Figure 2, panel b). Remarkably, the smooth transitions in groove dimensions over the Exd/Pbx1 binding site are abruptly interrupted at the critical adenine residue. CURVES analysis report that the A-T basepair is dramatically perturbed and may be displaced toward the major groove, thereby decreasing the depth of this groove. The authors of CURVES note that such sudden decreases in groove depth are indicative of a distorted helix (45). To better understand the nature of this perturbation, we aligned the idealized B-form DNA with the DNA structures obtained from four crystal structures of Hox–Exd/Pbx1 complexes. For the sake of clarity we only display one strand of the duplex (Figure 2, panel c). In all four structures the trajectory of the DNA backbone deviates from B-form DNA at the core GA base

step within the Exd/Pbx1 binding site. This base step is critical for Exd/Pbx1 binding as observed in the crystal structure (32–35) as well as our own comprehensive CSI microarray studies (46).

Design and Synthesis of the Minor Groove Wedges.

To probe the contribution of indirect DNA structural perturbations toward Exd–DNA complex formation, we used minor groove binding hairpin polyamides. Although the subtle structural perturbations of DNA induced by polyamides are not sufficient to disrupt nucleosome formation (47) or binding by bZIP transcription factors (48, 49), the perturbations inflict a 13–35-fold decrease in affinity of zinc fingers for overlapping sites, even though there is no steric overlap between the minor groove binding polyamides and the major groove binding zinc fingers (37). We modeled the binding of polyamide and Exd on an overlapping site and found no detectable source of direct intermolecular interactions (Figure 3, panel a). Previous measurements of polyamide–DNA complexes show increases in minor groove width of 1–4 Å, with the extent of deformation varying along the length of the polyamide (48). Thus, these synthetic molecules could function as chemical wedges to probe the contribution of DNA structural modulations toward Exd binding.

We synthesized an Im,Im,Py,Py-γ-Py,Py*,Py,Py-βala-Dp polyamide designed to target a physiological Hox–Exd composite site (Figure 3, panel a; Py* denotes the pyrrole with the pendant linker, for affinity measurements by quantitative footprinting, see Supplementary Figure 2 and Supplementary Table 1). Examining the entire sequence space of a typical binding site by cognate site identifier (CSI) microarray studies (46), we find that this hairpin polyamide targets a 5'WWGGTWW3' site (W = A/T), consistent with polyamide pairing rules (Warren *et al.*, unpublished work). To determine the contribution of the YPWM Hox peptide in the enhancement of Exd binding to the physiological site, we conjugated this peptide to the polyamide at an internal pyrrole *via* a ~7 Å linker (1). As a test of specificity we also synthesized a control molecule lacking the key Trp and the flanking Met residues of the Hox YPWM peptide (2). Similar to what is shown for the natural proteins in Figure 1, panel d, the “mutated” conjugate (2) should not enhance Exd–DNA interactions. The active and inactive forms of the Hox docking peptide were conjugated to the polyamide *via* a propylamine linker at an internal pyrrole, a conformation that would display the peptide

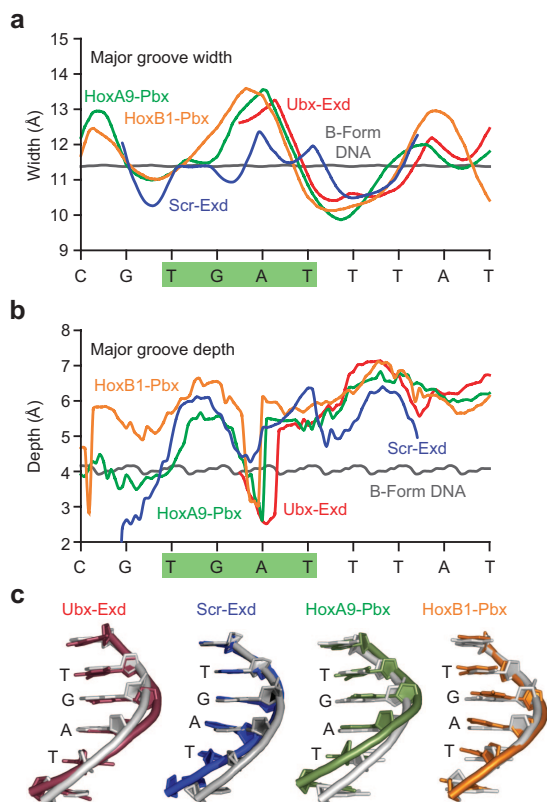


Figure 2. Structural analysis of Hox–TALE–DNA complexes. a,b) CURVES analysis of published Hox–Exd/Pbx1–DNA complexes, compared to B-form DNA. The DNA sequence for the Ubx–Exd complex is given, with the Exd/Pbx1 binding site highlighted. The curve for Ubx–Exd does not extend to the end of the binding site due to limitations of the CURVES algorithm close to the ends of the shorter Ubx–Exd DNA. a) The major grooves of the ternary complexes (color-coded) are wider than the B-form DNA (in gray) over the center of the Exd/Pbx1 binding site, transitioning to narrower as one enters the first few nucleotides of the Hox binding site. b) The major grooves of the complexes are generally deeper than those of the B-form, likely to accommodate the DNA recognition helices of the proteins. Of note is the sudden, sharp decrease in groove depth at the central adenine of the Exd/Pbx1 binding site, indicative of a structural deformation at that location. c) Alignment of the four DNAs against B-form (gray). Only a single strand is shown for clarity. All four show a bend away from B-form structure at the central 5'GA3' base step of the Exd/Pbx1 binding site.

