

COMMENTARY

The DNA-binding homeo domain of the Oct-2 protein

The specific interactions of a ubiquitous nuclear factor Oct-1 (also called OTF-1, NF-A1, NFIII, or OBP100) and of a lymphocyte-specific cellular protein Oct-2 (also called OTF-2 or NF-A2) with the regulatory octamer DNA element ATGCAAAT have been characterized extensively (Landolfi et al. 1986; Singh et al. 1986; Sive and Roeder 1986; Staudt et al. 1986; Bohmann et al. 1987; Eaton and Calame 1987; Fletcher et al. 1987; Scheidereit et al. 1987; Sturm et al. 1987). The octamer element is found in a large number of promoters and enhancers that are served by either RNA polymerase II or RNA polymerase III (Falkner and Zachau 1984; Parslow et al. 1984; for reviews, see Folk 1988 and Sollner-Webb 1988), and it has been shown to be a functional transcription regulatory element by deletion analysis of immunoglobulin promoters (Bergman et al. 1984; Mason et al. 1985). Mutagenesis experiments have demonstrated that the octamer element is responsible for the lymphoid-specific activation of immunoglobulin promoters (Wirth et al. 1987), and the B-lymphocyte-specific expression of Oct-2 has implicated it as a cell-type-specific *trans*-activator of these genes (Landolfi et al. 1986; Staudt et al. 1986; Gerster et al. 1987; Scheidereit et al. 1987; LeBowitz et al. 1988). Recently, we (Clerc et al. 1988; Staudt et al. 1988) and others (M.M. Müller et al. 1988; Scheidereit et al. 1988) have isolated cDNAs that encode Oct-2, and have demonstrated by Northern analysis that mRNAs encoding the cloned protein are restricted to cells of lymphoid origin (Clerc et al. 1988; M.M. Müller et al. 1988; Staudt et al. 1988). Furthermore, expression of the Oct-2 cDNA in non-B lymphocytes activates promoters containing the octamer motif (M.M. Müller et al. 1988; for review, see Schreiber et al. 1989). Finally, analysis of the Oct-2 predicted amino acid sequence has revealed a series of conserved sequences: a homeo domain (Gehring 1987), a POU box (Herr et al. 1988), a leucine repeat (Landschultz et al. 1988), and several regions rich in either acidic amino acids or glutamines (Courey and Tjian 1988; Hope et al. 1988; Ptashne 1988).

The region of the protein that is involved in DNA binding to the octamer DNA element has been shown to include a 60-amino-acid segment that is 31% homologous to the *Antennapedia* homeo domain (Schneuwly et al. 1986; Clerc et al. 1988; Ko et al. 1988). The homeo domain is thought to contain a helix–turn–helix structure similar to that of bacterial repressor proteins (Laughon and Scott 1984). We will argue here that this structural homology can be extended to the tri- α -helical

DNA-binding domain of prokaryotic repressors. This argument is based on the comparison of conserved amino acid sequences among homeo domains and several bacterial repressors, and on the atomic structure of repressor–operator co-crystals. The implication is that the entire homeo domain is the DNA-binding domain.

Comparison of Oct-2 with well-characterized prokaryotic proteins containing helix–turn–helix structures

The homeo domains, including less conserved versions such as those found in the yeast mating type (Mat) proteins, show remarkable conservation of hydrophobic amino acids known to be important in folding and stabilizing the helix–turn–helix structure of prokaryotic repressors (Laughon and Scott 1984; Pabo and Sauer 1984; Figs. 1 and 2). Computer analyses of homeo domains also have predicted α -helices in agreement with those seen in prokaryotic repressors (Buerklin 1988; Scott et al. 1989), and more recent proton nuclear magnetic resonance (NMR) studies on a purified homeo domain from the *Antennapedia* protein (Fig. 2) suggest that the homeo domain includes a helix–turn–helix structure (Otting et al. 1988). Chou–Fasman computer analysis predicts a similar structure in Oct-2 (Fig. 2)—not a surprising finding given the high sequence conservation with other homeo domains.

An alignment of the primary sequence of the homeo domain of Oct-2 with the DNA-binding domains of three prokaryotic repressors—the 434 repressor, the λ repressor, and the *Escherichia coli trp* repressor—is shown in Figure 1. These prokaryotic domains all have been shown to include three α -helices; the latter two are the helix–turn–helix structure (Pabo and Sauer 1984; Ptashne 1986; Fig. 1) demonstrated in protein crystals by X-ray crystallography (Pabo and Sauer 1984). Moreover, the interactions of these domains with their respective operator sites have been demonstrated in protein : DNA co-crystals by X-ray crystallography (Aggarwal et al. 1988; Jordan and Pabo 1988; Otwinowski et al. 1988), and these interactions fit well with predictions from genetic data (Pabo and Sauer 1984). A summary of the crystallographic findings is presented in Figure 1.

