Cell Reports, Volume *24*

Supplemental Information

Intrinsic DNA Shape Accounts for Affinity Differences between Hox-Cofactor Binding Sites

Tim Zeiske, Nithya Baburajendran, Anna Kaczynska, Julia Brasch, Arthur G. Palmer III, Lawrence Shapiro, Barry Honig, and Richard S. Mann

Supplemental Tables and Figures:

Table S1. Crystalization conditions, related to Experimental Procedures

Oligomer name	Oligomer sequence	Final condition
red14	GCATGATTTATGAC	200 mM $MgCl_2$, 0.1M Tris pH:5.8, 17.5% PEG 4000 5% glycerol
blue14	GCATGATTAATGAC	200 mM $MgCl_2$, 0.1 M Tris pH:5.8, 17.5% PEG 3350, 2.5% glycerol
magenta14*	GCATGATTTACGAC	90 mM $MgCl_2$ 0.1 M Tris pH:9.0 22% PEG 3350,
black14*	GCATGATAAATGAC	0.1 M Sodium citrate pH 5.3, 25% PEG 3350,

* The original protein sample contained 50 mM $MgCl_2$

Table S2. Data collection and refinement statistics, related to Figure 1

	Red	Magenta	Blue	Black
DATA COLLECTION				
Resolution range (Å)*	29.9 - 2.4 (2.5 - 2.4)	44.9 - 3.0 (3.1 - 3.0)	28.1 - 2.9 (3.0 - 2.9)	63.5 - 2.4 (2.5 - 2.4)
Space group	C 1 2 1	C 1 2 1	C 1 2 1	P 1
Unit cell				
a, b, c (Å)	77.1 49.4 95.2	77.3 49.5 95.1	77.7 49.7 96.8	45.4 45.6 66.9
α, β, γ (◦)	90 109.3 90	90 109.2 90	90 109.0 90	99.1 100.4 114.2
Unique reflections*	12413 (980)	6678 (624)	7524 (640)	16923 (1662)
Multiplicity	3.4	3.6	1.9	2.0
Completeness (%)*	97.33 (77.0)	98.61 (92.6)	94.74 (83.2)	93.06 (92.8)
Mean I/sigma(I)*	18.49 (1.92)	6.0 (1.70)	9.16 (1.75)	6.8 (2.00)
R-merge*	0.065 (0.539)	0.237 (0.629)	0.172 (0.852)	0.169 (0.706)
REFINEMENT				
R-work*	0.223 (0.319)	0.269 (0.344)	0.253 (0.376)	0.232 (0.294)
R-free*	0.252 (0.358)	0.288 (0.373)	0.281 (0.411)	0.276 (0.364)
Number of non-hydrogen atoms	1731	1660	1586	2852
macromolecules	1688	1660	1583	2772
water	43	0	3	80
Protein residues	164	166	158	249
RMS(bonds)	0.007	0.003	0.004	0.003
RMS(angles)	0.90	0.55	0.70	0.47
Ramachandran favored (%)	98	95	95	96
Ramachandran outliers (%)	0	0	0	0

Statistics for the highest-resolution shell are shown in parentheses.

∗ Values in parentheses are for the highest resolution shell.

‡ Ions were modeled as waters

Table S3: AbdB and Exd interaction surface areas and H-bonds, related to Figure 1

Chain interface	Area (Å^2); H-bonds (#)				
	Red	**Magenta**	**Blue**	**BlackF**	**BlackR**
AbdB:DNA	910.1; 21	943.6; 18	885.1; 23	918; 20	972; 20
Exd:DNA	1069.5; 25	1073.1; 20	863.5; 8	1052.4; 19	NA
AbdB:Exd	249.1; 1	163.9; 0	138.2; 0	278.3; 1	NA

All values were obtained with the PISA program within CCP4 (PISA; (Krissinel and Henrick, 2007)) (Collaborative Computational Project, 1994; Winn et al., 2011).

Table S4. Protein expression and purification conditions, related to Experimental Procedures.

