

COMMUNICATION

The DNA Binding Specificity of Engrailed Homeodomain

Alexandra Draganescu and Thomas D. Tullius*

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

The *engrailed* gene of *Drosophila melanogaster* is an integral member of the highly complex cascade which results in a fully developed fruitfly. The gene product of *engrailed* contains a homeodomain which is responsible for DNA binding *via* a helix-turn-helix motif. The crystal structure of this 60 amino acid residue domain complexed to DNA is analogous to structures of other homeodomain-DNA complexes, consistent with the high degree of sequence conservation within both protein and DNA. Despite the high degree of homology, homeodomains do exhibit distinct preferences for certain DNA sequences. Such specificity may be at least partly responsible for the interactions necessary for normal development. Using the hydroxyl radical as a chemical probe, we have examined complexes of Engrailed homeodomain with several DNA sequences to determine the protein's binding specificity in solution. We find that Engrailed forms a single, specific complex with a unique DNA binding site which is analogous to the complex seen in the co-crystal structure. In contrast, our chemical probe experiments show that the binding site of Engrailed that was determined by *in vitro* selection and that also was present in the co-crystal structure contains two possible binding sites. Modification of the sequence of this site to yield single binding sites removes the ambiguity, and results in two different, well-behaved Engrailed-DNA complexes. Our results underscore the utility of chemical probe experiments for defining the variety of modes of interaction of proteins with DNA that can occur in solution, but that might not be apparent in a crystal structure.

© 1998 Academic Press Limited

Keywords: hydroxyl radical footprinting; missing-nucleoside experiment; methylation interference

*Corresponding author

The homeodomain is a highly conserved sequence motif often found in proteins which regulate development in eukaryotes (Scott *et al.*, 1989; Morata, 1993). Many homeodomain-containing proteins have been shown to bind to DNA and thereby regulate transcription. The 60 amino acid residues of the homeodomain fold into three alpha helices and an unstructured amino-terminal "arm". Three-dimensional structural studies of homeodomain-DNA complexes (Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Hirsch &

Aggarwal, 1995; Li *et al.*, 1995; Wilson *et al.*, 1995) have revealed that helix 3 of the homeodomain forms part of a helix-turn-helix motif and binds in the DNA major groove. Several amino acid residue side-chains of the homeodomain make specific contacts with the DNA backbone and bases. The N-terminal "arm" of the homeodomain interacts with DNA bases in the adjacent minor groove.

An enigmatic aspect of the homeodomain is the role of the sequence-specificity of DNA binding in specifying developmental programs (Manak & Scott, 1993; Chan & Mann, 1996; Li *et al.*, 1996). Homeodomains are closely related at the level of protein sequence, and many homeodomains recognize similar DNA sites having the common core sequence TAAT (Ekker *et al.*, 1991, 1992). Given these similarities, then, how can one homeodomain specify a particular developmental pathway at the

Present addresses: A. Draganescu, Kimmel Cancer Institute, Thomas Jefferson University, Philadelphia PA 19107, USA; T. D. Tullius, Department of Chemistry, Boston University, Boston, MA 02215, USA.

Abbreviations used: EDTA, ethylenediaminetetraacetic acid; Dfd, Deformed protein; Ubx, Ultrabithorax protein.

expense of another? Since the DNA-binding homeodomain constitutes only part of the whole developmental regulatory protein, a mechanism for increasing specificity might involve the interaction of other proteins with the remainder of the homeodomain-containing protein (Manak & Scott, 1993; Chan & Mann, 1996). Indeed, for some homeoproteins, protein-protein interactions have been shown to be important for developmental regulation (Peltenburg & Murre, 1996, 1997). But other experiments (homeodomain swaps, for example (Lin & McGinnis, 1992)) have provided evidence that the DNA binding specificity of the homeodomain itself (Ekker *et al.*, 1992; Li *et al.*, 1996) is crucial to developmental specificity.