What can we learn from such a summary? At first glance it is clear that the DNA-binding domain of these repressors includes at least three α -helices, not just the helix 2–turn–helix 3 motif (Fig. 1). Moreover, in 434 and λ repressors, there are many amino acids outside of

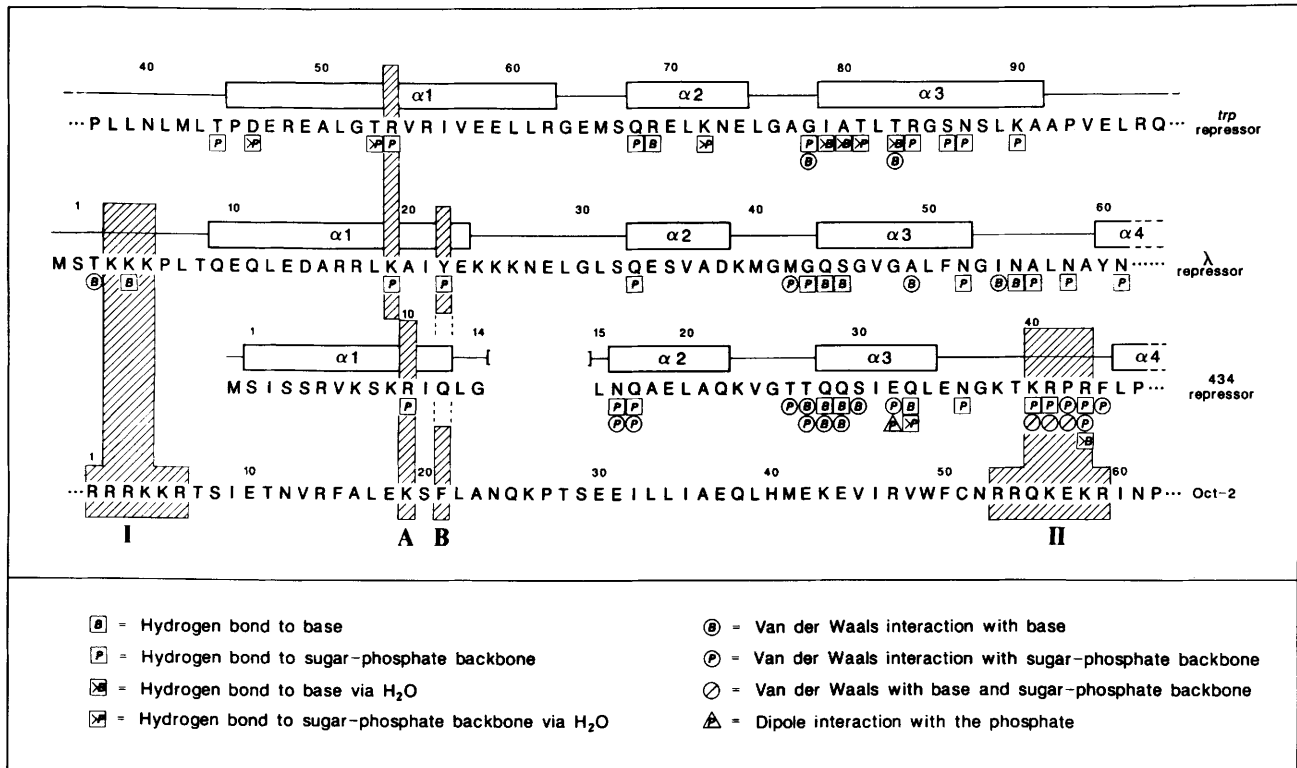


Figure 1. Comparison of the structures of DNA-binding domains of prokaryotic repressors with the Oct-2 homeo domain. The primary amino acid sequences of the *trp* repressor (Gunsalus and Yanofsky 1980; Otwinowski et al. 1988), λ repressor (Sauer and Anderregg 1978), and 434 repressor (Kuziel and Tucker 1987) were aligned based on the secondary structure and the position of conserved hydrophobic amino acids, as has been described (Pabo and Sauer 1984). The alignment of the Oct-2 was based on the previous alignment of the homeo domain putative helix 2–turn–helix 3 motif with the same motif in the λ repressor (Laughon and Scott 1984). The only gap necessary in this alignment was made in the 434 repressor sequence. Elements of secondary structure in the repressor proteins are indicated above the primary sequence; α -helices are represented by open boxes; the number within the open boxes indicates helix number in the DNA-binding domain. Numbers above the structural representation indicate residue numbers in the protein, except for Oct-2 where they represent residue numbers in the homeo domain. Shaded areas I, II, A, and B indicate conserved elements found in Oct-2 and at least one prokaryotic repressor protein (see text). The symbols shown under certain amino acid residues indicate that the amino acid is in contact with the operator site and detail the type of interaction (see identification of symbols at bottom of figure). This information was derived from the description of *trp* repressor–(Otwinowski et al. 1988), λ repressor–(Jordan and Pabo 1988), and 434 repressor–(Aggarwal et al. 1988) operator co-crystals. When an amino acid was shown to be involved in multiple interactions with the operator, not all interactions are necessarily indicated. The interactions shown are, for the most part, found in both operator half-sites. The interaction between T2 of the λ repressor and base pair 8 of the operator was assumed to be hydrophobic; however, this was not clear from the original paper (Jordan and Pabo 1988). The solvent bridged hydrogen bond of D46 in *trp* repressor may involve a Ca²⁺ ion (Otwinowski et al. 1988). Interestingly, the tryptophan bound to the *trp* repressor also contributes a hydrogen bond to the operator; this is not indicated in the figure (Otwinowski et al. 1988).

these three helices that are also in contact with the DNA.