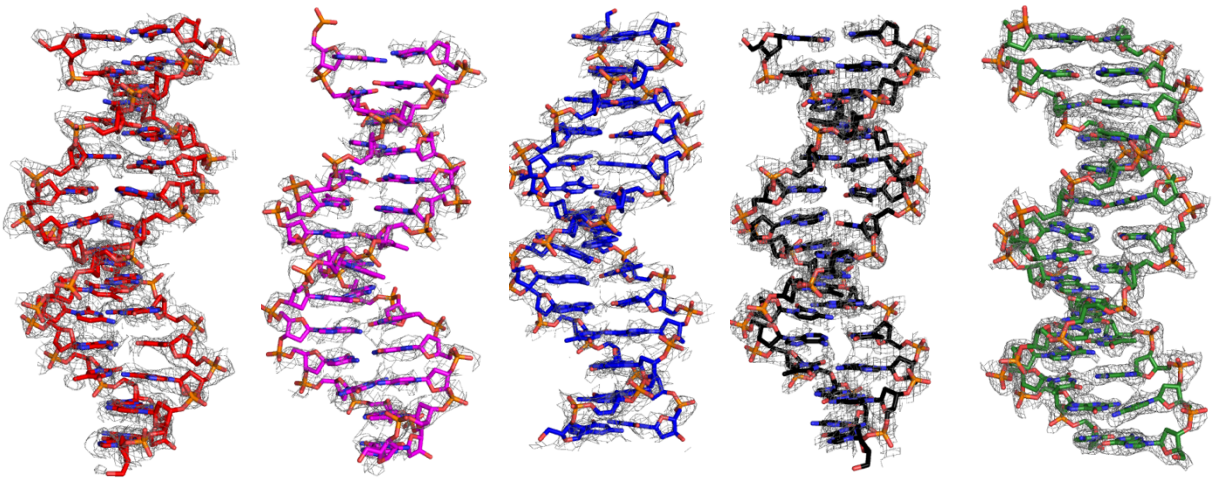
	Batch I	Batch II
Lysis buffer	50mM Tris pH 8.0, 300mM NaCl, 1% Triton X-100, 10% glycerol, 1mM PMSF, protease inhibitor (Roche)	50 mM Tris pH 7.5, 500 mM NaCl, 1 mM TCEP, 20 mM Imidazole, lysozyme, DNAse I, protease inhibitor (Roche)
Ni equilibration buffer	50mM Tris pH 8.0, 100mM NaCl, 1mM PMSF, protease inhibitor (Roche)	50 mM Tris pH 7.5, 500 mM NaCl, 1 mM TCEP, 20 mM Imidazole
Ni elution buffer	250mM Imidazole pH 8.0, 10mM Tris pH 8.0, 100 mM NaCl, protease inhibitor (Roche)	50 mM Tris pH 7.5, 500 mM NaCl, 1 mM TCEP, 500 mM Imidazole
Gel filtration buffer/Final storage buffer	10 mM Tris pH.8.0, 200 mM Nacl and 2 mM TCEP	10 mM TRIS pH 7.5, 200 mM NaCl, 2 mM TCEP, 50 mM $MgCl_2$

Table S5. Oligos used in the competition EMSA experiments, related to Experimental Procedures.

Core motif	Oligo sequence
Red	acgctctggGCATGATTTATGACccacgtctc
Magenta	acgctctggGCATGATTTACGACccacgtctc
Blue	acgctctggGCATGATTAATGACccacgtctc
Black	acgctctggGCATGATAAATGACccacgtctc

Figure S1. Composite omit maps confirm key structural features, related to Figure 1.

(A) Stick representations of the Red, Magenta, Blue, BlackF and BlackR DNAs present in the respective crystal structures. The composite omit maps (with anneal mode) are displayed in grey mesh and contoured at 1.0 σ. The coordinates of the DNA were deleted while making the composite omit maps to avoid model bias.



Red Magenta Blue BlackF BlackR

(B) NTA residues of AbdB from the Red structure as seen in Figure 2, panel C, in stick representation with the composite omit (with anneal mode) map displayed in grey mesh contoured at 1.0σ.

