Modes and Regulation of CDP/Cux DNA Binding

Nam Sung Moon

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirement for the degree of Doctor of Philosophy

Department of Biochemistry
McGill University, Montreal
Quebec, Canada
January 2002

ABSTRACT

The CDP/Cux transcription factor contains four DNA binding domains: three Cut repeats (CR1, CR2 and CR3) and the Cut homeodomain (HD). The CCAAT-displacement protein (CDP) was first identified as it represses transcription of certain promoters by competing for the occupancy of the CCAAT sequence. CDP was then found to be the ortholog of the Drosophila Cut protein (Cux: Cut homeobox). The goal of my project was to define the modes of interactions of CDP/Cux with DNA. I demonstrated that high affinity DNA binding requires the cooperation between at least two of the CDP/Cux DNA binding domains. Among all combinations of domains, Cut repeats 1 and 2 (CR1CR2) and Cut repeat 3 and the Cut homeodomain (CR3HD) exhibited the highest DNA binding affinities, but with different kinetics and specificities. Whereas CR1CR2 bound with fast kinetics to dimers of CAAT or CGAT, CR3HD bound to ATCGAT with slow kinetics. CR1CR2 was shown to be responsible for the CCAAT displacement activity of CDP/Cux. Surprisingly, the full-length CDP/Cux protein, p200, exhibited DNA binding properties similar to that of CR1CR2, indicating that CR3HD is inactive in this context. However, an amino terminally truncated 110 KDa CDP/Cux isoform capable of stable interaction with DNA was identified in cellular extracts. The p110 isoform was shown to be generated by proteolytic cleavage of p200 at the G1/S transition. The p200 and p110 isoforms displayed different transcriptional activities in reporter assays, as only p110 was able to activate transcription from the DNA polymerase a gene promoter. Interestingly, expression of p110 was found to be increased in human uterine leiomyomas as compared to the adjacent normal myometrium, raising the possibility that proteolytic processing of CDP/Cux is activated in cancer cells. Finally, cyclin D/CDK4 was shown to interact with CDP/Cux,

phosphorylate it on several serine residues and inhibit proteolytic processing as well as DNA binding by CR3HD. Replacement of these serines for alanine residues caused proteolytic processing to occur earlier in the cell cycle. These results demonstrate that cyclin D/CDK4 is responsible for the inhibition of CDP/Cux activity during the G1 phase of the cell cycle.

Résumé

Le facteur de transcription CDP/Cux contient quatre domaines de liaison à l'ADN : trois répétitions Cut (RC1, RC2 et RC3) et l'homéodomaine Cut (HD). Le "CCAAT-displacement protein" (CDP) réprime certains promoteurs en compétitionnant avec des activateurs pour la liaison avec la séquence CCAAT. Le clonage du CDP humain a révélé que celui-ci était l'orthologue de la protéine Cut de la Drosophile, d'où l'origine du nom Cux (Cut homeobox). Le but de mon projet était de définir les modes d'interaction de CDP/Cux avec l'ADN. J'ai montré que la liaison à haute affinité nécessitait la coopération entre au moins deux domaines. Parmi toutes les combinaisons possibles, les répétitions Cut 1 et 2 (RC1RC2) ainsi que la répétition Cut 3 et l'homéodomaine (RC3HD) liaient l'ADN avec la plus haute affinité mais avec des spécificités et cinétiques différentes. Alors que RC1RC2 liait de façon transitoire des dimères des séquences CAAT ou CGAT, RC3HD liait de façon stable la séquence ATCGAT. J'ai démontré que RC1RC2 était responsable de la compétition pour la séquence CAAT. Curieusement, la protéine CDP/Cux entière, p200, liait l'ADN de la même manière que RC1RC2, ce qui suggère que l'activité des domaines RC3HD soit inhibée dans le contexte de la protéine entière. Toutefois, j'ai identifié dans des extraits cellulaires une protéine CDP/Cux de 110 kDa dont l'extrémité amino-terminale est tronquée et qui est capable d'interagir avec l'ADN de façon stable. J'ai démontré que l'isoforme p110 est produite par clivage protéolytique de la p200 au moment de la transition entre les phases G1 et S du cycle cellulaire (G1/S). Les isoformes p200 et p110 semblent avoir des activités de régulation transcriptionnelle différentes puisqu' en transfections transitoires seule la p110 était capable de stimuler la transcription du promoteur de la polymérase de l'ADN alpha. Par ailleurs, l'analyse de léiomyomes utérins humain a montré

que l'expression de l'isoforme p110 était augmentée dans la majorité des tumeurs. Ce résultat suggère que le clivage protéolytique de CDP/Cux soit activé dans certaines cellules cancéreuses. Finalement, j'ai montré que le complexe kinasique cyclin D/CDK4 interagit avec CDP/Cux, le phosphoryle sur plusieurs résidus sérines et inhibe le clivage protéolytique ainsi que l'activité de liaison à l'ADN des domaines RC3HD. Le remplacement de ces résidus sérines par des résidus alanine a permis au clivage protéolytique de CDP/Cux de se produire plus tôt dans le cycle cellulaire. Ces résultats indiquent que la kinase cycline D/CDK4 est responsable de l'inhibition de CDP/Cux pendant la phase G1 du cycle cellulaire.

Acknowledgment

I would like to thank my supervisor, Dr. Alain Nepveu, for his intellectual and moral support for my scientific development. He has encouraged and guided me throughout my graduate study.

I wish to thank also my thesis committee for their help: Dr. Jerry Pelletier, Dr. Mark Featherstone.

My appreciation to all the current and former members of Dr. Nepveu's laboratory: Dr. Lelia Raynal, Sherry Hebert, Brigitte Goulet, Mary Truscott, Ginette Bérubé, Lam Leduy, Madeline Poirier, Marianne Santaguida, Oriana Muzzin, Dr. Olivier Coqueret, Dr. Wendy Zeng, Dr. Nathalie Martin, Dr. France Mailly, Dr. Peter Premdas.

I also appreciate friendship and support from members of Molecular Oncology Group, especially, Lisa Lock, Caroline Saucier, Louie Lamorte, Pascal Peschard, Darren Kamikura, Dr. Morag Park, Dr. Robert Sladeck.

Thanks to Mary Truscott, Sherry Hebert, Marianne Santaguida, Brigitte Goulet and Dr. Alain Nepveu for the correction of this thesis and, especially, Brigitte Goulet and Dr. Alain Nepveu for French translation.

I also would like to express my appreciation to my family, friends and God for their support throughout my graduate study.

Financial support during these studies was provided by the Canadian Institute of Health Research (CIHR).

Preface

"The Guidelines for Thesis preparation" issued by The Faculty of Graduate Studies And Research at McGill University reads as follows:

> " Candidates have the option of including, as part of the thesis, the text of one or more papers submitted, or to be submitted, for publication, or the clearly-duplicated text (not the reprints) of one or more published papers. These texts must conform to the "Guidelines for Thesis Preparation" with respect to font size, line spacing and margin sizes and must be bound together as an integral part of the thesis.
>
> The thesis must be more than a collection of manuscripts. All components must be integrated into a cohesive unit with a logical progression from one chapter to the next. In order to ensure that the thesis has continuity, connecting texts that provide logical bridges between the different papers are mandatory.
>
> In addition to the manuscripts, the thesis must include the following: (a) a table of contents, (b) an abstract in English and French, (c) an introduction which clearly states the rational and objectives of the research, (d) a comprehensive review of the literature (in addition to that covered in the introduction to each paper), (e) a final conclusion and summary
>
> As manuscripts for publication are frequently very concise documents, where appropriate, additional material must be provided (e.g., in appendices) in sufficient detail to allow a clear and precise judgment to be made of the importance and originality of the research reported in the thesis.
>
> In general, when co-authored papers are included in a thesis, the candidate must have made a substantial contribution to all papers included in the thesis. In addition, the candidate is required to make an explicit statement in the thesis as to who contributed to such work and to what extent. This statement should appear in a single section entitled "Contributions of Authors" as a preface to the thesis. The supervisor must attest to the accuracy of this statement at the doctoral oral defense. Since the task of the examiners is made more difficult in these cases, it is in the candidate's interest to clearly specify the responsibilities of all the authors of the co-authored papers. When previously published copyright material is presented in a thesis, the candidate must include signed waivers from the co-authors and publishers and submit these to the Thesis Office with the final deposition, if not submitted previously."

I have chosen to write my thesis according to these guidelines, with two published papers and two submitted manuscripts. The thesis is organized in seven chapters: (I) Literature review, (II – V) Manuscripts with its own abstract, introduction, material and method, results, and discussion, (VI) General discussion, and (VII) contribution to original research.

Publications arising from work of the thesis

First-author publication

1. **Moon, NS.**, Bérubé, G., and Nepveu, A. (2000) CCAAT displacement activity involves CUT repeats 1 and 2, not the CUT homeodomain . J. Biol. Chem. 275(40):31325-34

2. **Moon, NS.**, Premdas, P.D., Truscott, M., Leduy, L., Bérubé, G., and Nepveu A. (2001). S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by CDP/Cut. Mol. Cell. Biol. 21(18):6332-45.

3. **Moon, NS.**, Rong Zeng, W., Premdas, P.D., Santaguida, M., Bérubé, G., and Nepveu, A. Expression of N-Terminally Truncated Isoforms of CDP/Cux is Increased in Human Uterine Leiomyomas. Oncogene. (Submitted)

4. **Moon, NS.**, Santaguida, M., Coqueret, O., Bérubé, G., and Nepveu, A. Phosphorylation of CDP/Cux in G1 Inhibits Its Proteolytic processing and DNA binding Activity C. Mol. Cell. Biol. (Submitted)

Other publications

5. Nirodi, C., Hart, J., Dhawan, P., **Moon, NS.**, Nepveu, A., Richmond, A. (2001) The Role of CDP in The Negative Regulation of CXCL1 Gene Expression. J. Biol. Chem.13;276(28):26122-31

6. Jackson, RJ., Antonia, SJ., Wright, KL., **Moon NS.,** and Nepveu A, Munoz-Antonia T. Human cut-like repressor protein binds TGF-ß type II receptor gene promoter. (1999) Arch. Biochem. Biophys. 371(2):290-300

Contributions of Authors

Bérubé, G.

In manuscript number 1, she created plasmid constructs pET-15b CR1, CR2, CR2HD, CR1-L-L-HD, and CR2-L-HD. In manuscript number 2, she generated plasmid constructs pXJ Myc-CDP/Cux-HA, pXM 659, 878, 958, 1029, 1-1302, 1-1192, 1-1109, Δ659-878. In manuscript 3, she cloned CDP/Cux cDNA from human uterine leiomyomas. In manuscript 4, she made TrieEx 612-1328 wt and mutant plasmid constructs

Premdas, P.

In manuscript number 2, he performed reporter assay of figure 8. In manuscript number 3, he helped with grinding several tissue samples of human uterine leiomyomas and myometrium.

Truscott, M.

In manuscript number 2, she helped to set up the reporter assay condition.

Leduy, L.

In manuscript number 2, he provided technical support for figures 4 and 6 where he performed the Western blot analysis of synchronized NIH3T3 protein extracts and pulse chasing experiment.

Zeng, W.

In manuscript number 3, she collected and made library of human uterine leiomyomas and myometrium.

Santaguida, M.

In manuscript number 3, she helped sequencing cDNA of CDP/Cux from several human uterine leiomyomas. In manuscript number 4, she performed phospho-amino acid analysis of figure 5.

Coqueret, O.

In manuscript number 4, he made first observation on interaction between CDP/Cux and cyclin D and performed co-immunoprecipitation analysis of Figure 1.

In manuscript number 5 and 6, I provided pEt-15b CR3HD and CR1+2 constructs for the *in vitro* DNA binding study

Table of Content

List of Figures and tables

Chapter I Literature Review

Chapter II The CCAAT-Displacement Activity Involves Cut Repeats and 2, not the Cut Homeodomain.

Chapter III S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by the CDP/Cux Homeodomain Protein.

Chapter IV Expression of N-Terminally Truncated Isoforms of CDP/Cux Is Increased in Human Uterine Leiomyomas.

LITERATURE REVIEW

1 Protein-DNA interactions in transcription

Sequence-specific DNA binding by different transcription factors controls the selective gene expression in all cell types of an organism. Insight into the physiological significance of different transcription factors can be gained via the identification of target genes and the resulting effects on transcription. For example, identification of Apaf-1 as a target of the E2F transcription factor provided a mechanism by which E2F induces apoptosis (124). Recent advance in DNA array technology and the development in the chromatin immunoprecipitation (ChIP) assay have facilitated the identification of the *in vivo* targets of transcription factors. Alternatively, some studies have focused on identifying the important regulatory elements of a promoter to understand how the gene of interest is regulated. For example, studies of the TNF-α enhancer sequence have identified important regulatory transcription factors such as NF-κB and HMG-1. In this section, some of the DNA binding motifs that are shared by different transcription factors will be discussed.

1.1 DNA binding motifs in transcription factors

Transcription factors contain at least one DNA binding domain that enables them to bind to DNA. Often, structurally conserved DNA binding motifs are shared by unrelated transcription factors. Examples include helix-turn-helix (HTH), zinc finger, basic-helix-loop-helix (bHLH), and basic leucine zipper (bZip). These motifs are frequently identified by their amino acid sequence homology.