Upon alignment of the Oct-2 homeo domain with the DNA-binding domain of three prokaryotic repressors there are, in addition to the helix 2–turn–helix 3, three elements conserved between Oct-2 and one or all of the repressors: (1) A cluster of basic amino acids (K or R) immediately following the putative helix 3 in Oct-2 (II in Fig. 1); (2) a cluster of basic amino acids (K or R) preceding the predicted helix 1 in Oct-2 (I in Fig. 1); and (3) a conserved pair of amino acids in the predicted helix 1 of Oct-2 (A and B in Fig. 1). By analogy to similar elements in one or all of the prokaryotic repressors in Figure 1, these elements in Oct-2 may be in contact with the DNA and may contribute to the binding specificity.

A cluster of basic amino acids carboxy terminal to the recognition helix 3 is present in 434 repressor and is the most highly conserved element in the homeo domain

Oct-2 contains the amino acids RRQKEKR (II in Fig. 1) immediately following the predicted helix 3. The 434 repressor contains the sequence KRPRF (residues 40–44 in Fig. 1) in the turn following helix 3. Each of these five amino acids in the 434 repressor was shown to make contact with the operator DNA in protein : DNA co-crystals; all but F44 make more than one contact (Aggarwal et al. 1988). It is clear from Figure 1 that these clusters of basic amino acids in Oct-2 and the 434 repressor align well. Although the λ repressor does not contain such a cluster of basic amino acids, residues in

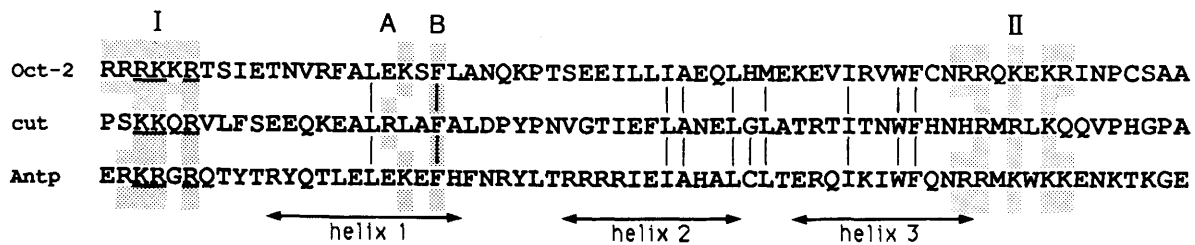


Figure 2. Alignment of the homeo domain amino acid sequences of Oct-2 (Clerc et al. 1988), *cut* (Blochlinger et al. 1988), and *Antp* (McGinnis et al. 1984). *Antp* is considered the consensus homeo domain and *cut* a distantly related homeo domain. α -Helices 1, 2, and 3 (arrows) are depicted as described for the *Antp* homeo domain (Otting et al. 1988). Oct-2, *cut*, and *Antp* amino acid sequences are aligned based on the predicted helix 2–turn–helix 3 motif (Scott et al. 1989). Basic clusters I and II are in shaded boxes; A and B indicate positions of conserved amino acids (shaded boxes) in the predicted helix 1. Vertical lines indicate highly conserved hydrophobic residues in predicted helices (Laughon and Scott 1984; Pabo and Sauer 1984).

equivalent positions make many contacts with the λ operator (Fig. 1). The *trp* repressor has a longer recognition helix 3; the carboxy-terminal end of this helix 3 has a lysine residue (K90) that aligns with R54 in Oct-2. This lysine in the *trp* repressor hydrogen-bonds to a phosphate group in the *trp* operator. Two other prokaryotic DNA-binding proteins, 434 Cro and λ Cro, have a cluster of basic amino acids following helix 3, and there are structural and genetic data that also suggest the involvement of these residues in DNA binding (Pabo and Sauer 1984; Wolberger et al. 1988).

Basic amino acids in cluster II are the most conserved element within the homeo domain (Fig. 2; Buerklin 1988). In the 86 homeo domains other than Oct-2 reviewed by Scott et al. (1989), the five underlined positions in RRQKEKR are either K or R in 84, 86, 86, 86, and 80 cases, respectively. Thus, this cluster of basic amino acids is probably a conserved DNA-binding element within the homeo domain.

A cluster of basic amino acids amino terminal to the putative helix 1 is present in the λ repressor and conserved in the homeo domain

Oct-2 has the sequence RRRKRR (I in Fig. 1) at the amino terminus of the homeo domain, and this sequence aligns well with the sequence TKKKPL in the λ repressor amino-terminal arm. Both T2 and K4 in this arm make contact with the λ operator (Fig. 1; Jordan and Pabo 1988). By analogy to the λ repressor, these basic residues in Oct-2 may be involved in DNA contacts.

Three basic residues in cluster I of Oct-2 (underlined in Fig. 2) are conserved in 76, 76, and 84 of 86 homeo domains, respectively (Fig. 2; Scott et al. 1989). This cluster of basic residues in all homeo domains may be a DNA-binding element outside the helix–turn–helix structure.