Insight into how homeodomains recognize specific DNA sequences has come from high-resolution X-ray and NMR structures (Kissinger *et al.*, 1990; Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Hirsch & Aggarwal, 1995; Li *et al.*, 1995; Wilson *et al.*, 1995), mutational studies (Hanes & Brent, 1991), and *in vitro* selection experiments (Ekker *et al.*, 1991, 1992; Ades & Sauer, 1994). We have become interested in a related problem, how a homeodomain recognizes and binds to a set of similar DNA sequences, since this is how the homeodomain must function in a biological system.

Engrailed is a segment-polarity gene found in *Drosophila melanogaster* whose product directs proper segmental subdivisions in the embryo (Hidalgo, 1996). The *Engrailed* protein acts as a transcriptional repressor (Han & Manley, 1993). *Engrailed* regulates the activity of other homeobox genes such as *ultrabithorax* (Mann, 1994), and *in vitro* transcription mediated by *fushi tarazu* or transcription factor IID (Ohkuma *et al.*, 1990). *Engrailed* and the *Drosophila* extradenticle protein can form a heterodimer and bind cooperatively to DNA (Peltenburg & Murre, 1996). At the carboxyl terminus of the *Engrailed* protein is a homeodomain. The *Engrailed* homeodomain has been crystallized alone (Clarke *et al.*, 1994) and in the presence of DNA (Kissinger *et al.*, 1990). Subsequent *in vitro* selection experiments found that the optimal binding site for *Engrailed* contains the sequence 5'-TAATTA-3' (Ades & Sauer, 1994), which also was present in the co-crystal DNA binding site (Kissinger *et al.*, 1990).

Since *Engrailed* was the first homeodomain for which the co-crystal structure was determined, it has become a model for homeodomain-DNA complexes. Indeed the complex of the *Engrailed* homeodomain with DNA has a similar overall structure to those observed for other homeodomain-DNA complexes (Wolberger *et al.*, 1991; Billeter *et al.*, 1993; Hirsch & Aggarwal, 1995; Li *et al.*, 1995; Wilson *et al.*, 1995). We became interested in characterizing how *Engrailed* binds to DNA in solution, in order to use it as a structurally characterized model for other homeodomain-DNA complexes we were studying using solution probe methods (Draganescu *et al.*, 1995). Our results,

though, reveal a surprising complexity in the selection of a DNA binding site in solution by *Engrailed*, which is not reflected in the X-ray co-crystal structure.

We first used hydroxyl radical footprinting to determine the solution structure of the complex of *Engrailed* homeodomain with a DNA sequence related to the one used in the co-crystal (which we call here the "co-crystal binding site"). This DNA molecule contains the high-affinity central binding site found in the *Engrailed* homeodomain-DNA co-crystal (Kissinger *et al.*, 1990). The sequence of the whole DNA molecule we studied is not precisely the same as that used in the co-crystal structure. We eliminated the weaker second binding site seen in the co-crystal by putting a G·C base-pair at position 2 and a C·G base-pair at position 4 (see Figure 1 for base-pair numbering). A gel mobility-shift experiment showed that *Engrailed* makes a single complex with this sequence, with a dissociation constant of ~15 nM (data not shown; Senear & Brenowitz, 1991).

Densitometric scans of both DNA strands from the hydroxyl radical footprinting experiment are shown in Figure 1(a). We were surprised to find that the positions of strongest protection did not correspond exactly to phosphate contacts assigned in the co-crystal structure, G5-T7 on the top strand and A10-T12 on the bottom strand (Kissinger *et al.*, 1990). The protected regions also are inconsistent with those observed by hydroxyl radical footprinting for the complex of either the Deformed (Dfd) or Ultrabithorax (Ubx) homeodomain with its optimal site (Draganescu *et al.*, 1995).

The *Engrailed* co-crystal binding site contains overlapping 5'-TAAT-3' and 5'-ATTA-3' sequences, which in principle could constitute parts of two separate binding sites. Our footprinting experiment suggests that in solution the *Engrailed* homeodomain is unable to distinguish between the two overlapping TAAT core sequences in the inversion-symmetric sequence 5'-TAATTA-3' in the co-crystal site, even though the sequences flanking these overlapping sites are different (Figure 1). Instead, *Engrailed* appears to be capable of binding to one or the other of the two sites in a given DNA molecule, yielding a composite footprint which is inconsistent with the backbone contacts observed for this sequence in the co-crystal structure.