1.1.1 Helix-turn-helix (HTH)

The HTH motif can be found in many different DNA binding proteins in prokaryotes (λCro, λ repressor, and CAP) and in eukaryotes (c-Myb, Oct-1, and Homeobox). The structure of many HTH motifs has been determined (8, 53, 95, 122, 123, 133, 180). The most basic structure of the HTH motif consists of two α-helices separated by a turn, which consists of three or more amino acids (Figure 1 A). While the length of the turn can vary, it serves to position the two α-helices in a perpendicular orientation (177). The second α-helix is called the recognition helix, because it allows the sequence specific interaction with DNA. The recognition helix is inserted into the major groove of DNA and the side chains of the helix contact specific bases of DNA. A family of eukaryotic proteins called homeobox proteins is defined by the presence of a unique HTH motif. The structure and the function of the homeobox will be discussed later.

1.1.2 Zinc finger

The zinc finger motif is present in transcription factors such as GAL4, Sp1, WT1 and steroid receptors (65, 88-90). A typical zinc finger is composed of two anti-parallel β-sheets and a recognition α-helix (Figure 1 B). Cysteine residue near the turn of each β-sheet and two histidine residues from the α-helix coordinate the zinc ion. This interaction between the peptide and the zinc ion stabilizes the globular structure of the motif. Usually, a cluster of two or more zinc fingers forms a DNA binding domain of a transcription factor. Steroid receptors contain atypical zinc fingers in that the coordination of the zinc ion is mediated by four cysteine residues from two perpendicular α-helices (154). Similar to the HTH motif, a recognition α-helix fits into the major groove and makes specific contacts with DNA.

1.1.3 Basic helix-loop-helix (bHLH) and basic leucine zipper (bZip)

bHLH and bZip motifs are similar in that they mediate two types of interactions. The first interaction is the protein-protein interaction, which results in either homo- or hetero-dimerization of the transcription factor. This dimerization allows the protein to bind DNA, constituting the second interaction. The bHLH motif consists of two α-helices connected by a loop. The basic region that binds to the major groove of DNA is an extension of the second α-helix (Figure 1 C) (109). The bZip motif has a similar structure except that the dimerization is mediated by a leucine zipper (Figure 1 D). Some proteins, such as Myc, contain a combination of both bHLH and bZip (bHLHZip) motifs. Formation of the homo- or hetero-dimers between different bHLH- and bZip-containing transcription factors is an important aspect in the regulation of transcription. An example of such regulation can be seen from the interaction between the Myc, Max, and Mad transcription factors (58, 108). While the Myc/Max heterodimer forms a transcriptional activator complex, the Mad/Max heterodimer represses transcription, antagonizing the action of the Myc/Max heterodimer.

1.1.4 Others

Many transcription factors have DNA binding motifs that are different from those mentioned above. For example, the p53 tumor suppressor protein has a DNA binding motif with a loop-sheet-helix secondary structure that has not been found in other transcription factors (26). Further more, several transcription factors are shown to bind to the minor groove of DNA, such as High mobility group-1 (HMG-1) and the TATA box binding protein (TBP). Often, these minor groove binding proteins induce a significant conformational change in the structure of DNA (10, 93, 94, 130, 166). For example, TBP binding to TATAA sequence results in DNA bending at a 90 degree angle and partial melting of the double strand DNA.

1.2 Homeodomain

The homeobox containing transcription factors can be found among different eukaryotes, ranging from yeast to human. The term homeotic gene was first given from the genetic studies in *drosophila melanogaster*, where mutations of homeotic genes resulted in the abnormal development of various body parts. For example, the Antennapedia (Ant) homeotic mutation results in the growth of a leg where an antenna should grow. Analysis of genes that result in homeotic mutations lead to the discovery of a conserved 180 bp sequence motif called the homeobox. This motif is translated into a 60 amino acid sequence, the homeodomain that can bind to a specific DNA sequence. Many homeobox containing genes were found by their sequence homology. Homeotic genes play an important role during embryonic development and their structures and functions are evolutionarily conserved among different species (118).

1.2.1 Structure of the homeodomain

The first homeodomain structure was determined with the Antennapedia homeodomain by nuclear magnetic resonance (NMR) spectroscopy (144). Since then, the structures of many other homeodomains have been identified by either X-ray crystallography or NMR. The homeodomain contains three α-helices that are folded together to form a globular structure (Figure 1 E). Helix I is connected to helix II by a loop making them anti-parallel to each other. Helices II and III are connected by a turn, resulting in the formation of a HTH motif. However, the recognition helix of the homeodomain is longer by two turns of helix than a typical HTH. Several amino acid side chains of homeodomain make contact with DNA. Like other HTH motifs, the recognition helix of the homeodomain, helix III, is inserted into the major groove, where it makes the majority of the intermolecular contacts with DNA. In addition, amino acids located at the

N-terminus of helix I were shown to bind in the minor groove and the linker region between helix I and helix II makes contact with the DNA backbone (53, 54).

1.2.2 Sequence specificity

The DNA recognition sequence of most homeodomains contains an ATTA (TAAT) core sequence (54). Conserved amino acids from helix III and N-terminal of helix I make contact with the major and the minor grooves of this sequence respectively. In addition, the 9^{th} amino acid of helix III was shown to make an important contact immediately upstream of the core sequence and contributes to the sequence specificity among different homeodomains. Mutating this particular residue was shown to be sufficient to change the DNA binding specificity of certain homeodomains. (66, 139, 163). However, the extent of sequence specificity provided by this particular amino acid could not account for the versatile functional specificity of different homeobox genes. One additional element of specificity was suggested by the observation that some homeodomain proteins interact with another transcription factor to bind to DNA. For example, the Extradenticle homeodomain protein was shown to cooperate with other homeodomain proteins, such as Ultrabithorax and abdominal-A, to increase their DNA binding specificity (165). Also, Fushi Tarazu (ftz) has been shown to interact with a nuclear orphan receptor, ftz-F1, and bind to DNA cooperatively (62). In addition, some homeobox proteins are shown to encode for an additional DNA binding domain. This has been found in the POU, Paired, Lim and Cut classes of homeodomain proteins, which encode a second DNA binding domain called POU specific, Paired, Lim, and Cut Repeat domain, respectively (53, 54). The presence of a second DNA binding domain was shown to increase the sequence specificity (4, 33, 160).

1.3 Regulating the DNA binding activity of a transcription factor. .

The activity of a transcription factor can be regulated at various levels: intracellular localization, protein-protein interaction and/or protein-DNA interaction, etc. In particular, the ability to bind DNA is an essential activity of transcription factors that is often targeted for regulation. Generally, a transcription factor that cannot bind to DNA is considered inactive. The DNA binding activity of a transcription factor can be modulated at a different level. In this section, some of various methods of controlling the DNA binding activity of a transcription factor will be discussed.

1.3.1 Protein-protein interaction

Most bHLH or bZip containing transcription factors require a partner protein in order to bind to DNA. It forms either a homodimer or a heterodimer. This requirement allows the opportunity to regulate DNA binding activity and/or specificity of bHLH or bZip containing transcription factors. For example, Id, a transcription factor with HLH without DNA binding domain, can inhibit the DNA binding activity of others (83). Also, Jun/Fos family of heterodimers typically binds to AP-1 sites (TGAC/GTCA) of promoters. However, they are able to bind to other sequence when dimerized with other family of transcription factors such as ATF and Maf (25). Another example of a transcription factor that requires protein-protein interaction is the p53 tumor suppressor protein. The active form of the p53 tumor suppressor protein that binds to DNA with high affinity was found to be a tetramer (49). It is suggested that in the absence of DNA damage, the concentration of the p53 is kept low to prevent tetramerization. When cells accumulate DNA damage, an increase in p53 protein concentration and the phosphorylation of p53 occurs to stimulate the tetramerization, thereby increasing DNA binding affinity (149).

1.3.2 Post-translational modifications

The DNA binding activity of many transcription factors has been shown to be modulated by phosphorylation (76). In addition, several transcriptional co-factors such as CBP/p300 and P/CAF have been shown to acetylate transcription factors and modify their DNA binding activity (61, 142). The mechanisms of such modifications can vary from one protein to another. For example, as shown in the cases of Max and E2F-1, the DNA binding domain or the nearby region can be directly modified and alter the DNA binding activity (9, 114). Alternatively, the modifications can occur outside of the DNA binding domain and alter the DNA binding activity indirectly (149). As will be described below, several proteins contain an auto regulatory domain, which is often regulated by a post-translational modification.

1.3.3 Intrinsic regulatory domain

The C-terminus of the p53 tumor suppressor protein was found to negatively regulate its DNA binding activity. It is believed that an allosteric interaction between the C-terminal domain and the sequence specific DNA binding domain of p53 inhibits DNA binding activity (61). Removal of this domain can increase the DNA binding activity and post-translational modifications, such as phosphorylation and acetylation of this domain, were shown to increase in the DNA binding activity of p53 (96). Ets-1 and IRF-3 transcription factors also contain an auto-inhibitory domain that modulate DNA binding activity (104, 141). Similar to p53, both Ets-1 and IRF-3 inhibitory domains affect the DNA binding activity by an allosteric interaction and can be modified by phosphorylation under different physiological conditions (31) (104).

2 The cell cycle control in eukaryotes.

In tissue culture, eukaryotic cells will duplicate themselves in time, resulting in two identical daughter cells from one parent cell. Exponentially growing cells go through cycles of different phases, called gap 1 (G1), synthesis of DNA (S), gap 2 (G2), and mitosis (M). Thus, the successive repetition of these phases is called the cell cycle. Specific events occur in each phase of the cell cycle. In G1 phase, cells increase their rate of metabolism and mass. In S phase, DNA replication occurs. In G2 phase, cells will again increase in size to be able to divide into two cells. In M phase, cellular division, mitosis, occurs. Different external signals such as nourishment, growth factors, and cellular adhesion influence the cell cycle progression. Cells can also exit from the cell cycle when some of the above requirements are not met. For example, withdrawal of serum from the medium will prevent NIH3T3 cells from dividing and drive them into the quiescent state (G0). These cells can reenter the cell cycle when the serum is added back to the medium. Cell cycle progression is controlled by sequential phosphorylation, transcription, and proteolysis of various genes and their products. In higher eukaryotes, the G1 to S phase transition is tightly regulated. There are two checkpoints where the cell cycle can be arrested if not all the requirements are met. For example, accumulation of DNA damage by either UV or chemical stress in early G1 will prevent cells from going into S phase. During cancer development, the control mechanisms of G1/S transition is often deregulated, allowing cancer cells to bypass the G1/S cell cycle checkpoint in the absence of all the necessary requirements. In this section, different processes that govern the G1 to S transition will be reviewed.

2.1 G1 Cyclin/Cdk and Cdc25A

A cyclin dependent kinase (Cdk) is a kinase that requires a cyclin as a positive regulator. Cyclins and Cdks were first identified in a yeast genetic study as important

regulatory genes for cell cycle progression (147). Since then, many different families of Cdks have been identified in higher eukaryotes and shown to be important for the cell cycle progression from G1 to M. In the G1 phase of mammalian cells, there are two families of Cdks that regulate the G1 to S phase transition: Cdk4 and Cdk6 that interact with cyclin D family members (D1, D2 and D3) in G1 and Cdk2 that complexes with cyclin E from late G1 to early S (43). Cdc25 is a dual specificity phosphatase that was also initially identified in yeast as a regulator of mitosis (148). In mammals, three Cdc25s (Cdc25A, Cdc25B, and Cdc25C) have been identified and Cdc25A was shown to play an important role in the G1/S transition (86). Both Cdk4/6 and Cdk2 kinase complexes and Cdc25A are capable of shortening G1 phase when over expressed in cells (16, 146). In addition, the level of the activity of theses enzymes is often up regulated in various cancers (30, 178). Interestingly, even though the physiological relevance of cyclins, Cdks, and Cdc25s has been established, the targets of their enzymatic action have been poorly understood.

2.1.1 Cyclin D dependent kinase

When tissue culture cells come out of quiescence and reenter the cell cycle, cyclin D dependent kinase activity, Cdk4/6, is detected before any other Cdk activity (116, 119). The kinase activity of Cdk4/6 is closely associated with the protein level of cyclin D, although other regulatory mechanisms are also at play. In the absence of serum, the level of cyclin D is kept low and increases when serum is added back. In exponentially growing cells, while the steady state level of other cyclins oscillates during the cell cycle, the expression level of cyclin D stays relatively constant (116). The steady state level of cyclin D can be regulated at the level of both synthesis and degradation. Transcription of cyclin D is activated by serum responsive transcription factors such as Myc and AP-1 (19, 71, 140). Cyclin D protein is degraded via ubiquitin dependent proteolysis and has been shown to be

regulated by GSK-3 phosphorylation (40). Besides the expression level of cyclin D, other factors are shown to modulate Cdk4/6 activity. Cyclin activating kinase (CAK) can phosphorylate different Cdks, including CDk4, and activates the kinase (48, 115). Furthermore, the Cip/Kip and ink4 family of inhibitor proteins can physically interact with cyclin D/Cdk4/6 complexes and Cdk4/6 alone, respectively, and inhibit the kinase activity (156). The role of the Cip/Kip and Ink4 families of inhibitors will be discussed later. As mentioned above, not many targets of cyclin D/Cdk4/6 have been identified. Rb family proteins are the known targets of cyclin D/Cdk4/6. Cyclin D/Cdk4/6 phosphorylates Rb protein to inactivate it (46). In fact, the effects on cell cycle progression by cyclin D/Cdk4/6 are tightly linked to their ability to phosphorylate Rb family members. For example, cyclin D/Cdk4/6 is unable to overcome the cell cycle arrest induced by Rb protein containing mutations in the Cdk phosphorylation sites (106). Also, drugs that specifically inhibit Cdk4/6 kinases cannot arrest the cell cycle in cell lines where Rb proteins are non-functional (50). In addition to their ability to phosphorylate Rb family members, a recent study demonstrated that cyclin D/Cdk4/6 complexes contributes to the activation cyclin E/Cdk2 by sequestering the Cip/Kip inhibitory proteins away from cyclin E/Cdk2 complex (140). Recently, beside its role in cell cycle progression, it was found that cyclin D dependent kinase may also play a role in cell growth (35).

