A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins

Vicente Andrés, María D. Chiara, and Vijak Mahdavi

Department of Cardiology, Children's Hospital, and Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115 USA

The recently cloned Clox (Cut-like homeo box) and CDP (CCAAT displacement protein), two mammalian counterparts of the Drosophila Cut homeo protein, correspond to alternatively spliced products of the same gene (mClox, for mammalian Cut-like homeo box). Although these proteins reportedly bind to apparently unrelated DNA sequences, we show by in vitro selection of optimal binding sites that both Clox and CDP have the same preferred DNA-binding specificity. The palindromic consensus target sequence, 5'-(t/a)(a/t)tATCGATTAt(t/c)(t/g)(t/a)-3', contains a bona fide homeo domain binding motif (ATTA). In addition, 37% of the in vitro-selected sequences have a CCAAT box, the canonical target for members of the family of CCAAT-binding factors. A characteristic feature of the cut homeo proteins is the presence of three evolutionarily conserved 73-amino-acid repeats of unknown function, the so-called cut repeats. We present evidence that the cut repeat II binds to mClox consensus targets independently of the DNA-binding activity of the homeo domain. In vitro selection of binding sites shows that the optimal targets for the cut repeat II contain one or more CCAAT boxes and, like the homeo domain, an ATTA core. These results indicate that the DNA-binding activity of the second cut repeat can account for the suggested role of CDP mClox as CCAAT displacement protein, a putative repressor of gene expression. We also report that the mClox homeo domain and cut repeat II interact in vitro in the absence of DNA. This interaction, which greatly enhances the DNA-binding activity of the binary complex, is specific to the cut homeo proteins. No cooperativity was observed between the cut repeat II and the homeo domains of Oct-1 and Gtx. Furthermore, the Drosophila cut repeat II, which does not appear to bind to DNA, also enhances the DNA-binding activity of the mClox homeo domain. Thus, the bifunctional cut repeat II, which defines a new family of bipartite DNA-binding proteins, is likely to play an important role in the function of the cut homeo proteins.

[Key Words: cut repeat; mClox; Cut; CCAAT displacement protein; homeo domain; cooperative protein–protein interactions]

Received June 16, 1993; revised version accepted November 24, 1993.

The establishment of the general body plan, the specification of cell identity in precursor cells, and differentiation of committed cells are crucial processes during embryogenesis. Among the number of developmentally important genes, homeo box-containing genes have proven to play a key role in a broad variety of organisms, from Drosophila to mammals [for review, see Holland and Hogan 1988, Shashikant et al. 1991, McGinnis and Krumlauf 1992]. The Drosophila Cut homeo domain proteins [Blochlinger et al. 1988] specify cell diversification in at least two tissues of different embryonic origin: the sensory organs [Dawson et al. 1987; Blochlinger et al. 1990, 1991; Jack et al. 1991], and the Malpighian tubules [Liu et al. 1991]. However, this function might be broader, as cut expression was reported in other tissues, including muscle, central nervous system, and ovarian follicle cells of embryos and adult flies [Blochlinger et al. 1990, 1993]. Two mammalian counterparts of Cut, named Clox (Cut-like homeo box) [Andrés et al. 1992] and CDP (CCAAT displacement protein) [Neufeld et al. 1992], recently have been cloned from dog heart and Hela cell expression libraries, respectively. A common feature of the cut homeo proteins is the presence of three evolutionarily conserved cut repeats of unknown function [Blochlinger et al. 1988, Andrés et al. 1992, Neufeld et al. 1992]. Whereas the Drosophila and mammalian cut homeo proteins are divergent outside their homeo domains and cut repeats, Clox and CDP sequences are conserved almost in their entirety. Yet a complete divergence noted
in two coding regions indicates that they represent alternatively spliced products from the same mammalian gene [Fig. 1].

The regulated expression of the mammalian cut homeo proteins in a broad range of tissues during murine embryonic development, and during myogenic and chondrogenic differentiation in culture (Andrés et al. 1992), indicates that as in Drosophila, these proteins are candidate regulators of cell-fate specification in divergent differentiation programs. The diversity of cut homeo proteins raises the question as to the identity or the specificity of their target genes in different cell lineages. For these reasons, we have characterized the unbiased preferred binding sites for Clox and CDP using anti-Clox antibodies and a pool of oligonucleotides containing a random core [Wright et al. 1991; Funk and Wright 1992]. Our results show that both mammalian cut homeo proteins have the same optimal binding site, despite previous observations that their target sites were apparently unrelated [Andrés et al. 1992; Neufeld et al. 1992]. The consensus binding site reported here is an AT-rich degenerated palindrome that contains an ATTA core motif recognized by the majority of the homeo proteins (for review, see Scott et al. 1989, Laughon 1991, Shashikant et al. 1991; Treisman et al. 1992). We demonstrate that the cut repeat II and the homeo domain of mClox bind to DNA independently and present evidence for a cooperative interaction between these two domains, which, as a binary complex, have a greatly enhanced DNA-binding activity. This study indicates that the interplay between the bifunctional cut repeat and homeo domain constitutes a novel regulatory mechanism likely to be a key determinant of the specificity of function of the cut homeo proteins.

Results

Clox and CDP select identical optimal DNA target sequences

Although Clox and CDP are presumably alternatively spliced isoforms of the mammalian cut homolog gene product, hereafter named mClox [mammalian Cut-like homeo box] [Fig. 1], comparison of their reported binding sites did not reveal a consensus motif. This observation raised the possibility that different isoforms of mClox may have distinct DNA-binding specificities in the variety of tissues and cell types in which they are expressed [Andrés et al. 1992; Neufeld et al. 1992]. To test this hypothesis, we sought to determine the unbiased preferred DNA-binding sites for these two proteins according to established procedures [Wright et al. 1991; Funk and Wright 1992]. Because the partial Clox cDNA clone lacks approximately the first 557 amino-terminal residues, as compared with the CDP sequence, we engineered a chimeric full-length protein, named CDP/Clox, comprising the amino-terminal 878 amino acids of CDP and the rest from Clox. Importantly, CDP/Clox contains the carboxy-terminal divergent putative alternatively spliced exon from Clox [see Fig. 1]. It is noteworthy that the three cut repeats and the homeo domain of Clox and CDP are identical. A pool of doublestranded oligonucleotides that contained 15 randomized bases flanked by primer sequences was incubated with extracts from COS cells overexpressing either CDP/Clox or CDP [see Materials and methods]. After preclearing the binding reaction with preimmune serum, the specific DNA–mClox complexes were immunoprecipitated with α-FP antibodies, an antiserum raised against a glutathione S-transferase [GST]–Clox fusion protein [Andrés et al. 1992]. The selected oligonucleotides were eluted and amplified by polymerase chain reaction [PCR] and used as substrate in subsequent rounds of selection. Two control reactions were carried through the selection. First, selection was performed using the same amount of protein extract from mock-transfected COS cells, which showed almost negligible DNA-binding activity using the reported mClox target sequences, as compared with COS cells overexpressing CDP/Clox or CDP [data not shown]. Second, selection was done in the absence of any cell extract to control for nonspecific immunoprecipitation. After nine cycles, the pools of selected oligonucleotides were labeled and used in gel mobility shift assays [GMSAs]. The oligonucleotides selected with CDP/Clox and CDP overexpressing cells displayed a DNA–protein complex that comigrated with the retarded band seen using previously reported mClox targets, whereas no retarded band could be detected using the pool of oligonucleotides selected in the reactions with control COS extract or with no extract [data not shown]. These results demonstrated that the selection with the transfected COS cells was specific and due to the binding activity of the exogenously expressed Clox and CDP cloned products and not to the endogenous proteins.