The average conservation, among homeo domains, of the underlined amino acids of these two basic clusters II and I (RRQKEKR and RRRKRR) is well over 95%, whereas the average conservation of the functionally equivalent amino acids (e.g., K, R, or H and E or D) in the homeo domain is less than 80%. Therefore, the degree of

conservation within clusters I and II is remarkable even within the conserved homeo domain.

Helix 1 may be involved in DNA binding

There is NMR evidence for the existence of an α -helix in the *Antennapedia* homeo domain that aligns well with helix 1 in the prokaryotic repressors (Otting et al. 1988), and it is reasonable to assume that this α -helix exists in all homeo domains, including Oct-2 (in Oct-2, an α -helix 1 is predicted by Chou–Fasman analysis; R.G. Clerc, unpubl.). Given that in prokaryotic repressors this helix interacts directly with the operator, a similar interaction between the octamer DNA element and the predicted helix 1 in Oct-2 may occur. Within the predicted Oct-2 helix 1, there are two residues that show an intriguing level of conservation with DNA-binding residues in helix 1 of the bacterial repressors. First, the Oct-2 K19 could be equivalent to K19 in the λ repressor, R10 in the 434 repressor, and R54 in the *trp* repressor, all of which contact the prokaryotic operators (Fig. 1). Second, F21 in Oct-2 aligns with Y22 in the λ repressor, an amino acid that contacts the operator.

Position 19 in the family of homeo domains is K or R in 72 of 86 cases, and in six cases in which homeo domain amino acid 19 is neither K nor R, neighboring amino acids (positions 18 or 20) are K or R. F21 is highly conserved in the homeo domain, being either F or Y in 86 of 86 cases (Scott et al. 1989).

Given the similarities of structural elements, basic cluster I, helix 1 followed by the helix 2–turn–helix 3 structure, and cluster II, the entire homeo domain shares structural homology with the tri- α -helical DNA-binding domains of prokaryotic repressors shown in Figure 1. Therefore, the DNA-binding domain of Oct-2 and all other homeo box-containing proteins is probably encoded by the 60-amino-acid-long homeo domain and not merely by the helix–turn–helix structure. A schematic of a hypothetical interaction is shown in Figure 3.

It is possible that Oct-2 embraces the DNA using clusters I and II. These basic amino acids could contribute to specific DNA binding by stabilizing the interactions between the helix–turn–helix structure and

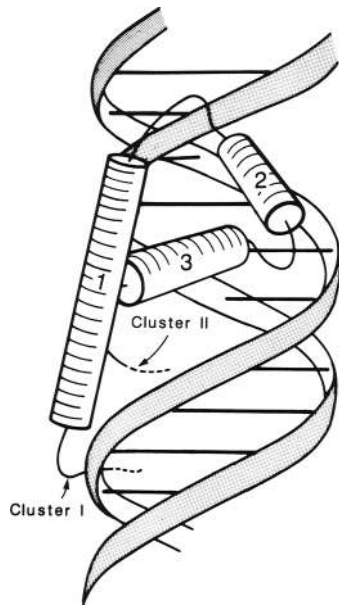


Figure 3. Schematic for the homeo domain (DNA-binding domain) of Oct-2. This schematic was derived by analogy to the structure of the phage λ repressor-operator complex. The figure is meant to convey a general trihelical structure in contact with DNA, and not the precise spatial coordinates of the α -helices. It is very possible that the relative positions of the three helices will vary significantly from that of the λ repressor.

the DNA helix. Whereas the λ , 434, and *trp* repressors bind DNA as dimers, Oct-2 binds the octamer sequence as a monomer (Clerc et al. 1988 and unpubl.). Thus, the presence of two basic clusters might be more important for helix-turn-helix proteins binding as monomers than for those binding as dimers.

Recently, Desplan et al. (1988), Mihara and Kaiser (1988), and M. Müller et al. (1988) showed that the 60-amino-acid homeo domain was sufficient for *engrailed* (*en*) and for *Antennapedia* DNA-binding activity. In both cases, the proteins probably bind to approximately 8–16 bp of specific sequences.

The question remains as to whether or not the full homeo domain is essential for DNA binding, as is suggested by the model presented in Figure 3. Results obtained recently from *Drosophila even-skipped* (*eve*) null mutant embryos have been shown that a single amino acid substitution in either homeo domain position 10 or 53 [respectively, T \rightarrow I [helix 1] and R \rightarrow H [cluster II]] disrupts the normal *eve* expression pattern in vivo (Frasch et al. 1988). These two point mutants also show significantly decreased DNA binding in vitro (M. Levine, pers. comm.). These data are consistent with a requirement for a full homeo domain for efficient DNA binding.

It has been suggested that the 74-amino-acid region POU box conserved among all POU domain proteins, Oct-2, Oct-1, Pit-1, and Unc-86 (Herr et al. 1988), is required in addition to the 60-amino-acid homeo box re-

gion for binding to the octamer sequence (Sturm and Herr 1988). This conclusion was based on the observations that mutation or deletion of the POU box region produced proteins that failed to bind specifically in vitro. These mutant proteins had an intact homeo domain. These data, however, do not exclude the possibility that the POU box region could enhance the stability of interactions between the homeo domain and the octamer sequence. In this case, the binding specificity to the octamer sequence would be specified solely by the homeo domain structure.