2.1.2 Cyclin E dependent kinase

After Cdk 4 kinase activity is first detected in early G1, kinase activity of Cdk2 can be detected next in late G1 when cyclin E begins to be expressed (41). The expression level of cyclin E oscillates in a cell cycle dependent manner and is highest from late G1 to early S phase. Similar to cyclin D, the steady state level of cyclin E is regulated both at the level of synthesis and degradation. Furthermore, it can also phosphorylate Rb family

members. It is believed that, Rb protein is sequentially phosphorylated by cyclin D/Cdk4/6 and cyclin E/Cdk2 from G1 to early S (107). The important role of the Rb family proteins will be discussed later. However, contrary to cyclin D/Cdk4/6, cyclin E/Cdk2 can overcome cell cycle arrest imposed by Rb that cannot be phosphorylated by cyclin E/Cdk2 (106), suggesting that there are other targets of cyclin E/Cdk2 that contribute to the G1/S transition. Recent studies showed that NPAT, a protein that is important for histone biosynthesis, is a downstream target of cyclin E/Cdk2 (187, 188). Also, some studies suggested that cyclin E/Cdk2 might be more directly involved in initiation of DNA replication (2).

2.1.3 Cdc25A

Cdc25A is a dual specificity phosphatase that can dephosphorylate phospho serine, threonine, and tyrosine. One of the known functions of the Cdc25 family of phosphatase is the removal of inhibitory phosphates on different Cdks (153, 155). In particular, Cdc25A removes the inhibitory phosphate at tyrosine 15 of Cdk2 (51, 155). Transcription of Cdc25A is activated by cell cycle dependent transcription factors such as Myc and E2F (152, 170). For optimal phosphatase activity, Cdc25A is phosphorylated by different kinases such as Raf1 and Pim1 (52, 121), including cyclin E/Cdk2 (73). Thus, it is believed that cyclin E/Cdk2 and Cdc25A are involved in a positive feedback loop that leads to an increase in their activity at the G1/S transition. Recent report that CDP/Cux is a novel substrate of Cdc25A raises the possibility that Cdc25A might have other targets that contribute to the G1 to S transition (28).

2.2 Cyclin-dependent inhibitors

There are two groups of cyclin dependent kinase inhibitors, Cip/Kip family proteins (p21, p27, and p57) and Ink4 family proteins (p16, p15, p18, and p19). The Cip/Kip family

can form complexes with both cyclin D/Cdk4/6 and cyclin E/Cdk2 (156). Interaction between the cyclin D dependent kinase complex and Cip/Kip can have both positive and negative effects on kinase activity depending on the stoichiometry of these molecules. It is believed that at a lower concentration, Cip/Kip proteins can act as an assembly factor for cyclin D dependent kinase complexes, and at a higher concentration, they can inhibit the kinase activity (98). However, the interaction between Cip/Kip proteins and cyclin E/cdk2 results exclusively in inhibition of the kinase activity (143). The steady state level of theses inhibitors is tightly controlled by various mechanisms under different circumstances. For example, transactivation of the $p21^{waf1/cip1/sdi1}$ gene by p53 is shown to be required for cell cycle arrest induced by DNA damage (179). Also, the stability of p27 protein is controlled by an ubiquitin dependent proteolysis that can be triggered by cyclin E/Cdk2 phosphorylation in a cell cycle dependent manner (164). Ink4 family proteins can only form a complex with Cdk4/6, and, not with Cdk2. They inhibit the kinase activity by preventing the association with cyclin D (72, 85). However, recent evidence suggests that the interaction of Ink4 family proteins with Cdk 4/6 also serves to relocate Cip/Kip family inhibitors from cyclin D dependent kinase complex to cyclin E dependent kinase complexes, resulting in inactivation of both cyclin D and E dependent kinases (117). Because of their ability to inhibit the Cdk activity, the Cip/Kip and Ink4 have a negative effect on cell cycle progression and are deregulated in human tumors.

2.3 Proteolysis

Many genes that are shown be important in the cell cycle progression have been first identified from yeast genetic studies. Often the biological importance of these genes was revealed before their mechanism of action. It is biochemical studies that have demonstrated that some of these cell cycle regulatory genes identified in yeast are involved

in ubiquitin dependent proteolysis (56). In particular, a trimeric ubiquitin ligase protein complex called SCF (Skp-Cullin-F-box protein) is found to be a key regulator of the G1/S transition (105). An F-box protein is responsible for the substrate specificity of SCF complexes (37). It physically interacts with substrates and targets it to the SCF complex for the ubiquitination. Often, the F-box protein selectively recognizes a phosphorylated substrate Families of different SCF complexes were identified in different organisms. In mammals, SCF complex that contains Skp2 as the F-box protein was shown to be important for the G1/S transition (127). There are two main targets of SCF^{Skp2} that have been identified, cyclin E and p27. In addition, a recent study demonstrated that SCF^{Skp2} can also ubiquitinate E2F-1(112). Ubiquitination of p27 by SCF^{Skp2} is phosphorylation dependent (164). Only p27 that has been phosphorylated by cyclin E/Cdk2 on threonine 187 is recognized by Skp2. It is believed that the full activation of cyclin E/Cdk2 is achieved when cyclin E/Cdk2 has accumulated enough to be able to phosphorylate p27 and target it for degradation. Cyclin E itself is believed to be a target of Skp2. Contrary to p27, SCF^{Skp2} does not recognize a phosphorylated form of cyclin E selectively. However, Skp2 interacts with the free form of cyclin E (126). Cyclin E that is complexed with Cdk2 is protected from SCF^{Skp2} dependent ubiquitination. Indeed, a Skp2 knock-out mouse shows an increase in the steady state level of free cyclin E and p27 (126). Recently, a human homologue of yeast cdc4, an F-box protein that recognizes a phosphorylated form of cyclin E has been identified (159). Other molecules, such as cyclin D, have also been shown to be degraded in an ubiquitin dependent manner (40). However, the F-box protein that targets these proteins for ubiquitination has not been identified. It is important to note that the proteolysis mentioned above is involved in a degradation pathway of the proteins.

2.4 Transcription

When serum deprived quiescent cells reenter the cell cycle by supplementing serum back to the medium, a large increase in the expression of immediate early gene such as the Myc transcription factor can be detected as early as 30 min. post-stimulation (18, 32). Myc has been shown to activate the transcription of genes that are important for cell growth (64). However, expression level of Myc does not change in continuously dividing cells ((67, 162), reviewed in (18)). Over-expressing Myc can shorten the G1phase of the cell cycle. It is one of the first identified proto-oncogenes that are frequently over-expressed in many cancers. Recent evidence suggests that the Myc transcription factor is an interdependent parallel pathway to the E2F transcription factor, which can stimulate G1 to S progression by activating transcription of important cell cycle regulators such as cyclin D and Cdc25A (19, 140, 152). The Rb/E2F/DP trimeric complex is the most studied cell cycle dependent transcription factor that controls the G1/S transition (42). In higher eukaryotes, the Rb/E2F/DP complex was shown to control the transcription of cyclin E, Cyclin A, Cdc25A, etc (34, 134, 138, 170). In mammals, three Rb family proteins (pRb, p107, and p130) and six E2F family proteins (E2F-1 to –6) and two DP family proteins (DP-1 and DP-2) have been identified so far. E2Fs and DPs heterodimerize and bind to DNA. Except for E2F-6, which does not have a transactivation domain, E2F/DP acts as a strong transcriptional activator. Rb family proteins do not bind to DNA directly, however, they can be recruited to DNA through E2F/DP and repress transcription. Recent studies showed that active repression by Rb is achieved, in part, by the association with histone deacetylase and chromatin remodeling factor (20, 42). In early G1, the hypophosphorylated form of Rb proteins are predominant and interact with the E2F/DP heterodimer. The Rb/E2F/DP protein complexes repress the transcription of the genes mentioned above. As cells

approach S phase, Rb family proteins are phosphorylated by cyclin D/Cdk4/6 and cyclin E/Cdk2. The hyperphosphorylated form of Rb dissociates from E2F/DP, which can then activate transcription (42). For transcription of some genes like cyclin E, repression by Rb/E2F/DP seems to have a predominant effect since the mutation of the E2F/DP binding site in their promoter results in derepression of transcription (99). For other promoter, the reverse seems to be true (171). Rb is often inactivated in cancers. In fact, Rb is the first identified tumor suppressor genes (101) and a target of DNA tumor virus proteins such as SV40 Large T and adenovirus E1A (36, 176). However, neither E2F nor DP family proteins have been found to function as oncogenes in human cancers thus far. A potential explanation for this is the fact that E2F can induce apoptosis under certain situations (145). Recent studies using ChIP and DNA array assays demonstrated that specific members of the E2F and Rb families regulate selective sets of genes (77) and some of previously identified targets might not be the primary targets of the Rb/E2F/DP complex (170).

2.5 Links between cell cycle regulation and cancer.

The proteins mentioned above work together to control the sequential events that are necessary for normal G1 to S transition. Since the activity of each protein can regulate each other's activity, deregulation of any component of the cell cycle regulation can potentially affect the activity of the other factors. Some of the positive and negative feedback network among different factors are described in figure 2. For example, E2F can activate transcription of cyclin E, which, in turn, can phosphorylate and inactivate Rb and p27. Also, increases in the level of p16 will inactivate Cdk 4, resulting in the activation of Rb which can inhibit E2F activity. Interestingly, in most human cancers, at least one component of the cell cycle regulation network is altered to favor uncontrolled cell cycle progression. In human retinoblastoma and osteosarcoma, inactivation of Rb can be

frequently found (11), whereas $p16^{Ink4}$ is frequently inactivated in melanoma (172). It is still unclear why the inactivation of a certain gene is manifested in a tissue-specific manner. The network of positive and negative regulation between different cell cycle regulators is closely linked to other important biological processes such as cell growth, apoptosis, development, and senescence. Some of the same cell cycle regulatory components are commonly used to regulate these biological processes. For example, exit from the cell cycle by over-expression of p21 is a necessary step for muscle development (63) and E2F and Myc can induce apoptosis under different condition (145).

3 Transcription factors and proteolysis.

In addition to the proteolysis that degrades a transcription factor (22, 60, 112), site-specific proteolysis can be used to regulate the activity of a transcription factor. In this section, I will discuss some examples of such proteolysis that can either activate or modify the activity of a transcription factor. The cleavage of a transcription factor often results in the migration of the protein from an extra-nuclear compartment to the nucleus. It is often triggered by a specific signal that will either activate or repress the transcription of a specific set of genes.

3.1 Cytoplasmic transcription factors

Many transcription factors such as Smad, Stat, and nuclear receptor are shown to migrate from the cytoplasm to the nucleus when cells receive a specific signal (87, 110). These proteins however do not require a proteolytic event for their localization to the nucleus. Nuclear factor κ-B (NF-κB) and Cubitus interruptus (Ci) transcription factors are two examples where translocation to the nucleus requires a proteolytic event.

3.1.1 NF-κB

Two distinct proteolytic processes are involved in the activation of NF-κB. NF-κB is a heterodimeric transcription factor that plays an essential role in the immune system. NF-κB can be activated by different stimuli including TNF-α, IL-2, and lipopolysaccharide (LPS) (55). In the absence of stimulus, NF-κB exists in a complex with IκB (inhibitor of κB). IκB binds to a conserved motif called the Rel homology region (RHR) of NF-κB. RHR is responsible for heterodimerization, DNA binding and nuclear localization (78). It is believed that when IκB binds to NF-κB, it masks the nuclear localization signal and retains NF-κB in the cytoplasm. When cells receive the signals mentioned above, the N-terminus of IκB is rapidly phosphorylated and ubiquitinated for degradation (91). The free form of NF-κB is then trans-located to the nucleus and activates a specific panel of genes. Recently, enzymes that are responsible for the phosphorylation and the ubiquitination of IκB have been identified (39). A large kinase complex call IKK (IκB kinase) was shown to phosphorylate serines 32 and 36 of Iκ-B. This allows recognition by a F-box protein called $E3RS^{I\text{-}\kappa B}$/β-TrCP (181). Interestingly, a subunit of NF-κB, p50, is itself a product of ubiquitin dependent partial proteolysis of the precursor protein p105. The exact mechanism of this process is still unclear (136).