To test this hypothesis, we constructed two new sequences which differ at a single base-pair from each other and from the co-crystal site. One sequence contains a C·G base pair at position 8, replacing an A·T base pair in the co-crystal site, resulting in a binding site having just the single 5'-TAAT-3' core to which *Engrailed* homeodomain is bound in the co-crystal. We call this the "co-crystal core" site. The other new sequence contains a G·C instead of a T·A base-pair at position 11, leaving the other potential TAAT core intact. We call this the "other TAAT core" site.

Hydroxyl radical footprints clearly show that the *Engrailed* homeodomain binds to only a single site

in each of the two new sequences (Figure 1(b) and (c)). There are three regions of protection for each sequence, two on one strand and one on the other. For the co-crystal core site (Figure 1(b)), the strongest protection occurs on the bottom strand at positions T9–T12. On the top strand, moderate protection is centered at positions C4–G6 in region I and A12–C13 in region II. A similar pattern is observed in the footprint of Engrailed on the other TAAT core site (Figure 1(c)). Strong protection is apparent in region III, this time on the top strand, at positions T7–A9. There also is strong protection of positions C6–A7 in region II, and G13 in region I. Each of these footprints is consistent with the expected protection pattern based on the co-crystal structure and with the protection patterns previously observed for Ubx and Dfd (Draganescu *et al.*, 1995). We note that the footprinting patterns for the two new sequences are inverted relative to each other, as would be expected since the TAAT sequence is on opposite strands for the two sequences.

To further study the interaction of Engrailed with these three binding sites, we used the missing-nucleoside experiment to determine the nucleoside-specific contacts made by Engrailed. In this experiment (Hayes & Tullius, 1989), radiolabeled DNA containing the binding site of interest is randomly gapped with hydroxyl radical and then incubated with the DNA binding protein. The bound and unbound DNA populations are separated using non-denaturing gel electrophoresis. DNA from the bound and unbound fractions is isolated and then subjected to electrophoresis on a denaturing polyacrylamide gel. Nucleosides important for protein binding are revealed in the denaturing gel as intense bands in the unbound DNA fraction. These important nucleosides are correspondingly depleted in the bound fraction.

Missing-nucleoside experiments with the co-crystal binding site (Figure 2(a)) reveal an extensive interference pattern which is inconsistent with the specific contacts assigned in the X-ray co-crystal structure. Removal of any nucleoside, on either strand, between positions 4 and 11 is detrimental to protein binding. This result is quite different from that observed for the Ubx and Dfd homeodomains binding to their optimal DNA sites (Draganescu *et al.*, 1995). Those interference patterns closely matched specific protein-DNA interactions assigned by NMR and crystallography for homologous homeodomain-DNA complexes.

We therefore judge it to be unlikely that a single Engrailed homeodomain actually makes specific contacts with both DNA strands over a nine base-pair region, as might be concluded from a cursory examination of the missing nucleoside pattern in Figure 2(a). This unusual interference pattern, like the complex footprint discussed above (Figure 1(a)), instead suggests to us that the Engrailed homeodomain is unable to distinguish between the two overlapping binding sites within the 5'-TAATTA-3' sequence in the co-crystal binding site.

We also used the missing-nucleoside experiment to examine the binding of Engrailed homeodomain individually to the two TAAT sites contained within the co-crystal binding site. The interference signals observed for the co-crystal core (Figure 2(b)) and other TAAT core (Figure 2(c)) binding sites are different from those found for the co-crystal site (Figure 2(a)), but similar to the patterns previously observed for Ubx and Dfd (Draganescu *et al.*, 1995).

Since our initial experiments showed that Engrailed was apparently capable of binding in a single orientation to a site containing a single TAAT sequence (see Figure 1(b) and (c)), we decided to examine the complex of Engrailed homeodomain with an unrelated homeodomain binding site which also had one TAAT sequence. This sequence had been shown by *in vitro* selection to be the optimal binding site for the Ubx homeodomain (Ekker *et al.*, 1991, 1992). We previously found that this sequence forms well-behaved complexes with both the Ubx and Dfd homeodomains (Draganescu *et al.*, 1995).

