MOLECULAR AND CELLULAR BIOLOGY, Oct. 1996, p. 5346–5357 Vol. 16, No. 10
0270-7306/96/$04.00+0

The Human Cut Homeodomain Protein Can Repress Gene Expression by Two Distinct Mechanisms: Active Repression and Competition for Binding Site Occupancy

FRANCE MAILLY,[1,2,3] GINETTE BÉRUBÉ,[1,2,3] RYOKO HARADA,[1,2,3] PEI-LIN MAO,[1,2,3] SHERLYN PHILLIPS,[3,4] AND ALAIN NEPVEU[1,2,3,4]*

Departments of Medicine,[1] Oncology,[2] and Biochemistry[4] and Molecular Oncology Group,[3] McGill University, Royal Victoria Hospital, Montreal, Quebec, Canada H3A 1A1

Received 22 November 1995/Returned for modification 18 January 1996/Accepted 9 July 1996

By analogy with other homeodomain proteins conserved in evolution, mammalian Cut proteins are believed, as in *Drosophila melanogaster*, to play an important role in determining cell type specificity in several tissues. At the molecular level, Cut proteins appear to serve as transcriptional repressors. In this study, we have examined the mechanism by which the human Cut (hCut) protein down-regulates gene expression. The homeodomain and the three regions called Cut repeats are evolutionarily conserved and were previously shown to function as DNA binding domains. The carboxy-terminal region, although it does not show amino acid sequence homology per se, in all cases is enriched in alanine and proline residues, a distinctive feature of some transcriptional repression domains. Our results reveal two distinct modes of repression: competition for binding site occupancy and active repression. On one hand, the composite DNA binding domain formed by Cut repeat 3 and the Cut homeodomain was shown to bind to CCAAT and Sp1 sites within the *tk* gene promoter and to reduce gene expression, presumably by preventing activation by the corresponding transcription factors. On the other hand, the carboxy-terminal region of mammalian Cut proteins was found to function as an active repression domain in a distance-independent manner. We have further narrowed this activity to two subdomains that can independently repress activated transcription. Finally, we present a model to illustrate the two mechanisms by which Cut proteins repress gene expression.

The cDNAs for several mammalian homologs of the *Drosophila* Cut homeodomain protein (7) have recently been isolated (2, 22, 56, 66, 69). One human cDNA was isolated by using antibodies raised against a purified preparation of the human CCAAT displacement protein (CDP) (56). Another human cDNA with an identical sequence was obtained by screening of an expression library with an oligonucleotide encoding an Sp1-like site present in the promoter of the c-*myc* proto-oncogene (22). Other mammalian Cut homologs isolated from dog, mouse, and rat cells were termed Clox (Cut-like homeobox), Cux (Cut homeobox), and CDP-2, respectively (2, 66, 69). The terms human Cut and mouse Cut (hCut and mCut) will be used hereafter.

Sequence homology between *Drosophila* and mammalian Cut proteins is limited to five evolutionarily conserved domains: a region predicted to form a coiled-coil structure, three related regions called Cut repeats, and a distinctive homeodomain with a histidine residue at the ninth amino acid of the third helix (7, 56). The homeodomain and the three Cut repeats were shown to be capable of DNA binding (1, 4, 30, 31). Cut proteins therefore contain four DNA binding regions. Cut repeat 3 and the homeodomain form a bipartite DNA binding domain which can specifically bind to DNA with high affinity (1, 4, 30, 31). The mode of interaction of Cut repeats 1 and 2 with DNA is less clear. Either region was found to specifically bind to DNA when expressed as part of a fusion protein together with glutathione *S*-transferase but not with the maltose-binding protein (4, 25). Since glutathione *S*-transferase fusion proteins exist as dimers and maltose-binding protein fusion proteins exist as monomers, it was concluded that only dimers of Cut repeat 1 or 2 can interact with DNA (30). While the nature of the interaction between native Cut proteins and DNA remains to be fully investigated, studies involving PCR-mediated random oligonucleotide selection have confirmed that Cut can bind to sequences closely related to the CCAAT and Sp1 consensus binding sites (2, 4, 30).

Although the biological function of Cut in mammals remains to be defined, the phenotypic expression of lethal and viable mutations at the *cut* locus in *Drosophila melanogaster* suggests that this locus is involved in cell type specification in several tissues (7–11, 34, 46, 47). Cut mutants present with defects in a variety of tissues, including the wings, legs, external sense organs, Malpighian tubules, tracheal system, and central nervous system (8, 11, 34, 46, 47). In flies carrying the viable *cut wing* mutation, which maps to a regulatory region upstream of the gene, cells that should differentiate to form the wing margin instead undergo apoptosis, thereby producing the truncated wing phenotype (34, 47). Embryonic lethal *cut* mutations, which alter the coding region, result in the transformation of external sensory organs and Malpighian tubules into internal sensory (chordotonal) organs and gut tissue, respectively (47). Thus, in *D. melanogaster*, the lack of functional Cut protein causes some cells to embark on the wrong developmental pathway. Conversely, forced expression of Cut in embryos resulted in the replacement of internal sensory organs by external sensory organs (9). By analogy with other homeodomain proteins also conserved in evolution, it is expected that mammalian Cut proteins will be found to play an equally important role in determining cell type specificity in mammals.

At the molecular level, Cut proteins are believed to function as transcription factors. In line with this view, in both *D. mela-*

* Corresponding author. Mailing address: Royal Victoria Hospital, 687 Pine Ave. West, Montreal, Quebec, Canada H3A 1A1. Phone: (514) 842-1231, ext. 5832. Fax: (514) 843-1478. Electronic mail address: alain@lan1.molonc.mcgill.ca.

nogaster and mice, Cut proteins have been localized to the nucleus (7, 66). Moreover, mammalian Cut proteins have been found to act as transcriptional repressors in tissue culture models. CDP/hCut was found to bind to upstream regulatory sequences of the gp91-phox gene (62), and the expression of this gene was shown to coincide with down-regulation of CDP/hCut binding activity upon differentiation of myeloid cells (44, 62). In cotransfection experiments, recombinant mammalian Cut proteins repressed transcription of reporter genes driven by the promoter of either the c-*myc*, Ncam, or gp91-phox gene, as well as by a promoter in which Cut consensus binding sites had been inserted (21, 22, 62, 66, 69).

In recent years, the study of negative transcriptional regulators has uncovered multiple mechanisms by which repression can be exerted (reviewed in reference 17). Some repressors act by interacting with specific DNA binding sites overlapping those of transcriptional activators, thereby competing for DNA site occupancy and preventing activation. Other repressors appear to negatively regulate transcription through a region of the protein distinct from the DNA binding domain (17, 18, 29, 49, 61), by a mechanism that has been termed active repression. Although little is known about their mechanism of action, it is assumed that active repression domains contact one or another component of the transcriptional machinery in a manner analogous to that of transactivators, albeit with opposite consequences (17). In support of this mode of action, the unliganded thyroid hormone receptor was found to prevent assembly of a functional preinitiation complex (24), and the Dr-1 and Eve proteins were shown to interact, in vivo and in vitro, with the TATA-binding protein (65, 68). That the TATA-binding protein can be a target for repression was also demonstrated by in vitro transcription studies using purified fractions of the RNA polymerase II general transcription factors (5, 68). Thus, active repression domains appear to inhibit the formation of a functional preinitiation complex, but the exact step that is inhibited, DNA binding or protein-protein interactions, is still a matter of debate (5, 14, 65). Active repressors may also interact with an activator bound at an adjacent position on DNA, thereby inhibiting its function, a mechanism referred to as quenching (36). Finally, it was proposed that some transcriptional repressors may alter the local chromatin structure to cause a reduction in gene expression (12).

In this study, we have examined the mechanism by which mammalian Cut proteins down-regulate gene expression. Our results have revealed two distinct modes of repression by Cut proteins: active and passive repression. We present evidence that the carboxy-terminal regions of the human and murine Cut proteins can function as active repression domains. The repression activity was mapped within the human Cut protein to two subdomains of 81 and 29 amino acids (aa) that independently repressed gene expression from an activated promoter. In addition, we found that the composite DNA binding domain designated CR3HD (Cut repeat 3 and homeodomain) can bind to a CCAAT and an Sp1 site within the *tk* gene promoter and can down-regulate expression of this promoter in transient transfection assays. We present a model to illustrate the modes of action of Cut proteins in transcriptional repression.