3.1.2 Cubitus interruptus (Ci)

Ci is a transcription factor that is best know as a downstream effecter molecule of the Hedgehog (Hh) signaling pathway in *Drosophila melanogaster* (5). The Ci transcript encodes for a protein with zinc finger motifs as the DNA binding domain (184). However, the full length form of Ci is a cytoplasmic protein because of the C-terminal sequence that retains it in the cytoplasm (6, 125). In the absence of Hh signal, Ci is proteolytically

cleaved. The resulting N-terminal half of Ci, which contains the DNA binding domain, migrates to the nucleus (6). In contrast to the full length Ci, which is believed to be a transcriptional activator, the proteolytically cleaved N-terminal part of Ci is a transcriptional repressor. The exact mechanism of transactivation by full length Ci is still unclear. Similarly to IκB, there is evidence that the processing of Ci is also regulated by phosphorylation (24). Furthermore, an F-box protein, Slimb, was shown to modulate the processing of Ci (84). However, it is still unknown whether Ci is a direct substrate of Slimb.

3.2 Membrane-bound transcription factors

Among the transcription factors that are found in the cytoplasm as inactive forms, some transcription factors exist as transmembrane proteins. These molecules can sense extracellular signals or homeostasis of cells and directly transmit the message to the nucleus for a selective and immediate response.

3.2.1 SREBP

The steroid responsive elements binding protein (SREBP) is a transmembrane protein that can be found in the endoplasmic reticulum (ER) (75). SREBP regulates the transcription of genes that are important for fatty acid biosynthesis and lipid uptake by cells (21). Both the amino and carboxy termini of SREBP are oriented toward the cytoplasmic side of the ER membrane because the central region of the protein spans the ER membrane twice (75). At a low level of cholesterol, SREBP is sequentially cleaved; first, in the sequence exposed to the lumen and second, in the transmembrane region (150, 173). The resulting N-terminal portion that contains the bHLHZip motif is then released from the ER membrane and migrates to the nucleus. It activates transcription of genes such as 3-hydroxy-3methylglutaryl coenzyme A reductase (57). The sterol responsive cleavage of

SREBP is regulated by SREBP cleavage activating protein (SCAP), which can sense the level of sterol in cells (151). Interestingly, the protease that cleaves the first site is active in the Golgi and SCAP is responsible for trafficking SREBP from the ER to the Golgi compartment upon a decrease in the cellular sterol level (132). Recently, the same protease that is responsible for cleaving SREBP was shown to cleave ATF6, another ER membrane bound transcription factor (182).

3.2.2 Notch

In *Drosophila melanogaster*, Notch is a developmentally important protein that plays an important role in cell fate determination (59). It is a cell surface receptor protein capable of transmitting extracellular signals directly to the nucleus (97). Delta and Serrate are the ligands of Notch, and are themselves cell surface molecules (175). Similarly to SREBP, upon binding to ligands, Notch is sequentially cleaved in the extracellular domain by metalo-proteases and the transmembrane domain by a transmembrane protease called presenilin (12). Even though Notch does not bind DNA directly, the cleaved cytoplasmic domain of Notch migrates to the nucleus and interacts with a transcription factor called Suppressor of hairless (Su[H]), (82). In the absence of Notch, Su[H] represses transcription, whereas it is capable of activating transcription when complexes with Notch.

4 The CDP/Cux transcription factor

CDP/Cux is an evolutionarily conserved transcription factor that is found among all metazoans. The gene is named *cut* in *Drosophila*, *cux* (Cut homeobox) in mouse, *clox* (Cut like homeobox) in canine and in human, the protein was cloned as the CCAAT displacement protein (CDP) (Figure 3). In this thesis, it will be referred to as CDP/Cux.

4.1 *Drosophila* Cut

The nomenclature of CDP/Cux provides the information about how this gene was identified in different species. The name *cut* was given to this gene as a result of genetic studies in *Drosophila melanogaster* where a mutation in the *cut* locus resulted in a scalloping wing phenotype (79, 80). In *Drosophila melanogaster*, Cut is expressed in many tissues including the central and peripheral nervous system, the wing and leg discs, and the malphighian tubules (13, 14, 128). The tissue specific transcription of *cut* is controlled by specific enhancer elements located up to 100kb upstream of the gene. Mutations in specific tissues were caused by insertion of a *gypsy* transposable element (an insulator of transcription) between the tissue specific enhancer and the core promoter of *cut*. Thus, closer the *gypsy* integration site to the *cut* gene, more tissues were affected. Furthermore, mutations in the coding region of Cut resulted in embryonic lethality (reviewed in (128)). In tissues where it is expressed, Cut seems to play a role in cell type specification. This was best demonstrated in the nervous tissues where the absence of Cut expression transformed external sensory organs into internal sensory organs and the forced expression of Cut resulted in differentiation of all precursor cells into external sensory organs (15, 17). Furthermore, absence of Cut expression in tissues such as the wing and leg margin resulted in truncated wing and kinked femur. In particular, in the wing margin and the ovary, Cut was shown to have non cell-autonomous effects during development (81, 120). These studies also demonstrated that Cut genetically interacts with developmentally important genes such as *notch* and *wingless*.

4.2 Mammalian CDP/Cux

In mammals, the *cut* homologue was first cloned as a transcription factor capable of repressing transcription by competing with transcriptional activators for binding to the CCAAT elements of a gene promoter (7, 129). Thus, the protein was named CCAAT-

displacement protein (CDP). The target genes of CDP can be divided into two groups: genes that are repressed in proliferating precursor cells such as β-globin, gp92-phox, CD8α, etc, and genes that are expressed in proliferating cells such as histone genes, thymidine kinase, c-myc, etc (128). From the sequence analysis of these promoters, CDP recognizes a wide range of DNA sequences such as CCAAT, ATCGAT, Sp1-like site, and AT-rich sequences. The finding that CDP contains four DNA binding domains provided some hints as to how CDP has such a broad range of sequence specificity (1, 3, 69) (Figure 3). However, the exact mechanism by which CDP interacts with DNA is still unclear. Besides its ability to compete away transcriptional activators such as CBF, CDP was shown to contain two active repression domains at the C-terminal end of the protein (111) (Figure 3). Recently, this region was shown to interact with the histone deacetylase HDAC1 (103). A *cux*$^{-/-}$ mouse equivalent to *Drosophila* with a *cut* mutation in the coding region, has been generated by two different groups (45, 157). The *cux*$^{-/-}$ mice displayed defects in hair follicle and lung, B and T cell development and often died shortly after birth. There were some phenotypic similarities between *cux*$^{-/-}$ mice and *cut* mutants in *Drosophila*, such as defects in leg development However, the *cux*$^{-/-}$ did not result in embryonic lethality and in some case, was able to grow into an adult mouse. Reduced severity of the *cux*$^{-/-}$ phenotype could be explained by the existence of a second *cut* homologue in mouse called *cux-2*, which is specifically expressed in nervous system. It is unclear whether cux-2 can compensate for the loss of the *cux-1* activity, or, it contributes to distinct developmental defects.

4.3 CDP/Cux DNA binding activity and its regulation during cell cycle..

In this section, the issues to be presented in the thesis will be introduced. A detailed introduction for each subject will be presented for each manuscript.

4.3.1 The alternative modes of the DNA binding by CDP/Cux

Understanding how CDP/Cux recognizes DNA is not a trivial task. There are two reasons for this complexity. First, as mentioned above, CDP/Cux was found to bind to a variety of sequences, which do not share much sequence similarity. Second, unlike most transcription factors, CDP/Cux contains four distinct DNA binding domains: three Cut Repeats (CR) and one Cut Homeodomain (HD) (1, 3, 69) (Figure 3). Interestingly, all other genes that encode protein containing for the Cut HD such as SATB-1 and HNF-6 have at least one CR as an additional DNA binding domain, suggesting that the Cut HD always functions with CR (38, 100, 102). It is believed that having the four DNA binding domains enables CDP/Cux to bind to a wide range of sequences. There have been several studies performed with individual DNA binding domains of CDP/Cux. PCR-mediated site selection study using GST/Cut repeat fusion proteins demonstrated that CR prefers ATNNAT and CCAAT sequences (68). Also, a similar study using cell extracts transfected with vectors expressing the CDP/Cux proteins yielded similar results (1). However, not much effort has been made to understand how the four DNA binding domains of CDP/Cux function in the context of the full-length molecule. Furthermore, the mechanism of the CCAAT-displacement activity by CDP has not been studied in detail. Interestingly, the diversity of the CDP/Cux DNA binding activity was also reflected in studies where cellular extracts were used. Often, in electrophoretic mobility shift assays, CDP/Cux DNA retardation complexes with different mobilities were observed (28, 131, 167). Even though there have been reports that CDP/Cux can form a complex with other protein, no clear explanation has been given for this observation.

4.3.2 The cell cycle dependent regulation of the CDP/Cux DNA binding activity

Generally, the DNA binding activity of CDP/Cux correlates with the level of cellular proliferation. When proliferating precursor cells stop dividing and start to differentiate, a decrease in the DNA binding activity of CDP/Cux is observed (23, 137, 158). However, there is an exception to this rule whereby the DNA binding activity of CDP/Cux is sustained during megakaryocyte differentiation (113). This finding, in general, agrees with the idea that CDP/Cux represses genes that are differentiation specific in proliferating precursor cells. However, some genes that are expressed in proliferating cells have also been reported to be regulated by CDP/Cux (44, 92, 167, 168). In particular, CDP/Cux was found to be a component of HiNF-D complex. HiNF-D complex was previously identified as transcriptional activator complex whose activity is shown to be important for S phase specific expression of various Histone genes (169). The DNA binding activity of HiNF-D itself is regulated in a cell cycle dependent manner and constitutively active in tumors (74). Even though CDP/Cux alone was unable to activate Histone gene transcription (168), these findings indicate that CDP/Cux could be part of an activator complex in a cell cycle dependent manner. Furthermore, the DNA binding activity of CDP/Cux was shown to be modulated in a cell cycle dependent manner and highest in G1/S (28). The same study demonstrated that the cell cycle dependent phosphorylation in the region of the Cut HD was responsible for the inhibition of the DNA binding activity by CR3HD in G1. This inhibitory phosphorylation was counteracted by the Cdc25A phosphatase when cells enter S phase. Phosphorylation of CRs by CKII and PKC was reported previously, however, these kinases did not phosphorylate the Cut HD nor the kinase regulated in a cell cycle dependent manner (27, 29).

4.3.3 CDP/Cux in human cancer

No clear evidence has been provided to demonstrate the involvement of CDP/Cux in human cancer. The region of chromosome 7q22, where CUTL1gene is located, is frequently rearranged in uterine leiomyomas (135), acute myeloid leukemia (47, 161) and myelodysplastic syndrome (70, 183). Loss of heterozygocity (LOH) of CUTL1 gene was analyzed in human uterine leiomyomas and breast cancer using polymorphic markers within CUTL1 gene. LOH in CUTL1 locus was detected in 15% of uterine leiomyomas and 18 % of breast cancer (185, 186). In conjunction with the study showing molecular interaction between CDP/Cux protein and Polyoma Large T oncoprotein in an animal model (174), these studies introduced the possible involvement of CDP/Cux in cancer development..

4.3.4 Issues addressed in this thesis

The mechanism and the regulation of the divers DNA binding activity of CDP/Cux are investigated in this thesis. *In vitro* DNA binding studies using different fusion proteins of CDP/Cux will demonstrate the alternative CDP/Cux DNA binding characteristics by different CDP/Cux DNA binding domains. Furthermore, studies using cellular extracts will uncover novel mechanisms of regulating DNA binding characteristics of CDP/Cux. The investigation of the regulatory mechanism of CDP/CuxDNA binding will also reveal how DNA binding activity of CDP/Cux is regulated in a cell cycle dependent manner. The regulation during G1/S transition will be discussed in detail. The potential involvement of CDP/Cux activity in tumor development will be also discussed in the thesis. The expression level and the activity of CDP/Cux in the human uterine leiomyomas versus normal tissues will be presented.

REFERENCE

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:**245-57.

2. **Arata, Y., M. Fujita, K. Ohtani, S. Kijima, and J. Y. Kato.** 2000. Cdk2-dependent and -independent pathways in E2F-mediated S phase induction. J Biol Chem **275:**6337-45.

3. **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:**7757-61.

4. **Aurora, R., and W. Herr.** 1992. Segments of the POU domain influence one another's DNA-binding specificity. Mol Cell Biol **12:**455-67.

5. **Aza-Blanc, P., and T. B. Kornberg.** 1999. Ci: a complex transducer of the hedgehog signal. Trends Genet **15:**458-62.

6. **Aza-Blanc, P., F. A. Ramirez-Weber, M. P. Laget, C. Schwartz, and T. B. Kornberg.** 1997. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. Cell **89:**1043-53.

7. **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:**347-59.

8. **Beamer, L. J., and C. O. Pabo.** 1992. Refined 1.8 A crystal structure of the lambda repressor-operator complex. J Mol Biol **227:**177-96.

9. **Berberich, S. J., and M. D. Cole.** 1992. Casein kinase II inhibits the DNA-binding activity of Max homodimers but not Myc/Max heterodimers. Genes Dev **6:**166-76.

10. **Bewley, C. A., A. M. Gronenborn, and G. M. Clore.** 1998. Minor groove-binding architectural proteins: structure, function, and DNA recognition. Annu Rev Biophys Biomol Struct **27:**105-31.

