J|A|C|S
COMMUNICATIONS

Published on Web 10/09/2003

Toward Artificial Developmental Regulators

Hans-Dieter Arndt,[†] Karl E. Hauschild,[‡] David P. Sullivan,[‡] Kara Lake,[‡] Peter B. Dervan,*[,†] and Aseem Z. Ansari*[,‡]

California Institute of Technology, Division of Chemistry and Chemical Engineering, Pasadena, California 91125, and University of Wisconsin-Madison, Department of Biochemistry and the Genome Center, Madison, Wisconsin 53706

Received July 8, 2003; E-mail: ansari@biochem.wisc.edu; dervan@caltech.edu

Hox (*Ho*meobo*x*) proteins belong to a family of transcriptional regulators that bear the "homeodomain": a trihelical DNA-binding domain that is conserved across vast evolutionary distances.[1] These proteins play a vital role in anterior−posterior patterning of the embryo.[2] The role of Hox proteins in controlling the developmental fate of an organism has been most dramatically illustrated in the case of *Drosophila melanogaster*: for example, mutations in the Hox gene Ultrabithorax (Ubx) led to the transformation of halteres (balancing organs) into a set of wings.[2a,3]

Hox proteins in general display poor affinity and sequence specificity for DNA in vitro. However, Hox proteins often interact with the TALE (*T*hree *A*mino Acid *L*oop *E*xtension) class of homeodomain proteins and together bind as heterodimers to target DNA sequences with very high affinity and specificity.[4] Recent crystal structures of one such ternary complex shows that Ubx interacts with the *Drosophila* TALE protein extradenticle (Exd) via a short docking YPWM peptide (Figure 1a).[5] This conserved peptide is a feature of all Hox proteins, and it occurs N-terminal to the homeodomain separated by an intervening "linker" peptide of varying lengths in the different Hox proteins.[6] Recently, it has been argued that additional interactions between Hox and TALE proteins may play a key role in their ability to bind DNA cooperatively.[7] Here, we describe a small-molecule polyamide−peptide conjugate that mimics the ability of Hox-family developmental regulators to bind cooperatively with Exd to a cognate DNA site.

The Ubx docking peptide is presented to Exd via a flexible linker (disordered in the crystal structure). Therefore, we decided to recreate this interface by presenting the docking peptide off a minor groove DNA-binding polyamide (PA). PAs bind to B-form DNA in a highly sequence-specific manner.[9] Using a simple set of "pairing rules", these compounds can be tailored to target various DNA sequences with nanomolar affinities.[9c] Recently, it was shown that regulatory peptides could be conjugated to a PA to generate small molecules that can upregulate gene expression in vitro.[10] The success of modular artificial transcription factors (ATFs) prompted us to design PAs that would cooperatively recruit Exd to DNA.

Further inspection of the Ubx−Exd−DNA crystal structure[5] revealed that the minor groove dimensions next to the Exd binding site compare well with a high-resolution structure of a PA bound to the minor groove of DNA (Figure 1b). Furthermore, the C-terminal residue of the docking peptide is located above the minor groove (Figure 1c). A propyl linker projecting off the ring in the PA could connect both moieties without significantly perturbing polyamide or peptide binding to their respective targets (Figure 1). The residue over the base closest to the C terminus served as the connection point. We used this structure-based design to generate the PA−docking-peptide conjugates **2** and **3** (Figure 2). The conjugates display a functional (**2**) or a mutant (**3**) hexapeptide

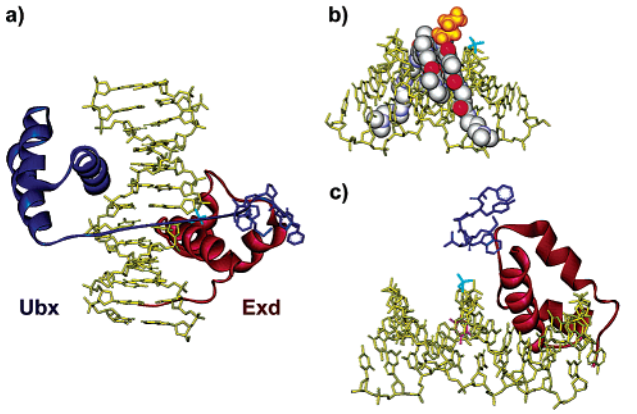

Figure 1. Structural rationale for Exd-recruiting polyamides. (a) X-ray crystal structure of the Ubx−Exd ternary complex (PDB code 1B8I).[5] Only side chains of the Ubx docking hexapeptide are shown. (b) X-ray crystal structure of a polyamide dimer bound to DNA (PDB code 407D).[8] with a propylamine chain (orange) docked to a pyrrole ring nitrogen. (c) Side view of the minor groove next to the Exd binding site. The hexapeptide bound to Exd is shown in blue, the Exd homeodomain is colored red. The nearest phosphate is indicated in bright blue, the closest base (T) in the minor groove is highlighted in violet.