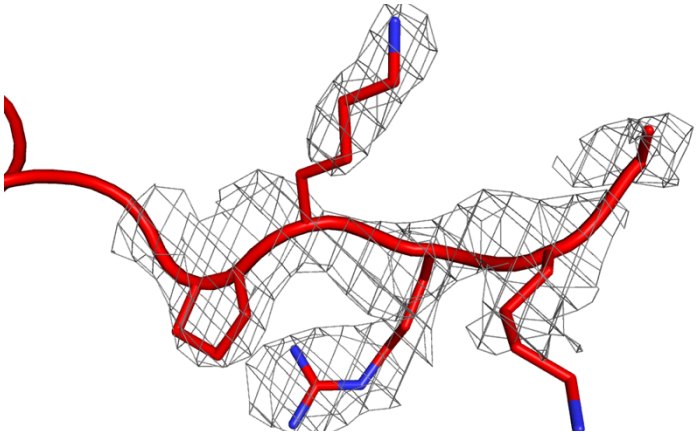


Figure S2. Schematic summarizing protein-DNA contacts for all five complexes using DNAproDB (Sagendorf et al., 2017), related to Figure 1.

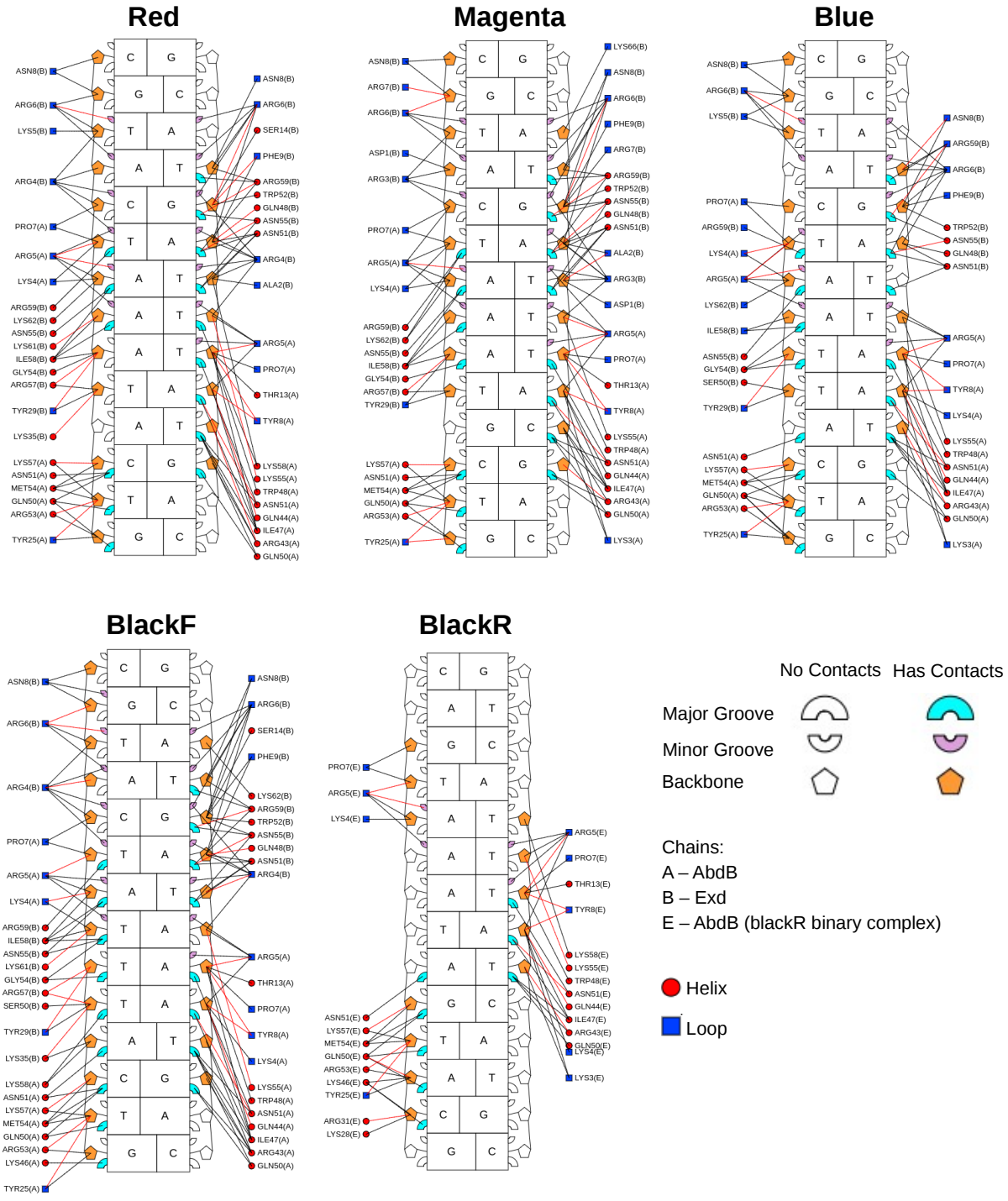


Figure S3. Comparison of NTA backbone paths of the red complex with HoxB1 and Ubx, related to Figure 2.