MATERIALS AND METHODS

Plasmid construction. The sequence of the hCut protein has been published as the human CDP sequence, and the cDNA sequence (HSCDP) can be obtained from GenBank under accession number M74099 (56). The nucleotide and amino acid numbers used hereafter are taken from this cDNA sequence and its deduced amino acid sequence. To localize the transcriptional repression activity within the hCut protein, various segments of the hCut cDNA were inserted in frame within the expression vector pSG424 or pGALM (39, 58), thereby generating a series of yeast GAL4 (aa 1 to 147)hCut fusion proteins. For the following constructs, we started from the 76.2 cDNA clone comprising nucleotides (nt) 1605 to 5376 of HSCDP and inserted into the pBluescribe KS vector (Stratagene). An *Eco*RI site had been added to nt 1605 during cDNA cloning. The following hCut fragments were inserted into the pSG424 vector: Cut repeat 1, *Eco*RV (nt 1605)-*Bam*HI (nt 2019) fragment treated with Klenow enzyme and inserted into the *Sma*I site; Cut repeat 1 plus linker, *Eco*RV (nt 1605)-*Acc*I (nt 2853) fragment treated with Klenow enzyme and inserted into the *Sma*I site; Cut repeats 2 and 3, *Rsa*I (nt 2861)-*Rsa*I (nt 3737) fragment treated with T4 DNA polymerase and inserted into the *Sac*I site; CR3HD, *Sau*96I (nt 3379)-*Sau*96I (nt 3982) fragment treated with Klenow enzyme and inserted into the Klenow enzyme-treated *Bam*HI site; Cut homeodomain, *Pvu*II (nt 3772)-*Apo*I (nt 3963) fragment treated with Klenow enzyme and inserted into the Klenow enzyme-treated *Bam*HI site; carboxy-terminal region (aa 1299 to 1505), *Bam*HI (nt 3936)-*Sfi*I (nt 4603) fragment (treated with T4 DNA polymerase following *Sfi*I digestion and prior to *Bam*HI digestion) inserted into the *Bam*HI and Klenow enzyme-treated *Sal*I sites; aa 1299 to 1336, *Bam*HI (nt 3936)-*Pvu*II (nt 4053), and aa 1299 to 1362, *Bam*HI (nt 3936)-*Bsr*bI (nt 4129) fragments inserted into the *Bam*HI and Klenow enzyme-treated *Sal*I sites; aa 1299 to 1380, *Sma*I (in vector) (Cut nt 3936)-*Sma*I (nt 4185) fragment inserted into the *Sma*I site; aa 1316 to 1459, *Bgl*I (nt 3990)-*Bgl*I (nt 4422) fragment treated with T4 DNA polymerase and inserted into the *Sma*I site; aa 1338 to 1505, *Pvu*II (nt 4053)-*Sfi*I (nt 4603) fragment treated with T4 DNA polymerase and inserted into the *Sma*I site; aa 1363 to 1505, *Bsr*bI (nt 4129)-*Sfi*I (nt 4603) fragment treated with T4 DNA polymerase and inserted into the Klenow enzyme-treated *Bam*HI site; aa 1380 to 1505, *Ppu*mI (nt 4179)-*Sfi*I (nt 4603) fragment with Klenow enzyme and T4 DNA polymerase treatments, respectively, and inserted into the *Sma*I site; aa 1461 to 1505, *Bgl*I (nt 4422)-*Bgl*I (nt 4602) fragment treated with T4 DNA polymerase and inserted into the *Sma*I site; and aa 1399 to 1468, *Sma*I (nt 4239)-*Sma*I (nt 4448) fragment inserted into the *Sma*I site. The fragment containing Cut repeats 1, 2, and 3 and the homeodomain (nt 1605)-*Apo*I (nt 3963) was cloned into the *Eco*RI site; the fragment containing Cut repeats 2 and 3, the homeodomain, and the carboxy terminus was obtained following *Xho*I (nt 2861)-*Bam*HI (nt 3936) digestion of a cDNA clone isolated from a cDNA library in pMX and was inserted into the corresponding sites. The aa 1316 to 1380 *Bgl*I (nt 3990)-*Sma*I (nt 4185) construct was derived from plasmid pSG424 1316-1459 after digestion with *Sma*I and *Xba*I, treatment with T4 DNA polymerase, and religation on itself. Four of the pSG424 hCut constructs described above (aa 1338 to 1505, aa 1363 to 1505, aa 1380 to 1505, and aa 1399 to 1468) were cut with *Eco*RI (vector) and *Bgl*I (T4 DNA polymerase treated), inserted into the pGALM expression vector digested with *Xba*I, treated with Klenow enzyme, then further digested with *Eco*RI. This generated the aa 1338 to 1459, aa 1363 to 1459, aa 1380 to 1459, and aa 1399 to 1459 constructs.

The following hCut fragments were inserted into the *Eco*RI-*Xba*I sites of pSG424 following DNA amplification via PCR and digestion of the corresponding restriction sites added to the 5′ end of the primers: aa 1451 to 1505 (nt 4393 to 4567), aa 1424 to 1505 (nt 4307 to 4567), aa 1476 to 1505 (nt 4470 to 4567), aa 1451 to 1485 (nt 4393 to 4500), and divergent region (aa 1345 to 1449, nt 4079 to 4393). The hCut repression domain without the GAL4 DNA binding domain (aa 1299 to 1505) was amplified with an upper strand primer containing a Kozak consensus sequence (40) and a nuclear localization signal (25) (5′-AAATCGA TCCACCATGGGTTCCTAAAAAGAAGCGCAAGGTTGAATTC-3′) followed by nt 3941 to 3960 of the hCut cDNA sequence. The PCR product was cloned directly into the PCRII vector (Invitrogen, San Diego, Calif.). This construct was then digested with *Eco*RI and subcloned into the similarly digested vector pSG5 (Stratagene). The following fragments from the Cux/mCut cDNA (GenBank accession number X75013) were cloned into the *Eco*RI-*Bam*HI sites of pGALM following PCR amplification: aa 1114 to 1332 (nt 3662 to 4321) and divergent region (aa 1158 to 1300, nt 3794 to 4162). The integrity of clones and their reading frames was verified by sequencing.

Six different reporter constructs with the chloramphenicol acetyltransferase (CAT) gene were used in the transfection experiments. The pBL2CAT (containing the −105 to +51 region of the *tk* promoter) (48), GAL4E1bCAT (minimal promoter containing a TATA box) (45), and GAL4InrCAT (initiator element) (55) plasmids have been described previously. The GAL4*tk*CAT plasmid was constructed by the insertion of five GAL4 binding sites into the *Hin*dIII-*Xba*I sites of pBLCAT2 (51, 61). To generate the GAL4–1608-kb spacer–*tk*-CAT (GAL4Spacer*tk*CAT) reporter plasmid, the last construct was further modified by inserting a genomic *Bam*HI 1608-bp fragment originating from the intergenic region downstream of the murine c-*myc* gene into the *Bam*HI site located between the GAL4 binding sites and the *tk* promoter. The GAL4*tk*ΔClaICAT plasmid was obtained by digestion of the GAL4*tk*CAT plasmid isolated from Dam$^-$ bacteria with *Cla*I, followed by Klenow enzyme fill-in, digestion with *Sma*I (immediately downstream of the *Cla*I site), and religation.

PCR amplification. Plasmid DNA containing portions of the hCut cDNA was used as the template for the amplification of various portions of the human or murine *cut* gene. Oligonucleotide primers were either obtained from the McGill Biotechnology Center (Montreal, Quebec, Canada) or synthesized on a Pharmacia Gene Assembler. Amplification was performed on a Coy Tempcycler unit, using 1 U of *Taq* polymerase (Gibco-BRL) as specified by the manufacturer of the enzyme. Following amplification, DNA fragments were digested with relevant restriction endonucleases, purified, and cloned in the pSG424 or pGALM

vector. Each construct obtained by PCR cloning was verified by DNA sequencing.

Transient-transfection and CAT assays. COS and NIH 3T3 cells were grown in Dulbecco modified Eagle medium supplemented with 10% fetal bovine and calf serum, respectively. Cells were plated at a density of 0.6×10^6 cells per 90-mm-diameter plates 24 h prior to the transfection procedure. Plasmid DNA was introduced by the calcium phosphate precipitation technique (NIH 3T3) or the DEAE-dextran method (53, 67). All transfection experiments were repeated at least three times, more often in the case of small effector effects. Except for the results presented in Fig. 4, in which variable amounts of effector plasmid were used to determine the lowest amount needed for repression, typically 10 μg of effector plasmid DNA and 5 μg of reporter CAT plasmid were introduced into cells by the DEAE-dextran method (53). Essentially similar results were obtained with 2 and 1 μg of effector and reporter plasmids, respectively. The amount of transfected DNA was kept constant by addition of appropriate amounts of the parental expression plasmid. As an internal control for transfection efficiency, 1 μg of pCH110 (pSV40-βgal; Pharmacia) was also included, and β-galactosidase activity was assayed essentially as described previously (32). Cells were treated with chloroquine (5 mg/ml of medium) on the next day and harvested with 1 ml of TEN buffer (40 mM Tris [pH 7.5], 1 mM EDTA, 150 mM NaCl) 48 h later. After resuspension in 100 μl of Tris 0.25 M (pH 7.5), cells were subjected to three cycles of freeze-thaw with dry ice-ethanol and a 37°C water baths and centrifuged; the cytoplasmic extracts were recovered and used directly in CAT assays or stored at −80°C. CAT assays were performed as described previously (26); the results were visualized by autoradiography and quantitated by using Fuji imaging plates and a PhosphoImager analyzer.

Preparation of nuclear extracts. The pellet that remained following the preparation of cytoplasmic extracts was used for the preparation of nuclear extracts by the method of Lee et al. (41), with some modifications. Briefly, the pellet was resuspended in 30 μl of buffer C (20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid [HEPES; pH 7.9], 25% glycerol, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA), incubated at 4°C for 30 min with gentle rocking, and spun down at 12,000 rpm for 5 min. The protein extract was either used immediately or quickly frozen in a dry ice-ethanol bath.

DNase footprinting assay. The GAL4*tk*CAT vector, which contains two Sp1 sites and a CCAAT box in the *tk* promoter (−105 to +51 relative to the transcription start site), was used for this analysis. The plasmid was ^{32}P end labeled at the *Bgl*II site (+51) with T4 polynucleotide kinase and cleaved with *Hin*dIII or *Xba*I. The fragments thus generated contained (*Hin*dIII) or did not contain (*Xba*I) the five GAL4 binding sites. After electrophoresis through a 5% polyacrylamide gel, the labeled fragments were purified by passive elution in 10 mM Tris HCl (pH 7.5)–1 mM EDTA. DNase footprinting was carried out essentially as described previously (23, 30). End-labeled DNA (100,000 cpm per reaction) was incubated with 20 μg of protein from nuclear extracts (prepared as outlined above), in the presence of 1 μg of poly(dI-dC) or with 50 ng of bacterially expressed fusion proteins previously purified by affinity chromatography, for 15 min at room temperature in a final volume of 25 μl in 10 mM Tris (pH 7.5)–25 mM NaCl–1 mM $MgCl_2$–1 mM dithiothreitol–5% glycerol–4% (wt/vol) polyvinyl alcohol. Fifty microliters of 10 mM $MgCl_2$–5 mM $CaCl_2$ was added, and tubes were incubated for 90 s. Various dilutions of DNase I were added, and samples were incubated for 90 s. Ninety microliters of DNase stop solution (20 mM EDTA, 1% sodium dodecyl sulfate [SDS], 0.2 M NaCl) was added and mixed by vortexing. Following phenol-chloroform extraction and ethanol precipitation, samples were electrophoresed through an 8% denaturing polyacrylamide gel (30:1) in 1× Tris borate-EDTA. Gels were dried and visualized by autoradiography.