over the minor groove and deliver the Trp residue into the docking site on Exd (46, 50–52). Conjugate 3 is the parent polyamide containing only the propylamine linker at the internal pyrrole. Finally, an additional polyamide lacking the propylamine linker was also synthe-

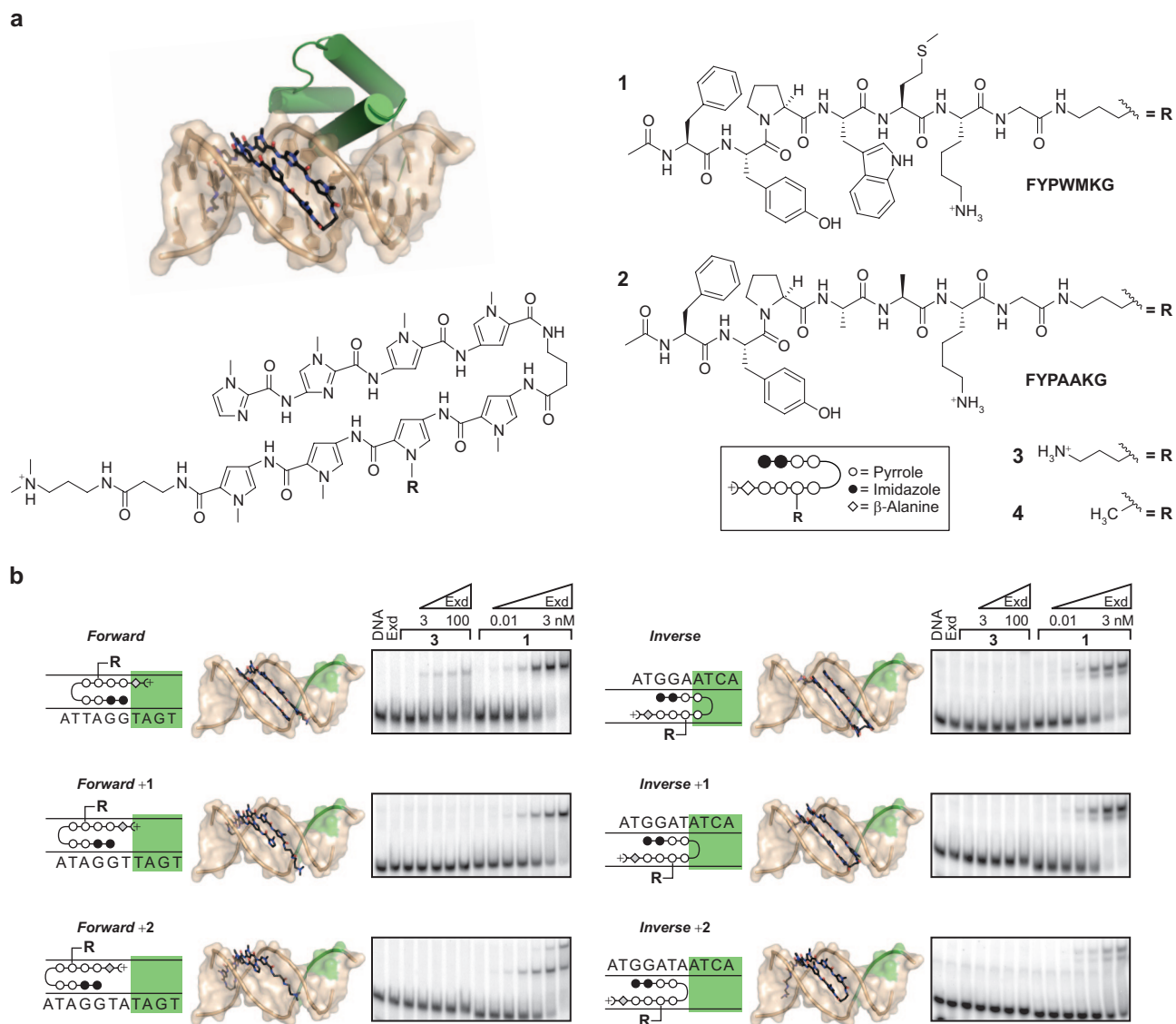


Figure 3. Ternary complex assembly by small molecules. **a)** Polyamides bind to the minor groove and are known to increase the width and depth of the minor groove in comparison to B-form DNA. An Im,Im,Py,Py- γ -Py,Py*,Py,Py- β -ala-Dp polyamide can bind site specifically to the minor groove of DNA at a physiological site. In a model of joint binding, there are no steric contacts between the minor groove bound polyamide and the major groove bound Exd (green). The four compounds used in this study are a YPWM Exd interaction peptide (1) or a nonfunctional derivative (2) conjugated to the polyamide *via* a propylamine linker, which was also present in the unconjugated form (3). In later experiments the linker and peptide was replaced by a methyl group (4). **b)** The ability of the compounds to recruit Exd to different sites was qualitatively assessed by EMSA (in 1x TBE). The effect of the compounds varied on the basis of the relative orientation and position of the polyamide and Exd binding sites (Exd binding site highlighted). The polyamide displaying the Exd interaction peptide (1) was able to recruit Exd to DNA in all binding orientations tested. The compound with no peptide (3) was still able to recruit Exd to DNA when properly oriented with respect to the Exd binding site (the *Forward* orientation).

sized to eliminate adventitious electrostatic interactions between the positively charged linker amine and the negative phosphodiester backbone (4). This set of

compounds was used in exploring the role of DNA structural perturbations in cooperative binding of Exd to its site.

Polyamide Binding Register and Its Effect on Exd Binding.

The type and extent of DNA structural alteration by polyamides on the Exd binding site may depend greatly on the relative orientation and separation of the two binding sites. Thus, we designed composite DNA sites where the polyamide binding site ($5'$ WWGGW-WW $3'$) is positioned in various registers with respect to the Exd binding site ($5'$ TGAT $3'$), depicted in green in Figure 3, panel b). In the *Forward* orientation the polyamide is positioned such that the positively charged dimethylpropylamine (Dp) tail lies in the minor groove of the Exd binding site (Figure 3, panel b). The positive charge is in close proximity to the exocyclic amine of the guanine residue of the $5'$ TGAT $3'$ site. In this orientation the polyamide would widen the minor groove at the very edge of the Exd binding site while delivering the β -alanine and Dp tail within the cognate site minor groove that is compressed by Exd binding. Polyamide binding in this orientation will not perturb minor groove interactions by the N-terminal arm of Exd (Arg 5 contact with the distal T residue of the Exd site). Remarkably, the parent polyamide (**3**) lacking the YPWM docking peptide enhances Exd binding to its site (Figure 3, panel b). Consistent with this result the FYPAAKG-polyamide conjugate (**2**) is also capable of enhancing Exd binding at this site (Supplementary Figure 3). The polyamide conjugated to the functional FYPWMK peptide (**1**) is most active and it strongly enhances Exd binding to the composite site (Figure 3, panel b; note lower concentration range for the active conjugate).

Whereas the polyamide conjugated to an active Hox peptide (**1**) is able to enhance Exd binding in various orientations and registers, this ability does not extend to the parental polyamide (**3**) or the polyamide conjugated to the peptide lacking the critical Trp residue (**2**). It is im-