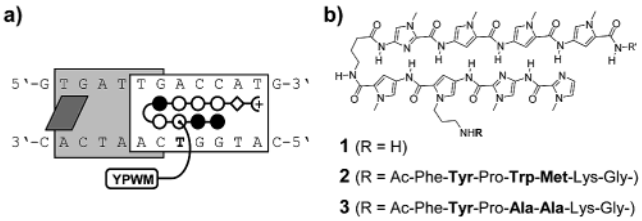

Figure 2. (a) Composite Exd−polyamide binding site. Major groove occupancy by the Exd protein is shaded in light gray, minor groove occupancy by the Exd N-terminal residues in dark gray. (b) Chemical structure of the compounds investigated in this study, R′ = β-Dp. Key residues for Exd binding are in bold type.

attached to the PA−propylamine side chain via a glycine spacer. The conjugates were synthesized by solution-phase coupling of protected peptide acid fragments to the parent PA **1** (see Supporting Information).

The DNA-binding properties of the compounds **1**−**3** were investigated by quantitative DNase I footprinting assays.[11] The equilibrium binding constants of each of the compounds for a matched versus three single base pair mismatch sites is compiled in Table 1. The conjugation of the peptides to the parent polyamide **1** leads to a reduction in binding affinity, not unlike previous observations.[10] Interestingly, the functional peptide sequence in **2** has a much greater influence on binding affinity and specificity than the mutant peptide in **3**. This behavior may be indicative of additional properties of the docking peptide (under investigation). Equally unexpected was the ability of the conjugates **2** and **3** to

[†] California Institute of Technology.
[‡] University of Wisconsin-Madison.

10.1021/ja0371395 CCC: $25.00

Table 1. Equilibrium Dissociation Constants K_d [nM] for **1**, **2**, and **3**[a,b]

	5′-TGGTCA-3′	5′-TGGCCA-3′	5′-TGGGCA-3′	5′-AGCTCA-3′
1	0.048 ± 0.015	0.97 ± 0.41 [20]	0.76 ± 0.39 [16]	3.1 ± 0.9 [65]
2	5.8 ± 0.8	7.9 ± 1.6 [1.4]	≥100 [≥17]	6.4 ± 1.1 [1.1]
3	0.86 ± 0.32	12 ± 5 [14]	≥100 [≥116]	14 ± 5 [16]

[a] Lower strand sequence shown. Mismatched bp are underlined, the residue under the YPWM peptide is in bold face. [b] Rel. specificities are given in square brackets.

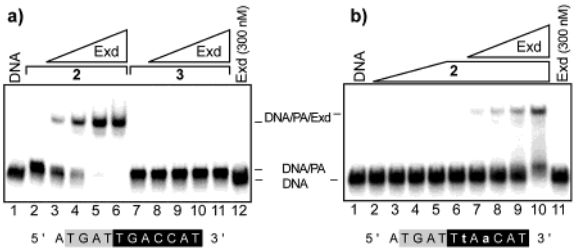

Figure 3. Specific recruitment of Exd by **2** or **3**. (a) 50 nM PA (**2** in lanes 2−6 and **3** in lanes 7−11) was incubated with a ^{32}P-labeled DNA containing a single Exd−PA composite site (see Supporting Information) at increasing concentrations of Exd (3, 10, 30, 100 nM in lanes 3−6 or 8−11). (b) DNA with a 2 bp mutated PA binding site (lower case) was incubated with increasing concentrations of **2** (1, 3.3, 10, 33, 100 nM in lanes 2−6). At 100 nM of **2** increasing concentrations of Exd were added (10, 30, 100, 300 nM Exd in lanes 7−10).

discriminate between the CG and the GC mismatch base pair—a property not shown by the parent **1**. The mismatch lies directly under the polyamide ring to which the peptide is attached (Figure 2a). It is striking that the protruding exocyclic amine of guanine is not tolerated by the conjugates.

The ability of the compounds **1**−**3** to enhance Exd binding to its adjacent cognate site was tested using electrophoretic mobility shift assays (EMSA). A 47-base-pair duplex DNA (Figure 3) with one cognate site was incubated with saturating concentrations (50 nM) of each compound. As shown in Figure 3 polyamide−peptide conjugates bind DNA and slightly decrease its mobility (compare lanes 1 and 2 in Figure 3a). In the presence of conjugate **2**, Exd (residues 238−324) binds DNA with very high affinity. No binding was observed with the mutant peptide conjugate **3**. The K_d of Exd for the DNA·**2** complex is 4.4 ± 2 nM, 2-fold higher than that for DNA−Ubx complex. However, the affinity of Ubx for its DNA site is at least ∼40-fold lower than that of **2** for its respective site.[12] Indeed, a 20-fold lower concentration of compound **2** is required to recruit Exd as compared to its natural Hox partner. In our experiments, neither **3** nor **1** showed any ability to recruit Exd to its cognate site (Figure 3 and Supporting Information). At 1 μM, Exd binds DNA, but the mobility of the band suggests multiple Exd molecules bind to DNA nonspecifically (Supporting Information). Thus, conjugate **2** improves the affinity of Exd for its cognate site by at least ∼200-fold, and far more importantly, it enhances specific binding of Exd to a target site.