The pools of selected sequences were then cloned and individually sequenced. All clones analyzed contained different sequences \( n = 138 \), demonstrating that they

Figure 1. Scheme of the predicted sequence of the cut homeo proteins. The numbers indicate amino acid residues. The three internal cut repeats are in roman numerals. No sequence similarities between the Drosophila Cut protein and its mammalian homolog could be found outside the cut repeats (open boxes) and homeo domain (solid boxes). Conversely, Clox and CDP sequences are conserved almost in their entirety. The predominant divergences correspond to two alternative coding exons (hatched boxes). The chimeric CDP/Clox protein contains the first 878 residues from CDP and the rest from Clox.
were derived from independently selected oligonucleotides. Generation of the consensus binding sites shown in Figure 2 was greatly facilitated by the presence within most sequences of the palindrome 5'-ATCGATA-3'. Only 3% of the oligonucleotides selected with the COS-CDP/Clox or COS-CDP cell extracts showed a random sequence. In contrast, 80% of the sequences selected with the control COS extract were apparently random, with the remaining 20% fitting only portions of the consensus (n = 45 sequences, data not shown). These results further demonstrate the specificity of the selection with the overexpressing COS-CDP and COS-CDP/Clox systems.

As shown in Figure 2A, the consensus binding sites for CDP/Clox and CDP were essentially identical, especially in the highly constrained octamer sequence ATCGATA [positions 4-11; see Fig. 2B]. Therefore, we conclude that two different mClox isoforms have the same optimal DNA target sequence, 5'-(t/a)|a|t|t|ATCGATA&t|c|t|g|t|a|-3', which constitutes a degenerated palindrome [Fig. 2B]. As for most homeo proteins, this consensus binding site is an AT-rich sequence with a conserved ATTA core motif (for review, see Scott et al. 1989; Laughon 1991; Shashikant et al. 1991; Treisman et al. 1992). Of note, only 6 nucleotides within the previously reported 62-bp-long Cbs-1 [André et al. 1992] and gpd [Neufeld et al. 1992] target sequences could be opti-

Figure 2. Consensus DNA-binding sequence of mClox. The pool of oligonucleotides recovered after nine cycles of selection was cloned, and individual clones were sequenced. The consensus sequences shown in A and B were obtained after alignment and quantification of the relative frequency (%) of the bases at each position. Uppercase bases indicate high constrain (58% or more); lowercase bases indicate moderate constrain. The number of independent sequences analyzed in each case is indicated (n). [A] Consensus sequence derived from the selections using extracts from COS cells overexpressing CDP/Clox and CDP. [B] Compiled consensus sequence of mClox, the mammalian cut homolog. (C) Alignment of the sequences containing a CCAAT box (boldface type) in addition to the mClox core-binding site [B, positions 4-11]. The sequences selected from the random core are indicated in uppercase, lowercase indicates bases within the flanking primers (see Materials and methods). The ATTA core is boxed, and the sequences derived from the selection with CDP/Clox are indicated with an asterisk (*). The occurrence of CCAAT box-containing sequences within the oligonucleotides selected with CDP/Clox and CDP was 33% and 38%, respectively.
Andrés et al.

mally aligned with the left half and right half of the consensus, respectively [see Fig. 3D]. Interestingly, 37% of the selected sequences also included a CCAAT box motif [Fig. 2C], the canonical target for members of the family of CCAAT-binding factors (Johnson and McKnight 1989; Umek et al. 1991). The CCAAT box and the mClox core-binding site [Fig. 2B, positions 4–11] overlapped in some selected targets and were as many as 3 bp apart in others. In addition, the CCAAT box was found on either side of the mClox consensus binding site, and in either orientation [Fig. 2C].

To examine the relative affinity of the endogenous mClox proteins for individual oligonucleotides derived from the consensus binding site, we chose two representative cloned sequences, one with and one without the CCAAT box motif, named D3 and E36, respectively [Fig. 3C]. As shown in Figure 3A, binding to the D3 probe in untransfected COS cells [lane 1] was specifically competed with a 50 M excess of unlabeled D3 [lane 2] or E36 [lane 3] oligonucleotides. As expected, a retarded band of the same electrophoretic mobility could be detected in the reverse experiment in which the E36 site was used as a probe [data not shown]. The involvement of endogenous mClox in the complex formed with the D3 probe was demonstrated by the specific competition produced by the α-P or α-FP antisera, two specific anti-Clox antibodies [Andrés et al. 1992], whereas no effect was noted with the corresponding preimmune sera [Fig. 3B]. The consensus mClox target sequence reported here constitutes a higher-affinity binding site than previously reported ones [Andrés et al. 1992; Neufeld et al. 1992]. The gpd and the Cbs-1 probes did not efficiently compete binding to D3, even in a 100–250 M excess [Fig. 3A, cf. lanes 1–3 vs. 5–8]. Yet, D3 effectively competed binding to the Cbs-1 [lanes 12–14] and gpd [lanes 16–18] probes, whereas D3 mutant [D3 mt] did not compete binding to either Cbs-1 [lane 15] or gpd [lane 19].

The D3 mt oligonucleotide, which includes a mutated mClox core consensus binding site but an intact CCAAT box [Fig. 3C], did not compete binding to the D3 probe [Fig. 3A, lane 4], nor was it bound by mClox [data not shown]. Although these results suggest that the CCAAT box is not required for mClox DNA-binding activity, they do not rule out that mClox might bind to the CCAAT box with low affinity.