11. **Blanquet, V., C. Turleau, M. S. Gross-Morand, C. Senamaud-Beaufort, F. Doz, and C. Besmond.** 1995. Spectrum of germline mutations in the RB1 gene: a study of 232 patients with hereditary and non hereditary retinoblastoma. Hum Mol Genet **4:**383-8.

12. **Blobel, C. P.** 2000. Remarkable roles of proteolysis on and beyond the cell surface. Curr Opin Cell Biol **12:**606-12.

13. **Blochlinger, K., R. Bodmer, J. Jack, L. Y. Jan, and Y. N. Jan.** 1988. Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila. Nature **333:**629-35.

14. **Blochlinger, K., R. Bodmer, 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:**1322-31.

15. **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:**1124-35.

16. **Blomberg, I., and I. Hoffmann.** 1999. Ectopic expression of Cdc25A accelerates the G(1)/S transition and leads to premature activation of cyclin E- and cyclin A-dependent kinases. Mol Cell Biol **19:**6183-94.

17. **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 D. melanogaster. Cell **51:**293-307.

18. **Bouchard, C., P. Staller, and M. Eilers.** 1998. Control of cell proliferation by Myc. Trends Cell Biol **8:**202-6.

19. **Bouchard, C., K. Thieke, A. Maier, R. Saffrich, J. Hanley-Hyde, W. Ansorge, S. Reed, P. Sicinski, J. Bartek, and M. Eilers.** 1999. Direct induction of cyclin D2 by Myc contributes to cell cycle progression and sequestration of p27. EMBO J **18:**5321-33.

20. **Brehm, A., and T. Kouzarides.** 1999. Retinoblastoma protein meets chromatin. Trends Biochem Sci **24:**142-5.

21. **Brown, M. S., and J. L. Goldstein.** 1997. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell **89:**331-40.

22. **Campanero, M. R., and E. K. Flemington.** 1997. Regulation of E2F through ubiquitin-proteasome-dependent degradation: stabilization by the pRB tumor suppressor protein. Proc Natl Acad Sci U S A **94:**2221-6.

23. **Chattopadhyay, S., C. E. Whitehurst, and J. Chen.** 1998. A nuclear matrix attachment region upstream of the T cell receptor beta gene enhancer binds Cux/CDP and SATB1 and modulates enhancer-dependent reporter gene expression but not endogenous gene expression. J Biol Chem **273:**29838-46.

24. **Chen, Y., N. Gallaher, R. H. Goodman, and S. M. Smolik.** 1998. Protein kinase A directly regulates the activity and proteolysis of cubitus interruptus. Proc Natl Acad Sci U S A **95:**2349-54.

25. **Chinenov, Y., and T. K. Kerppola.** 2001. Close encounters of many kinds: Fos-Jun interactions that mediate transcription regulatory specificity. Oncogene **20:**2438-52.

26. **Cho, Y., S. Gorina, P. D. Jeffrey, and N. P. Pavletich.** 1994. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science **265:**346-55.

27. **Coqueret, O., G. Berube, and A. Nepveu.** 1996. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Protein Kinase C. J Biol Chem **271:**24862-68.

28. **Coqueret, O., G. Berube, and A. Nepveu.** 1998. The mammalian Cut homeodomain protein functions as a cell-cycle dependent transcriptional repressor which downmodulates p21WAF1/CIP1/SDI1 in S phase. EMBO J **17:**4680-94.

29. **Coqueret, O., N. Martin, G. Berube, M. Rabbat, D. W. Litchfield, and A. Nepveu.** 1998. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Casein Kinase II. J Biol Chem **273:**2561-66.

30. **Courjal, F., G. Louason, P. Speiser, D. Katsaros, R. Zeillinger, and C. Theillet.** 1996. Cyclin gene amplification and overexpression in breast and ovarian cancers: evidence for the selection of cyclin D1 in breast and cyclin E in ovarian tumors. Int J Cancer **69:**247-53.

31. **Cowley, D. O., and B. J. Graves.** 2000. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. Genes Dev **14:**366-76.

32. **Curran, T., R. Bravo, and R. Muller.** 1985. Transient induction of c-fos and c-myc in an immediate consequence of growth factor stimulation. Cancer Surv **4:**655-81.

33. **Czerny, T., and M. Busslinger.** 1995. DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). Mol Cell Biol **15:**2858-71.

34. **Dalton, S.** 1992. Cell cycle regulation of the human cdc2 gene. EMBO J **11:**1797-804.

35. **Datar, S. A., H. W. Jacobs, A. F. de la Cruz, C. F. Lehner, and B. A. Edgar.** 2000. The Drosophila cyclin D-Cdk4 complex promotes cellular growth. EMBO J **19:**4543-54.

36. **DeCaprio, J. A., J. W. Ludlow, J. Figge, J. Y. Shew, C. M. Huang, W. H. Lee, E. Marsilio, E. Paucha, and D. M. Livingston.** 1988. SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene. Cell **54:**275-83.

37. **DeSalle, L. M., and M. Pagano.** 2001. Regulation of the G1 to S transition by the ubiquitin pathway. FEBS Lett **490:**179-89.

38. **Dickinson, L. A., C. D. Dickinson, and T. Kohwi-Shigematsu.** 1997. An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. J Biol Chem **272:**11463-70.

39. **DiDonato, J. A., M. Hayakawa, D. M. Rothwarf, E. Zandi, and M. Karin.** 1997. A cytokine-responsive IkappaB kinase that activates the transcription factor NF-kappaB. Nature **388:**548-54.

40. **Diehl, J. A., M. Cheng, M. F. Roussel, and C. J. Sherr.** 1998. Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. Genes Dev **12:**3499-511.

41. **Dulic, V., E. Lees, and S. I. Reed.** 1992. Association of human cyclin E with a periodic G1-S phase protein kinase. Science **257:**1958-61.

42. **Dyson, N.** 1998. The regulation of E2F by pRB-family proteins. Genes Dev **12:**2245-62.

43. **Ekholm, S. V., and S. I. Reed.** 2000. Regulation of G(1) cyclin-dependent kinases in the mammalian cell cycle. Curr Opin Cell Biol **12:**676-84.

44. **el-Hodiri, H. M., and M. Perry.** 1995. Interaction of the CCAAT displacement protein with shared regulatory elements required for transcription of paired histone genes. Mol Cell Biol **15:**3587-96.

45. **Ellis, T., L. Gambardella, M. Horcher, S. Tschanz, J. Capol, P. Bertram, W. Jochum, Y. Barrandon, and M. Busslinger.** 2001. The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev **15:**2307-19.

46. **Ewen, M. E., H. K. Sluss, C. J. Sherr, H. Matsushime, J. Kato, and D. M. Livingston.** 1993. Functional interactions of the retinoblastoma protein with mammalian D- type cyclins. Cell **73:**487-97.

47. **Fenaux, P., C. Preudhomme, J. L. Lai, P. Morel, R. Beuscart, and F. Bauters.** 1989. Cytogenetics and their prognostic value in de novo acute myeloid leukaemia: a report on 283 cases. Br J Haemat **73:**61-67.

48. **Fisher, R. P., and D. O. Morgan.** 1994. A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. Cell **78:**713-24.

49. **Friedman, P. N., X. Chen, J. Bargonetti, and C. Prives.** 1993. The p53 protein is an unusually shaped tetramer that binds directly to DNA. Proc Natl Acad Sci USA **90:**3319-23.

50. **Fry, D. W., D. C. Bedford, P. H. Harvey, A. Fritsch, P. R. Keller, Z. Wu, E. Dobrusin, W. R. Leopold, A. Fattaey, and M. D. Garrett.** 2001. Cell cycle and

biochemical effects of PD 0183812. A potent inhibitor of the cyclin D-dependent kinases CDK4 and CDK6. J Biol Chem **276:**16617-23.

51. **Gabrielli, B. G., M. S. Lee, D. H. Walker, H. Piwnica-Worms, and J. L. Maller.** 1992. Cdc25 regulates the phosphorylation and activity of the Xenopus cdk2 protein kinase complex. J Biol Chem **267:**18040-6.

52. **Galaktionov, K., C. Jessus, and D. Beach.** 1995. Raf1 interaction with Cdc25 phosphatase ties mitogenic signal transduction to cell cycle activation. Genes Dev **9:**1046-58.

53. **Gehring, W. J., M. Affolter, and T. Burglin.** 1994. Homeodomain proteins. Annu Rev Biochem **63:**487-526.

54. **Gehring, W. J., Y. Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A. F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, and K. Wuthrich.** 1994. Homeodomain-DNA recognition. Cell **78:**211-23.

55. **Ghosh, S., M. J. May, and E. B. Kopp.** 1998. NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol **16:**225-60.

56. **Goebl, M. G., J. Yochem, S. Jentsch, J. P. McGrath, A. Varshavsky, and B. Byers.** 1988. The yeast cell cycle gene CDC34 encodes a ubiquitin-conjugating enzyme. Science **241:**1331-5.

57. **Goldstein, J. L., and M. S. Brown.** 1990. Regulation of the mevalonate pathway. Nature **343:**425-30.

58. **Grandori, C., S. M. Cowley, L. P. James, and R. N. Eisenman.** 2000. The Myc/Max/Mad network and the transcriptional control of cell behavior. Annu Rev Cell Dev Biol **16:**653-99.

59. **Greenwald, I.** 1998. LIN-12/Notch signaling: lessons from worms and flies. Genes Dev **12:**1751-62.

60. **Gregory, M. A., and S. R. Hann.** 2000. c-Myc proteolysis by the ubiquitin-proteasome pathway: stabilization of c-Myc in Burkitt's lymphoma cells. Mol Cell Biol **20:**2423-35.

61. **Gu, W., and R. G. Roeder.** 1997. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. Cell **90:**595-606.

62. **Guichet, A., J. W. Copeland, M. Erdelyi, D. Hlousek, P. Zavorszky, J. Ho, S. Brown, A. Percival-Smith, H. M. Krause, and A. Ephrussi.** 1997. The nuclear receptor homologue Ftz-F1 and the homeodomain protein Ftz are mutually dependent cofactors. Nature **385:**548-52.

63. **Guo, K., J. Wang, V. Andres, R. C. Smith, and K. Walsh.** 1995. MyoD-induced expression of p21 inhibits cyclin-dependent kinase activity upon myocyte terminal differentiation. Mol Cell Biol **15:**3823-9.

64. **Guo, Q. M., R. L. Malek, S. Kim, C. Chiao, M. He, M. Ruffy, K. Sanka, N. H. Lee, C. V. Dang, and E. T. Liu.** 2000. Identification of c-myc responsive genes using rat cDNA microarray. Cancer Res **60:**5922-8.

65. **Haber, D. A., A. J. Buckler, T. Glaser, K. M. Call, J. Pelletier, R. L. Sohn, E. C. Douglass, and D. E. Housman.** 1990. An internal deletion within an 11p13 zinc finger gene contributes to the development of Wilms' tumor. Cell **61:**1257-69.

66. **Hanes, S. D., and R. Brent.** 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. Cell **57:**1275-83.

67. **Hann, S. R., C. B. Thompson, and R. N. Eisenman.** 1985. C-myc oncogene protein synthesis is independent of the cell cycle in human and avian cells. Nature **314:**366-69.

68. **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:**129-40.

69. **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:**2062-7.

70. **Heim, S.** 1992. Cytogenetic findings in primary and secondary MDS. Leukemia Research **16:**43-46.

71. **Herber, B., M. Truss, M. Beato, and R. Muller.** 1994. Inducible regulatory elements in the human cyclin D1 promoter. Oncogene **9:**1295-304.

72. **Hirai, H., M. F. Roussel, J. Y. Kato, R. A. Ashmun, and C. J. Sherr.** 1995. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. Mol Cell Biol **15:**2672-81.

73. **Hoffmann, I., G. Draetta, and E. Karsenti.** 1994. Activation of the phosphatase activity of human cdc25A by a cdk2-cyclin E dependent phosphorylation at the G1/S transition. EMBO J **13:**4302-10.

74. **Holthuis, J., T. A. Owen, A. J. van Wijnen, K. L. Wright, A. Ramsey-Ewing, M. B. Kennedy, R. Carter, S. C. Cosenza, K. J. Soprano, J. B. Lian, and et al.** 1990. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science **247:**1454-7.

75. **Hua, X., J. Sakai, Y. K. Ho, J. L. Goldstein, and M. S. Brown.** 1995. Hairpin orientation of sterol regulatory element-binding protein-2 in cell membranes as determined by protease protection. J Biol Chem **270:**29422-7.

76. **Hunter, T., and M. Karin.** 1992. The regulation of transcription by phosphorylation. Cell **70:**375-87.

77. **Hurford, R. K., Jr., D. Cobrinik, M. H. Lee, and N. Dyson.** 1997. pRB and p107/p130 are required for the regulated expression of different sets of E2F responsive genes. Genes Dev **11:**1447-63.

78. **Israel, A.** 2000. The IKK complex: an integrator of all signals that activate NF-kappaB? Trends Cell Biol **10:**129-33.

79. **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:**735-47.

80. **Jack, J. W.** 1985. Molecular organization of the cut locus of Drosophila melanogaster. Cell **42:**869-76.