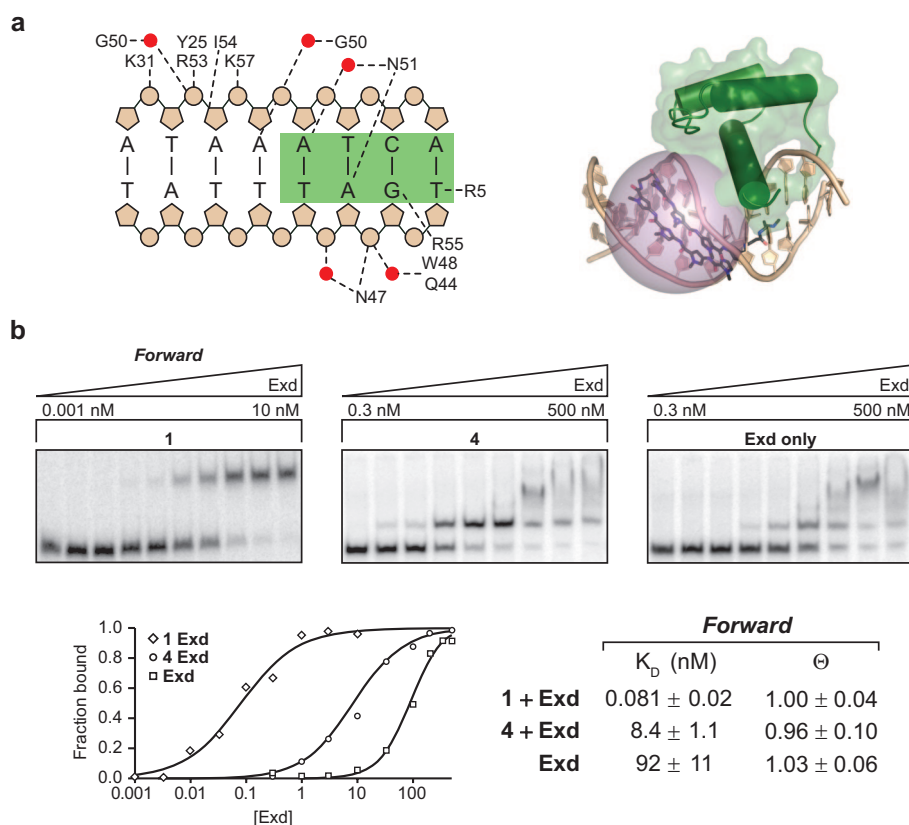


Figure 4. Contribution of direct and indirect interactions to polyamide–Exd ternary complex formation. **a**) Exd makes a number of contacts to DNA, including many to the backbone phosphates (figure redrawn from ref 32). These contacts have the potential to be disrupted by the binding of the positively charged propylamine of compound **3** (potential range shown as purple sphere) The propylamine linker may make direct or water-mediated contacts with the DNA or protein. **b**) Quantitation of Exd binding to *Forward* DNA by optimized (0.5x TBE) EMSA. Although Exd alone shows binding at high concentrations, the presence of **1** and **4** improves the Exd–DNA interaction. The fraction of bound DNA (including the nonspecific, slower migrating species) was calculated for each Exd concentration, plotted, and fit to the Hill binding equation. The dissociation constant (K_d) and the fraction occupancy at saturation (θ) of each condition are shown. In the *Forward* binding orientation, addition of **4** improves the binding of Exd 10-fold. The addition of the Exd interaction peptide improves Exd recruitment a further 100-fold.

portant to note that in the *Inverse* orientation the polyamide γ turn and a Py ring pair are positioned within the Exd binding site, yet no enhancement of Exd binding is observed. The structures indicate that the minor groove in the Ubx–Exd bound structure is sufficiently wide over these base steps to accommodate binding by the stacked Py rings and the γ turn of the hairpin polyamide (see Results and Discussion and Figure 2). It is possible that simply increasing groove width or decreasing conformational entropy is insufficient for enhancing the cooperative association of Exd with its site.

Quantifying the Extent of Binding Afforded by Allosteric Perturbation. A closer examination of the contacts between Exd and the phosphate backbone within the Hox site led us to question whether the terminal amine of the propylamine linker on the parent polyamide (**3**) would disrupt these Exd interactions. The enhancement of Exd binding by the parent polyamide lacking the Hox docking peptide (**3**) hints at the contribution of allostery in cooperative assembly of Hox–Exd–DNA complexes. However, Exd makes a number of contacts to the phosphate backbone, and the positively charged propylamine linker of the parent polyamide (**3**) may disrupt these contacts, either directly or through water-mediated interactions with the same phosphates (Figure 4, panel a). The propylamine linker on the polyamide is also physically capable of direct or indirect interactions with Exd itself (Figure 4, panel a). To eliminate this potential source of interactions, a polyamide lacking the propylamine linker was synthesized (compound **4** in Figure 3). The “linker-less” polyamide should not perturb DNA backbone contacts made by Exd nor should it contact the protein itself. In the absence of the propylamine linker, the polyamide (**4**) further enhanced Exd binding to the *Forward* composite site.