The Cut homeo proteins are bipartite DNA-binding proteins

Two lines of evidence indicated that mClox, in addition to its homeo domain, might include another DNA-bind-
ing domain: [1] 37% of the oligonucleotides recovered in the PCR selection assay contained a CCAAT box [see Fig. 2C]; and [2] the simultaneous addition of the gpd and Cbs-1 oligonucleotides greatly enhanced their intrinsic ability to compete the binding of mClox to the D3 probe [Fig. 3A, lanes 5–8 vs. 9–11]. A good candidate to function as a DNA-binding domain is the evolutionary conserved cut repeat, the ~73-amino-acid motif occurring in three copies in all cut homeo proteins characterized to date [see Fig. 1]. To test this hypothesis, distinct GST fusion proteins, including the respective cut repeats [GST–Cut I, II, and III] and the homeo domain [GST–Hd] of mClox, were produced in bacteria and used for GMSA. As illustrated in Figure 4A, both GST–Cut II and GST–Hd bound the two probes derived from the mClox consensus binding site, E36 [lanes 7,14] and D3 [lanes 8,15]. On the other hand, GST–Hd bound both Cbs-1 and gpd [lanes 17,19], whereas binding of GST–Cut II to gpd or Cbs-1 could not be detected under the same conditions [lanes 10,12], indicating that these two subdomains have similar but distinct DNA-binding specificities. Surprisingly, however, despite the striking amino acid conservation among the three cut repeats [Blochlinger et al. 1988; Andrés et al. 1992; Neufeld et al. 1992], we have not been able to demonstrate the binding of GST–Cut I [Fig. 4B, lanes 10–12] and GST–Cut III [Fig. 4B, lanes 13–15] to the E36 probe, using the same conditions that allowed efficient binding by GST–Cut II [Fig. 4B, lanes 7–9].

In addition to the results presented above, several lines of evidence demonstrated the specificity of the DNA-binding activity of GST–Cut II and GST–Hd. First, equivalent or even higher amounts of GST protein did not bind the probes used in these assays [Fig. 4A, lanes 1–6; Fig. 4B, lanes 1–3]. Second, mutations in the target sequences that ablated binding by the full-length protein also abrogated binding by GST–Cut II [D3 mt, Fig. 4A, lane 9] and by GST–Hd [D3 mt and Cbs-1 mt, Fig. 4A, lanes 16 and 18, respectively]. Third, neither GST–Cut II nor GST–Hd bound to the unrelated AT-rich MEF2-binding site [Yu et al. 1992] [Fig. 4A, lanes 13,20], further indicating that these subdomains interact with DNA in a selective, sequence-specific manner and do not bind indiscriminately to AT-rich sequences.

Homeo domain and cut repeat II select distinct but overlapping optimal DNA target sequences

We then sought to determine the preferred binding sites of each DNA-binding subdomain of mClox by using the PCR selection assay. The pool of oligonucleotides recovered after six cycles of selection with GST–Hd was cloned, and individual clones were sequenced. As shown in Figure 5A, these targets could be categorized in two classes, which contained either a single ATTA core within the consensus cTTGATTAg, consensus 1, or two and three repeats of the GATC motif, consensus 2 and 2′, respectively. The spacing between adjacent GATC motifs was 8, 6, 5, or 2 bp. Of note, the homeo domain of mClox also binds the Cbs-1 probe, which contains an unspaced duplicated GATC motif, and mutations within this sequence abolish binding activity [see Fig. 4A, lanes 17,18]. Therefore, each consensus optimally recognized by the homeo domain is represented in the previously reported mClox target sites, gpd and Cbs-1, as they include a TGATTA core or a duplicated GATC motif, respectively [see Fig. 3C].

Figure 5B shows the sequence alignment of the oligonucleotides recovered after six cycles of selection with GST–Cut II. Remarkably, the cut repeat II consensus includes an ATTA core, like the homeo domain target DNA [see Fig. 5A, consensus 1], in addition to 1–4
Figure 5. Optimal binding sites of mClox homeo domain and cut repeat II. After six cycles of selection/amplification, the selected oligonucleotides were cloned and individual clones were sequenced. The sequences selected from the random core are indicated in uppercase, lowercase indicate bases within the flanking primers (see Materials and methods). Upper- and lowercase bases within the consensus sequences indicate high (>57%) and moderate constraint, respectively. (A) Alignment of individual clones selected with GST–Hd. The N15 random oligonucleotide was used. The spacing between adjacent GATC motifs in consensus 2 and 2’ is indicated. The first sequence shown in consensus 1 was found in six independent clones; the second sequence in consensus 2’ was found in three independent clones. Finally, two clones showed an apparently random sequence. [B] Alignment of individual clones selected with GST–Cut II. The original length of the random core was 15 bp (oligonucleotides 1–12) and 40 bp (oligonucleotides 13–25). Note that the length of some selected oligonucleotides (1, 9–12, 16, 21, 22, 25) varied from that of the original random library. Some sequences were found in several independent clones (oligonucleotide 1, four clones; 2, nine clones; 5, five clones; 9, two clones; 11, three clones). The 5’-ATTA-3’ and the 5’-CCAAT-3’ (5’-AT-TGG-3’) motifs are boxed. Sequences containing single mutations within the CCAAT motif are also boxed. (C) Alignment of the optimal targets of mClox homeo domain, cut repeat II, and full-length protein (mClox). The numbers are as in Fig. 2, A and B.

CCAAT boxes. Like the CCAAT box-containing oligonucleotides selected with the full-length mClox protein [Fig. 2C], the CCAAT box was found on either side of the ATTA core, in either orientation, and with variable spacing to the ATTA core. Furthermore, the cut repeat II, like mClox, also fails to bind to the D3 mt probe, which included a mutated ATTA core but an intact CCAAT motif [Fig. 4A, lane 9]. Thus, the cut repeat II recognizes the CCAAT box, yet this sequence is not sufficient for binding. The most likely explanation to address this seemingly paradoxical observation is that binding of the cut repeat II to the ATTA core is required for this domain to recognize the CCAAT box.

Alignment of the consensus targets selected by mClox and by its two subdomains shows a high degree of identity [Fig. 5C]. Interestingly, the homeo domain targets overlap positions 4–11 within the mClox consensus, whereas the cut repeat II target overlaps positions 8–15 and contains a CCAAT box, like some of the oligonucleotides selected with mClox. Taken together, the above results demonstrate that the mClox homeo domains are bipartite DNA-binding proteins, in which the isolated homeo domain and second cut repeat contribute significantly to the overall binding specificity of mClox. Of note, whereas GST–Hd bound to the pool of oligonucleotides selected with GST–Cut II, the targets selected with GST–Hd were not bound by GST–Cut II [data not shown]. These results further emphasize that the DNA-
binding specificity of the cut repeat II is more restricted than that of mClox homeo domain, in spite of the partial overlap within their respective targets.