81. **Jackson, S. M., and K. Blochlinger.** 1997. cut interacts with Notch and protein kinase A to regulate egg chamber formation and to maintain germline cyst integrity during Drosophila oogenesis. Development **124:**3663-72.

82. **Jarriault, S., C. Brou, F. Logeat, E. H. Schroeter, R. Kopan, and A. Israel.** 1995. Signalling downstream of activated mammalian Notch. Nature **377:**355-8.

83. **Jen, Y., H. Weintraub, and R. Benezra.** 1992. Overexpression of Id protein inhibits the muscle differentiation program: in vivo association of Id with E2A proteins. Genes Dev **6:**1466-79.

84. **Jiang, J., and G. Struhl.** 1998. Regulation of the Hedgehog and Wingless signalling pathways by the F- box/WD40-repeat protein Slimb. Nature **391:**493-6.

85. **Jin, X., D. Nguyen, W. W. Zhang, A. P. Kyritsis, and J. A. Roth.** 1995. Cell cycle arrest and inhibition of tumor cell proliferation by the p16INK4 gene mediated by an adenovirus vector. Cancer Res **55:**3250-3.

86. **Jinno, S., K. Suto, A. Nagata, M. Igarashi, Y. Kanaoka, H. Nojima, and H. Okayama.** 1994. Cdc25A is a novel phosphatase functioning early in the cell cycle. EMBO J **13:**1549-56.

87. **Johnson, H. M., B. A. Torres, M. M. Green, B. E. Szente, K. I. Siler, J. Larkin, 3rd, and P. S. Subramaniam.** 1998. Hypothesis: ligand/receptor-assisted nuclear translocation of STATs. Proc Soc Exp Biol Med **218:**149-55.

88. **Johnston, M.** 1987. Genetic evidence that zinc is an essential co-factor in the DNA binding domain of GAL4 protein. Nature **328:**353-5.

89. **Johnston, M., and J. Dover.** 1988. Mutational analysis of the GAL4-encoded transcriptional activator protein of Saccharomyces cerevisiae. Genetics **120:**63-74.

90. **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:**1079-90.

91. **Karin, M., and Y. Ben-Neriah.** 2000. Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. Annu Rev Immunol **18:**621-63.

92. **Kim, E. C., J. S. Lau, S. Rawlings, and A. S. Lee.** 1997. Positive and Negative Regulation Of the Human Thymidine Kinase Promoter Mediated By CCAAT Binding Transcription Factors Nf-Y/Cbf, Dbpa, and Cdp/Cut. Cell Growth & Differentiation **8:**1329-38.

93. **Kim, J. L., D. B. Nikolov, and S. K. Burley.** 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature **365:**520-7.

94. **Kim, Y., J. H. Geiger, S. Hahn, and P. B. Sigler.** 1993. Crystal structure of a yeast TBP/TATA-box complex. Nature **365:**512-20.

95. **Klemm, J. D., M. A. Rould, R. Aurora, W. Herr, and C. O. Pabo.** 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell **77:**21-32.

96. **Ko, L. J., and C. Prives.** 1996. p53: puzzle and paradigm. Genes Dev **10:**1054-72.

97. **Kopan, R., E. H. Schroeter, H. Weintraub, and J. S. Nye.** 1996. Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. Proc Natl Acad Sci U S A **93:**1683-8.

98. **Labaer, J., M. D. Garrett, L. F. Stevenson, J. M. Slingerland, C. Sandhu, H. S. Chou, A. Fattaey, and E. Harlow.** 1997. New Functional Activities For the P21 Family Of Cdk Inhibitors. Genes Dev **11:**847-62.

99. **Lam, E. W., and R. J. Watson.** 1993. An E2F-binding site mediates cell-cycle regulated repression of mouse B- myb transcription. EMBO J **12:**2705-13.

100. **Lannoy, V. J., T. R. Burglin, G. G. Rousseau, and F. P. Lemaigre.** 1998. Isoforms Of Hepatocyte Nuclear Factor-6 Differ In Dna-Binding Properties, Contain a Bifunctional Homeodomain, and Define the New Onecut Class Of Homeodomain Proteins. J Biol Chem **273:**13552-62.

101. **Lee, W. H., R. Bookstein, F. Hong, L. J. Young, J. Y. Shew, and E. Y. Lee.** 1987. Human retinoblastoma susceptibility gene: cloning, identification, and sequence. Science **235:**1394-9.

102. **Lemaigre, F. P., S. M. Durviaux, O. Truong, V. J. Lannoy, J. J. Hsuan, and G. G. Rousseau.** 1996. Hepatocyte Nuclear Factor 6, a Transcription Factor That Contains a Novel Type Of Homeodomain and a Single Cut Domain. Proc Natl Acad Sci USA **93:**9460-64.

103. **Li, S., L. Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, and M. J. Walsh.** 1999. Transcriptional Repression of the Cystic Fibrosis Transmembrane Conductance Regulator Gene, Mediated by CCAAT Displacement Protein/cut Homolog, Is Associated with Histone Deacetylation. J Biol Chem **274:**7803-15.

104. **Lin, R., Y. Mamane, and J. Hiscott.** 1999. Structural and functional analysis of interferon regulatory factor 3: localization of the transactivation and autoinhibitory domains. Mol Cell Biol **19:**2465-74.

105. **Lisztwan, J., A. Marti, H. Sutterluty, M. Gstaiger, C. Wirbelauer, and W. Krek.** 1998. Association of human CUL-1 and ubiquitin-conjugating enzyme CDC34 with the F-box protein p45(SKP2): evidence for evolutionary conservation in the subunit composition of the CDC34-SCF pathway. EMBO J **17:**368-83.

106. **Lukas, J., T. Herzinger, K. Hansen, M. C. Moroni, D. Resnitzky, K. Helin, S. I. Reed, and J. Bartek.** 1997. Cyclin E-induced S phase without activation of the pRb/E2F pathway. Genes Dev **11:**1479-92.

107. **Lundberg, A. S., and R. A. Weinberg.** 1998. Functional inactivation of the retinoblastoma protein requires sequential modification by at least two distinct cyclin-cdk complexes. Mol Cell Biol **18:**753-61.

108. **Luscher, B., and L. G. Larsson.** 1999. The basic region/helix-loop-helix/leucine zipper domain of Myc proto- oncoproteins: function and regulation. Oncogene **18:**2955-66.

109. **Ma, P. C., M. A. Rould, H. Weintraub, and C. O. Pabo.** 1994. Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. Cell **77:**451-9.

110. **Maduzia, L. L., and R. W. Padgett.** 1997. Drosophila MAD, a member of the Smad family, translocates to the nucleus upon stimulation of the dpp pathway. Biochem Biophys Res Commun **238:**595-8.

111. **Mailly, F., G. Berube, R. Harada, P. L. Mao, S. Phillips, and A. Nepveu.** 1996. The human cut homeodomain protein can repress gene expression by two distinct

mechanisms: active repression and competition for binding site occupancy. Mol Cell Biol **16:**5346-57.

112. **Marti, A., C. Wirbelauer, M. Scheffner, and W. Krek.** 1999. Interaction between ubiquitin-protein ligase SCFSKP2 and E2F-1 underlies the regulation of E2F-1 degradation. Nat Cell Biol **1:**14-9.

113. **Martin-Soudant, N., J. G. Drachman, K. Kaushansky, and A. Nepveu.** 2000. CDP/Cut DNA binding activity is down-modulated in granulocytes, macrophages and erythrocytes but remains elevated in differentiating megakaryocytes. Leukemia **14:**863-73.

114. **Martinez-Balbas, M. A., U. M. Bauer, S. J. Nielsen, A. Brehm, and T. Kouzarides.** 2000. Regulation of E2F1 activity by acetylation. EMBO J **19:**662-71.

115. **Matsuoka, M., J. Y. Kato, R. P. Fisher, D. O. Morgan, and C. J. Sherr.** 1994. Activation of cyclin-dependent kinase 4 (cdk4) by mouse MO15-associated kinase. Mol Cell Biol **14:**7265-75.

116. **Matsushime, H., D. E. Quelle, S. A. Shurtleff, M. Shibuya, C. J. Sherr, and J. Y. Kato.** 1994. D-type cyclin-dependent kinase activity in mammalian cells. Mol Cell Biol **14:**2066-76.

117. **McConnell, B. B., F. J. Gregory, F. J. Stott, E. Hara, and G. Peters.** 1999. Induced expression of p16(INK4a) inhibits both CDK4- and CDK2- associated kinase activity by reassortment of cyclin-CDK-inhibitor complexes. Mol Cell Biol **19:**1981-9.

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

119. **Meyerson, M., and E. Harlow.** 1994. Identification of G1 kinase activity for cdk6, a novel cyclin D partner. Mol Cell Biol **14:**2077-86.

120. **Micchelli, C. A., E. J. Rulifson, and S. S. Blair.** 1997. The function and regulation of cut expression on the wing margin of Drosophila: Notch, Wingless and a dominant negative role for Delta and Serrate. Development **124:**1485-95.

121. **Mochizuki, T., C. Kitanaka, K. Noguchi, T. Muramatsu, A. Asai, and Y. Kuchino.** 1999. Physical and functional interactions between Pim-1 kinase and Cdc25A phosphatase. Implications for the Pim-1-mediated activation of the c-Myc signaling pathway. J Biol Chem **274:**18659-66.

122. **Mondragon, A., S. Subbiah, S. C. Almo, M. Drottar, and S. C. Harrison.** 1989. Structure of the amino-terminal domain of phage 434 repressor at 2.0 A resolution. J Mol Biol **205:**189-200.

123. **Mondragon, A., C. Wolberger, and S. C. Harrison.** 1989. Structure of phage 434 Cro protein at 2.35 A resolution. J Mol Biol **205:**179-88.

124. **Moroni, M. C., E. S. Hickman, E. L. Denchi, G. Caprara, E. Colli, F. Cecconi, H. Muller, and K. Helin.** 2001. Apaf-1 is a transcriptional target for E2F and p53. Nat Cell Biol **3:**552-8.

125. **Motzny, C. K., and R. Holmgren.** 1995. The Drosophila cubitus interruptus protein and its role in the wingless and hedgehog signal transduction pathways. Mech Dev **52:**137-50.

126. **Nakayama, K., H. Nagahama, Y. A. Minamishima, M. Matsumoto, I. Nakamichi, K. Kitagawa, M. Shirane, R. Tsunematsu, T. Tsukiyama, N. Ishida, M. Kitagawa, and S. Hatakeyama.** 2000. Targeted disruption of Skp2 results in accumulation of cyclin E and p27(Kip1), polyploidy and centrosome overduplication. EMBO J **19:**2069-81.

127. **Nakayama, K. I., S. Hatakeyama, and K. Nakayama.** 2001. Regulation of the cell cycle at the G1-S transition by proteolysis of cyclin E and p27Kip1. Biochem Biophys Res Commun **282:**853-60.

128. **Nepveu, A.** 2001. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. Gene **270:**1-15.

129. **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. Nature Genetics **1:**50-5.

130. **Nikolov, D. B., H. Chen, E. D. Halay, A. Hoffman, R. G. Roeder, and S. K. Burley.** 1996. Crystal structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad Sci USA **93:**4862-7.

131. **Nirodi, C. S., J. G. Hart, P. Dhawan, N. Moon Ns, A. Nepveu, and A. Richmond.** 2001. The 170 kDa CCAAT displacement protein (CDP/Cut) selectively binds the IUR cis-acting element in the CXCL1 promoter : The role of CDP in the negative regulation of CXCL1 gene expression. J Biol Chem **22:**22.

132. **Nohturfft, A., D. Yabe, J. L. Goldstein, M. S. Brown, and P. J. Espenshade.** 2000. Regulated step in cholesterol feedback localized to budding of SCAP from ER membranes. Cell **102:**315-23.

133. **Ogata, K., S. Morikawa, H. Nakamura, A. Sekikawa, T. Inoue, H. Kanai, A. Sarai, S. Ishii, and Y. Nishimura.** 1994. Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. Cell **79:**639-48.

134. **Ohtani, K., J. DeGregori, and J. R. Nevins.** 1995. Regulation of the cyclin E gene by transcription factor E2F1. Proc Natl Acad Sci USA **92:**12146-50.

135. **Ozisik, Y. Y., A. M. Meloni, U. Surti, and A. A. Sandberg.** 1993. Deletion 7q22 in uterine leiomyoma. A cytogenetic review. Cancer Genetics & Cytogenetics **71:**1-6.

136. **Palombella, V. J., O. J. Rando, A. L. Goldberg, and T. Maniatis.** 1994. The ubiquitin-proteasome pathway is required for processing the NF- kappa B1 precursor protein and the activation of NF-kappa B. Cell **78:**773-85.

137. **Pattison, S., D. G. Skalnik, and A. Roman.** 1997. CCAAT displacement protein, a regulator of differentiation-specific gene expression, binds a negative regulatory element within the 5' end of the human papillomavirus type 6 long control region. J Virol **71:**2013-22.

138. **Pearson, B. E., H. P. Nasheuer, and T. S. Wang.** 1991. Human DNA polymerase alpha gene: sequences controlling expression in cycling and serum-stimulated cells. Mol Cell Biol **11:**2081-95.