The extent to which polyamide **4** enhanced Exd binding was quantified using EMSA. In the absence of a natural or synthetic Hox partner, Exd binds to its site with extremely low affinity. More vexingly, the Exd rapidly aggregates on the DNA at concentrations where initial binding to the TGAT site is detected. We optimized electrophoresis conditions to best detect stoichiometric binding between Exd and its cognate site (see Methods for details). Under these conditions we measured the contribution of polyamide **4** to the recruitment of Exd to its site (Figure 4, panel b). In the *Forward* orientation we find that simply docking the polyamide in an overlapping site enhanced Exd binding to its site by ~11-fold (Figure 4, panel b). This effect requires the appropriate register of the Exd cognate site and simply binding a polyamide to DNA at a distance from a functional Exd site does not improve Exd binding (data not shown). The presence of the Hox docking peptide (**1**) further enhances Exd binding by ~100-fold. Together using allosteric as well as direct interactions the FYPWMK-polyamide conjugate (**1**) enhances Exd binding by >3 orders of magnitude. The increase is very much within the

realm of the natural Hox–Exd complex at the relevant cognate site (Figure 1).

Rationale for the Allosteric Effect. While the efforts to re-engineer protein–DNA interactions have focused on side chains that contact DNA, several other forces contribute to specificity in molecular recognition. Indirect and direct interactions combined with the inherent microstructural properties of the DNA sequence are responsible for the overall affinity and sequence specificity of DNA ligands (proteins and small molecules) (9–22). Such specificity is critical for targeting 2–4 nm sized cognate sites in metazoan genomes that are a billion times larger. Despite such daunting requirements for specific recognition of a target site, metazoan transcription factors display limited sequence specificity, binding short DNA sequences (4–10 bp) and tolerating significant deviations from the optimal DNA binding sequence (13, 23). The Hox family of transcription factors is a particularly important example. Each Hox protein regulates genes that define different, specific, anterior–posterior patterns in most metazoan organisms (24, 25, 27, 39–41, 53–55). Yet *in vitro* each member of the Hox family binds a common consensus site with near identical affinity (24, 26, 27). Thus, the exquisite *in vivo* specificity of such metazoan transcription factors must rely on additional specificity determinants. In the case of Hox proteins it is argued that cooperative assembly with Exd and other transcription factors plays a critical role in targeting different genes in different segments of the organism (28, 29). A short YPWM peptide conserved in all Hox proteins interacts weakly with Exd/Pbx1 to promote cooperative binding to subtly different sites (24, 28–35). The importance of such interaction has been questioned as a result of data where cooperative binding to DNA was observed between Hox–Exd proteins even in the absence of the YPWM peptide (28, 36, 41). In addition, it was recently shown that Hox–Exd can recognize the sequence-dependent structure of its minor groove binding site (35). The implications of these data are that alternate direct contacts between proteins, DNA structure, and perhaps allosteric modulation of DNA play a role in cooperative assembly of the ternary complex at subtly different Hox–Exd sites.