Cooperative interaction between the homeo domain and cut repeat II enhances their intrinsic DNA-binding activity

The observation that the two DNA-binding subdomains of mClox have partially overlapping sequence specificities raised the question as to whether they would compete for binding to mClox target sequences. To investigate this hypothesis, the E36 probe was chosen for further analysis. As shown in Figure 6A, the simultaneous presence of the cut repeat II and the homeo domain of mClox in the binding reaction led to the formation of a slow migrating ternary complex (lane 4), as compared to reactions with the cut repeat II (lane 1) or the homeo domain (lane 3) alone, which bound to DNA as monomers. Remarkably, the DNA-binding activity of the ternary complex was greatly enhanced, strongly indicating that the two subdomains interact with each other and cooperate to bind to DNA. Several independent lines of evidence supported this interpretation. First, when a truncated portion of Clox that contained the homeo domain was produced by in vitro translation and incubated with a GST–Cut II resin, bound protein could be recovered from the beads [Fig. 6B]. This result indicates that the homeo domain and cut repeat II can interact in the absence of DNA, thus demonstrating a direct protein–protein interaction. Second, the cut repeat II from Drosophila Cut, which apparently did not bind to E36 [Fig. 6A, lane 2], also directed ternary complex formation and significantly enhanced the DNA-binding activity of mClox homeo domain [Fig. 6A, lane 5]. Finally, ternary complex formation and binding cooperativity could be observed using the Cbs-1 probe [data not shown], which was bound by the mClox homeo domain but not by cut repeat II (see Fig. 4A, lanes 10, 17).

The specificity of the binding cooperativity produced by the interaction of the second cut repeat II and the mClox homeo domain was demonstrated in binding assays using other cloned homeo domains. Gtx [Komuro et al. 1993] and Oct-1 [Herr et al. 1988] homeo domains bound to E36, yet addition of GST–Cut II did not have any significant effect on their binding avidity [Fig. 6A, lanes 6–9]. Furthermore, neither GST–Cut I [lane 10], GST–Cut III [lane 11], nor even GST [data not shown] had any effect on the avidity of the homeo domain for the mClox consensus target.

To delineate the regions of each subdomain involved in DNA and protein interactions, several cut repeat II and homeo domain peptides, deleted at their carboxy-terminal end, were expressed as GST fusion proteins and tested in the GMSA shown in Figure 7A. As compared to that of the full-length peptide [1–73], the intrinsic affinity of cut repeat II 1–64 for the E36 probe was reduced, and cut repeat II 1–54 no longer bound DNA [lanes 4–6]. Cooperative binding produced by the addition of the homeo domain was only slightly reduced with cut repeat II 1–64, and no ternary complex was formed with cut repeat II 1–54 [lanes 4 and 14 vs. 11–13]. The cooperativity between the two subdomains was even more obvious using the E36 mt probe [cf. lanes 1 and 7 vs. 8 and 9], to which the deleted cut repeats did not bind [lanes 2, 3], and for which both full-length subdomains had a reduced affinity [cf. lane 1 vs. 4 and 7 vs. 14]. Deletion of the carboxy-terminal portion of the homeo domain (GST–Hd 1–54) abolished both its intrinsic DNA-binding activity and its ability to cooperatively interact with cut
repeat II [lanes 14–17]. These results, which are summarized in Figure 7B, show the correlation between the extent of cooperativity and respective apparent DNA-binding affinity of the cut repeat II and the homeo domain. They also demonstrate that both domains contribute to the overall cooperativity.

Discussion

The Drosophila Cut and the mammalian Clox homeo proteins appear to be involved in the specification of cell fate in diverse lineages of precursor cells [Bodmer et al. 1987; Blochlinger et al. 1988, 1990, 1991, 1993; Jack et al. 1991; Liu et al. 1991; Andrés et al. 1992; Neufeld et al. 1992]. Two alternatively spliced isoforms of mClox have been reported to date: Clox [Andrés et al. 1992] and CDP [Neufeld et al. 1992]. In addition, a DNA-binding activity that is attributable to mClox (CDP) appears to recognize upstream regulatory sequences of various unrelated tissue-specific genes, such as the human gp91–phox [Skalnick et al. 1991; Neufeld et al. 1992] and y-globin [Supterti-Furga et al. 1988, 1989] genes, and the sea urchin sperm histone H2B gene [Barberis et al. 1987]. These observations raised the question as to the mechanisms that establish the functional specificity of the cut homeo proteins. To gain insight into this question, we have examined the DNA-binding characteristics of Clox and CDP to assess the identity of their target DNA sequences. A hallmark feature of the cut homeo proteins is the cut repeat, a ~73-amino acid triple internal repeat of unknown function. We present evidence that the mammalian cut repeat II is a new DNA-binding domain, which binds DNA independently of the homeo domain and significantly contributes to the binding specificity of mClox. Although both subdomains optimally recognize sequences containing a canonical ATTA core, they have extended and distinct binding requirements. Importantly, the cut repeat II also binds the CCAAT box motif. This DNA-binding activity of the second cut repeat might account for the suggested role of the mammalian cut homolog as CCAAT displacement protein, a putative repressor of gene expression [Neufeld et al. 1992]. Most striking is that the homeo domain and cut repeat II of mClox interact with each other and form a highly cooperative ternary DNA–protein complex. From an evolutionary point of view, it is remarkable that the cut repeat II from Drosophila Cut does not appear to bind to DNA alone, yet it cooperatively enhances the binding activity of mClox homeo domain. The implications of these results on the function of the cut homeo proteins, a new family of bipartite DNA-binding proteins, is discussed below.

The cut repeat II and the homeo domain correlate with their intrinsic DNA-binding affinity. [A] GMSA using full-length and carboxy-terminal deletions of cut repeat II and homeo domain (0.1–0.2 µg each fusion protein). The numbers indicate the amino acid residues contained in the fusion protein that correspond to either cut repeat II or homeo domain (full-length peptides are 1–73 and 1–60, respectively). In a shorter exposure, a slightly more intense retarded band in lane 11 compared with lane 12 could be noticed. Note that the retarded band seen in lanes 10 and 13 corresponds to the DNA/GST–Hd complex, and that in lane 17 to DNA/GST–Cut II. The cartoon schematizes the protein/DNA complexes, which contained both GST–Hd and GST–Cut II (A), GST–Hd (B), or GST–Cut II (C). [B] Schematic summarizing the results of A. (n.t.) Not tested.