The above discussion is directed at the construction of a model for the structure of the entire homeo domain as a DNA-binding domain. It is of interest to ask whether or not an examination of the *cis*-acting elements bound by these related proteins sheds any light on the above discussion.

Oct-2 binds a family of related sites

The octamer binding transcription factors Oct-1 and Oct-2 (which have identical residues in 54 of the 61 positions of the homeo domain) recognize the octameric DNA sequence (ATGCAAAT) with high affinity and specificity. Oct-1 and Oct-2 also bind specifically, but with diminished affinities, several DNA recognition sites that deviate from the octamer site canonical sequence (ATGCAAAT) (see Fig. 4). Both proteins bind several related sequences in SV40 (e.g., ATGCAAAG) (Sturm et al. 1987), the adenovirus replication protein NFIII binding site (TATGATAATGA) (Prujijn et al. 1988), the recently recognized heptamer sequence in immunoglobulin genes (CTCATGA) (Poellinger and Roeder 1989), the herpes simplex Octa-TAATGARAT *cis*-element (Gerster and Roeder 1988), the *en* site

	A	B
OCTAMER	ATGC	AAAT
SV40 OCTA1	ATGC	AAAAG
NFIII	TATGA	TAATGA
HEPTAMER BC11	NNNN	TCATGA
HEPTAMER V1	NNNN	CTAATGA
OCTA-TAATGARAT	ATGC	TAATGATAT
<i>en</i>	TC	AAATTAATGA
<i>Ubx</i>	TAATAG	GTGTAATA
<i>Mat-a2</i>	ATGT	ACTTACC

Figure 4. Oct-2 binds a family of related sites. Alignment of the Oct-2 DNA-binding site (Parslow et al. 1984) together with degenerate octamer DNA motifs [SV40 Octa1, Sturm et al. 1987; NFIII, Pruijn et al. 1988; heptamer BC11, Poellinger and Roeder 1989 (same as BCL1 in Eaton and Calame 1987), and heptamer V1, Eaton and Calame 1987; OCTA-TAATGARAT, Gerster and Roeder 1988; *en*, Scheiderei et al. 1988; *Ubx*, Thali et al. 1988; *Mat-a2*, Ko et al. 1988] that have been shown to bind specifically the Oct-2 protein. The DNA-binding sites are divided into a 5' region (box A) and into a conserved AT-rich 3' region (box B). It is important to point out that Baumruker et al. (1988) showed that the highly related Oct-1 bound directly to the TAATGARAT motif in the herpes simplex virus immediate early promoter.

(TCAATTAATGA) (Scheidereit et al. 1988), the *Ultra-bithorax* (*Ubx*) homeo domain protein target sequence (TAATAGGTGTAAAT) (Thali et al. 1988), and a Mat-a2 binding site (ATGTACCTTACC) (Ko et al. 1988). These recognition sites show an intriguing degree of diversity, with the only obvious conserved element being an AT-rich box found in the 3' half of the canonical recognition site (B in Fig. 4). It should be noted, however, that flanking sequences affect the binding of the highly related Oct-1 protein to degenerate octamer sites (Baumruker et al. 1988). Thus, all of the protein-DNA interactions important for binding of the Oct-2 protein may not be confined to the consensus 8-bp sequence.

Drosophila homeo domain proteins show in vitro binding promiscuity similar to that observed for Oct-2 protein. The physiologically important binding sites for most *Drosophila* proteins have not been determined, and deletion analysis of binding sites identified in vitro often have failed to show a corresponding in vivo phenotype (R. Warrior and M. Levine, pers. comm.). Interestingly enough, the DNA targets of several homeo domain proteins, *Antennapedia* (Mihara and Kaiser 1988; Müller et al. 1988), *Ubx* (Beachy et al. 1988), *en* (Desplan et al. 1988; Hoey and Levine 1988), *eve* (Hoey et al. 1988), and *zerknüllt* (Hoey et al. 1988) also show a significantly high content of AT nucleotides. A murine homeo domain protein, *Hox-1.3*, similarly binds AT-rich sequences (Odenwald et al. 1989).

It is possible that the binding of *Drosophila* homeo domain proteins to these sites reflects a general affinity for a type of duplex structure commonly formed by sequences rich in AT base pairs. These specific sites may not all be binding sites for the homeo domain proteins in vivo. In this case, sites with a higher affinity for the homeo domain protein than that of these AT-rich sequences or, alternatively, sites where the homeo domain protein might interact with other DNA-bound proteins would probably constitute the physiologically important regulatory elements.

The probability that most homeo domain proteins have a general affinity for a particular type of DNA structure creates a major difficulty for studies in vivo with genes encoding these proteins. The window of level of expression of the homeo domain protein where physiologically significant regulation can be studied might be narrow. Too low a level of expression would result in little occupancy of the regulatory element. Too high a level of expression would result in the binding of the homeo domain to the regulatory elements of other homeo domain-containing proteins. As a potential example of this, the *Drosophila* homeo domain proteins *en* and *Antennapedia* have been shown to bind the prototype immunoglobulin octamer motif in vitro (cited in Scheidereit et al. 1988). Moreover, when two plasmid vectors—one encoding one of these two *Drosophila* homeo domain proteins and the other a reporter gene with an octamer-binding site—are cotransfected into HeLa cells, the *Drosophila* proteins *trans*-activated the reporter gene (Thali et al. 1988). This suggests that high levels of expression of one homeo domain protein can

result in productive binding to the regulatory site of other homeo domain proteins. If this binding to heterologous sites in vivo reflects the relative binding affinities of the various sites in vitro, then a 10-fold range will distinguish interactions with relevant sites versus heterologous sites.

