Homeodomains: together again for the first time

The first three co-crystal structures of homeodomain dimers show that there is more than one way for homeodomains to associate with each other on DNA.

Structure 15 November 1995, **3**:1143–1145

Development in higher organisms is a process of great complexity, but a pattern emerges when one considers the genes that regulate morphological fate. These genes were first discovered in *Drosophila*, through study of homeotic mutations [1]. More recently, remarkably similar clusters of genes (the HOM-C genes) have been found in organisms all the way from flies to humans [2]. The strong family resemblance of the products of these genes lies in the homeodomain, a segment of 60 amino acids that folds into three α helices; the part containing helices 2 and 3 closely resembles a helix-turn-helix motif. The homeodomain constitutes the DNA-binding unit of these proteins. Helix 3 is a 'recognition helix' which binds in the major groove of DNA. The N-terminus of the homeodomain forms an 'arm' (extending from helix 1), that binds to the adjacent minor groove.

The fact that it is possible to recognize a highly-conserved domain in a large class of DNA-binding proteins that regulate gene expression in developmental pathways, leads to an intriguing enigma. How can such closely-related proteins, that bind to very similar DNA sequences *in vitro* [3], enforce specific developmental programs *in vivo*? Although it might be thought that it is the rest of a homeoprotein that really provides specificity, homeodomain swap experiments have shown, in several cases, that it is the identity of the homeodomain itself that is important in selecting which developmental path is taken [4].

The three-dimensional structures of a handful of homeodomain–DNA complexes are now known [5–7]. One might be forgiven for saying "seen one, seen 'em all" after having a look these structures; they are remarkably alike, from the fold of the helix-turn-helix, to the contacts with the phosphates of the DNA backbone, to the hydrogen bonds made by the conserved residue Asn51 with an adenine in the DNA major groove. So, although there are differences among these structures, it is not easy to see where developmental specificity might lie.

Protein–protein interaction is the other likely possibility for distinguishing one homeodomain–DNA complex from another, in the cell [8]. Indeed, although many homeodomains are perfectly capable of binding to DNA with high affinity as monomers (and the previous three-dimensional structures are all of monomers), in living systems it is clear that homeoproteins often bind to DNA in clusters. Cooperative interaction among DNA-binding proteins is clearly a mechanism for ensuring specificity. So it is with great anticipation that we take our first high-resolution look at homeodomain dimers bound to DNA. In contrast to the uniformity of homeodomain monomer–DNA complexes, we find, from these three structures [9–11], that there is more than one way to arrange two homeodomains on the surface of DNA.

The three new co-crystal structures are of the DNA complexes of the Paired [9], Eve [10], and a1/α2 [11] homeodomains. Two structures are of homo-homeodomain dimers [9,10], whereas the third [11] provides the first example of a heterodimer of homeodomains bound to DNA. First, let's dispense with the similarities. The folds of the helix-turn-helix units of each homeodomain are nearly identical; superposition of the Cα positions of Eve or Paired with those of Engrailed [6], or of α2 with a1, leads to an rms difference of 1 Å or less. Similar contacts are found with the phosphates on either side of the major groove, into which the recognition helix of the homeodomain is inserted. Each homeodomain makes the Asn51–adenine contact that is, by now, expected; these could hardly be homeodomain co-crystal structures if these hydrogen bonds were absent! Also, in three of the four homeodomains (or five of the six, if we consider monomers to be distinguishable) Val47 apparently interacts hydrophobically with a thymine methyl group in the major groove, a feature seen in some previous structures [5,6].

So what are the differences? A look at Figure 1 reveals the most obvious one; each of the three structures shows an alternative arrangement of the two homeodomains on the DNA. The Eve complex (Fig. 1b), in some ways, is the simplest. Two Eve homeodomains bind in tandem to a 10 base-pair DNA site, consisting of the nearly perfect direct repeat AATTAAATTC. The two protein monomers are related by a screw axis, and bind on opposite faces of the DNA. The DNA remains notably straighter than it is in most other homeodomain complexes, probably as a consequence of having two homeodomains bound only a half-turn of the helix apart. The two Eve homeodomains do not seem to interact with each other, although Hirsch and Aggarwal speculate that such an interaction may involve parts of the complete Eve protein that are not included in the crystallized homeodomain [10].