The Cut homeo proteins are bipartite DNA-binding proteins

Clox and CDP optimally recognize in vitro the same DNA target, 5'-(t/a)(a/t)tATCGATTAt(t/c)(t/g)(t/a)-3', which constitutes a degenerate palindromic sequence. Thus, putative functional differences between these two isoforms of mClox may not ensue from their DNA-binding specificity. Their optimal binding target contains the canonical ATTA homeo protein binding motif [Scott et al. 1989; Hayashi and Scott 1990; Laughon 1991; Shashikant et al. 1991; Treisman et al. 1992]. Mutations within this core abolish mClox binding. In addition, other unrelated homeo proteins, such as Gtx [Komuro et al. 1993] and Oct-1 [Herr et al. 1988], bind to the mClox consensus binding site. These observations raise the question as to the binding selectivity of distinct homeo proteins for their cognate target genes in vivo. Regulatory specificity of mClox and other homeo proteins, such as Hox-1.5, Hox-7.1, and En-1 [Catron et al. 1993], might be achieved in part through imposed interactions with highly constrained nucleotides adjacent to the common consensus core. An additional mechanism for the selection of discrete targets is the coupling of the homeo domain to another DNA-binding domain, like the POU-specific [Sturm and Herr 1988], the paired [Treisman et al. 1991], and the LIM (lin-11 isl-1 mec-3) [Freyd et al. 1990; Karlsson et al. 1990] domains. We show here that the cut homeo pro-


teins represent a new family of bipartite DNA-binding proteins, in which the cut repeat II and the homeo domain bind specifically and independently to DNA. Comparison between the consensus targets of mClox and each of its two DNA-binding subdomains indicates that the overall binding specificity of the cut homeo proteins comprises the coupled activity of the homeo domain and the second cut repeat. Similarly, both the POU homeo domain and the POU-specific domain contribute to the DNA-binding specificity of Oct-1 [Ingraham et al. 1990; Verrijzer et al. 1990, 1992; Aurora and Herr 1992].

It is noteworthy that the mClox homeo domain is able to recognize not only the ATTA core but also a duplicated GATC motif. The ability to recognize two unrelated sequences has been similarly reported for other homeo domains, including Eve [Hoey and Levine 1988], Abd-A [Thali et al. 1988], Caudal [Dearolf et al. 1989], MATa2, Porter and Smith 1986, and Prd [Treisman et al. 1989]. Mutational analysis has shown that different residues within these homeo domains affected differently the binding to the two classes of sites, thus suggesting that these "bivalent" homeo domains can bind to DNA using either of two distinct sets of contact residues [Treisman et al. 1992]. This hypothesis, together with the ability of the cut repeat II to bind to DNA, might account for the diversity of putative targets for mClox [see below].

The high degree of amino acid identity among the three cut repeats (55–70%) is difficult to reconcile with the apparent incapacity of the mammalian cut repeats I and III and the Drosophila cut repeat II to bind to mClox consensus targets. We cannot exclude that these repeats have an intrinsic but distinct DNA-binding activity, which might increase the diversity of mClox targets. However, we have failed to select binding sites for the fruit fly cut repeat II, under the same conditions that allowed selection of optimal binding sites by the mammalian cut repeat II.

The interaction between the homeo domain and cut repeat II: a new mechanism of modulation of homeo protein function

The observation that the mClox homeo domain and cut repeat II interact, and that such interaction increases cooperatively their DNA-binding affinity, is striking. The ability of the two subdomains to interact in vitro, even in the absence of DNA, is unprecedented and provides an additional mechanism for the regulation of mClox function. Unlike the cooperative DNA binding by the POU-specific and the POU-homeo domains [Sturm et al. 1988], the cooperativity between mClox subdomains does not require a covalent linkage via a peptide linker. In addition, this cooperative effect is highly specific, as demonstrated by the absence of cooperativity between the cut repeat II and the homeo domains of Oct-1 and Gtx, which nevertheless can bind to the mClox consensus target. Of note, the Oct-1 homeo domain mediates interaction with the viral VP16 protein [Lai et al. 1992; Pomerantz et al. 1992].

Figure 8 shows a model for the mechanism of the binding of mClox to DNA. We have not directly addressed whether the interplay between the cut repeat II and the homeo domain represents an intramolecular interaction or, alternatively, mediates homodimerization. Nevertheless, the lack of a band of intermediate electrophoretic mobility in GMSA in which extracts from COS cells overexpressing full-length and truncated [amino acids 979–1505; see Fig. 1] mClox were mixed seems to exclude homodimerization [data not shown]. Moreover, the optimal binding sites selected with the cut repeat II and the homeo domain overlap the consensus target of the full-length protein, arguing that mClox must fold in such a way that its two DNA-binding subdomains are in close contact when docked onto DNA, thus favoring the hypothesis of an intramolecular interaction. Similarly, in the bifunctional Prd protein, the paired domain can act through intramolecular cooperativity with the homeo domain to enhance binding specificity [Treisman et al. 1991]. Further investigation of the molecular basis of the cooperativity between the second cut repeat and mClox homeo domain would ascertain whether the interaction between the cut-homeo proteins and their target sequences follows the general proposed model for homeo domain–DNA interactions [Otting et al. 1988, 1990; Qian et al. 1989; Kissinger et al. 1990; Harrison 1991; Wolberger et al. 1991].

From an evolutionary perspective, it appears that the DNA-binding function of cut repeat II is more recent than its ability to interact with the homeo domain. Two lines of evidence indicate that these two functions are independent. First, the Drosophila cut repeat II does not
appear to bind to DNA alone, yet it interacts with and greatly enhances the DNA-binding activity of the mClox homeo domain. Second, the shorter Cut II peptide (1–64) displayed reduced apparent DNA-binding affinity yet retained its ability to interact cooperatively with the homeo domain. The identity of the residues involved in each of these apparently separable functions should be elucidated by site-directed mutagenesis.