In conclusion, given the similarities of the entire Oct-2 homeo domain with the tri- α -helical DNA-binding domain of λ , 434, and *trp* repressors, we speculated above that the DNA-binding domain of Oct-2 and, by analogy of other homeo-box-containing proteins, is encoded by the highly conserved 60-amino-acid-long homeo domain and not merely by the helix 2–turn–helix 3 structure. The cross-binding properties of these regulatory proteins with a family of related DNA targets probably reflect the common structure of the DNA-binding domains. The differences in binding affinities for different DNA sites, however, most likely reflect subtle differences in protein structure. These differences together with tight physiological control over homeo domain protein concentrations should result in the strong in vivo specificity attributed to regulation by these homeo domain proteins.

Acknowledgments

We thank M.A. Brown, S. Buratowski, D. Chang, J. LeBowitz, R. Sauer, and H. Singh for critically reading the manuscript. We also thank M. Levine, R. Warrior, C. Pabo, and M. Scott for providing us with manuscripts prior to their publication. We are indebted to M. Sifaca for the preparation of the typescript. M.G.B. is funded by a postdoctoral fellowship from the National Institutes of Health (No. 1F32-CA08342) and R.G.C. acknowledges support from the Schweizerischer National Fonds. This work was supported by a grant from the National Science Foundation under the Engineering Research Center Initiative to the Biotechnology Process Engineering Center (Cooperative Agreement No. CDR-8802014), by Public Health Service grants RO1-GM34277 and RO1-GM32467 from the NIH, and, partially, by National Cancer Institute Cancer Center Core Grant P30-CA14051 to P.A.S.

Mariano A. García-Blanco,¹ Roger G. Clerc,¹ and Phillip A. Sharp

Center for Cancer Research and Department of Biology,
Massachusetts Institute of Technology, Cambridge,
Massachusetts 02139 USA

¹This review represents an equal contribution of both first authors.

References

- Aggarwal, A.K., D.W. Rodgers, M. Drottler, M. Ptashne, and S.C. Harrison. 1988. Recognition of a DNA operator by the repressor of phage 434: A view at high resolution. *Science* **242**: 899–907.
- Baumruker, T., R. Sturm, and W. Herr. 1988. OPB 100 binds remarkably degenerate octamer motifs through specific in-