The hydroxyl radical footprint of Engrailed bound to the Ubx-optimal sequence (data not shown; see Figure 3) is highly similar to those we observed in our earlier studies of the Ubx and Dfd homeodomains (Draganescu *et al.*, 1995). This footprint also is very similar to those seen with the "single sites" produced by mutating the co-crystal sequence (Figure 1(b) and (c)), and very different from the footprint that Engrailed makes with the co-crystal site itself (Figure 1(a)).

A "signature" hydroxyl radical footprinting pattern is thus apparent for homeodomain-DNA complexes (Sauer *et al.*, 1988; Wolberger *et al.*, 1991; Draganescu *et al.*, 1995). The homeodomain footprint consists of two protected regions, separated by seven or eight nucleotides, on one DNA strand, and a third footprint on the other strand between the first two. This third footprint is offset by five or six nucleotides in the 5' direction from the 5'-most footprint on the other strand, and by two or three nucleotides in the 3' direction from the 3'-most footprint on the opposite strand.

We also examined the missing nucleoside interference pattern for Engrailed homeodomain binding to the Ubx-optimal site (data not shown; see Figure 3). Removal of any nucleoside, on either strand, within the TAAT core sequence affects protein binding. There is a moderate effect upon removal of T9 and T10, but removal of any other nucleoside within the core has a disastrous effect on protein binding. Strong signals are also observed at C7 and A13.

Summaries of the chemical probe data for Engrailed binding to the Ubx-optimal site, and the contacts found in the Engrailed co-crystal structure, are shown in Figure 3. One footprint corresponds to phosphate contacts that are made by Arg31, Arg53, Tyr25, and Lys57. This footprint corresponds to region I in the footprints shown in Figure 1. A second footprint (corresponding to

region III) fits well with the backbone contacts made by Thr6, Trp48, and Lys55. This comparison suggests that the amino acid residues of Engrailed which make phosphate contacts in the co-crystal structure are responsible for the DNA backbone protections observed in the footprinting experiment with the Ubx-optimal site.

As a final test of our hypothesis that the footprint we observe for Engrailed on the co-crystal site (Figure 1(a)) is really the superposition of two

footprints, we decided to attempt to simulate this footprint by combining the footprints we determined for the two individual TAAT sites (Figure 1(b) and (c)). The simulated footprint is an excellent match for the experimental footprint (compare Figure 1(a) and Figure 4). We emphasize that we made no attempt to optimize the simulated footprint by varying the contributions of the two component footprints; we merely added them in equal proportions. Nonetheless, the correspon-