The role of the Cut homeo proteins as CCAAT displacement proteins

Neufeld and co-workers (1992) have suggested that the human cut homolog is the CCAAT displacement protein (CDP), a candidate repressor of gene expression that functions by competing the binding of CCAAT-binding enhancer proteins, a family of positively acting transcription factors [Johnson and McKnight 1989; Umeek et al. 1991]. Whereas the sea urchin CDP displaces in vitro a CCAAT-binding factor from the histone H2B gene [Barberis et al. 1987], there is no direct evidence that the human CDP displaces CCAAT-binding factors from its reported targets, the γ-globin [Superti-Furga et al. 1988, 1989] and gp91–phox [Skalnik et al. 1991, Neufeld et al. 1992] genes. CDP footprints over a region of the histone H2B and gp91–phox genes that contains a duplicated CCAAT box and differently spaced GATTAT and GATN motifs, which are comprised in the selected mClox-binding sites. However, the footprint generated by CDP on the human γ-globin promoter [Superti-Furga et al. 1988, 1989] includes only two CCAAT motifs. Whereas mutations within both CCAAT boxes in the γ-globin gene strongly reduced CDP binding [Superti-Furga et al. 1988], neither CCAAT box seems to be required for the binding of CDP to the histone H2B [Barberis et al. 1987] and gp91–phox [Neufeld et al. 1992] genes. We report here that the second cut repeat, like full-length mClox, optimally selects oligonucleotides containing a CCAAT box motif in addition to the ATTA core sequence. The CCAAT box is neither necessary nor sufficient for DNA-binding activity, indicating that binding to the ATTA core is required for the cut repeat II to recognize the CCAAT motif. Taken together, these results suggest that the proposed role of the mammalian cut homolog as CCAAT displacement protein can be attributed to the DNA-binding activity of the cut repeat II. Figure 8 shows an integrative model for mClox function. Support to this model comes in part from previous observations that Clox and CDP DNA-binding activity is high in committed undifferentiated cells, and is downregulated in terminally differentiated cells [Barberis et al. 1987, Skalnik et al. 1991, Andrés et al. 1992, Neufeld et al. 1992]. Furthermore, mClox failed to trans-activate in transient transfection assays reporter gene constructs containing one or three copies of the mClox consensus binding site [data not shown], in agreement with its proposed role as repressor of gene expression. Further investigation is required to ascertain the role of the first and third cut repeats on mClox function.

Materials and methods

Eukaryotic expression systems

Eukaryotic expression vectors were constructed by subcloning Clox and CDP cDNAs into pMT2 [Kaufman et al. 1989]. The pMT2–CDP plasmid is described in Neufeld et al. (1992). The chimeric CDP/Clox cDNA was constructed by ligating a ~2.7-kb Xhol–ScaI fragment excised from pBluescript II SK-CDP [Neufeld et al. 1992] and a ~2.5-kb SacI–NsiI fragment excised from pGEM7–Clox (Andrés et al. 1992) into XhoI–NsiI-digested pGEM7zf(+) [Promega]. The ~5.2-kb CDP/Clox cDNA was then excised by digestion with Asp718 and XbaI (polynucleotides, made blunt-ended with Klenow polymerase, and purified on an agarose gel. Ligation of EcoRI–NotI adaptors to the blunt-ended CDP/Clox cDNA was done as suggested by the manufacturer [Pharmacia]. The resulting DNA was ligated into EcoRI-digested pMT2. Restriction analysis was performed to select a clone with sense orientation. The procedure for cell culture and transient transfection of COS cells, and the preparation of whole cell extracts have been described previously [Thompson et al. 1991; Thompson et al. 1992]. Western blot analysis with anti-Clox antiserum [a-P, Andrés et al. 1992] showed high levels of expression in transfected cells compared with endogenous levels in untransfected cells [data not shown].

In vitro selection of optimal binding sites

The procedure for the selection of optimal binding sites was essentially as described by Wright and co-workers [Wright et al. 1991, Funk and Wright 1992]. Two libraries of degenerated oligonucleotides were constructed by synthesizing oligonucleotides bearing a stretch of either 15 or 40 random bases flanked by 20-mer PCR priming and cloning sequences (5′-GCC-SalI-HindIII-XbaI-NsiI/40/EcoRI-BamHI-Xhol-CG-3′). The random sequence was generated using a premixed solution of all four phosphoramidites to ensure an equal representation of all 4 bases at each position. 20-Mer P1 and P2 primers complementary to either the 5′- or 3′-flanking sequences, respectively, were synthesized for PCR purposes. The degenerated oligonucleotide [10 μg] was extended by one cycle of PCR (94°C for 1 min, 65°C for 1 min, 72°C for 20 min] in 10 μl of PCR buffer (10 mm Tris-HCl at pH 8.5, 50 mm KCl, 1.5 mm MgCl2, 0.1% Triton X-100) containing 0.2 mm each dNTP, 1 nmol of P1 primer, and 10 units of Taq DNA polymerase [Perkin-Elmer Cetus]. The resulting double-stranded oligonucleotide was gel purified in a 10% acrylamide gel, eluted, and precipitated with ethanol.

All of the incubations described below were done in polypropylene microcentrifuge tubes at room temperature in a rocking platform, and collection of the beads was done by a ~3-sec spin in a microcentrifuge set at maximum speed. For the first cycle of selection with the full-length mClox protein, 3.5 μg of doublestranded random [Ni4] oligonucleotide was incubated with 4 μg of COS cells transfected with either CDP/Clox or CDP (see above) in 20 μl of binding buffer (20 mm HEPES at pH 7.8, 20% glycerol, 0.2 mm EDTA, 1 mm DTT, 1 mm MgCl2 1 BSA). After 20 min, the reaction was brought to 0.45 ml with binding buffer, and 30 μl of a 50% slurry of protein A beads [Repligen] saturated with preimmune serum (85 mg/ml of immunoglobulins) was added. After 30 min, the beads were pelleted and the supernatant was incubated for 60 min with 30 μl of protein A beads saturated with α-FP antibodies (50 mg/ml of immunoglobulins), a specific anti-Clox polyclonal antibody [Andrés et al. 1992]. The beads were then washed three times with washing buffer (0.1 M KCI in binding buffer), once with PCR buffer, and resuspended with 95 μl of PCR buffer. DNA was eluted by incubation at
95°C for 5 min and, after harvesting the beads, the supernatant was added to 9 μl of PCR mixture containing all four dNTPs (0.2 mM final concentration), 100 pmoles each P1 and P2 primer, and 5 units of Taq DNA polymerase. Twenty-microliter aliquots were taken after 10, 15, and 20 cycles of amplification (94°C for 15 sec, 65°C for 1 min, 72°C for 15 sec), and 6 μl was analyzed in a 4% agarose gel stained with ethidium bromide. DNA from the aliquot that gave the first visible band was ethanol precipitated and subjected to a new cycle of selection/amplification, in which sonicated calf thymus DNA (5 μg, cycles 2–6) or poly[d(I-C)] (5 μg, cycles 7–9) was used as nonspecific competitors.

For the selection with GST–Hd and GST–Cut II (see below), the double-stranded random oligonucleotide (5 μg of N15, or 15 μg of N30) was incubated with 5 ng of protein in 15 μl of binding buffer containing 2 μg of poly[d(I-C)]. After 20 min, 5 μl of a 50% slurry of glutathione–agarose beads ([Sigma] was added and incubated for 5 min. The washing of the beads, DNA elution, and PCR amplification were performed as indicated above. The amount of protein was reduced to 2.5 ng in cycles five and six to increase the specificity of the selection.

After selection, the PCR products were digested with HindIII and BamHI and cloned into equally digested p Bluescript II KS (Stratagene). Individual clones were sequenced using M13 universal primers and an automatic DNA sequencer (Applied Biosystems, model 370A).