- teractions with flanking sequences. *Genes Dev.* **2**: 1400–1413.
- Beachy, P.A., M.A. Krasnow, E.R. Gavis, and D.S. Hogness. 1988. An *Ultrabithorax* protein binds sequences near its own and the *Antennapedia* P1 promoters. *Cell* **55**: 1069–1081.
- Bergman, Y., D. Rice, R. Grosschedl, and D. Baltimore. 1984. Two regulatory elements for immunoglobulin κ light chain gene expression. *Proc. Natl. Acad. Sci.* **81**: 7041–7045.
- Blochlinger, K., R. Bodmer, J. Jack, L.Y. Jan, and Y.N. Jan. 1988. Primary structure and expression of a product from *cut*, a locus involved in specifying sensory organ identity in *Drosophila*. *Nature* **333**: 629–635.
- Bohmann, D., W. Keller, T. Dale, H.R. Scholes, G. Tebb, and I.W. Mattaj. 1987. A transcription factor which binds to the enhancers of SV40 immunoglobulin heavy chain and U2 snRNA genes. *Nature* **325**: 268–272.
- Buerklin, T.R. 1988. The yeast regulatory gene *pho2* encodes a homeo box. *Cell* **53**: 339–340.
- Clerc, R.G., L.M. Corcoran, J.H. LeBowitz, D. Baltimore, and P.A. Sharp. 1988. The B-cell-specific Oct-2 protein contains Pou box- and homeo box-type domains. *Genes Dev.* **2**: 1570–1581.
- Courey, A.J. and R. Tjian. 1988. Analysis of Sp1 in vivo reveals multiple transcriptional domains, including a novel glutamine-rich activation motif. *Cell* **55**: 887–898.
- Desplan, C., J. Theis, and P.H. O'Farrell. 1988. The sequence specificity of homeodomain DNA interaction. *Cell* **54**: 1081–1090.
- Eaton, S. and K. Calame. 1987. Multiple DNA sequence elements are necessary for the function of an immunoglobulin heavy chain promoter. *Proc. Natl. Acad. Sci.* **84**: 7634–7638.
- Falkner, F.G. and H.Z. Zachau. 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. *Nature* **310**: 71–74.
- Fletcher, C., N. Heintz, and R.G. Roeder. 1987. Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2B gene. *Cell* **51**: 773–781.
- Folk, W.R. 1988. Changing directions in Pol III transcription. *Genes Dev.* **2**: 373–375.
- Frasch, M., R. Warrior, J. Tugwood, and M. Levine. 1988. Molecular analysis of *even-skipped* mutants in *Drosophila* development. *Genes Dev.* **2**: 1824–1838.
- Gehring, W.J. 1987. Homeo boxes in the study of development. *Science* **236**: 1245–1252.
- Gerster, T., P. Matthias, M. Thali, J. Jiricny, and W. Schaffner. 1987. Cell type-specific elements of the immunoglobulin heavy chain gene enhancer. *EMBO J.* **6**: 1323–1330.
- Gerster, T. and T.G. Roeder. 1988. A herpesvirus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins. *Proc. Natl. Acad. Sci.* **85**: 6347–6351.
- Gunsalus, R.P. and C. Yanofsky. 1980. Nucleotide sequence and expression of *Escherichia coli trpR*, the structural gene for the *trp* aporepressor. *Proc. Natl. Acad. Sci.* **77**: 7117–7121.
- Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, P.A. Sharp, H.A. Ingraham, M.G. Rosenfeld, M. Finney, G. Ruvkin, and H.R. Horvitz. 1988. The Pou domain: A large conserved region in the mammalian *pit-1*, *oct-1*, *oct-2* and *C. elegans unc-86* gene products. *Genes Dev.* **2**: 1513–1516.
- Hoey, T. and M. Levine. 1988. Divergent homeo box proteins recognize similar DNA sequences in *Drosophila*. *Nature* **332**: 858–861.
- Hoey, T., R. Warrior, J. Manak, and M. Levine. 1988. DNA binding activities of the *Drosophila melanogaster even-skipped* protein are mediated by its homeo domain and influenced by protein context. *Mol. Cell. Biol.* **8**: 4598–4607.
- Hope, I.A., S. Mahadevan, and K. Struhl. 1988. Structural and functional characterization of the short acidic transcriptional activation region of yeast GCN4 protein. *Nature* **333**: 635–640.
- Jordan, S.R. and C.O. Pabo. 1988. Structure of the λ complex at 2.5Å resolution: Details of the repressor-operator interactions. *Science* **242**: 893–898.
- Ko, H.-S., P. Fast, W. McBride, and L.M. Staudt. 1988. A human protein specific for immunoglobulin octamer DNA motif contains a functional homeobox domain. *Cell* **55**: 135–144.
- Kuziel, W.A. and P.W. Tucker. 1987. Determination of vector:insert junctions in λ gt10 cDNAs that do not recut with EcoRI. Nucleotide sequence of the λ imm434 HindIII–EcoRI DNA fragment encoding part of the cI protein. *Nucleic Acids Res.* **15**: 3181.
- Landolfi, N.F., J.D. Capra, and P.W. Tucker. 1986. Interaction of cell-type-specific nuclear proteins with immunoglobulin V_H promoter region sequences. *Nature* **323**: 548–551.
- Landschultz, W.H., P.F. Johnson, and S.L. McKnight. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**: 1759–1764.
- Laughon, A. and M.P. Scott. 1984. Sequence of a *Drosophila* segmentation gene: protein homology with DNA binding proteins. *Nature* **310**: 25–31.
- LeBowitz, J.H., T. Kobayashi, L. Staudt, D. Baltimore, and P.A. Sharp. 1988. Octamer-binding proteins from B or HeLa cells stimulate transcription of the immunoglobulin heavy chain promoter in vitro. *Genes Dev.* **2**: 1227–1237.
- Mason, J.O., G.T. Williams, and M.S. Neuberger. 1985. Transcription cell-type specificity confused by an immunoglobulin N_H gene precursor that includes a functional consensus sequence. *Cell* **41**: 479–487.
- Mihara, H. and E.T. Kaiser. 1988. A chemically synthesized *Antennapedia* homeodomain binds to a specific DNA sequence. *Science* **242**: 925–927.
- Müller, M.M., S. Rupert, W. Schaffner, and P. Matthias. 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. *Nature* **336**: 544–551.
- Müller, M., M. Affolter, W. Leupin, G. Otting, K. Wüthrich, and W.J. Gehring. 1988. Isolation and sequence-specific DNA binding of the *Antennapedia* homeodomain. *EMBO J.* **7**: 4299–4304.
- McGinnis, W., R.L. Garber, J. Wirz, A. Kuroiwa, and W.J. Gehring. 1984. A homologous protein coding sequence in *Drosophila* homeotic genes and its conservation in other metazoans. *Cell* **37**: 403–408.
- Odenwald, W.F., J. Garbern, H. Arnheiter, E. Tourmier-Lasserve, and R.A. Lazzarini. 1989. The *Hox-1.3* homeobox protein is a sequence-specific DNA-binding phosphoprotein. *Genes Dev.* **3**: 158–172.
- Otting, G., Y.-Q. Qian, M. Müller, M. Affolter, W.J. Gehring, and K. Wüthrich. 1988. Secondary structure determination for the *Antennapedia* homeodomain by nuclear magnetic resonance and evidence for a helix-turn-helix motif. *EMBO J.* **7**: 4305–4309.
- Otwinowski, Z., R.W. Schevitz, R.G. Zhang, C.L. Lawson, A. Joachimiak, R.Q. Marmorstein, B.F. Luisi, and P.B. Sigler. 1988. Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature* **335**: 321–329.