We employed polyamides to dissect the energetic contributions of the Hox YPWM peptide and the allosteric modulation of the DNA binding site toward cooperative assembly of the ternary complex. In the *Forward* register the positively charged Dp tail of the polyamide is

in close proximity to the exocyclic amine of the guanine ($5'$ TGAT $3'$) in the Exd cognate site. We surmise that the proximity and the consequent charge and/or steric repulsion between Dp and the amine in the minor groove would favorably displace the GA base step toward Exd in the major groove. CURVES analysis indicates that multiple Hox–Exd/Pbx1 bound DNA sequences deviate from idealized B-form DNA in a manner compatible with the perturbations that may be induced by the Dp tail (Figure 2). The premise is further supported by the fact that simply decreasing conformational flexure of the Exd binding site, as would happen as a result of polyamide binding in the *Inverse* register, does not enhance binding to the same degree as it does when the Dp tail is in close proximity of the GA base step. The effect of minor groove located charges in Hox–three amino acid loop extension homeodomain (TALE) complexes has been borne out in structural studies of Scr–Exd–DNA complexes, where a minor groove bound arginine from Scr has been correlated with allosteric Scr–Exd interactions (35). However the studies with Scr posited that this effect was due to pre-existing DNA conformational preferences between different DNAs. While we cannot rule out an effect of sequence-dependent structural effects in modulating Hox–TALE interaction, the ability of polyamides to increase Exd–DNA binding affinity on the same piece of DNA points to a potential role of induced DNA structural changes in cooperative Exd binding.

Taken together, our results indicate that at a physiological Hox–Exd binding site the tethered YPWM peptide enhances Exd–DNA interaction by ~ 100 -fold and allosteric modulation of DNA sites adds another ~ 10 -fold to the binding affinity (see Figure 4). Thus, the synthetic molecules deconvolute the direct and allosteric components of Exd–DNA complex formation.

METHODS

Synthesis of Polyamides. Polyamides and polyamide–peptide conjugates 1–4 were synthesized according to previously described procedures (50–52, 56, 57).

Purification of Exd, Ubx, and Mutant Ubx. The *Drosophila melanogaster* protein Exd (residues 236–320) and Ubx IVa (residues 233–313) were purified using standard protocols (32, 50, 51). Residues YPWM of Ubx IVa were mutated to alanines using the Quikchange site directed mutagenesis kit (Stratagene) via the manufacturer's protocols. The mutant Ubx was confirmed by sequencing and expressed and purified using the same protocols as the wild-type Ubx. Purification consisted of ammonium

Global Implications. The role of allosteric modulation of DNA structure in cooperative assembly of multiple proteins at a composite binding site has been elegantly dissected at the interferon- β enhancer (15). At this enhancer eight proteins modulate binding site conformation and thereby stabilize cooperative assembly without significant contribution from interprotein interactions (15). Oct1 presents a simpler example, where loss of protein–protein connectivity does not eliminate cooperative DNA binding by two of its DNA binding modules (18). In this protein, the two contiguous DNA binding modules when expressed as separate peptides are able to reciprocally increase the affinity of each other at the composite DNA binding site, even in the absence of any direct contact (18). Intriguingly, in these two cases, as well as several others, allosteric modulation often positively contributes ~ 10 -fold or ~ 1.5 kcal/mol to assembly of the partner proteins (9, 11, 14–22). This is the same scale at which DNA deformation caused by hairpin polyamides contribute to Exd binding. Thus, here we have utilized a sequence-specific chemical wedge as an architectural cofactor and harnessed allosteric DNA modulation to enhance assembly of a protein–DNA complex. It is possible that such noncontact effects would contribute significantly to overall enhancer function by assisting assembly of multiple proteins and disfavoring binding in undesired orientations and could be an under-appreciated general mechanism of sequence recognition. In the future, we intend to incorporate these indirect effects to enhance sequence-specificity of rationally engineered DNA binding molecules. Such precision-tailored synthetic molecules will be instrumental in the design of a novel class of functional regulators that will greatly contribute to molecular medicine, functional genomics, and synthetic biology.

sulfate precipitation followed by cation exchange and size exclusion chromatography as previously described (32, 50, 51). After purification, the concentrations of Exd, wild-type Ubx, and mutant Ubx were determined by using an extinction coefficient at 280 nm of 12,090, 16,620, and 9,970 $\text{cm}^{-1} \text{M}^{-1}$, respectively.