EMSAs. Electrophoretic mobility shift assays (EMSAs) were performed with 5 μg of proteins from nuclear extracts prepared as described above. Samples were incubated at room temperature for 5 min in 10 mM Tris (pH 7.5)–25 mM NaCl–1 mM $MgCl_2$–5 mM EDTA (pH 8.0)–5% glycerol–1 mM dithiothreitol, in a final volume of 40 μl, with 1 μg of poly(dI-dC) and 2 μg of sheared salmon sperm DNA as nonspecific competitors. A double-stranded oligonucleotide containing a GAL4 binding site (upper strand, 5′-CGACTCGGAGGACTGTCCT CCGAG-3′) was end labeled by standard methods. The probe (20,000 cpm, ~10 pg) was then added to the protein mixture, and samples were further incubated for 20 min. Samples were loaded on a 5% polyacrylamide gel (30:1) and separated by electrophoresis at 8 V/cm for 2 h in 0.5× Tris borate-EDTA. Gels were dried and visualized by autoradiography.

Western blot (immunoblot) analysis. Nuclear extracts were prepared as described above and mixed with 3× Laemmli buffer. Equal amounts of protein (5 to 10 μg) were boiled for 5 min and loaded on an SDS–10% polyacrylamide gel. The gel was soaked for 10 min in a solution of 0.1 M Tris–0.192 M glycine–20% (vol/vol) methanol, and proteins were electrotransferred to a nylon membrane for 1 h at 4°C. The membrane was then rinsed twice in phosphate-buffered saline (PBS) supplemented with 0.1% Tween (PBS-Tween) and incubated for 1 h at room temperature in PBS-Tween containing 3% bovine serum albumin to prevent nonspecific binding of the antibody. Following a brief wash, the blot was incubated in PBS-Tween with a pool of three monoclonal antibodies directed against the carboxy-terminal region of hCut. The membrane was washed three times for 10 min each time in 100 ml of PBS-Tween and then incubated with a second antibody conjugated to horseradish peroxidase for 1 h at room temperature. After brief washes in PBS-Tween and PBS, proteins were visualized by using the enhanced chemiluminescence (ECL) system (Amersham) according to the manufacturer's instructions.

RESULTS

Two domains in the carboxy-terminal region of hCut can independently act as active repression domains. Mammalian Cut homeodomain proteins have been found to repress expression of target genes containing a Cut binding site. Interestingly, while no direct sequence homology is apparent between the carboxy termini of *Drosophila* and mammalian Cut proteins, both are enriched in alanine residues, a distinctive feature of previously characterized transcriptional repressors such as the Engrailed, Krüppel, Even-skipped, and Msx-1 homeodomain proteins. These proteins appear to decrease transcription by a mechanism termed active repression (17, 18, 29, 35, 49) whereby the regulatory activity of the protein is conferred by a region residing outside the DNA binding domain. We therefore sought to verify whether the carboxy-terminal region of hCut exhibits any repression activity. We prepared a series of vectors expressing chimeric proteins with various portions of hCut linked to aa 1 to 147 of the yeast GAL4 transcription factor. This region of GAL4 contains a specific DNA binding and dimerization domain. As shown in Fig. 1, the carboxy-terminal region of hCut fused to the GAL4 DNA binding domain can function as an active repression domain.

To define more precisely the amino acids that constitute the active repressors, we prepared a series of constructs with amino-terminal and carboxy-terminal deletions of the hCut carboxy-terminal region. Two regions were found to function as active repression domains: a region of 81 aa immediately downstream of the homeodomain (aa 1299 to 1380 [Fig. 1, lanes 5 and 14]) and a region of 54 aa at the carboxy terminus (aa 1451 to 1505 [lane 18 and Table 1]). Removal of 17 and 18 aa at either end of the 1299-1380 protein completely abolished its repression capability (Fig. 1, lanes 13 to 16). In contrast, deletion of few amino acids from either end of the 1451-1505 protein diminished but did not eliminate its repression effect (Fig. 1, lanes 17 to 21). In fact, this region could be reduced to 29 aa, from aa 1476 to 1505, without much reduction in potency (Fig. 1, lane 20). To ensure that the inability of some fusion proteins to repress was not due to a lack of expression or greatly reduced stability, EMSAs were performed with oligonucleotides encoding the GAL4 binding site together with nuclear extracts from transfected COS cells (Fig. 2). All extracts were found to contain GAL4 DNA binding activity, and although some variation was observed in the intensity of the signals, it did not account for the differences in the ability to repress the *tk*CAT promoter. Several proteins exhibiting stronger band shift signals were in fact those with the lowest repression potential (Fig. 2, lanes 5 and 6). One exception is the 1424-1505 fusion protein: DNA binding is clearly weaker than for the 1451-1505 protein (compare lanes 7 and 8). It is possible, therefore, that the absence of repression by the 1424-1505 fusion protein is due to its lower expression, stability, or DNA binding activity. In summary, the carboxy-terminal region of hCut contains two subdomains of 81 and 55 aa that can function as repressors. The smallest peptide with repression activity contained 29 aa from the hCut carboxy terminus.

The carboxy-terminal region from the mCut protein (Cux) also functions as an active repression domain. The hCut and mCut are nearly identical except in two regions that show marked sequence divergence: upstream of Cut repeat 2 and in the carboxy-terminal region (56, 66). A sequence comparison of their carboxy termini is presented in Fig. 3A. The two proteins are virtually identical at the beginning and the end of

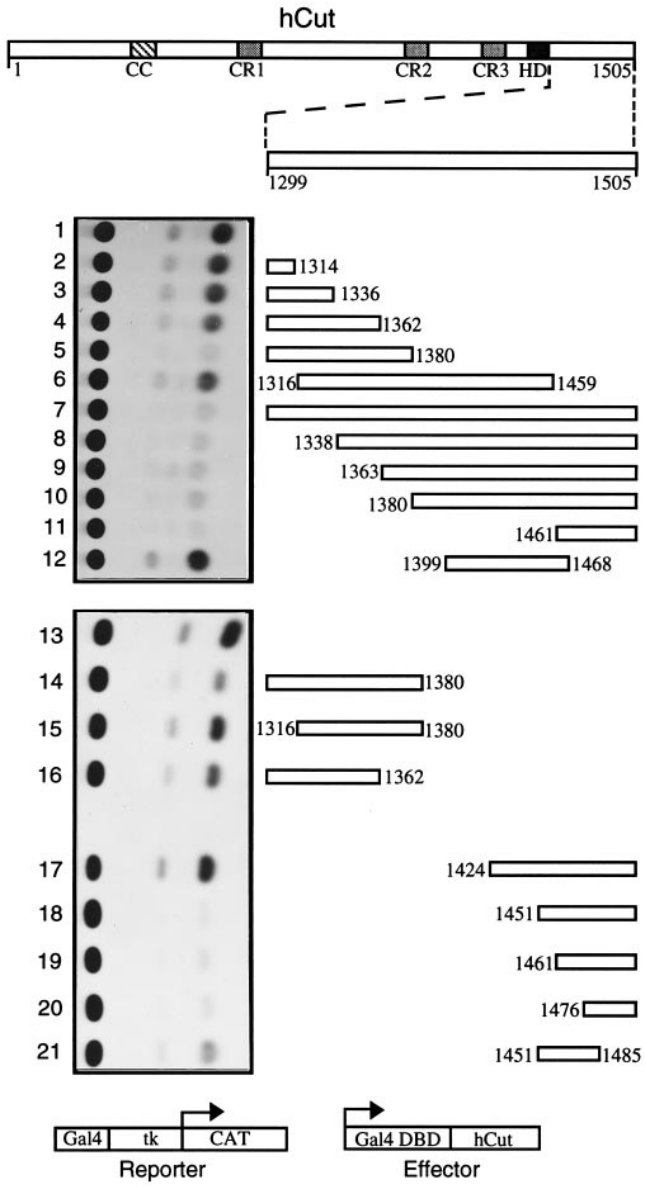

FIG. 1. Identification of two active repression domains in the human Cut carboxy-terminal region. Transfection assays were performed in COS cells, using the GAL4*tk*CAT reporter plasmid. The effector plasmids contained sequences encoding various portions of the hCut carboxy-terminal region fused to the yeast GAL4 DNA binding domain (DBD). Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. A schematic representation of the hCut protein is displayed at the top, with the evolutionarily conserved regions depicted as boxes: CC, coiled coil; CR, Cut repeat; HD, homeodomain. Schematic diagrams of the hCut segments encoded by the effector plasmids are shown on the right, with numbers identifying the first and last amino acids. Lanes 1 to 12 show the constructs used for the initial identification of the two autonomous repression domains; the refined delineation of the domains is presented in lanes 13 to 21. All transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown. Values representing the average of three separate experiments are provided in Table 1 for some of the effector plasmids.

this region but differ markedly in the central portion. To verify whether sequence divergences may impart different regulatory properties, we tested in parallel the two carboxy-terminal regions in cotransfection experiments using the same reporter construct. The carboxy-terminal region from the mCut protein was found to repress expression of the reporter construct to the same extent as the corresponding region of hCut (Fig. 3B and Table 1). We have also made and tested several chimeric repression domains containing various parts from the human and murine repression domains. All of these fusion proteins behaved like the repression domain from the hCut protein (data not shown). Finally, we tested the ability of the hCut and mCut divergent regions to repress or activate various reporter constructs. Neither fusion protein had any clear effect on gene expression (data not shown). We conclude that in spite of their sequence divergence, the carboxy-terminal domains of hCut and mCut behave similarly as active repression domains.

The hCut repression domain must be tethered to DNA for activity. Two approaches were used to verify whether the hCut repression domain must be tethered to DNA in order to carry its function. First, we measured the extent of repression in NIH 3T3 cells, using two reporter plasmids with or without binding sites for the GAL4 DNA binding domain. As shown in Fig. 4A, expression of the reporter with GAL4 binding sites was reduced to less than 25% of the control level when 0.5 μg of the effector construct encoding the GAL4-hCut fusion protein was used. In contrast, with the reporter that does not contain GAL4 binding sites, 10 times more (5 μg) of the effector plasmid was necessary to achieve the same level of repression. We hypothesize that the weak repression imparted by the GAL4-hCut fusion protein on the reporter without GAL4 binding sites must be due to nonspecific DNA binding. We cannot, however, exclude that the *tk*CAT plasmid contains a cryptic suboptimal GAL4 binding site. Second, we compared the repression activities of the hCut repression domain fused and not fused to the GAL4 DNA binding domain. To ensure that the hCut repression domain would be transported to the nucleus, a nuclear localization signal was fused to its amino terminus. No repression was observed with this effector in COS cells, although the same levels of proteins were present in the nuclei of transfected cells (Fig. 4B and C). Together, these results indicate that the hCut repression domain must be tethered to DNA for efficient repression to occur.