Preparation of GST proteins

Specific primers containing BamHI (5’ primer) and EcoRI (3’ primer) sites were used to synthesize by PCR the three cut repeat cDNAs. The Clox cDNA (Andrés et al. 1992) was used as template for mClox cut repeats II and III. CDP cDNA (Neufeld et al. 1992) was used for mClox cut repeat I because the partial Clox cDNA lacks the first 6 residues of this repeat. It is noteworthy that Clox and CDP cut repeats are identical at the amino acid level. To generate the Drosophila cut repeat II, the Cut cDNA was used as template (Blochlinger et al. 1988). The amplified fragments were digested with BamHI and EcoRI and cloned into equally digested pGEX2-TK (Glutagene). Because the Clox homeo domain has a naturally occurring BamHI site, the two primers used to amplify this region contained an EcoRI site. The PCR product was digested with EcoRI, cloned into equally digested pGEX2-T, and a clone with the sense orientation was chosen for expression. Carboxy-terminal deletions of the cut repeat II and homeo domain were similarly generated using nested 3’ primers for the PCR reaction, except that in this case the 5’ primer used to synthesize the homeo box cDNA, which overlapped the sequence immediately upstream of the naturally occurring BamHI site, also had a BamHI site. The resulting PCR products were digested with EcoRI and BamHI and cloned into equally digested pGEX2-TK. pGEX–3XGtx is described in Komuro et al. (1993). GST fusion proteins were overexpressed in Escherichia coli (HB101) and purified essentially as described by Smith and Johnson (1988).

In vitro binding assays

Synthetic oligonucleotides used for GMSA were end-labeled with T4 polynucleotide kinase. Cbs-1 sequence has been reported previously (Andrés et al. 1992), and gpd corresponds to the CDP-binding site in which the distal CCAAT box has been mutated to CCggT to abolish binding of an endogenous CCAAT-binding factor (DIST, see Neufeld et al. 1992). Other oligonucleotide sequences are shown in Figure 3C. For GMSA, the indicated proteins were preincubated for 5 min with 100 μg/ml of poly[d(I-C)] in binding buffer (see above). Then, 10–20 pmoles of probe together with cold competitor where indicated, was added for a final reaction volume of 20 μl. Incubation was continued for 15 min, and the samples were electrophoresed at 4°C in 0.5× TBE, 3.5–5% native polyacrylamide gels (80:1 acryl/ bis). GMSA involving the addition of antibodies was carried out as described previously (Andrés et al. 1992).

For the in vitro binding assays involving GST beads and [35S]methionine-labeled Clox protein, RNA synthesis and in vitro translation with rabbit reticulocyte lysate was performed using 1 μg of Clox 1 cDNA template (Andrés et al. 1992) linearized with SacI, and 10 units of SP6 RNA polymerase, as suggested by the manufacturers [Promega, Stratagene]. Labeled in vitro-translated protein (4 μl) was diluted to 0.425 ml binding buffer (see above). Then, either GST or GST–CloxII glutathione–Sepharose beads [Sigma] were added at the indicated amounts, and the mix was incubated overnight on a rotary platform at 4°C. Beads were then washed four times with ice-cold NETN buffer (20 mm Tris-HCl at pH 8, 100 mm NaCl, 1 mm EDTA, 0.5% NP-40). The bound protein was eluted by boiling in Laemmli’s buffer and separated by 10% SDS-PAGE [Laemmli 1970]. The gels were fixed, dried, and analyzed by autoradiography.

Acknowledgments

We especially thank Bernardo Nadal-Ginard for constant support and invaluable discussion throughout the course of this study. We also thank Ellis Neufeld for the CDP cDNA; Joel Pomerantz and Phillip Sharp for the gift of PA–Oct-1 homeo domain, Issie Komuro and Seigo Izumo for pGEX–3XGtx homeo domain; Karen Blochlinger for the Cut cDNA; Tom O’Keeffe and Roger McCarrick for DNA sequencing, Clara Hwang and Maria J. Andrés for technical assistance; Steve Harrison and members of our laboratory for suggestions and helpful discussions; Roger Breitbart and John McDermott for critical reading of the manuscript; and Greta Knaaschaefeer for excellent graphical work. V.A. has been supported in part by research fellowships from the Spanish M.E.C. and from the American Heart Association [Massachusetts Affiliate Inc.], and M.D.C. by the Spanish M.E.C. This work was supported in part by grants from the National Institutes of Health to V.M.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 USC section 1734 solely to indicate this fact.

References

Andrés, V., B. Nadal-Ginard, and V. Mahdavi. 1992. Clox, a mammalian homeobox gene related to Drosophila cut, encodes DNA-binding regulatory proteins differentially expressed during development. Development 116: 321–334.

Aurora, R. and W. Herr. 1992. Segments of the POU domain homeo gene promoter. Mol. Cell. Biol. 12: 455–467.

Barberis, A., G. Superti-Furga, and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT-binding factor and of a displacement protein with overlapping sequences of a histone gene promoter. Cell 50: 347–359.

Blochlinger, K., R. Bodmer, J.W. 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.

Blochlinger, K., R. Bodmer, L.Y. Jan, and Y.N. Jan. 1990. Pat-
terns of expression of Cut, a protein required for external sensory organ development in wild-type and cut mutant Drosophila embryos. Genes & Dev. 4: 1322–1331.

Blochinger, K., L.Y. Jan, and Y.N. Jan. 1991. Transformation of sensory organ identity by ectopic expression of Cut in Drosophila. Genes & Dev. 5: 1124–1135.

——. 1993. Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila. Development 117: 441–450.

Bodmer, R., S. Barbel, S. Sheperd, J.W. Jack, L.Y. Jan, and Y.N. Jan. 1987. Transformation of sensory organs by mutations of the cut locus of D. melanogaster. Cell 51: 293–307.

Catron, K.M., N. Iler, and C. Abate. 1993. Nucleotides flanking a conserved TAAT core dictate the DNA binding specificity of three murine homeodomain proteins. Mol. Cell. Biol. 13: 2354–2365.

Dearolf, C.R., J. Topol, and C.S. Parker. 1989. The caudal gene of Drosophila melanogaster. Genes & Dev. 1: 1322–1331.

Herr, W., R.A. Sturm, R.G. Clerc, L.M. Corcoran, D. Baltimore, and M. Rosenfeld. 1990. Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut. Nature Genetics 2: 773–782.

Hollander, P.W.H. and B.L.M. Hogan. 1988. Expression of homeobox genes during mouse development: A review. Genes & Dev. 2: 773–782.

Ingraham, H., S. Flynn, J. Voss, V. Albert, M. Kapiloff, L. Wilson, and M. Rosenberg. 1990. The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA-binding and DNA-dependent Pit-1-Pit-1 interactions. Cell 61: 1021–1033.