Electrophoretic Mobility Shift Assays. EMSAs were run as previously described (50, 51). Synthetic oligonucleotides (for sequences, see Supplementary Table 1) were annealed and 5' phosphorylated with ^{32}P using standard methods. Reactions were performed in a buffer composed of 150 mM K-glutamate, 50 mM HEPES (pH 8.0), 2 mM DTT, 100 ng μL^{-1} BSA, 10% DMSO, and 10% glycerol. ^{32}P -DNA (0.1 nM estimated final concentration) was mixed with saturating amounts of either poly-

amide (50 nM final concentration; see Supplementary Figure 2) or Ubx (100 nM for wild-type and 10 nM for the mutant; see Supplementary Figure 1) and incubated at 4 °C for 30 min. Exd was added to bring the reaction volume to 20 μ L at the appropriate final concentration, and the reactions were incubated at 4 °C for a further 1 h. Fifteen microliters of the reaction was loaded onto prerun, nondenaturing, 10% acrylamide/3% glycerol gels in Tris-borate-EDTA buffer (TBE). Initial polyamide trials (Figure 3; Supplementary Figure 3) were run in 1x TBE, and quantitation of both protein and polyamide (Figures 1 and 4; Supplementary Figure 1) was performed on complexes resolved by nondenaturing PAGE in 0.5x TBE. The gels were loaded at 200 V and run for 2.5 h. The dried gels were exposed to a PhosphorImager screen overnight and visualized on a Typhoon 9410 PhosphorImager.

Quantitation and Analysis of EMSAs. The fraction of the total radioactivity in the free DNA band was quantitated with IMAGEQUANT TL and fit to the Hill binding equation with SIGMAPLOT 6.0 using nonlinear regression. Hill coefficients were allowed to vary freely but were near unity, except where noted. This analysis assumes that Ubx–DNA binding is saturated (cf. Supplementary Figure 1) and that Ubx and Exd are not interacting in the absence of DNA. (Previous experiments with the Hox–Exd interaction in the absence of DNA give a K_D in the 10–100 μ M range (31, 58).)

Groove Dimension Analysis. The experimentally determined structures of the Ubx–Exd–DNA (32), HoxA9–Pbx1–DNA (34), HoxB1–Pbx1–DNA (33), and Scr–Exd–DNA (35) complexes were obtained from the protein databank. B-Form DNA models were generated with Namot (59) and 3DNA (60) using the sequence of the Ubx–Exd complex. CURVES version 5.3 R.L. (43–45) was used to determine the groove dimensions at 10 points per basepair, using a combined helical axis. All other parameters were set to defaults. The alignment of the experimental and B-form models was generated using the align command of PyMol, using the first T–A basepair of the TGAT binding site, along with the two preceding base pairs, as the reference points. Both strands were used for the alignment, although only the TGAT bearing strand is shown.

Acknowledgment: We gratefully acknowledge the support of the NIH (GM069420 to A.Z.A., GM27681 to P.B.D., and AI552218 to J.S.T.), March of Dimes and W. M. Keck Foundations (to A.Z.A.), and the Shaw Scholar and Vilas Associate awards (to A.Z.A.). R.M. was supported by the NIH-MBTG predoctoral fellowship, L.D. was supported by the NLM-GSTP postdoctoral fellowship, and H.H. was an NSF-REU scholar.

Supporting Information Available: This material is available free of charge via the Internet.