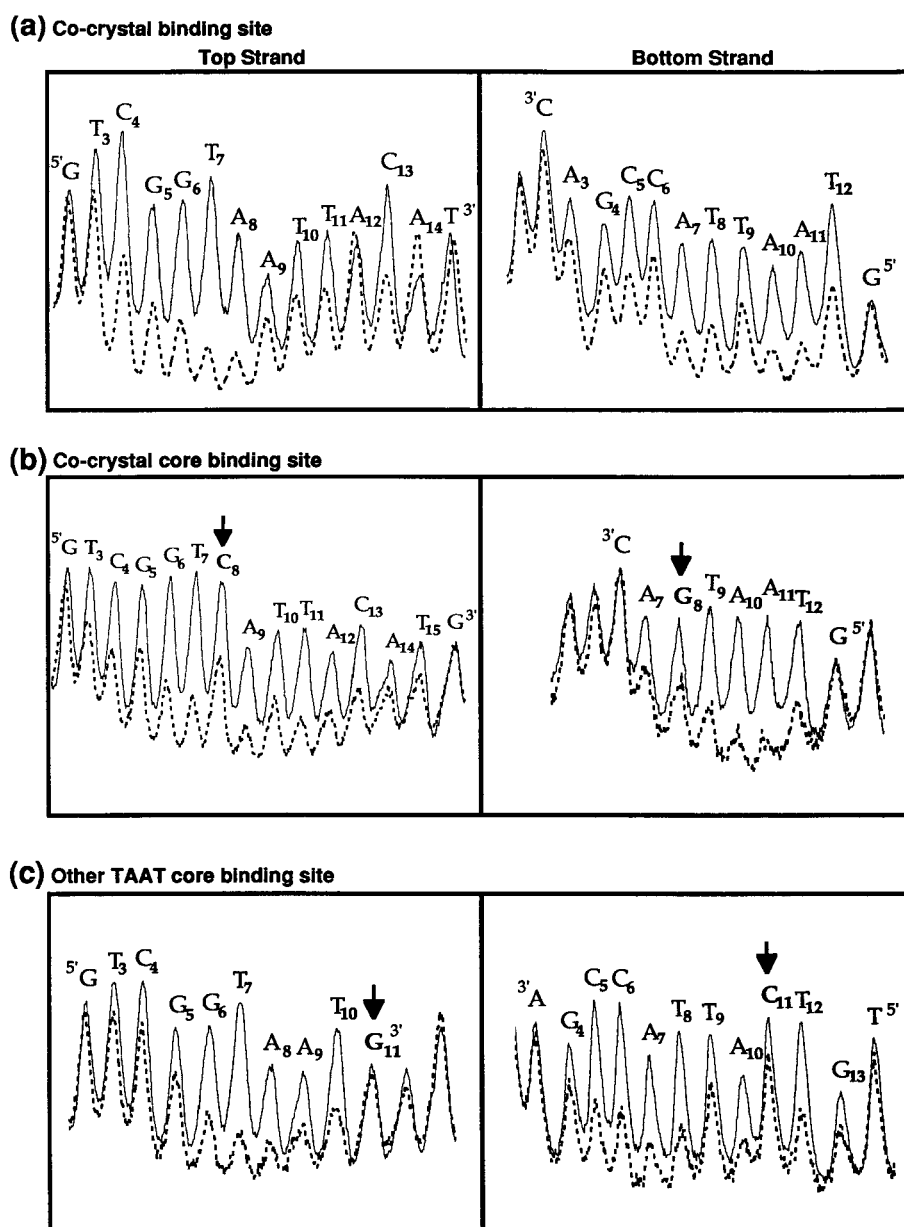


Figure 2. Scans from missing nucleoside experiments for Engrailed homeodomain binding to a DNA molecule containing (a) the co-crystal site; (b) the co-crystal core binding site; and (c) the other TAAT core binding site. Continuous line, hydroxyl radical cleavage pattern of free DNA (the input DNA in the missing nucleoside experiment); broken line, DNA recovered from the bound fraction. Base-pair 8 in the co-crystal core binding site, and base-pair 11 in the other TAAT core binding site, are marked by arrows. These are the sites of the base pair substitutions made to isolate each of the two possible binding sites in the co-crystal sequence. Missing nucleoside experiments (Hayes & Tullius, 1989) were carried out as described (Draganescu *et al.*, 1995). Native gel electrophoresis was performed as for the footprinting experiments described in the legend to Figure 1.

dence between the observed and simulated footprints demonstrates that the two alternative Engrailed complexes with the co-crystal site must be present in roughly equimolar amounts, despite differences in the sequences flanking the two possible TAAT sequences.

In vitro selection experiments identified the sequence 5'-TAATTA-3' as the optimal DNA binding site for Engrailed homeodomain (Ades & Sauer, 1994). However, results from footprinting and missing nucleoside experiments presented here show that two different complexes are possible with this sequence. How can these apparently conflicting results be reconciled? The *in vitro*-selected binding site sequence has inversion symmetry and could be the result of sequence selection by Engrailed protein bound in two different orientations. Once the 5'-TAATTA-3' binding site had been selected, all subsequent assays of binding affinity of Engrailed homeodomain for optimal and suboptimal sites used mobility shift gels (Ades & Sauer, 1994). Since the two possible TAAT core sequences overlap, only a single protein molecule may bind at a time. Therefore only a single shifted band would be observed, which corresponds to a 1:1 complex. We also observed only a single shifted band in binding studies using the co-crystal binding site. It is only when the results of footprinting or missing nucleoside experiments are

examined that the possibility of Engrailed binding in two orientations becomes apparent.