To investigate the contribution of the peptide−Exd interaction to the binding site specificity, the polyamide-binding site on the DNA template was eliminated. As shown in Figure 3b, **2** did not bind this mutated template even at 100 nM concentrations. However, upon addition of 300 nM Exd, an unstable ternary complex is formed (<10%). The ability of Exd to bind DNA nonspecifically at 1 μM concentration prevents us from determining a reliable binding constant.

Our results provide strong evidence that the YPWM peptide contributes significantly to the cooperative interaction between a Hox protein and its TALE partner on DNA. The data indicate that

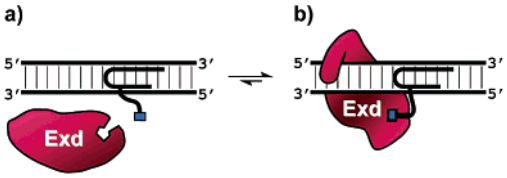

Figure 4. Exd recruitment by polyamide−peptide conjugate **2**.

the polyamide−peptide **2** binds to the DNA and presents a bidentate binding surface (peptide and DNA site) to Exd (Figure 4a). This increases the net affinity of the ternary assembly on DNA to ∼5 nM of Exd at saturating concentrations of PA (→ Figure 4b). Due to the low affinity of Exd for its DNA binding site, binding of Exd alone is not detected under our conditions. However, under conditions where the polyamide binding site has been abolished, Exd cooperatively assists the formation of an unstable ternary complex with **2** (Figure 3b).

In summary, our data demonstrate that cooperative interactions between a protein and a minor groove binding polyamide−peptide conjugate (Exd and **2**) specifically stabilize the formation of a ternary complex on a composite DNA site. The ability of conjugate **2** to recruit Exd more efficiently than its natural Hox protein partner illustrates that structure-based modular design is a valid strategy toward artificial transcription factors (ATFs). In extension, the data suggest that this approach is not limited to generating ATFs which mimic Hox factors. Joining two sequence-specific domains (one for DNA and one for a TF protein) should lead to cooperative protein−DNA dimerizers. Thereby, a new class of small-molecule ATFs could be obtained—those that function in concert with natural transcription factors.[13]

Acknowledgment. This work was supported in part by the UW-I & EDR and Steenbock Funds to A.Z.A., a NIH grant to P.B.D., and a DAAD postdoctoral fellowship to H.D.A.

Supporting Information Available: Experimental details for the synthesis of compounds **1**−**3**, the DNase I footprinting titrations, protein expression, and the EMSA studies (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Manak, J. R.; Scott, M. P. *Dev. Suppl.* **1994**, 61−71.
(2) (a) Lewis, E. B. *Nature* **1978**, *276*, 565−570. (b) Scott, M. P.; Carroll, S. B. *Cell* **1987**, *51*, 689−98. (c) McGinnis, W.; Krumlauf, R. *Cell* **1992**, *68*, 283−302.
(3) Weatherbee, S. D.; Halder, G.; Kim, J.; Hudson, A.; Carroll, S. B. *Genes Dev.* **1998**, *12*, 1474−82.
(4) (a) van Dijk, M. A.; Murre, C. *Cell* **1994**, *78*, 617−24. (b) Chan, S.-K.; Mann, R. S. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5223−5228. (c) Mann, R. S.; Chan, S. K. *Trends Gen.* **1996**, *12*, 258−262.
(5) Passner, J. M.; Ryoo, H. D.; Shen, L.; Mann, R. S.; Aggarwal, A. K. *Nature* **1999**, *397*, 714−719.
(6) Duboule, D. *Guidebook to the Homeobox Genes*; Oxford University Press: Oxford, 1994.
(7) Galant, R.; Walsh, C. M.; Carroll, S. B. *Development* **2002**, *129*, 3115−3126.
(8) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111−115.
(9) (a) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559−561. (b) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468−471. (c) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215−2235.
(10) (a) Mapp, A.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3930−3935. (b) Ansari, A. Z.; Mapp, A. K.; Nguyen, D.; Dervan, P. B.; Ptashne, M. *Chem. Biol.* **2001**, *8*, 583−592. (c) Arora, P. S.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 13067−13071. (d) Ansari, A. Z.; Mapp, A. K. *Curr. Opin. Chem. Biol.* **2002**, *6*, 765−772.
(11) Trauger, J. W.; Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450−466.
(12) Brezinski, M.; Warren, C. L.; Ansari, A. Z. Unpublished work.
(13) For optimizing sequence specificity and cell permeation of the conjugates, design variables such as points of attachment of the recruiting peptide as well as linker length are currently being investigated.

JA0371395