The repression subdomains of the hCut protein can repress at a large distance. The mode of action of active repression domains is not yet fully understood. However, some mechanisms have been described, and it is possible to confirm or rule out their involvement (17). For example, one type of repression, termed quenching, involves inhibition of an activator by a repressor that is bound in close proximity in the promoter. Thus, the position of the repressor relative to the activator is crucial for its regulatory effect, and moving away its binding site should abolish its function. We sought to determine whether the repression domains from the hCut protein would still function when bound at a large distance from the transcription initiation site. To test this, we modified the reporter construct by inserting a 1.6-kb fragment of neutral spacer DNA between the GAL4 binding sites and the *tk* promoter. Previous experiments had confirmed that this DNA region displayed no regulatory property and that it acted only as a spacer sequence (data not shown). When transient transfection assays were carried out with this reporter construct, the complete hCut carboxy-terminal domain (aa 1299 to 1505) was found to repress with the same relative strength as when the GAL4 binding sites were located closer to the initiation site (Fig. 5, lanes 1 and 2). Consistent with this observation, repression at a distance was also observed with the two hCut repression subdomains (aa 1299 to 1380 and 1451 to 1505), although to a lesser extent with the leftmost subdomain (Fig. 5, lanes 3 to 5). We conclude that the hCut repression domain can function at a large distance and that its mechanism(s) of action must be distinct from quenching.

TABLE 1. Relative CAT activities in COS cell extracts following cotransfection of CAT reporters and hCut or mCut effector constructs[a]

Effector	% CAT activity[b] with indicated reporter (range)			
	GAL4*tk*CAT	GAL4E1bCAT	GAL4InrCAT	GAL4Spacer*tk*CAT
GAL4 DBD	100	100	100	100
hCut CR1	68 (57–81)			
hCut CR1 linker	66 (59–76)			
hCut linker CR23	86 (81–93)			
hCut CR3HD	18 (9–28)	82 (65–107)	74 (65–80)	
hCut CR123HD	26 (15–40)			
hCut CR123HDCarb	6 (1–14)			
hCut CR23HDCarb	8 (3–17)			
hCut HD	89 (84–92)			
hCut HDCarb	14 (11–19)			
hCut 1299-1505	13 (5–21)	140 (123–155)	130 (90–152)	16 (11–19)
hCut 1299-1314	77 (67–86)			
hCut 1299-1336	68 (54–77)			
hCut 1299-1362	65 (56–80)			
hCut 1299-1380	22 (18–29)	80 (55–107)	90 (84–97)	18 (15–23)
hCut 1316-1380	73 (67–82)			
hCut 1316-1459	58 (52–77)			
hCut 1338-1505	12 (7–17)			
hCut 1363-1505	11 (6–19)			
hCut 1380-1505	18 (15–23)			
hCut 1399-1468	113 (105–130)			
hCut 1424-1505	87 (81–96)			
hCut 1451-1505	5 (1–11)	83 (73–93)	76 (71–79)	7 (3–14)
hCut 1451-1485	24 (22–29)			
hCut 1461-1505	8 (3–16)			
hCut 1476-1505	7 (4–12)			
NLS hCut 1299-1505	90 (74–108)			
mCut 1114-1332	19 (15–21)	86 (78–103)	83 (67–93)	

[a] The hCut and mCut segments are indicated by the name of the evolutionarily conserved regions or, for the carboxy-terminal region, by amino acid numbers. GAL4 DBD, yeast GAL4 DNA binding domain (aa 1 to 147); CR1, Cut repeat 1; CR2, Cut repeat 2; CR3, Cut repeat 3; CR123, Cut repeats 1, 2, and 3; HD, homeodomain; Carb, carboxy-terminal domain (aa 1299 to 1505); NLS, nuclear localization signal.

[b] Average activity obtained from at least three separate experiments after correction for transfection efficiencies, using plasmid pCH110 (pSV40-βgal; Pharmacia) as a control. Values were obtained by calculating the ratio of chloramphenicol conversion with the test construct relative to conversion in the presence of the appropriate control plasmid, pSG424 or pGALM, encoding only the GAL4 DNA binding domain. The NLS hCut 1299-1505 construct was compared with the empty pSG5 vector.

The carboxy-terminal region of the hCut and mCut proteins do not repress minimal promoters. The experiments described above indicated that hCut and mCut repression domains could down-regulate activated transcription. To verify whether these repression domains could also repress a minimal promoter, we used two reporter constructs: the E1bCAT construct, which contains only the TATA box from the adenovirus E1b promoter, and the InrCAT construct, which contains an initiator element but no TATA box (36, 47). Neither reporter construct was repressed by the hCut and mCut carboxy-terminal regions (Table 1). Consistent with these findings, the two repression subdomains of hCut (aa 1299 to 1380 and 1451 to 1505) did not repress transcription from a minimal promoter (data not shown). In view of recent reports demonstrating increased repression activity conferred by the homeodomain, we sought to verify whether a protein comprising both the homeodomain and the carboxy terminus of hCut could repress minimal transcription (5, 14, 65). This protein was found to have little impact on transcription from minimal promoters (data not shown).

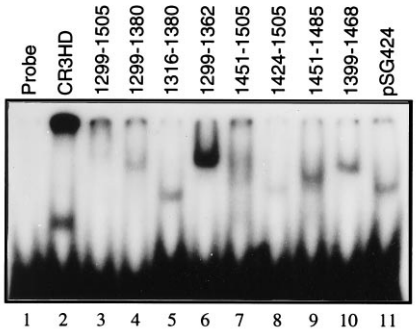

FIG. 2. EMSA confirming expression of GAL4-Cut fusion proteins. Nuclear extracts (5 μg of protein) prepared from cells transfected with various GAL4 (aa 1 to 147)-Cut effector plasmids were incubated with a labeled, double-stranded synthetic oligonucleotide probe containing a GAL4 binding site. Complexes were resolved from the unbound probe on a 5% acrylamide gel. The numbers above each lane identify the hCut segment (first and last amino acids) of each fusion protein as in the text and other figures.

Two regions of hCut contribute to down-regulate gene expression: the CR3HD DNA binding domain and the carboxy-terminal repression domain. To verify whether other regions of hCut are involved in transcriptional repression, we prepared a series of vectors expressing chimeric proteins with various portions of hCut linked to the GAL4 DNA binding domain. Maximal repression was obtained with a GAL4-Cut fusion protein that included Cut repeats 1, 2, and 3, the Cut homeodomain, and the carboxy-terminal region (Fig. 6, lanes 1 and 2). Removal of Cut repeat 1 and of the linker region between Cuts repeat 1 and 2 did not reduce the repression potential (compare lanes 2 and 4). When smaller segments of hCut were tested, two regions were found independently to repress gene expression: CR3HD and the carboxy-terminal region down-