Finally, we presume that symmetry-breaking which occurs during crystallization is the reason why only one orientation of Engrailed bound to DNA is observed in the co-crystal structure. In other words, once one particular Engrailed-DNA complex begins to crystallize, similar complexes are selected out of the mixture of orientations in solution during the formation of well-ordered crystals.

In summary, hydroxyl radical footprinting and missing-nucleoside experiments reveal that in solution the Engrailed homeodomain binds to a DNA molecule containing a single TAAT sequence in an analogous manner to that observed in the X-ray co-crystal structure (Figure 3). We also find that Engrailed binds to a more complex site, having overlapping TAAT sequences (5'-TAATTA-3'), in more than one orientation. This result is of particular significance since it is this second sequence which was used for the co-crystal structure of Engrailed, and which was selected as the optimum binding site by *in vitro* selection experiments. Our results illustrate how the interaction of a protein with DNA in solution can be more complicated than is apparent from a crystal structure. This work also demonstrates the value of chemical probe experiments in characterizing high-affinity

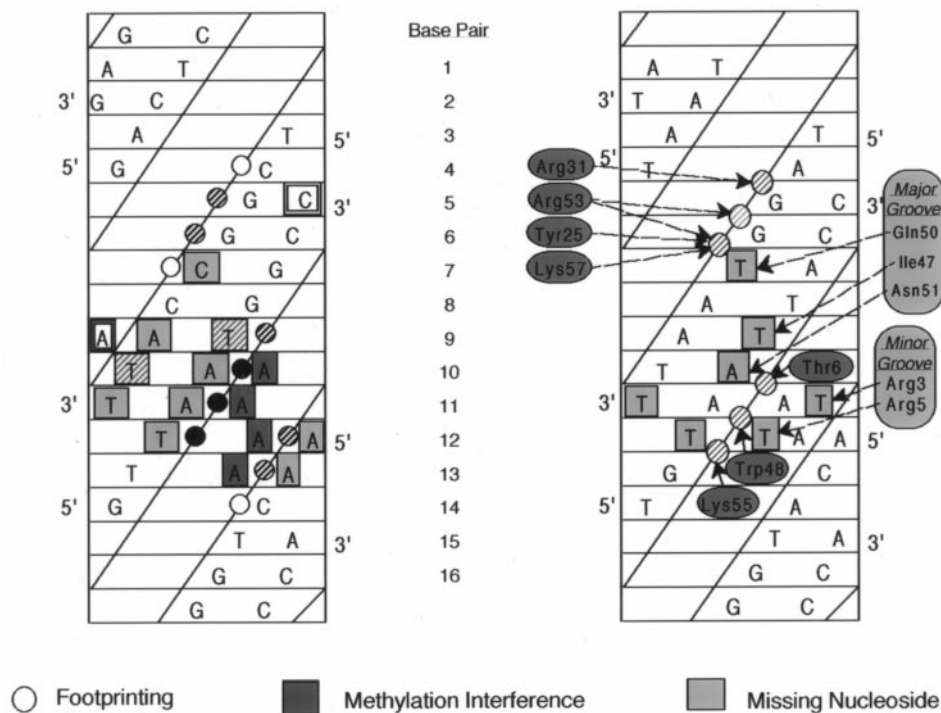


Figure 3. Left, helical projection summarizing hydroxyl radical footprinting (filled circles, strong; hatched circles, moderate; open circles, weak), missing nucleoside (light gray filled boxes, strong; hatched boxes, moderate; light gray outlined boxes, weak), and methylation interference (dark gray filled boxes, strong; dark gray outlined boxes, weak) data for Engrailed homeodomain bound to the Ubx-optimal sequence. Right, helical projection summarizing base-specific contacts (light gray filled boxes) and phosphate contacts (hatched circles) observed in the Engrailed homeodomain co-crystal structure (Kissinger *et al.*, 1990). The amino acid residue of Engrailed that was assigned as making a contact is connected by an arrow to the projection of the DNA helix.