- Pabo, C.O. and R.T. Sauer. 1984. Protein-DNA recognition. *Annu. Rev. Biochem.* **53**: 293-321.
- Parslow, T.G., D.L. Blair, W.J. Murphy, and D.K. Granner. 1984. Structure of the 5' ends of immunoglobulin genes: A novel conserved sequence. *Proc. Natl. Acad. Sci.* **81**: 2650-2654.
- Poellinger, L. and R.G. Roeder. 1989. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. *Mol. Cell. Biol.* **9**: 747-756.
- Pruijn, G.J.M., R.T. van Miltenburg, J.A.J. Claessens, and P.C. van der Vliet. 1988. Interaction between the octamer-binding protein nuclear factor III and the adenovirus origin of DNA replication. *J. Virol.* **62**: 3092-3102.
- Ptashne, M. 1986. *A Genetic Switch*. Cell Press, Cambridge.
- . 1988. How eukaryotic transcriptional activators work. *Nature* **335**: 683-689.
- Sauer, R.T. and R. Anderregg. 1978. Primary structure of the λ repressor. *Biochemistry* **17**: 1092-1100.
- Scheidereit, C., A. Heguy, and R.G. Roeder. 1987. Identification and purification of a human lymphoid specific octamer-binding protein (OTF-2) that activates transcription of an immunoglobulin promoter in vitro. *Cell* **51**: 783-793.
- Scheidereit, C., J.A. Cromlish, T. Gerster, K. Kawakami, C.G. Balmaceda, R.A. Currie, and R.G. Roeder. 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. *Nature* **336**: 551-557.
- Schneuwly, S., A. Kuroiwa, P. Baumgartner, and W.J. Gehring. 1986. Structural organization and sequence of the homeotic gene *Antennapedia* of *Drosophila melanogaster*. *EMBO J.* **5**: 733-739.
- Schreiber, E., M.M. Müller, W. Schaffner, and P. Matthias. 1989. Octamer transcription factors mediate B-cell specific expression of immunoglobulin heavy chain genes. In *Tissue specific gene expression* (ed. R. Renkawitz). VCH Publishers, Weinheim, FRG and Deer Field Beach, USA. (in press).
- Scott, M.P., J.W. Tamkun, and G.W. Hartzell III. 1989. The structure and function of the homeodomain. *BBA Reviews on Cancer* (in press).
- Singh, H., R. Sen, D. Baltimore, and P.A. Sharp. 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin genes. *Nature* **319**: 154-158.
- Sive, H.L. and R.G. Roeder. 1986. Interaction of a common factor with conserved promoter and enhancer sequences in histone H2B, immunoglobulin and U2 small nuclear RNA (snRNA) genes. *Proc. Natl. Acad. Sci.* **83**: 6382-6386.
- Sollner-Webb, B. 1988. Surprises in polymerase III transcription. *Cell* **52**: 153-154.
- Staudt, L.M., H. Singh, R. Sen, T. Wirth, P.A. Sharp, and D. Baltimore. 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. *Nature* **323**: 640-643.
- Staudt, L.M., R.G. Clerc, H. Singh, J.H. LeBowitz, P.A. Sharp, and D. Baltimore. 1988. Cloning of a cDNA encoding a B-cell restricted octamer binding factor. *Science* **241**: 577-580.
- Sturm, R.A., T. Baumruker, B.R. Franza, and W. Herr. 1987. A 100-kD HeLa cell octamer-binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. *Genes Dev.* **1**: 1147-1160.
- Sturm, R.A. and W. Herr. 1988. The POU domain is a bipartite DNA binding structure. *Nature* **336**: 601-604.
- Thali, M., M.M. Müller, M. Delorenzi, P. Matthias, and M. Bienz. 1988. *Drosophila* homeotic genes encode transcriptional activators similar to mammalian OTF-2. *Nature* **336**: 598-601.
- Wirth, T., L. Staudt, and D. Baltimore. 1987. An octamer oligonucleotide upstream of a TATA motif is sufficient for lymphoid-specific promoter activity. *Nature* **329**: 174-177.
- Wolberger, C., Y. Dong, M. Ptashne, and S.C. Harrison. 1988. Structure of a phage 434 Cro/DNA complex. *Nature* **335**: 789-795.



The DNA-binding homeo domain of the Oct-2 protein.

M A García-Blanco, R G Clerc and P A Sharp

Genes Dev. 1989, **3**:

Access the most recent version at doi:[10.1101/gad.3.6.739](https://doi.org/10.1101/gad.3.6.739)

References

This article cites 57 articles, 23 of which can be accessed free at:

<http://genesdev.cshlp.org/content/3/6/739.full.html#ref-list-1>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

The advertisement features a dark background with a colorful, abstract image of what appears to be a DNA double helix or a similar molecular structure. On the left, the text 'Dharmacon Reagents' is displayed in white, with a smaller line below it: 'Custom synthesis, RNAi, and CRISPR solutions'. In the center, the words 'Infinite Reliability' are written in a large, white, sans-serif font. To the right of this text is a small white box containing the word 'More'. On the far right, the 'horizon' logo is shown in white, with 'a PerkinElmer company' written in a smaller font below it.