Jack, J., D. Dorsett, Y. Delotto, and S. Liu. 1991. Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development 113: 735–747.

Johnson, P.F. and S.L. McKnight. 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. 58: 799–839.

Karlsson, O., S. Thor, T. Norberg, H. Ohlsson, and T. Edlund. 1990. Insulin gene enhancer binding protein Is1-1 is a member of a novel class of proteins containing both a homeo- and a Cys-His domain. Nature 344: 879–882.

Kaufman, R.J., M.W. Davies, V.K. Pathak, and J.W.V. Hershey. 1989. The phosphorylation state of eukaryotic initiation factor 2 alters translational efficiency of specific mRNAs. Mol. Cell. Biol. 9: 946–958.

Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. Pabo. 1990. Crystal structure of an engrailed homeodomain-DNA complex at 2.8 A resolution: A framework for understanding homeodomain-DNA interactions. Cell 63: 579–590.

Komuro, I., M. Schalling, L. Jahn, R. Bodmer, N.A. Jenkins, N.G. Copeland, and S. Izumo. 1993. Gtx: A novel murine homeobox-containing gene, expressed specifically in glial cells of the brain and germ cells of testis, has a transcriptional repressor activity in vitro for a serum inducible promoter. EMBO J. 12: 1387–1401.

Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.

Lai, J.-S., M.A. Cleary, and W. Herr. 1992. A single amino acid exchange transfers VP16-induced positive control from the Oct-1 to the Oct-2 homeo domain. Genes & Dev. 6: 2058–2065.

Laugon, A. 1991. DNA binding specificity of homeodomains. Biochemistry 30: 11357–11367.

Liu, S., E. McLeod, and J. Jack. 1991. Four distinct regulatory regions of the cut locus and their effect on cell type specification in Drosophila. Genetics 127: 151–159.

McGinnis, W. and R. Krumlauf. 1992. Homeobox genes and axial patterning. Cell 68: 283–302.

Neufeld, E.J., D.G. Skalnick, P.M.-J. Lievens, and S.H. Orkin. 1992. Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut. Nature Genetics 1: 50–55.

Otting, G., Y.Q. Qian, M. Müller, M. Affolter, W.J. Ghering, 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.

Otting, G., Y.Q. Qian, M. Biller, M. Müller, M. Affolter, W.J. Ghering, and K. Wüthrich. 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. EMBO J. 9: 3085–3092.

Pomerantz, J.L., T.M. Kristie, and P.A. Sharp. 1992. Recognition of the surface of a homeo domain protein. Genes & Dev. 6: 2047–2057.

Porter, S.D. and M. Smith. 1986. Homeo-domain homology in yeast MAT α2 is essential for repressor activity. Nature 320: 766–768.

Qian, Y.Q., M. Biller, G. Otting, M. Müller, W.J. Ghering, and K. Wüthrich. 1989. The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: Comparison with prokaryotic repressors. Cell 59: 573–580.

Scott, M.P., J.W. Tamkun, and G.W. Hartzell, III. 1989. The structure and function of the homeodomain. Biochim. Biophys. Acta 989: 25–48.

Shashikant, C.S., M.F. Utset, S.M. Violette, T.L. Wise, P. Einat, J.W. Pendleton, K. Schugart, and F.H. Ruddle. 1991. Homeobox genes in mouse development. Crit. Rev. in Euk. Gene Exp. 1: 207–245.

Skalnik, D.G., E.C. Strauss, and S.H. Orkin. 1991. CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter. J. Biol. Chem. 266: 16736–16744.

Smith, D.B. and K.S. Johnson. 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. Gene 67: 31–40.

Sturm, R.A. and W. Herr. 1988. The POU domain is a bipartite DNA-binding structure. Nature 336: 601–604.

Sturm, R.A., G. Das, and W. Herr. 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU domain with a
homeo box subdomain. *Genes & Dev.* 2: 1582–1599.
Superti-Furga, G., A. Barberis, G. Schaffner, and M. Busslinger. 1988. The -117 mutation in greek HPFH affects the binding of the three nuclear factors to the CCAAT region of the γ-globin gene. *EMBO J.* 7: 3099–3107.
Superti-Furga, G., A. Barberis, E. Schreiber, and M. Busslinger. 1989. The protein CDP, but not CP1, footprints on the CCAAT region of the γ-globin gene in unfractionated B-cell extracts. *Biochim. Biophys. Acta* 1007: 237–242.
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.
Thompson, W.R., B. Nadal-Ginard, and V. Mahdavi. 1991. A MyoD1-independent muscle-specific enhancer controls the expression of the β-myosin heavy chain gene in skeletal and cardiac muscle cells. *J. Biol. Chem.* 266: 22678–22688.
Treisman, J., P. Gonczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* 59: 553–562.
Treisman, J., E. Harris, and C. Desplan. 1991. The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes & Dev.* 5: 594–604.
Treisman, J., E. Harris, D. Wilson, and C. Desplan. 1992. The homeodomain: A new face for the helix-turn-helix? *BioEssays* 14: 145–150.
Umek, R.M., A.D. Friedman, and S.L. McKnight. 1991. CCAAT-enhancer binding protein: A component of a differentiation switch. *Science* 251: 288–292.
Verrijzer, C.P., A. Kal, and P.C. van der Vliet. 1990. The Oct-1 homeo domain contacts only part of the octamer sequence and full Oct-1 DNA-binding activity requires the POU-specific domain. *Genes & Dev.* 4: 1964–1974.
Verrijzer, C.P., M.J. Alkema, W.W. van Wepener, H.C. Van Leeuwen, M.J.J. Strating, and P.C. van der Vliet. 1992. The DNA binding specificity of the bipartite POU domain and its subdomains. *EMBO J.* 11: 4993–5003.
Wolberger, C., A.K. Vershon, B. Liu, A.D. Johnson, and C.O. Fabo. 1991. Crystal structure of a MATa2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. *Cell* 67: 517–528.
Wright, W.E., M. Binder, and W. Funk. 1991. Cyclic amplification and selection of targets [CASTing] for the myogenin consensus binding site. *Mol. Cell. Biol.* 11: 4104–4110.
Yu, Y.-Y., R.E. Breitbart, V. Mahdavi, and B. Nadal-Ginard. 1992. Human myocyte-specific enhancer factor 2 (MEF2) comprises a group of transcription factors belonging to the MADS gene family. *Genes & Dev.* 6: 1783–1798.
A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins.

V Andrés, M D Chiara and V Mahdavi

*Genes Dev.* 1994, 8:

Access the most recent version at doi:10.1101/gad.8.2.245

References

This article cites 54 articles, 22 of which can be accessed free at:

http://genesdev.cshlp.org/content/8/2/245.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.