139. **Percival-Smith, A., M. Muller, M. Affolter, and W. J. Gehring.** 1990. The interaction with DNA of wild-type and mutant fushi tarazu homeodomains. EMBO J **9:**3967-74.

140. **Perez-Roger, I., S. H. Kim, B. Griffiths, A. Sewing, and H. Land.** 1999. Cyclins D1 and D2 mediate myc-induced proliferation via sequestration of p27(Kip1) and p21(Cip1). EMBO J **18:**5310-20.

141. **Petersen, J. M., J. J. Skalicky, L. W. Donaldson, L. P. McIntosh, T. Alber, and B. J. Graves.** 1995. Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. Science **269:**1866-9.

142. **Polesskaya, A., A. Duquet, I. Naguibneva, C. Weise, A. Vervisch, E. Bengal, F. Hucho, P. Robin, and A. Harel-Bellan.** 2000. CREB-binding protein/p300 activates MyoD by acetylation. J Biol Chem **275:**34359-64.

143. **Polyak, K., M. H. Lee, H. Erdjument-Bromage, A. Koff, J. M. Roberts, P. Tempst, and J. Massague.** 1994. Cloning of p27Kip1, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals. Cell **78:**59-66.

144. **Qian, Y. Q., M. Billeter, G. Otting, M. Muller, W. J. Gehring, and K. Wuthrich.** 1989. The structure of the Antennapedia homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell **59:**573-80.

145. **Qin, X. Q., D. M. Livingston, W. G. Kaelin, Jr., and P. D. Adams.** 1994. Deregulated transcription factor E2F-1 expression leads to S-phase entry and p53-mediated apoptosis. Proc Natl Acad Sci USA **91:**10918-22.

146. **Resnitzky, D., M. Gossen, H. Bujard, and S. I. Reed.** 1994. Acceleration of the G1/S phase transition by expression of cyclins D1 and E with an inducible system. Mol Cell Biol **14:**1669-79.

147. **Richardson, H. E., C. Wittenberg, F. Cross, and S. I. Reed.** 1989. An essential G1 function for cyclin-like proteins in yeast. Cell **59:**1127-33.

148. **Russell, P., and P. Nurse.** 1986. cdc25+ functions as an inducer in the mitotic control of fission yeast. Cell **45:**145-53.

149. **Sakaguchi, K., H. Sakamoto, M. S. Lewis, C. W. Anderson, J. W. Erickson, E. Appella, and D. Xie.** 1997. Phosphorylation of serine 392 stabilizes the tetramer formation of tumor suppressor protein p53. Biochemistry **36:**10117-24.

150. **Sakai, J., E. A. Duncan, R. B. Rawson, X. Hua, M. S. Brown, and J. L. Goldstein.** 1996. Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. Cell **85:**1037-46.

151. **Sakai, J., A. Nohturfft, D. Cheng, Y. K. Ho, M. S. Brown, and J. L. Goldstein.** 1997. Identification of complexes between the COOH-terminal domains of sterol regulatory element-binding proteins (SREBPs) and SREBP cleavage- activating protein. J Biol Chem **272:**20213-21.

152. **Santoni-Rugiu, E., J. Falck, N. Mailand, J. Bartek, and J. Lukas.** 2000. Involvement of Myc activity in a G(1)/S-promoting mechanism parallel to the pRb/E2F pathway. Mol Cell Biol **20:**3497-509.

153. **Sebastian, B., A. Kakizuka, and T. Hunter.** 1993. Cdc25M2 activation of cyclin-dependent kinases by dephosphorylation of threonine-14 and tyrosine-15. Proc Natl Acad Sci USA **90:**3521-4.

154. **Severne, Y., S. Wieland, W. Schaffner, and S. Rusconi.** 1988. Metal binding 'finger' structures in the glucocorticoid receptor defined by site-directed mutagenesis. EMBO J **7:**2503-8.

155. **Sexl, V., J. A. Diehl, C. J. Sherr, R. Ashmun, D. Beach, and M. F. Roussel.** 1999. A rate limiting function of cdc25A for S phase entry inversely correlates with tyrosine dephosphorylation of Cdk2. Oncogene **18:**573-82.

156. **Sherr, C. J., and J. M. Roberts.** 1999. CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev **13:**1501-12.

157. **Sinclair, A. M., J. A. Lee, A. Goldstein, D. Xing, S. Liu, R. Ju, P. W. Tucker, E. J. Neufeld, and R. H. Scheuermann.** 2001. Lymphoid apoptosis and myeloid hyperplasia in CCAAT displacement protein mutant mice. Blood **98:**3658-67.

158. **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:**16736-44.

159. **Strohmaier, H., C. H. Spruck, P. Kaiser, K. A. Won, O. Sangfelt, and S. I. Reed.** 2001. Human F-box protein hCdc4 targets cyclin E for proteolysis and is mutated in a breast cancer cell line. Nature **413:**316-22.

160. **Sturm, R. A., and W. Herr.** 1988. The POU domain is a bipartite DNA-binding structure. Cell **64:**475-78.

161. **Swansbury, G. J., S. D. Lawler, G. Alimena, D. Arthur, R. Berger, H. Van Den Berghe, C. D. Bloomfield, A. de la Chappelle, G. Dewald, O. M. Garson, A. Hagemeijer, F. Mitelman, J. D. Rowley, and M. Sakurai.** 1994. Long-term survival in acute myelogenous leukemia: A second follow-up of the Fourth International Workshop on Chromosomes in Leukemia. Cancer Genet Cytogenet **73:**1-7.

162. **Thompson, C. B., P. B. Challoner, P. E. Neiman, and M. Groudine.** 1985. Levels of c-myc oncogene mRNA are invariant throughout the cell cycle. Nature **314:**363-66.

163. **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:**553-62.

164. **Tsvetkov, L. M., K. H. Yeh, S. J. Lee, H. Sun, and H. Zhang.** 1999. p27(Kip1) ubiquitination and degradation is regulated by the SCF(Skp2) complex through phosphorylated Thr187 in p27. Curr Biol **9:**661-4.

165. **van Dijk, M. A., and C. Murre.** 1994. extradenticle raises the DNA binding specificity of homeotic selector gene products. Cell **78:**617-24.

166. **van Houte, L., A. van Oers, M. van de Wetering, D. Dooijes, R. Kaptein, and H. Clevers.** 1993. The sequence-specific high mobility group 1 box of TCF-1 adopts a predominantly alpha-helical conformation in solution. J Biol Chem **268:**18083-7.

167. **van Wijnen, A. J., C. Cooper, P. Odgren, F. Aziz, A. De Luca, R. A. Shakoori, A. Giordano, P. J. Quesenberry, J. B. Lian, G. S. Stein, and J. L. Stein.** 1997. Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells. J Cell Biochem **66:**512-23.

168. **van Wijnen, A. J., M. F. van Gurp, M. C. de Ridder, C. Tufarelli, T. J. Last, M. Birnbaum, P. S. Vaughan, A. Giordano, W. Krek, E. J. Neufeld, J. L. Stein, and G. S. Stein.** 1996. CDP/cut is the DNA-binding subunit of histone gene transcription factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. Proc Natl Acad Sci USA **93:**11516-21.

169. **van Wijnen, A. J., K. L. Wright, J. B. Lian, J. L. Stein, and G. S. Stein.** 1989. Human H4 histone gene transcription requires the proliferation-specific nuclear factor HiNF-D. Auxiliary roles for HiNF-C (Sp1-like) and HiNF-A (high mobility group-like). J Biol Chem **264:**15034-42.

170. **Vigo, E., H. Muller, E. Prosperini, G. Hateboer, P. Cartwright, M. C. Moroni, and K. Helin.** 1999. CDC25A phosphatase is a target of E2F and is required for efficient E2F-induced S phase. Mol Cell Biol **19:**6379-95.

171. **Wade, M., T. F. Kowalik, M. Mudryj, E. S. Huang, and J. C. Azizkhan.** 1992. E2F mediates dihydrofolate reductase promoter activation and multiprotein complex formation in human cytomegalovirus infection. Mol Cell Biol **12:**4364-74.

172. **Wagner, S. N., C. Wagner, L. Briedigkeit, and M. Goos.** 1998. Homozygous deletion of the p16INK4a and the p15INK4b tumour suppressor genes in a subset of human sporadic cutaneous malignant melanoma. Br J Dermatol **138:**13-21.

173. **Wang, X., R. Sato, M. S. Brown, X. Hua, and J. L. Goldstein.** 1994. SREBP-1, a membrane-bound transcription factor released by sterol- regulated proteolysis. Cell **77:**53-62.

174. **Webster, M. A., N. Martinsoudant, A. Nepveu, R. D. Cardiff, and W. J. Muller.** 1998. The Induction Of Uterine Leiomyomas and Mammary Tumors In Transgenic Mice Expressing Polyomavirus (Pyv) Large T (Lt) Antigen Is Associated With the Ability Of Pyv Lt Antigen to Form Specific Complexes With Retinoblastoma and Cutl1 Family Members. Oncogene **16:**1963-72.

175. **Weinmaster, G.** 1997. The ins and outs of notch signaling. Mol Cell Neurosci **9:**91-102.

176. **Whyte, P., N. M. Williamson, and E. Harlow.** 1989. Cellular targets for transformation by the adenovirus E1A proteins. Cell **56:**67-75.

177. **Wintjens, R., and M. Rooman.** 1996. Structural classification of HTH DNA-binding domains and protein-DNA interaction modes. J Mol Biol **262:**294-313.

178. **Wu, W., Y. H. Fan, B. L. Kemp, G. Walsh, and L. Mao.** 1998. Overexpression of cdc25A and cdc25B is frequent in primary non-small cell lung cancer but is not associated with overexpression of c-myc. Cancer Res **58:**4082-5.

179. **Wyllie, F. S., M. F. Haughton, J. A. Bond, J. M. Rowson, C. J. Jones, and D. Wynford-Thomas.** 1996. S phase cell-cycle arrest following DNA damage is independent of the p53/p21(WAF1) signalling pathway. Oncogene **12:**1077-82.

180. **Xu, W., M. A. Rould, S. Jun, C. Desplan, and C. O. Pabo.** 1995. Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. Cell **80:**639-50.

181. **Yaron, A., A. Hatzubai, M. Davis, I. Lavon, S. Amit, A. M. Manning, J. S. Andersen, M. Mann, F. Mercurio, and Y. Ben-Neriah.** 1998. Identification of the receptor component of the IkappaBalpha-ubiquitin ligase. Nature **396:**590-4.

182. **Ye, J., R. B. Rawson, R. Komuro, X. Chen, U. P. Dave, R. Prywes, M. S. Brown, and J. L. Goldstein.** 2000. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. Mol Cell **6:**1355-64.

183. **Yunis, J. J., M. Lobell, M. A. Arnesen, M. M. Oken, M. G. Mayer, R. E. Rydell, and R. D. Brunning.** 1988. Refined chromosome study helps define prognostic subgroups in most patients with primary myelodysplastic syndrome and acute myelogenous leukemia. Br J Haemat **68:**189-94.

184. **Zarkower, D., and J. Hodgkin.** 1992. Molecular analysis of the C. elegans sex-determining gene tra-1: a gene encoding two zinc finger proteins. Cell **70:**237-49.

185. **Zeng, W. R., S. W. Scherer, M. Koutsilieris, J. J. Huizenga, F. Filteau, L. C. Tsui, and A. Nepveu.** 1997. Loss Of Heterozygosity and Reduced Expression Of the Cutl1 Gene In Uterine Leiomyomas. Oncogene **14:**2355-65.

186. **Zeng, W. R., P. Watson, J. Lin, S. Jothy, R. Lidereau, M. Park, and A. Nepveu.** 1999. Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene. Oncogene **18:**2015-21.

187. **Zhao, J., B. Dynlacht, T. Imai, T. Hori, and E. Harlow.** 1998. Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S- phase entry. Genes Dev **12:**456-61.

188. **Zhao, J., B. K. Kennedy, B. D. Lawrence, D. A. Barbie, A. G. Matera, J. A. Fletcher, and E. Harlow.** 2000. NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. Genes Dev **14:**2283-97.