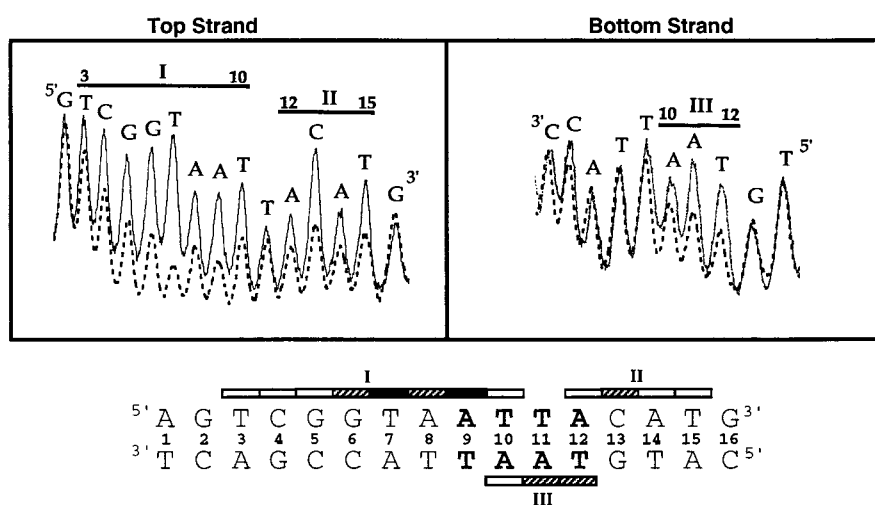


Figure 4. Simulated hydroxyl radical footprint for Engrailed homeodomain binding to the co-crystal site. Continuous line, experimental hydroxyl radical cleavage pattern of free DNA; broken line, simulated cleavage pattern of the Engrailed homeodomain-DNA complex. The simulated footprint is represented schematically below the scans. Densitometric scans of footprint gel lanes for Engrailed bound to the co-crystal core site and to the other TAAT core site (see Figure 1(b) and (c)) were scaled to overlay precisely in regions outside the protein binding site, using Microsoft Excel. The two scaled scans were then added together point by point to generate the simulated footprint (broken line).

DNA binding sites that are recovered by *in vitro* selection. We conclude that although the sequence 5'-TAA-3' might be sufficient to bind Engrailed homeodomain (Pan *et al.*, 1995), this protein can show differential recognition of highly similar binding sites.

Acknowledgments

We thank Dr Thomas Kornberg for generously providing the expression vector for the Engrailed protein. This research was supported by PHS grant GM 41930. We gratefully acknowledge the use of densitometry instrumentation maintained by the Institute for Biophysical Research on Macromolecular Assemblies at Johns Hopkins, which was supported by an NSF Biological Research Centers Award (DIR-8721059) and by a grant from the W. M. Keck Foundation.

References

- Ades, S. E. & Sauer, R. T. (1994). Differential DNA-binding specificity of the Engrailed homeodomain: the role of residue 50. *Biochemistry*, **33**, 9187–9194.
- 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.
- Bourbon, H. M., Martin-Blanco, E., Rosen, D. & Kornberg, T. B. (1995). Phosphorylation of the *Drosophila* Engrailed protein at a site outside its homeodomain enhances DNA binding. *J. Biol. Chem.* **270**, 11130–11139.
- Chan, S. K. & 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. USA*, **93**, 5223–5228.
- Clarke, N. D., Kissinger, C. R., Desjarlais, J., Gilliland, G. L. & Pabo, C. O. (1994). Structural studies of the Engrailed homeodomain. *Protein Sci.* **3**, 1779–1787.
- Dixon, W. J., Hayes, J. J., Levin, J. R., Weidner, M. F., Dombroski, B. A. & Tullius, T. D. (1991). Hydroxyl radical footprinting. *Methods Enzymol.* **208**, 380–413.
- Draganescu, A., Levin, J. R. & Tullius, T. D. (1995). Homeodomain proteins: what governs their ability to recognize specific DNA sequences? *J. Mol. Biol.* **250**, 595–608.
- 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.
- Ekker, S. C., von Kessler, D. P. & Beachy, P. A. (1992). Differential DNA sequence recognition is a determinant of the specificity in homeotic gene action. *EMBO J.* **11**, 4059–4072.
- Han, K. & Manley, J. L. (1993). Functional domains of the *Drosophila* Engrailed protein. *EMBO J.* **12**, 2723–2733.
- Hanes, S. D. & Brent, R. (1991). A genetic model for interaction of the homeodomain recognition helix with DNA. *Science*, **251**, 426–430.
- Hayes, J. J. & Tullius, T. D. (1989). The missing nucleoside experiment: a new technique to study recognition of DNA by protein. *Biochemistry*, **28**, 9521–9527.
- Hidalgo, A. (1996). The roles of *engrailed*. *Trends Genet.* **12**, 1–4.
- Hirsch, J. A. & Aggarwal, A. K. (1995). Structure of the even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. *EMBO J.* **14**, 6280–6291.
- 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.