REFERENCES

- Pabo, C. O., and Sauer, R. T. (1992) Transcription factors: structural families and principles of DNA recognition, *Annu. Rev. Biochem.* 61, 1053–1095.
- Choo, Y., and Klug, A. (1997) Physical basis of a protein–DNA recognition code, *Curr. Opin. Struct. Biol.* 7, 117–125.
- Harrison, S. C., and Aggarwal, A. K. (1990) DNA recognition by proteins with the helix–turn–helix motif, *Annu. Rev. Biochem.* 59, 933–969.
- Wharton, R. P., and Ptashne, M. (1985) Changing the binding specificity of a repressor by redesigning an α helix, *Nature* 316, 601–605.
- Pabo, C. O., Peisach, E., and Grant, R. A. (2001) Design and selection of novel Cys2His2 zinc finger proteins, *Annu. Rev. Biochem.* 70, 313–340.
- Beerli, R. R., and Barbas, C. F., III (2002) Engineering polydactyl zinc finger transcription factors, *Nat. Biotechnol.* 20, 135–141.
- Ashworth, J., Havranek, J. J., Duarte, C. M., Sussman, D., Monnat, R. J., Jr., Stoddard, B. L., and Baker, D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity, *Nature* 441, 656–659.
- Simon, M. D., and Shokat, K. M. (2004) Adaptability at a protein–DNA interface: Re-engineering the engrailed homeodomain to recognize an unnatural nucleotide, *J. Am. Chem. Soc.* 126, 8078–8079.
- Koudelka, G. B., and Carlson, P. (1992) DNA twisting and the effects of non-contacted bases on affinity of 434 operator for 434 repressor, *Nature* 355, 89–91.
- Travers, A. A. (1993) DNA–Protein Interactions, Chapman & Hall, London.
- Hamm, M. K., and Schepartz, A. (1995) Studies on the formation of DNA–protein interfaces—DNA specificity and straightening by CREB, *Bioorg. Med. Chem. Lett.* 5, 1621–1626.
- Lefstin, J. A., and Yamamoto, K. R. (1998) Allosteric effects of DNA on transcriptional regulators, *Nature* 392, 885–888.
- Pabo, C. O., and Nekudova, L. (2000) Geometric analysis and comparison of protein–DNA interfaces: why is there no simple code for recognition? *J. Mol. Biol.* 301, 597–624.
- Siggers, T., Silkov, T., and Honig, B. (2005) Bending in the right direction, *Structure* 13, 1400–1401.
- Panne, D., Maniatis, T., and Harrison, S. C. (2007) An atomic model of the interferon-beta enhanceosome, *Cell* 129, 1111–1123.
- Hogan, M., Dattagupta, N., and Crothers, D. M. (1979) Transmission of allosteric effects in DNA, *Nature* 278, 521–524.
- Ansari, A. Z., Chael, M. L., and O'Halloran, T. V. (1992) Allosteric underwinding of DNA is a critical step in positive control of transcription by Hg–MerR, *Nature* 355, 87–89.
- Klemm, J. D., and Pabo, C. O. (1996) Oct-1 POU domain–DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage, *Genes Dev.* 10, 27–36.
- Hines, C. S., Meghoo, C., Shetty, S., Biburger, M., Brenowitz, M., and Hegde, R. S. (1998) DNA structure and flexibility in the sequence-specific binding of papillomavirus E2 proteins, *J. Mol. Biol.* 276, 809–818.
- Escalante, C. R., Brass, A. L., Pongubala, J. M., Shatova, E., Shen, L., Singh, H., and Aggarwal, A. K. (2002) Crystal structure of PU.1/IRF-4/DNA ternary complex, *Mol. Cell* 10, 1097–1105.
- Zhang, Y., Xi, Z., Hegde, R. S., Shakked, Z., and Crothers, D. M. (2004) Predicting indirect readout effects in protein–DNA interactions, *Proc. Natl. Acad. Sci. U.S.A.* 101, 8337–8341.
- Segal, E., Fondulfe-Mittendorf, Y., Chen, L., Thastrom, A., Field, Y., Moore, I. K., Wang, J. P., and Widom, J. (2006) A genomic code for nucleosome positioning, *Nature* 442, 772–778.
- Ptashne, M., and Gann, A. (2002) *Genes & signals*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Mann, R. S., and Chan, S. K. (1996) Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins, *Trends Genet.* 12, 258–262.
- Moens, C. B., and Selleni, L. (2006) Hox cofactors in vertebrate development, *Dev. Biol.* 291, 193–206.
- Ekker, S. C., Jackson, D. G., von Kessler, D. P., Sun, B. I., Young, K. E., and Beachy, P. A. (1994) The degree of variation in DNA sequence recognition among four Drosophila homeotic proteins, *EMBO J.* 13, 3551–3560.
- Duboule, D., Ed. (1994) *Guidebook to the homeobox genes*, Oxford University Press, Oxford.
- Chan, S. K., Jaffe, L., Capovilla, M., Botas, J., and Mann, R. S. (1994) The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein, *Cell* 78, 603–615.
- Rauskolb, C., and Wieschaus, E. (1994) Coordinate regulation of downstream genes by extradenticle and the homeotic selector proteins, *EMBO J.* 13, 3561–3569.