© Current Biology Ltd ISSN 0969-2126

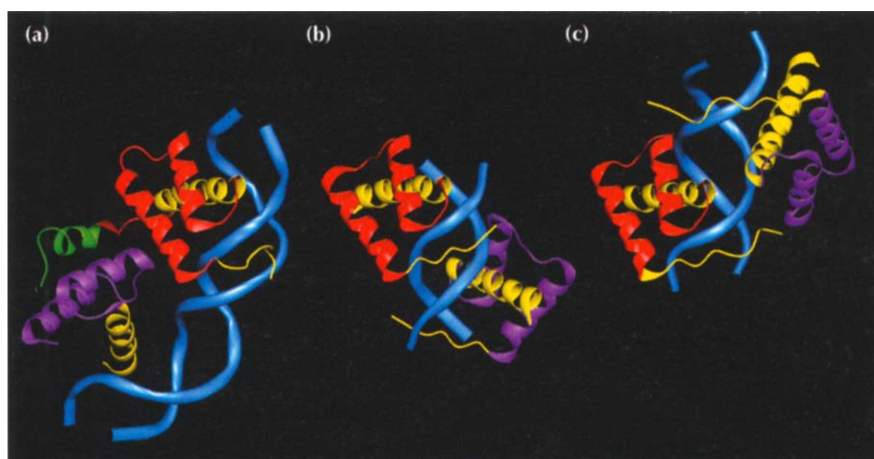

Fig. 1. Three homeodomain dimer–DNA complexes. **(a)** a1/α2; **(b)** Eve; **(c)** Paired. To make comparison easier, one homeodomain from each structure was superimposed. These homeodomains are colored red. The other homeodomain in each structure is colored purple. To highlight the DNA-binding portions of the homeodomains, the recognition helices and the N-terminal arms are colored yellow. In the a1/α2 structure, the C-terminal tail of α2 that binds to the back side of a1 is colored green.

In the Paired complex (Fig. 1c), the two homeodomains adopt a head-to-head orientation, with helix 2 of one Paired homeodomain linking up with the N-terminal arm of the other homeodomain, and *vice versa*. There is extensive interaction between the two protein monomers. Indeed, around 1000 $Å^2$ of surface area is buried in the complex. Dimerization is mediated by complementary surfaces in the two homeodomains, involving charged hydrogen bonds, hydrophobic interactions, and water-mediated hydrogen bonds. The DNA in the complex is distorted. In fact, it is not possible for the two Paired homeodomains to contact one another if the DNA is modeled as straight B-form DNA. Kuriyan and coworkers suggest that DNA distortion is likely to be induced by the first Paired monomer to bind [9]; they show that the TAAT sequence in one Paired-binding site is quite close in structure to that of the same sequence in the Engrailed co-crystal structure [6]. Even more interesting is the observation that outside the aligned TAAT sites the DNA in the two complexes also superimposes closely, even though there is only one homeodomain bound in the Engrailed complex.

The a1/α2 structure (Fig. 1a) belongs to the emerging class of co-crystal structures in which two different proteins are bound to DNA [12], and so gives us a view of one of the inherently asymmetric protein–DNA complexes that are commonly found in gene regulatory systems in higher organisms. Both a1 and α2 are homeodomains, but they bind on their own to DNA with, at best, modest affinity. Put them together, though, and a high-specificity and high-affinity complex results. These two homeodomains have come up with a third distinctive way to bind as a (hetero)dimer to DNA, that is, head-to-tail, as for Eve, but adjacent to each other on one side of the DNA helix.

The dimerization interface for the a1/α2 complex also is unique, and potentially a harbinger of protein–protein interfaces in other systems. Unlike Paired, there are no homeodomain–homeodomain contacts. Instead, dimerization is mediated by a segment of the α2 protein, C-terminal to the homeodomain, that is disordered both for the protein free in solution and bound to DNA as a monomer. In the a1/α2 complex, the C-terminal tail folds into a (partial) α helix that binds to a complementary surface on the a1 protein. Hydrophobic interactions and hydrogen bonds hold together a three-helix bundle that is formed from helices 1 and 2 of a1 and the C-terminal extension of α2. The structure might be thought of as α2 putting its (C-terminal) arm around the back of a1.

The DNA in the a1/α2 complex is even more distorted than in the Paired complex, taking on a smooth 60° curvature (see Fig. 1). Again, without this bend the two homeodomains could not interact. Remarkably, in the structure of the α2 monomer–DNA complex [7] the DNA is nearly straight, suggesting that the strong bend is the consequence of the formation of the a1/α2 heterodimer on DNA.