Left. Ubx's NTA has a similar path as other posterior Hox proteins such as AbdB's, but differs from Scr's NTA.

Right. HoxB1's NTA is unique, but most similar to Scr's NTA, and its Thr6 also contacts the phosphate backbone.

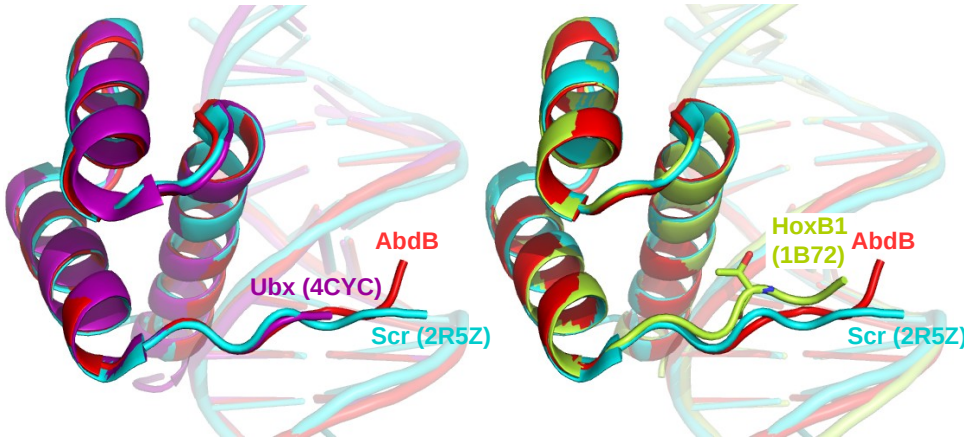
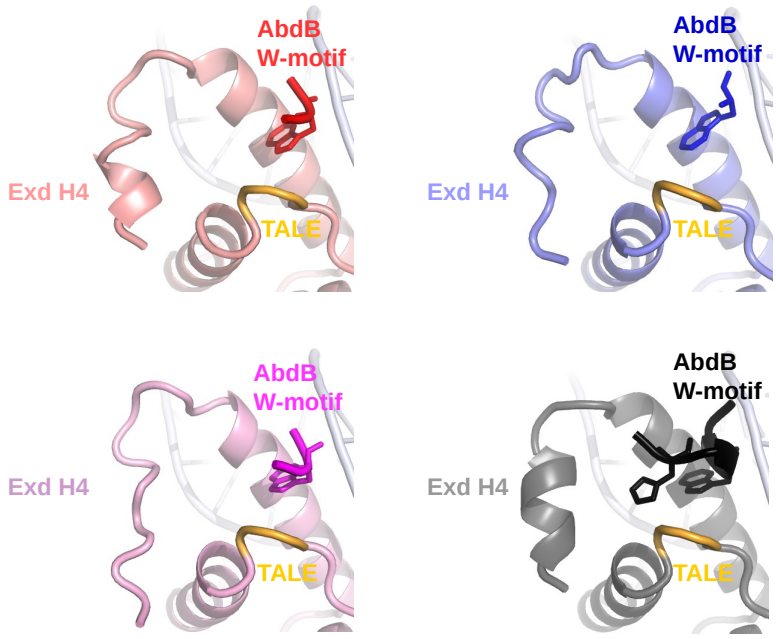


Figure S4. Comparison of helix 4 of Exd for all four AbdB-Exd ternary structures, showing its proximity to the W-motif-TALE interaction, related to Figure 5.



References:

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr *50*, 760-763.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

Sagendorf, J.M., Berman, H.M., and Rohs, R. (2017). DNAproDB: an interactive tool for structural analysis of DNA-protein complexes. Nucleic Acids Res *45*, W89-W97.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A.*, et al.* (2011). Overview of the *CCP* 4 suite and current developments. Acta Crystallographica Section D Biological Crystallography *67*, 235-242.