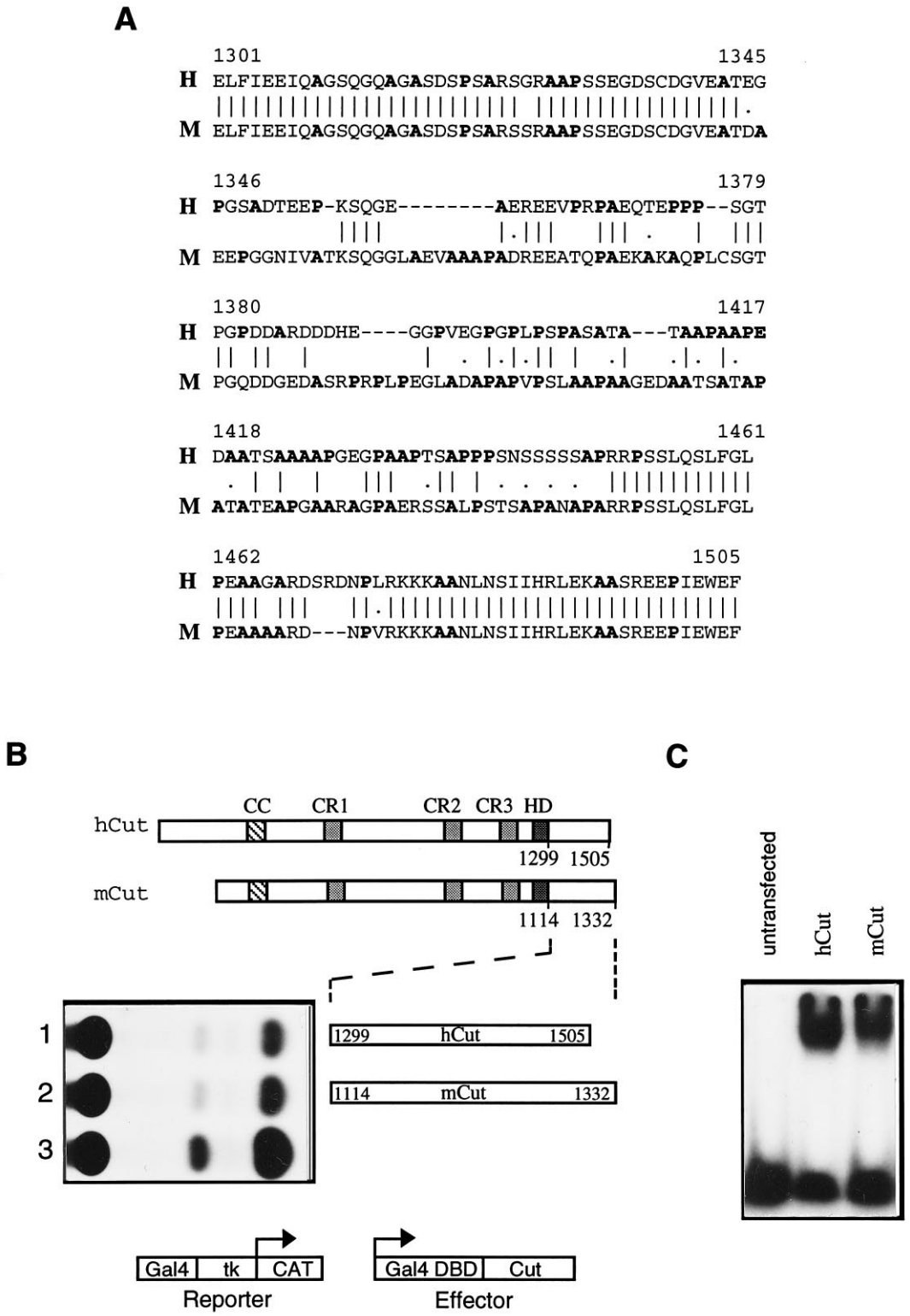

FIG. 3. Sequence comparison and repression activities of the human and murine Cut carboxy-terminal regions. (A) Alignment of amino acid sequences of human (aa 1301 to 1505) and murine (aa 1114 to 1332) C-terminal regions. The amino acid sequences are shown in the one-letter code. Vertical bars indicate identical amino acids, dots denote conservative substitutions, and dashes show gaps introduced to maximize the alignment. Alanine and proline residues are shown in boldface. (B) CAT activity in cellular extracts following transfection of COS cells with the GAL4*tk*CAT reporter plasmid and effector constructs encoding either the GAL4-hCut fusion protein (lane 1), the GAL4-mCut fusion protein (lane 2), or the GAL4 DNA binding domain (DBD) alone (lane 3; pSG424 control). Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. A schematic representation of the human and murine Cut proteins is displayed at the top, with the evolutionarily conserved regions depicted as boxes: CC, coiled coil; CR, Cut repeat; HD, homeodomain. Schematic diagrams of the hCut and mCut segments encoded by the effector plasmids are shown on the right, with numbers identifying the first and last amino acids. Transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown. Values representing the average of three separate experiments are provided in Table 1.

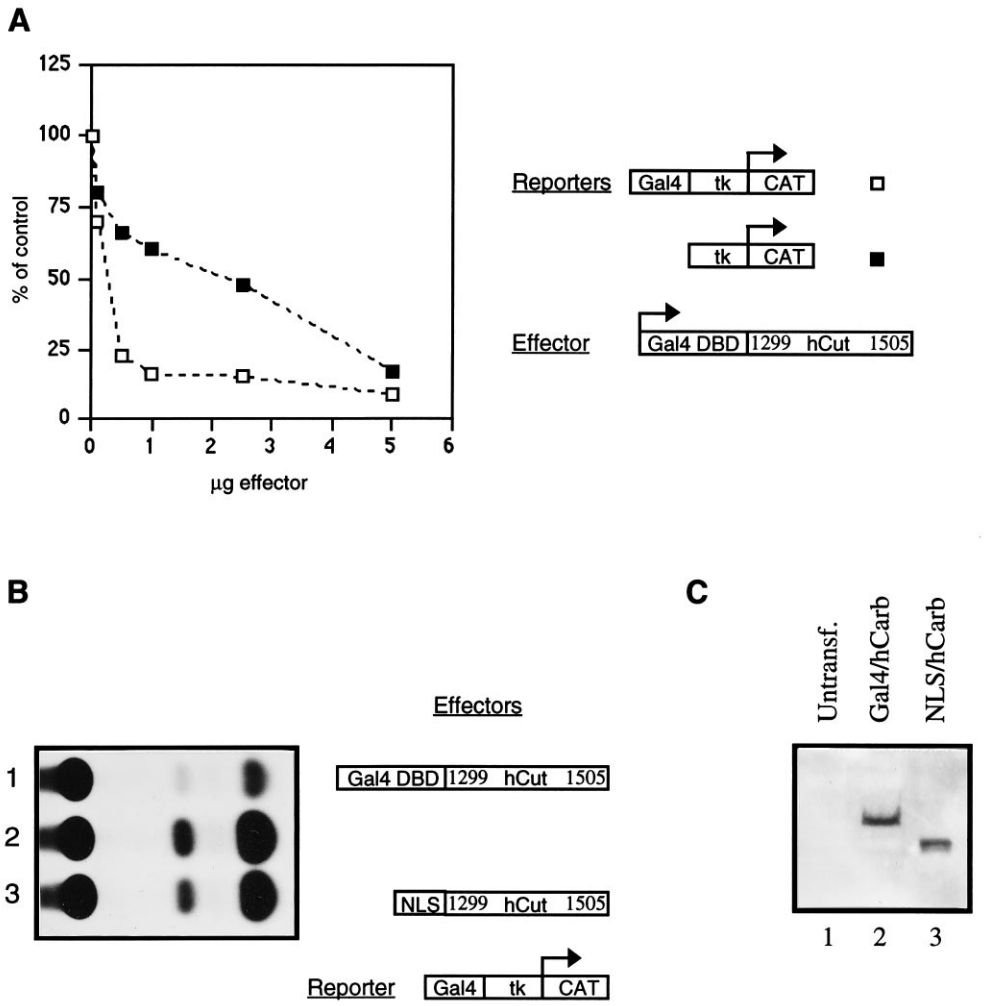

FIG. 4. The hCut repression domain must be tethered to DNA for activity. (A) Relative CAT activity in cellular extracts following transfection of NIH 3T3 cells with variable amounts of an effector construct encoding the GAL4-hCut fusion protein and *tk*CAT reporter plasmids containing (GAL4*tk*CAT) or not containing (*tk*CAT) binding sites for the GAL4 DNA binding domain (DBD). The level of CAT activity in cells transfected with the effector plasmid expressing the GAL4 DNA binding domain only has been set to 100%. Values representing the average of three separate experiments are provided. Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. (B) CAT activity in cellular extracts following transfection of COS cells with the GAL4*tk*CAT reporter plasmid and effector constructs encoding the hCut carboxy-terminal domain fused to either the GAL4 DNA binding domain or a nuclear localization signal (NLS). Transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown. Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. (C) Western blot analysis of nuclear extracts from untransfected COS cells (lane 1) and COS cells transfected with an effector construct encoding the hCut carboxy-terminal domain fused to either the GAL4 DNA binding domain (lane 2) or a nuclear localization signal (lane 3).

stream of the Cut homeodomain (Fig. 6, lanes 1, 10, 11, and 12). The effects of these domains on gene expression appeared to be additive since repression was maximal when both regions were present in the chimeric protein (Fig. 6, lanes 2 to 4, 10, and 12). The relative strengths of the constructs with repressor activity are presented in Table 1.

CR3HD binds to Sp1 and CCAAT sites within the *tk* promoter. Analysis of the *tk* promoter has revealed the presence of two and one binding sites for the transcription factors Sp1 and CCAAT-binding protein (CBP), and efficient expression was shown to be dependent on the presence of all three of these sites (37). Interestingly, cDNAs encoding hCut were originally isolated through their ability to interact with Sp1-like and CCAAT DNA binding sites (22, 56). We therefore hypothesized that the down-regulatory effect of CR3HD on the *tk*CAT reporter plasmid was due to competition for site occupancy with the sequence-specific transactivators Sp1 and CCAAT.

To test this hypothesis, we first performed DNase footprinting assays using a fragment from the reporter construct as a probe and crude nuclear extracts from transfected COS cells (Fig. 7A). In addition to the GAL4 sites, the GAL4-CR3HD fusion protein protected one Sp1 and one CCAAT binding site (Fig. 7A, lane 2). These sites were not protected in extracts from untransfected cells (Fig. 7A, lanes 7, 15, and 19). This was not unexpected since previous studies demonstrated that protection of these sites necessitates either partial purification of the cellular extract or isolation of unbound and complexed DNA following DNase digestion (12, 15, 16, 20, 22, 23, 27, 37, 38). Further analysis with a shorter DNA probe that does not include the GAL4 binding sites showed that the GAL4-CR3HD fusion protein could still interact directly with the Sp1

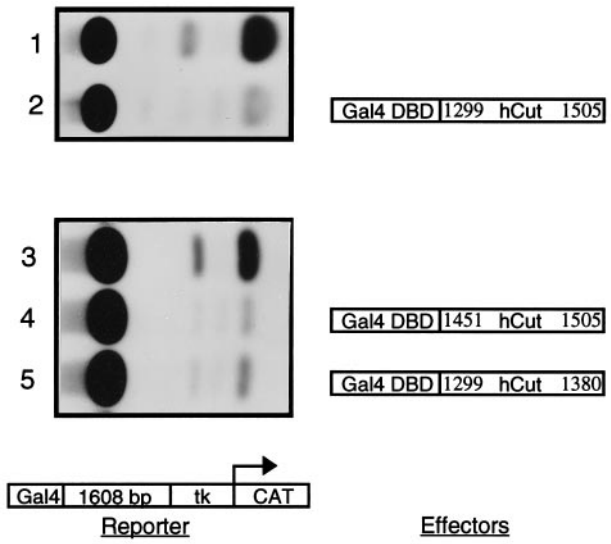

FIG. 5. Repression at a large distance from the transcription start site by GAL4-hCut fusion proteins. COS cells were cotransfected with the GAL4*tk*CAT reporter plasmid containing a 1,608-bp spacer DNA (see Materials and Methods) inserted between the GAL4 binding sites and the *tk* promoter and effector plasmids encoding either the GAL4 DNA binding domain (DBD) (lanes 1 and 3) or GAL4-hCut fusion proteins (lanes 2, 4, and 5). Amino acid numbering refers to the hCut sequence. Transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown. Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency.