30. Neuteboom, S. T., Peltenburg, L. T., van Dijk, M. A., and Murre, C. (1995) The hexapeptide LFPWMR in Hoxb-8 is required for cooperative DNA binding with Pbx1 and Pbx2 proteins, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 9166–9170.
31. Lu, Q., and Kamps, M. P. (1996) Structural determinants within Pbx1 that mediate cooperative DNA binding with pentapeptide-containing Hox proteins: proposal for a model of a Pbx1-Hox-DNA complex, *Mol. Cell. Biol.* **16**, 1632–1640.
32. Passner, J. M., Ryoo, H. D., Shen, L., Mann, R. S., and Aggarwal, A. K. (1999) Structure of a DNA-bound ultrabithorax-extradenticle homeodomain complex, *Nature* **397**, 714–719.
33. Piper, D. E., Batchelor, A. H., Chang, C. P., Cleary, M. L., and Wolberger, C. (1999) Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation, *Cell* **96**, 587–597.
34. LaRonde-LeBlanc, N. A., and Wolberger, C. (2003) Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior, *Genes Dev.* **17**, 2060–2072.
35. Joshi, R., Passner, J. M., Rohs, R., Jain, R., Sosinsky, A., Crickmore, M. A., Jacob, V., Aggarwal, A. K., Honig, B., and Mann, R. S. (2007) Functional specificity of a Hox protein mediated by the recognition of minor groove structure, *Cell* **131**, 530–543.
36. Galant, R., Walsh, C. M., and Carroll, S. B. (2002) Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites, *Development* **129**, 3115–3126.
37. Nguyen-Hackley, D. H., Ramm, E., Taylor, C. M., Joung, J. K., Dervan, P. B., and Pabo, C. O. (2004) Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands, *Biochemistry* **43**, 3880–3890.
38. Dervan, P. B., and Edelson, B. S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides, *Curr. Opin. Struct. Biol.* **13**, 284–299.
39. Weatherbee, S. D., Halder, G., Kim, J., Hudson, A., and Carroll, S. (1998) Ultrabithorax regulates genes at several levels of the wing-patterning hierarchy to shape the development of the *Drosophila* haltere, *Genes Dev.* **12**, 1474–1482.
40. Beachy, P. A. (1990) A molecular view of the ultrabithorax homeotic gene of *Drosophila*, *Trends Genet.* **6**, 46–51.
41. Mann, R. S., and Hogness, D. S. (1990) Functional dissection of Ultrabithorax proteins in *D. melanogaster*, *Cell* **60**, 597–610.
42. Shanmugam, K., Featherstone, M. S., and Saragovi, H. U. (1997) Residues flanking the HOX YPWM motif contribute to cooperative interactions with PBX, *J. Biol. Chem.* **272**, 19081–19087.
43. Lavery, R., and Sklenar, H. (1988) The definition of generalized helical parameters and of axis curvature for irregular nucleic acids, *J. Biomol. Struct. Dyn.* **6**, 63–91.
44. Lavery, R., and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: conventions and principles, *J. Biomol. Struct. Dyn.* **6**, 655–667.
45. Stofor, E., and Lavery, R. (1994) Measuring the geometry of DNA grooves, *Biopolymers* **34**, 337–346.
46. Warren, C. L., Kratochvil, N. C., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N., Jr., and Ansari, A. Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 867–872.
47. Gottesfeld, J. M., Melander, C., Suto, R. K., Raviol, H., Luger, K., and Dervan, P. B. (2001) Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides, *J. Mol. Biol.* **309**, 615–629.
48. Bremer, R. E., Baird, E. E., and Dervan, P. B. (1998) Inhibition of major-groove-binding proteins by pyrrole-imidazole polyamides with an arg-pro-arg positive patch, *Chem. Biol.* **5**, 119–113.
49. Oakley, M. G., Mrksich, M., and Dervan, P. B. (1992) Evidence that a minor groove-binding peptide and a major groove-binding protein can simultaneously occupy a common site on DNA, *Biochemistry* **31**, 10969–10975.
50. Arndt, H. D., Hauschild, K. E., Sullivan, D. P., Lake, K., Dervan, P. B., and Ansari, A. Z. (2003) Toward artificial developmental regulators, *J. Am. Chem. Soc.* **125**, 13322–13323.
51. Hauschild, K. E., Metzler, R. E., Arndt, H. D., Moretti, R., Raffaele, M., Dervan, P. B., and Ansari, A. Z. (2005) Temperature-sensitive protein-DNA dimers, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 5008–5013.
52. Stafford, R. L., Arndt, H. D., Brezinski, M. L., Ansari, A. Z., and Dervan, P. B. (2007) Minimization of a protein-DNA dimerizer, *J. Am. Chem. Soc.* **129**, 2660–2668.
53. Mann, R. S., and Carroll, S. B. (2002) Molecular mechanisms of selector gene function and evolution, *Curr. Opin. Genet. Dev.* **12**, 592–600.
54. Chan, S. K., and Mann, R. S. (1996) A structural model for a homeotic protein-extradenticle-DNA complex accounts for the choice of HOX protein in the heterodimer, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5223–5228.
55. Ryoo, H. D., Marty, T., Casares, F., Affolter, M., and Mann, R. S. (1999) Regulation of Hox target genes by a DNA bound Homothorax/Hox/Extradenticle complex, *Development* **126**, 5137–5148.
56. Baird, E. E., and Dervan, P. B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids, *J. Am. Chem. Soc.* **118**, 6141–6146.
57. Rucker, V. C., Foister, S., Melander, C., and Dervan, P. B. (2003) Sequence specific fluorescence detection of double strand DNA, *J. Am. Chem. Soc.* **125**, 1195–1202.
58. Sprules, T., Green, N., Featherstone, M., and Gehring, K. (2003) Lock and key binding of the HOX YPWM peptide to the PBX homeodomain, *J. Biol. Chem.* **278**, 1053–1058.
59. Tung, C. S., and Carter, E. S., 2nd (1994) Nucleic acid modeling tool (NAMOT): an interactive graphic tool for modeling nucleic acid structures, *Comput. Appl. Biosci.* **10**, 427–433.
60. Lu, X. J., and Olson, W. K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures, *Nucleic Acids Res.* **31**, 5108–5121.