- Levin, J. R., Burkhoff, A. M. & Tullius, T. D. (1991). Using the chemistry of the hydroxyl radical to determine structural details about DNA and protein-DNA complexes. In *A Laboratory Guide to In Vitro Studies of Protein/DNA Interactions* (Saluz, H. P. & Jost, J. P., eds), BioMethods Series, vol. 5, pp. 133–144, Birkhäuser Verlag, Basel.
- Li, L., von Kessler, D., Beachy, P. A. & Matthews, K. S. (1996). pH-dependent enhancement of DNA binding by the Ultrabithorax homeodomain. *Biochemistry*, **35**, 9832–9839.
- Li, T., Stark, M. R., Johnson, A. D. & Wolberger, C. (1995). Structure of the MATa1/MAT α 2 homeodomain heterodimer bound to DNA. *Science*, **270**, 262–269.
- Lin, L. & McGinnis, W. (1992). Mapping functional specificity in the Dfd and Ubx homeodomains. *Genes Dev.* **6**, 1071–1081.
- Manak, J. R. & Scott, M. P. (1993). Able assistants for homeodomain proteins. *Curr. Biol.* **3**, 318–320.
- Mann, R. S. (1994). *Engrailed*-mediated repression of *Ultrabithorax* is necessary for the parasegment 6 identity in *Drosophila*. *Development*, **120**, 3205–3212.
- Morata, G. (1993). Homeotic genes of *Drosophila*. *Curr. Opin. Genet. Dev.* **3**, 606–614.
- Ohkuma, Y., Horikoshi, M., Roeder, R. G. & Desplan, C. (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.
- Pan, C. Q., Landgraf, R. & Sigman, D. S. (1995). *Drosophila* *Engrailed*-1,10-phenanthroline chimeras as probes of homeodomain-DNA complexes. *Protein Sci.* **4**, 2279–2288.
- Peltenburg, L. T. & Murre, C. (1996). *Engrailed* and Hox homeodomain proteins contain a related Pbx interaction motif that recognizes a common structure present in Pbx. *EMBO J.* **15**, 3385–3393.
- Peltenburg, L. T. & Murre, C. (1997). Specific residues in the Pbx homeodomain differentially modulate the DNA-binding activity of Hox and *Engrailed* proteins. *Development*, **124**, 1089–1098.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sauer, R. T., Smith, D. L. & Johnson, A. D. (1988). Flexibility of the yeast α 2 repressor enables it to occupy the ends of its operator, leaving the center free. *Genes Dev.* **2**, 807–816.
- Scott, M. P., Tamkun, J. W. & Hartzell, G. W. (1989). The structure and function of the homeodomain. *Biochim. Biophys. Acta*, **989**, 25–48.
- Senear, D. F. & Brenowitz, M. (1991). Determination of binding constants for cooperative site-specific protein-DNA interactions using the gel mobility-shift assay. *J. Biol. Chem.* **266**, 13661–13671.
- 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.
- Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D. & Pabo, C. O. (1991). Crystal structure of a MAT α 2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell*, **67**, 517–528.

Edited by R. Ebright

(Received 13 January 1997; received in revised form 4 December 1997; accepted 4 December 1997)