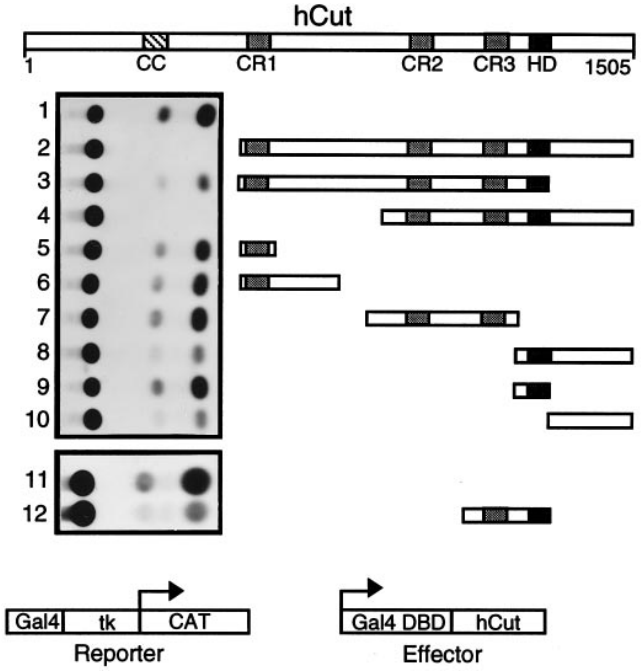

FIG. 6. Two regions of hCut contribute to down-regulate gene expression. Transfection assays were performed in COS cells, using the GAL4*tk*CAT reporter plasmid. The effector plasmids contained sequences encoding various portions of the hCut protein fused to the GAL4 DNA binding domain (DBD). A schematic representation of the hCut protein is displayed at the top, with the evolutionarily conserved regions depicted as boxes: CC, coiled coil; CR, Cut repeat; HD, homeodomain. Diagrams of the hCut segments encoded by the effector plasmids are shown on the right. Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. All transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown. Values representing the average of at least three separate experiments are provided in Table 1 for some of the effector plasmids.

and CCAAT sites but that the protected regions were not as wide (Sp1 site) or intense (CCAAT site) (Fig. 7A, lanes 5, 13, and 14). It is probable that the enhanced protection of these sites in the presence of the GAL4 sites involves cooperative DNA binding by the two DNA binding regions present in the GAL4-CR3HD fusion protein. To verify whether the GAL4-CR3HD was able alone to bind to these sites or needed to interact with other endogenous factors, this protein was expressed in bacteria with a short amino-terminal histidine tail. The purified fusion protein protected very well the Sp1 site and weakly the CCAAT site (Fig. 7, lanes 9 and 11). These results confirm that the Cut repeat 3 and homeodomain specifically recognize Sp1 and CCAAT sites and probably compete with the corresponding transcription factors for the occupancy of these sites.

To verify whether interaction with the GAL4 binding sites was essential for repression by the GAL4-CR3HD fusion protein, we performed CAT assays using derivatives of the *tk*CAT plasmid which lacked the GAL4 sites. We also tested repression with a reporter construct in which the single site for the *Cla*I restriction enzyme had been deleted. This control was important since the sequence recognized by this enzyme, ATC GAT, has been identified as a consensus binding site for Cut proteins. The results presented in Fig. 7B indicated that the three *tk*CAT reporter constructs were repressed to similar degrees, while basal promoter constructs were not affected. Thus, repression by the GAL4-CR3HD fusion protein does not require the presence of GAL4 or *Cla*I sites and most likely results from competition for site occupancy with the endogenous Sp1 protein and CBP.

DISCUSSION

In cotransfection experiments, mammalian Cut proteins were found to repress the expression of genes containing Cut-specific binding sites (21, 22, 62, 66, 69). In the present study, we investigated the mechanisms by which Cut proteins effect transcriptional repression. While some repressors function by competing with sequence-specific transactivators for binding to a common or overlapping binding site(s) (6, 62), others appear to negatively regulate transcription through active repression (14, 17, 18, 29, 35, 49, 61). Our results suggest that Cut proteins can repress gene expression through both mechanisms (Fig. 8).

We have demonstrated that the bipartite CR3HD DNA binding domain can bind to Sp1 and CCAAT binding sites within the *tk* gene promoter and that this in itself is sufficient to down-regulate expression of a *tk*CAT reporter construct. These results are in accordance with previous studies which independently demonstrated that Cut proteins can function as CDPs upstream of the human gp91-phox gene (44, 62) and of the sea urchin sperm histone H2B gene (6) or bind to an Sp1-like binding site within the promoter of the c-*myc* proto-oncogene (3, 19, 21, 22). We conclude that Cut proteins have the potential to prevent the Sp1 and CBP transactivators from exerting their stimulatory effect on target genes.

While the expression of a large number of genes is stimulated by Sp1 and CBP, it is not clear at this point whether all promoters containing sites for these transcriptional activators represent authentic in vivo targets for repression by Cut proteins. Two parameters mainly will determine whether a promoter containing Sp1 and/or CCAAT sites is effectively repressed by Cut proteins: the relative amounts of these activator

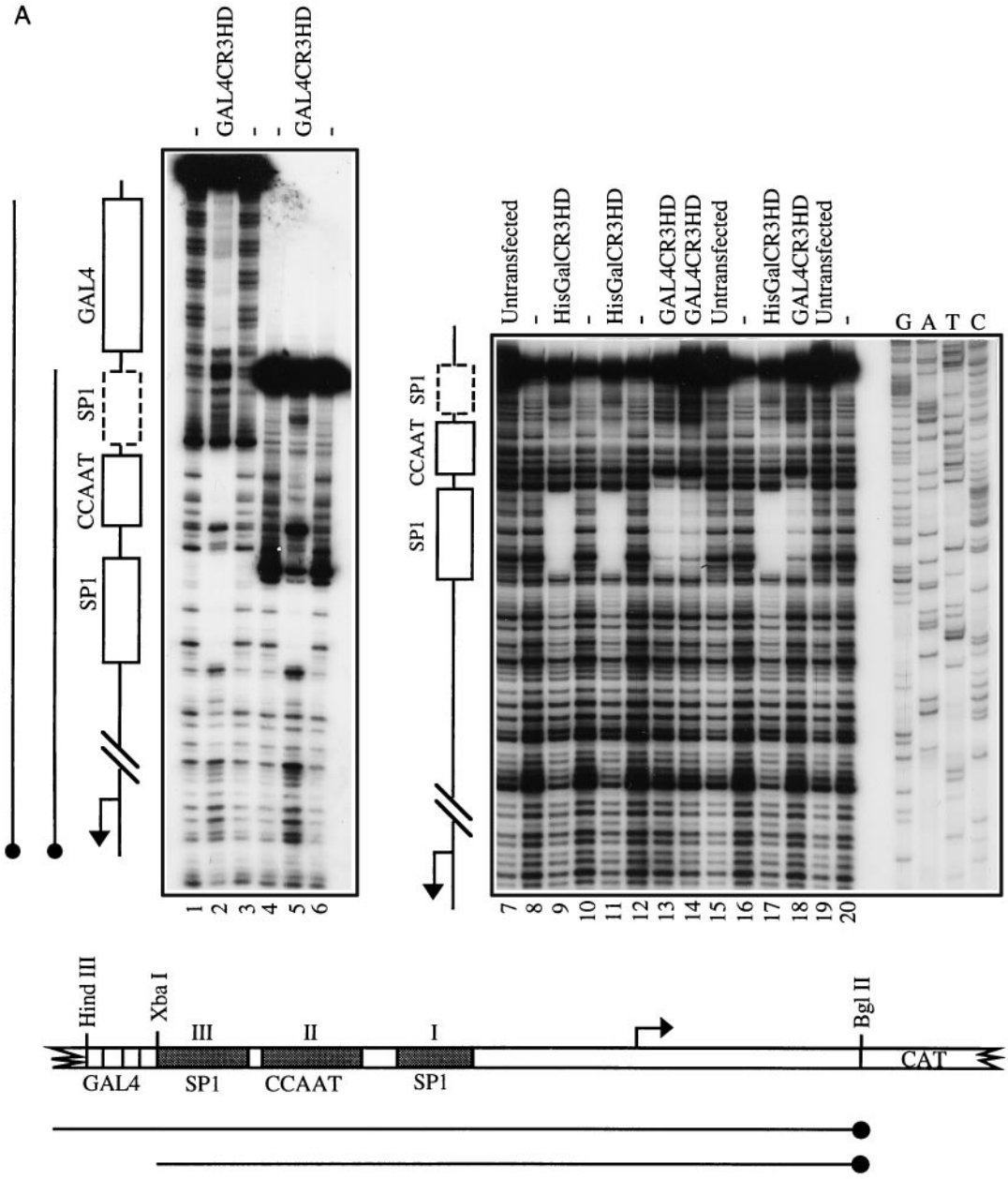

FIG. 7. Transcriptional repression by the CR3HD region involves binding to Sp1 and CCAAT sites within the *tk* promoter. (A) DNase I footprinting analysis of GAL4-CR3HD fusion protein. The GAL4*tk*CAT reporter plasmid was cleaved at the *Bgl*II site, end labeled, and then digested at either the *Hin*dIII or *Xba*I site. The radiolabeled fragments thus generated contained (lanes 1 to 3) or did not contain (lanes 4 to 20) the five GAL4 binding sites. The labeled DNA was incubated with 20 μg of crude nuclear extract from transfected or untransfected COS cells or 50 ng of purified bacterially expressed fusion protein as indicated. Lanes 1, 3, 4, 6, 8, 10, 12, 16, and 20 are control lanes in which no protein extract was added. Lanes 13 and 14 differ in the amount of DNase that was used: 1/200 and 1/500 dilutions in lanes 13 and 14, respectively. (B) Transcriptional repression by the GAL4-CR3HD fusion protein does not require the presence of GAL4 binding sites or of *Cla*I sites within the reporter plasmid. Transfection assays were performed in COS cells, using various CAT reporter plasmids and effector plasmids encoding the GAL4-CR3HD fusion protein (even-numbered lanes) or the GAL4 DNA binding domain (DBD) alone (odd-numbered lanes). Plasmid pCH110 (Pharmacia) encoding β-galactosidase was included as an internal control for transfection efficiency. A schematic representation of the hCut protein is displayed at the top, with the evolutionarily conserved regions depicted as boxes: CC, coiled coil; CR, Cut repeat; HD, homeodomain. Diagrams of the hCut segment encoded by the effector plasmids are shown on the right. Transfection assays were repeated at least three times. Results obtained from a single representative experiment are shown, and values representing the average of three separate experiments are provided in Table 1 for the GAL4*tk*CAT reporter plasmid.