There are, of course, other points of interest in the new structures once one goes beyond the overall arrangement of homeodomain dimers on DNA. For example, the higher resolution of these three structures compared with previous homeodomain co-crystal structures has made it possible to pick up a number of water-mediated protein–DNA interactions, that were first apparent in the NMR structure of the Antennapedia–DNA complex [5]. It seems that the DNA major groove, in which the recognition helix of the homeodomain lies, is rather wet, leading perhaps to alternative ways in which the amino acid side chains of the homeodomain can interact with the DNA bases. This is clearly true for the residue at position 50, which had been assigned, previously, as a source of sequence specificity [13]. In the Eve structure three different orientations of the Gln50 side chain are seen, making both direct and water-mediated hydrogen bonds to the DNA bases, whereas in the Paired complex a conformation similar to that found earlier, for Engrailed, is observed.

Implications of the new homeodomain dimer structures

These three complexes offer a structural paradigm not only for homeodomain dimers, but perhaps also for how other helix-turn-helix proteins might bind relative to one another on DNA. For example, the POU proteins consist of a homeodomain linked by a flexible peptide chain to a

helix-turn-helix domain related in structure to the helix-turn-helix domain of the phage repressors. The co-crystal structure of the POU protein Oct-1 [14] shows the two DNA-binding domains on opposite sides of the DNA helix, similar to the Eve structure. But there seems to be no compelling reason for this to be the only possible arrangement of the two domains; why not adopt a structure like Paired, in which the recognition helices of the POU homeodomain and POU-specific domain follow each other around the major groove of DNA (like zinc fingers do [15])? Only time, and new crystals, will tell.

References

1. Lewis, E.B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570.
2. McGinnis, W. & Krumlauf, R. (1992). Homeobox genes and axial patterning. *Cell* **68**, 283–302.
3. Ekker, S.C., Young, K.E., von Kessler, D.P. & Beachy, P.A. (1991). Optimal DNA sequence recognition by the Ultrabithorax homeodomain of *Drosophila*. *EMBO J.* **10**, 1179–1186.
4. Lin, L. & McGinnis, W. (1992). Mapping functional specificity in the Dfd and Ubx homeodomains. *Genes Dev.* **6**, 1071–1081.
5. Billeter, M., Qian, Y.Q., Otting, G., Müller, M., Gehring, W.J. & Wüthrich, K. (1993). Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain–DNA complex. *J. Mol. Biol.* **234**, 1084–1097.
6. Kissinger, C.R., Liu, B., Martin-Blanco, E., Kornberg, T.B. & Pabo, C.O. (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.
7. Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D. & Pabo, C.O. (1991). Crystal structure of a MATα2 homeodomain–operator complex suggests a general model for homeodomain–DNA interactions. *Cell* **67**, 517–528.
8. Manak, J.R. & Scott, M.P. (1993). Able assistants for homeodomain proteins. *Curr. Biol.* **3**, 318–320.
9. Wilson, D.S., Guenther, B., Desplan, C. & Kuriyan, J. (1995). High resolution crystal structure of a Paired (Pax) class cooperative homeodomain dimer on DNA. *Cell* **82**, 709–719.
10. Hirsch, J.A. & Aggarwal, A.K. (1995). Structure of even skipped homeodomain complexed to AT rich DNA: new perspectives on homeodomain specificity. *EMBO J.*, in press.
11. Li, T., Stark, M.R., Johnson, A.D. & Wolberger, C. (1995). Structure of the MATa1/MATα2 homeodomain heterodimer bound to DNA. *Science* **270**, 262–269.
12. Nikolov, D.B., *et al.*, & Burley S.K. (1995). Crystal structure of a TFIIB–TBP–TATA-element complex. *Nature* **377**, 119–128.
13. Hanes, S.D. & Brent, R. (1991). A genetic model for interaction of the homeodomain recognition helix with DNA. *Science* **251**, 426–430.
14. Klemm, J.D., Rould, M.A., Aurora, R., Herr, W. & Pabo, C.O. (1994). Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**, 21–32.
15. Pavletich, N.P. & Pabo, C.O. (1991). Zinc finger–DNA recognition: crystal structure of a Zif268–DNA complex at 2.1 Å. *Science* **252**, 809–817.

Tom Tullius, Department of Chemistry, The Johns Hopkins University, Baltimore, MD 21218, USA.