and repressor proteins in a given cell and their relative affinities for a specific promoter. Although Sp1 is expressed in nearly all tissues and is considered a ubiquitous factor, its level of expression vary greatly from one cell type to another (53, 57; reviewed in reference 14). At least five CCAAT/enhancer-binding proteins which also exhibit tissue-specific patterns of expression have been described (reviewed in reference 54). From RNase mapping analysis, the murine *cut* gene appears to be expressed at low levels in most tissues (56a), but we do not know whether, as in *D. melanogaster*, it is present only within a specific subset of cells in each tissues. It will be important to investigate the number and DNA binding activities of Cut

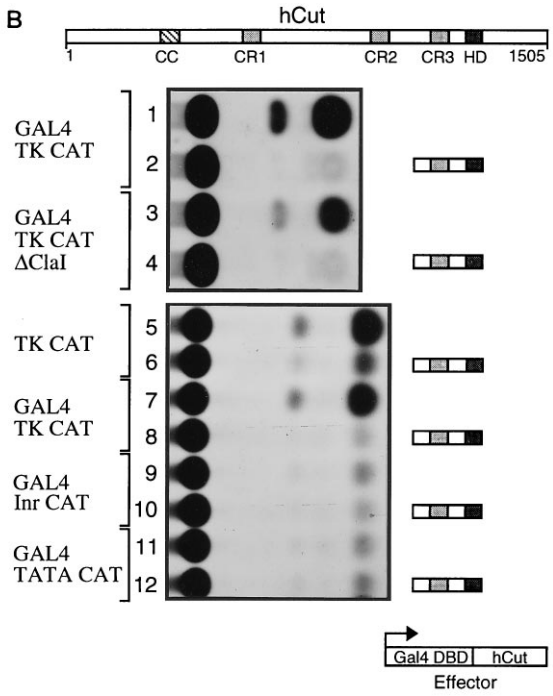

FIG. 7—*Continued.*

molecules in a variety of cell types and throughout the cell cycle. In addition, in light of the fact that Sp1 proteins have the potential to oligomerize and to bind cooperatively to several sites, the affinity for a given site will be, at least partially, determined by the presence and the position of other binding sites in the same DNA region (52, 63). The affinity for a given site must therefore be evaluated in the context of the entire promoter. It is likely that Cut proteins also are capable of cooperative DNA binding, since they contain several DNA binding domains. This property would make Cut proteins well suited to compete with Sp1.

Several active repression domains have been found in recent years, notably within the *Drosophila* homeodomain proteins Engrailed (28, 35), Even-skipped (5, 29, 64, 65), and Krüppel (42, 43, 60, 70), the mammalian proteins YY1 (33, 61), WT1 (49, 50, 57), and Msx-1 (14), and the *Drosophila* and mammalian Polycomb group proteins (13). Very little is known about the structure of repression domains, but it has been suggested that the relative paucity of charged amino acids together with an enrichment in alanine residues, and occasionally in proline and glycine residues, is a key feature of these domains (28). In contrast, Saha et al. (59) have shown that some protein segments with repressor activity contain a high proportion of basic residues. It is possible that, as for activation domains, there are several types of repression domains with different amino acid compositions. Interestingly, the carboxy-terminal region of Cut proteins is rich in alanine and proline residues, and this feature is conserved between *Drosophila* and mammalian Cut proteins, suggesting that it is important for function (Fig. 3).

Examination of the human and murine sequences revealed that the Cut carboxy-terminal region can be divided in three regions. The first 55 aa immediately downstream of the homeodomain and the 54 most carboxy-terminal residues are identical except for four substitutions (two of which are conservative) and a 3-residue insertion in hCut. In contrast, the central region diverges markedly, with less than 35% residue identity. We found that the hCut repression domain could be subdivided into two subdomains that more or less matched the region of sequence conservation between human and mouse cDNA sequences, while the central region had no activity. Curiously, while the two subdomains could function independently of one another, they did not have an additive effect when expressed together. This observation could suggest that the two subdomains have the same molecular target and therefore are redundant. However, we favor the hypothesis that each subdomain has a distinct function and therefore may have an additive effect but that this could not be revealed by using the *tk*CAT reporter construct. In this context, the central, divergent region could play the role of a flexible arm that links the two repression domains and allow their interactions with their respective targets.

The precise mode of action of active repression domains is not yet understood, but several mechanisms have been proposed. One repression mechanism, termed quenching, occurs when a repressor inhibits a transactivator that is bound next to it in the promoter (29, 36). This mechanism can be ruled out for Cut proteins since their repression domains can function at different positions and at a large distance from the transcription start site. Another mechanism involves interaction between a repression domain and the TATA-binding protein or other general transcription factors to prevent the assembly of the preinitiation complex. The ability of an active repression domain to repress a minimal promoter is generally taken as indirect evidence for interaction with a component of the basic transcription machinery. Our results indicate that the repression domain of Cut proteins has no effect on basal transcription but has the ability to repress activated transcription. We therefore propose that the Cut repression domain hinders the interactions between some activation domain and a component of the preinitiation complex. Whether this effect is achieved

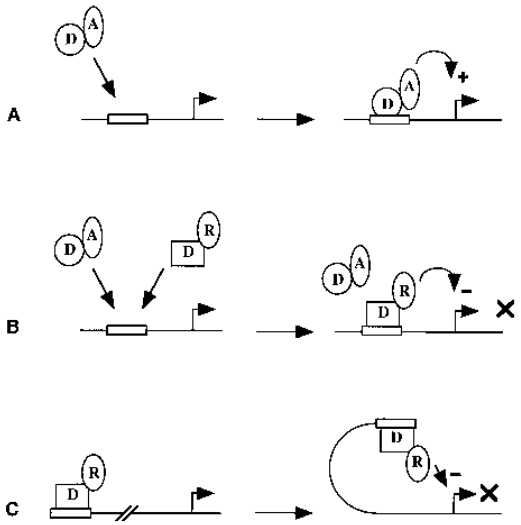

FIG. 8. Cut homeodomain proteins repress gene expression by two distinct mechanisms. (A) Binding of a sequence-specific transcriptional activator containing a DNA binding domain (D) and an activation domain (A) results in activation (+) of gene expression. (B) The hCut protein is depicted with a DNA binding domain (D) and a repression domain (R). The hCut protein competes with an activator, either Sp1 or a CCAAT-binding factor, for binding site occupancy. This prevents transcriptional activation. In addition, once tethered to DNA via the DNA binding domain, the active repression domain down-modulates (−) gene expression. (C) The hCut protein can bind to DNA at a large distance from the transcription start site. DNA looping allows the repression domain to down-modulate gene expression.

through binding with transactivation domains, the preinitiation complex, or an adapter molecule remains to be investigated.

ACKNOWLEDGMENTS

We are grateful to Ivan Sadowski for the gift of the pSG424 plasmid; Gregory J. Kato and Chi V. Dang for the gift of the pGALM plasmid; Thomas Shenk for the gift of the GAL4*tk*CAT plasmid; Bruno Luckow for the gift of the pBL2CAT, pBL3CAT, BL5CAT, and pBL6CAT plasmids; Tony Kouzarides for the gift of the GAL4InrCAT plasmid; and Michael Green for the gift of the GAL4E1bCAT plasmid.

A.N. and F.M. are the recipients of a scholarship and a fellowship from the Fonds de la Recherche en Santé du Québec. This research was supported by grant MT-11590 from the Medical Research Council of Canada and grant 3497 from the National Cancer Institute of Canada to A.N.

The first two authors contributed equally to this study.

REFERENCES

1. **Andres, V., M. D. Chiara, and V. Mahdavi.** 1994. A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins. Genes Dev. **8:**245–257.
2. **Andres, 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.
3. **Asselin, C., A. Nepveu, and K. B. Marcu.** 1989. Molecular requirements for transcriptional initiation of the murine c-myc gene. Oncogene **4:**549–558.
4. **Aufiero, B., E. J. Neufeld, and S. H. Orkin.** 1994. Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein. Proc. Natl. Acad. Sci. USA **91:**7757–7761.
5. **Austin, R. J., and M. B. Biggin.** 1995. A domain of the even-skipped protein represses transcription by preventing TFIID binding to a promoter: repression by cooperative blocking. Mol. Cell. Biol. **15:**4683–4693.
6. **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.
7. **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 identify in Drosophila. Nature (London) **333:**629–635.
8. **Blochlinger, K., R. Bodmer, J. W. Jack, L. Y. Jan, and Y. N. Jan.** 1990. Patterns 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.
9. **Blochlinger, 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.
10. **Blochlinger, K., L. Y. Jan, and Y. N. Jan.** 1993. Postembryotic patterns of expression of Cut, a locus regulating sensory organ identify in Drosophila. Development **117:**441–450.
11. **Bodmer, R., S. Barbel, S. Shepherd, J. W. Jack, L. Y. Jan, and Y. N. Jan.** 1987. Transformation of sensory organs by mutations of the cut locus of Drosophila melanogaster. Cell **51:**293–307.
12. **Briggs, M. R., J. T. Kadonaga, S. P. Bell, and R. Tjian.** 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science **234:**47–52.
13. **Bunker, C. A., and R. E. Kingston.** 1994. Transcriptional repression by Drosophila and mammalian polycomb group proteins in transfected mammalian cells. Mol. Cell. Biol. **14:**1721–1732.
14. **Catron, K. M., H. Zhang, S. C. Marshall, J. A. Inostroza, J. M. Wilson, and C. Abate.** 1995. Transcriptional repression by Msx-1 does not require homeodomain DNA-binding sites. Mol. Cell. Biol. **15:**861–871.
15. **Chodosh, L. A., A. S. Baldwin, R. W. Carthew, and P. A. Sharp.** 1988. Human CCAAT-binding proteins have heterologous subunits. Cell **53:**11–24.
16. **Chodosh, L. A., J. Olesen, S. Hahn, A. S. Baldwin, L. Guarente, and P. A. Sharp.** 1988. A yeast and a human CCAAT-binding protein have heterologous subunits that are functionally interchangeable. Cell **53:**25–35.
17. **Cowell, I. G.** 1994. Repression versus activation in the control of gene transcription. Trends Biochem. Sci. **19:**38–42.
18. **Cowell, I. G., and H. C. Hurst.** 1994. Transcriptional repression by the human bZIP factor E4BP4: definition of a minimal repression domain. Nucleic Acids Res. **22:**59–65.
19. **DesJardins, E., and N. Hay.** 1993. Repeated CT elements bound by zinc finger proteins control the absolute and relative activities of the two principal human c-*myc* promoters. Mol. Cell. Biol. **13:**5710–5724.
20. **Dorn, A., J. Bollekens, A. Staub, C. Benoist, and D. Mathis.** 1987. A multiplicity of CCAAT box-binding proteins. Cell **50:**863–872.
21. **Dufort, D., M. Drolet, and A. Nepveu.** 1993. A protein binding site from the murine c-myc promoter contributes to transcriptional block. Oncogene **8:**165–171.
22. **Dufort, D., and A. Nepveu.** 1994. The human cut homeodomain protein represses transcription from the c-*myc* promoter. Mol. Cell. Biol. **14:**4251–4257.
23. **Dynan, W. S., and R. Tjian.** 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell **35:**79–87.
24. **Fondell, J. D., A. L. Roy, and R. G. Roeder.** 1993. Unliganded thyroid hormone receptor inhibits formation of a functional preinitiation complex: implications for active repression. Genes Dev. **7:**1400–1410.
25. **Garcia-Bustos, J., J. Heitman, and M. N. Hall.** 1991. Nuclear protein localization. Biochim. Biophys. Acta **1071:**83–101.
26. **Gorman, C. M., L. F. Moffat, and B. H. Howard.** 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. Mol. Cell. Biol. **2:**1044–1051.
27. **Graves, B. J., P. F. Johnson, and S. L. McKnight.** 1986. Homologous recognition of a promoter domain common to the MSV LTR and the HSV tk gene. Cell **44:**565–576.
28. **Han, K., and J. L. Manley.** 1993. Functional domains of the Drosophila Engrailed protein. EMBO J. **12:**2723–2733.
29. **Han, K., and J. L. Manley.** 1993. Transcriptional repression by the Drosophila even-skipped protein: definition of a minimal repression domain. Genes Dev. **7:**491–503.
30. **Harada, R., G. Berube, O. J. Tamplin, C. Denis-Larose, and A. Nepveu.** 1995. DNA-binding specificity of the cut repeats from the human cut-like protein. Mol. Cell. Biol. **15:**129–140.
31. **Harada, R., D. Dufort, C. Denis-Larose, and A. Nepveu.** 1994. Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains. J. Biol. Chem. **269:**2062–2067.
32. **Herbomel, P., B. Bourachot, and M. Yaniv.** 1984. Two distinct enhancers with different cell specificities coexist in the regulatory region of polyoma. Cell **39:**653–662.
33. **Inouye, C. J., and E. Seto.** 1994. Relief of YY1-induced transcriptional repression by protein-protein interaction with the nucleolar phosphoprotein B23. J. Biol. Chem. **269:**6506–6510.
34. **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.
35. **Jaynes, J. B., and P. H. O'Farrell.** 1991. Active repression of transcription by the engrailed homeodomain protein. EMBO J. **10:**1427–1433.
36. **Johnson, A. D., and I. Herskowitz.** 1985. A repressor (MAT alpha 2 product) and its operator control expression of a set of cell type specific genes in yeast. Cell **42:**237–247.
37. **Jones, K. A., K. R. Yamamoto, and R. Tjian.** 1985. Two distinct transcription factors bind to the HSV thymidine kinase promoter in vitro. Cell **42:**559–572.
38. **Kadonaga, J. T., K. R. Carner, F. R. Masiarz, and R. Tjian.** 1987. Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. Cell **51:**1079–1090.
39. **Kato, G. J., J. Barrett, M. Villa-Garcia, and C. V. Dang.** 1990. An amino-terminal c-myc domain required for neoplastic transformation activates transcription. Mol. Cell. Biol. **10:**5914–5920.
40. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:**283–292.
41. **Lee, K. A., A. Bindereif, and M. R. Green.** 1988. A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. Gene Anal. Tech. **5:**22–31.
42. **Licht, J. D., M. J. Grossel, J. Figge, and U. M. Hansen.** 1990. Drosophila Kruppel protein is a transcriptional repressor. Nature (London) **346:**76–79.
43. **Licht, J. D., W. Hanna-Rose, J. C. Reddy, M. A. English, M. Ro, M. Grossel, and U. Hansen.** 1994. Mapping and mutagenesis of the amino-terminal transcriptional repression domain of the Drosophila Kruppel protein. Mol. Cell. Biol. **14:**4057–4066.
44. **Lievens, P. M. J., J. J. Donady, C. Tufarelli, and E. J. Neufeld.** 1995. Repressor activity of CCAAT displacement protein in HL-60 myeloid leukemia cells. J. Biol. Chem. **270:**12745–12750.
45. **Lillie, J. W., and M. R. Green.** 1989. Transcription activation by the adenovirus E1a protein. Nature (London) **338:**39–44.
46. **Liu, S., and J. Jack.** 1992. Regulatory interactions and role in cell type specification of the Malpighian tubules by the Cut, Kruppel, and caudal genes of Drosophila. Dev. Biol. **150:**133–143.
47. **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.
48. **Luckow, B., and G. Schutz.** 1987. CAT constructions with multiple unique restriction sites for the functional analysis of eukaryotic promoters and regulatory elements. Nucleic Acids Res. **15:**5490–5490.
49. **Madden, S. L., D. M. Cook, J. F. Morris, A. Gashler, V. P. Sukhatme, and I. F. Rauscher.** 1991. Transcriptional repression mediated by the WT1 Wilms tumor gene product. Science **253:**1550–1553.
50. **Madden, S. L., D. M. Cook, and F. J. Rauscher.** 1993. A structure-function analysis of transcriptional repression mediated by the WT1, Wilms' tumor suppressor protein. Oncogene **8:**1713–1720.
51. **Martin, K. J., J. W. Lillie, and M. R. Green.** 1990. Evidence for interaction

of different eukaryotic transcriptional activators with distinct cellular targets. Nature (London) **346:**147–152.
52. **Mastrangelo, I. A., A. J. Courey, J. S. Wall, S. P. Jackson, and P. V. Hough.** 1991. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. Proc. Natl. Acad. Sci. USA **88:**5670–5674.
53. **McCutchan, J. H., and J. S. Pagano.** 1968. Enchancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl-dextran. J. Natl. Cancer Inst. **41:**351–357.
54. **McKnight, S. L.** 1992. CCAAT/enhancer binding proteins, p. 771–796. *In* S. L. McKnight and K. R. Yamamoto (ed.), Transcriptional regulation, vol. 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
55. **Metz, R., A. J. Bannister, J. A. Sutherland, C. Hagemeier, E. C. O'Rourke, A. Cook, R. Bravo, and T. Kouzarides.** 1994. c-Fos-induced activation of a TATA-box-containing promoter involves direct contact with TATA-box-binding protein. Mol. Cell. Biol. **14:**6021–6029.
56. **Neufeld, E. J., D. G. Skalnik, P. M. Lievens, and S. H. Orkin.** 1992. Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut. Nat. Genet. **1:**50–55.
56a. **Phillips, S., and A. Nepveu.** Unpublished data.
57. **Rauscher, F. J., III.** 1993. Tumor suppressor genes which encode transcriptional repressors: studies on the EGR and Wilms' tumor (WT1) gene products. Adv. Exp. Med. Biol. **348:**23–29.
58. **Sadowski, I., and M. Ptashne.** 1989. A vector for expressing GAL4(1-147) fusions in mammalian cells. Nucleic Acids Res. **17:**7539–7539.
59. **Saha, S., J. M. Brickman, N. Lehming, and M. Ptashne.** 1993. New eukaryotic transcriptional repressors. Nature (London) **363:**648–652.
60. **Sauer, F., J. D. Fondell, Y. Ohkuma, R. G. Roeder, and H. Jackle.** 1995. Control of transcription by Kruppel through interactions with TFIIB and TFIIE beta. Nature (London) **375:**162–164.
61. **Shi, Y., E. Seto, L. S. Chang, and T. Shenk.** 1991. Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein. Cell **67:**377–388.
62. **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.
63. **Su, W., S. Jackson, R. Tjian, and H. Echols.** 1991. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. Genes Dev. **5:**820–826.
64. **TenHarmsel, A., R. J. Austin, N. Savenelli, and M. D. Biggin.** 1993. Cooperative binding at a distance by even-skipped protein correlates with repression and suggests a mechanism of silencing. Mol. Cell. Biol. **13:**2742–2752.
65. **Um, M., C. Li, and J. L. Manley.** 1995. The transcriptional repressor Even-skipped interacts directly with the TATA-binding protein. Mol. Cell. Biol. **15:**5007–5016.
66. **Valarche, I., J. P. Tissier-Seta, M. R. Hirsch, S. Martinez, C. Goridis, and BrunetJf.** 1993. The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. Development **119:**881–896.
67. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. Chasin.** 1979. DNA-mediated transfer of the adenine phospho-ribosyltransferase locus into mammalian cells. Proc. Natl. Acad. Sci. USA **76:**1373–1376.
68. **Yeung, K. C., J. A. Inostroza, F. H. Mermelstein, C. Kannabiran, and D. Reinberg.** 1994. Structure-function analysis of the TBP-binding protein Dr1 reveals a mechanism for repression of class II gene transcription. Genes Dev. **8:**2097–2109.
69. **Yoon, S. O., and D. M. Chikaraishi.** 1994. Isolation of two E-box binding factors that interact with the rat tyrosine hydroxylase enhancer. J. Biol. Chem. **269:**18453–18462.
70. **Zuo, P., D. Stanojevic, J. Colgan, K. Han, M. Levine, and J. L. Manley.** 1991. Activation and repression of transcription by the gap proteins hunchback and Kruppel in cultured Drosophila cells. Genes Dev. **5:**254–264.