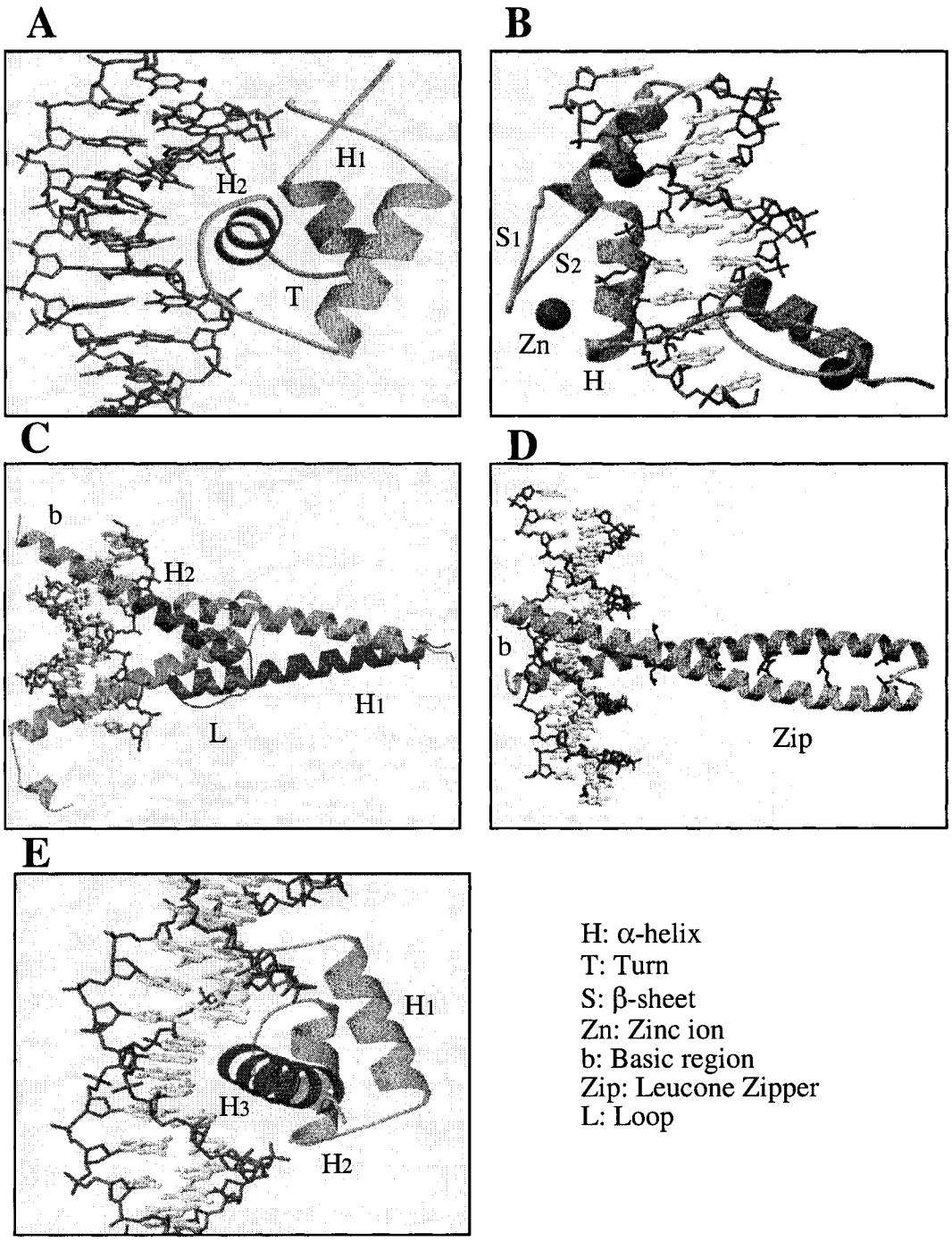


Figure 1. Structure of different DNA binding motifs. Structural representations of different DNA binding motifs bound to DNA are presented in the figure. (A) Helix-turn-helix, (B) Zinc finger, (C) Basic-leucine zipper, (D) Basic-helix-loop-helix and (E) Homeodomain.

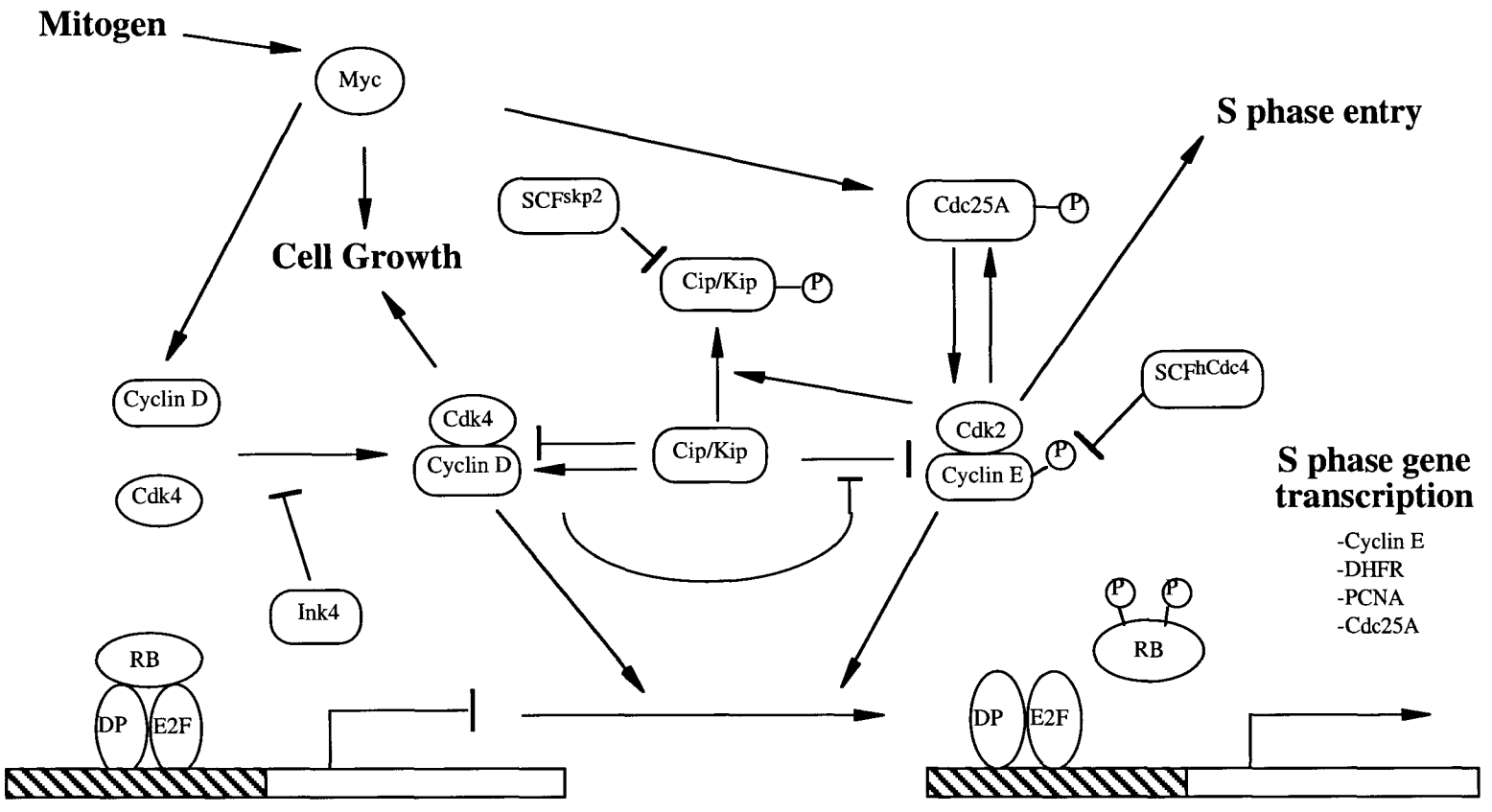


Figure 2. Interaction between different molecules that govern the G1 to S progression

The Cut Homeodomain Protein Family

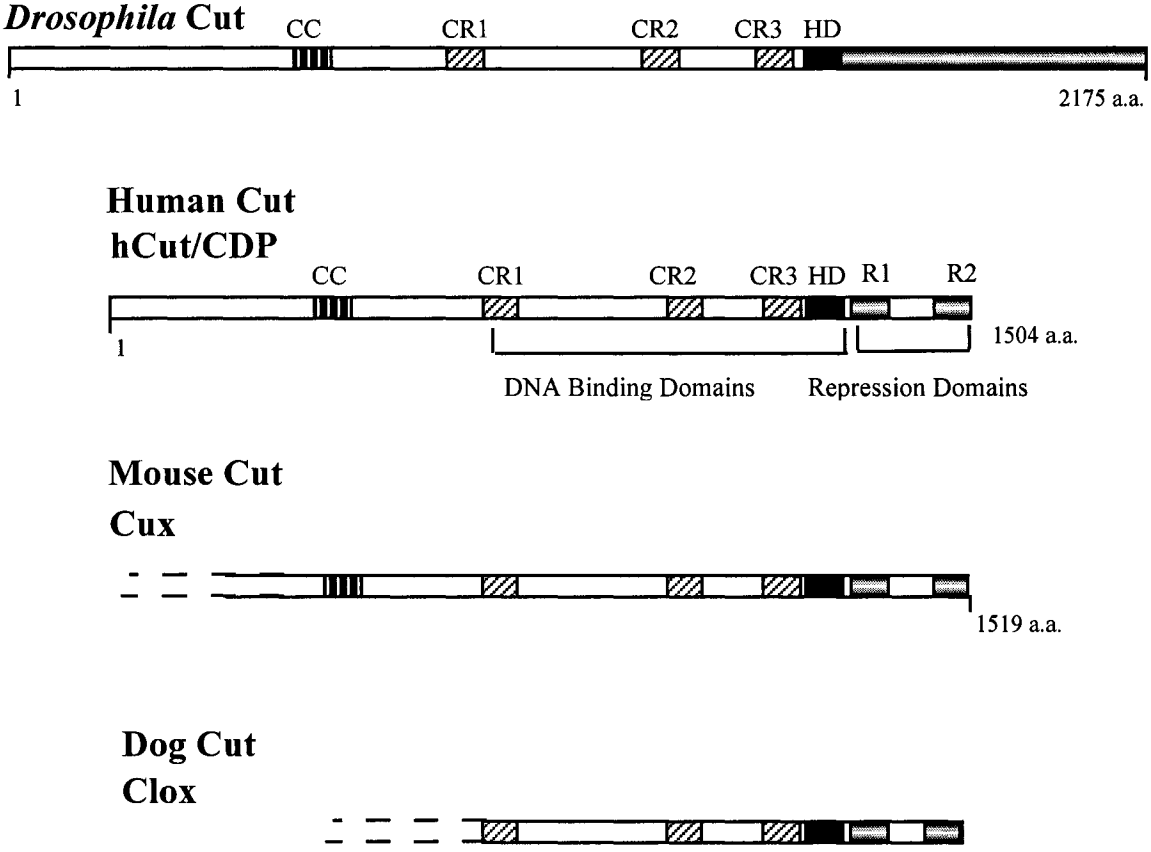


Figure 3. The Cut homeodomain protein family. Schematic representation of Cut homologue from different species is shown in the figure. Evolutionarily conserved domains are indicated as different shade of box. CC: Coiled Coil, CR: Cut Repeat, HD: Homeodomain, R: Repression domain.

Chapter II The CCAAT-Displacement Activity Involves Cut Repeats 1 and 2, not the Cut Homeodomain.

Preface

The CDP/Cux transcription factor is unique in a manner that it contains four distinct DNA binding domains, three Cut Repeats and one Cut homeodomain. It is known to bind to variety of sequences, such as CCAAT, ATCGAT, Sp1 site, and AT rich matrix attachment region. In particular, CDP/Cux can act as a repressor of transcription by preventing activators of transcription from binding to CCAAT site of a gene promoter, CCAAT-displacement activity. Previously, several studies have investigated DNA binding specificity and affinity by the individual DNA binding domain of CDP/Cux. However, they have failed to provide a molecular basis for how CDP/Cux bind to a wide range of DNA sequences. Furthermore, it was still unclear how the CCAAT-displacement activity is achieved by CDP/Cux. The following manuscript describes two most efficient bipartite DNA binding domains of CDP/Cux with different DNA binding characteristics. It discusses alternative modes of DNA binding by CDP/Cux and identification of the domain responsible for the CAAT-displacement activity.

ABSTRACT

The CCAAT-displacement protein (CDP), the homologue of the *Drosophila melanogaster* Cut protein, contains four DNA binding domains: three Cut repeats (CR1, CR2, CR3) and the Cut homeodomain (HD). Using a panel of fusion proteins, we found that a Cut repeat cannot bind to DNA as a monomer, but that certain combinations of domains exhibit high DNA binding affinity: CR1+2, CR3HD, CR1HD and CR2HD. One combination, CR1+2, exhibited strikingly different DNA binding kinetics and specificities. CR1+2 displayed rapid on and off-rates, and bound preferably to two CA/GAT sites, organized as direct or inverted repeats. Accordingly, only CR1+2 was able to bind to the CCAAT sequence, and its affinity was increased by the presence of a CA/GAT site at close proximity. A purified CDP/Cux protein exhibited DNA binding properties similar to those of CR1+2, and in nuclear extracts the CCAAT-displacement activity also required the simultaneous presence of a CA/GAT site. Moreover, CR1+2, but not CR3HD, was able to displace the NF-Y factor. Thus, the CCAAT-displacement activity requires the presence of an additional sequence, CAAT or CGAT, and involves Cut repeats 1 and 2, but not the Cut homeodomain.

INTRODUCTION

A CCAAT displacement activity was identified originally in the sea urchin and later in humans and other mammals (9, 51, 52). Sequence analysis of the cDNA encoding for the human CCAAT- displacement protein (CDP) revealed a high degree of conservation with the *Drosophila melanogaster* homeodomain-protein, Cut (12, 46). Several lethal and viable *cut* mutations have been reported in *Drosophila* (reviewed in (31)); (11, 14, 16, 26, 29, 30, 41, 42). Altogether, genetic studies in *Drosophila* suggested that the Cut protein plays an important role in determining cell-type specificity in several tissues late in development (13, 14, 41, 42). In mammals, CDP/Cux has been characterized primarily as a transcriptional repressor (38, 39, 43, 44). The bulk of the results indicated that CDP/Cux expression or activity may be restricted to proliferating cells. Accordingly, many of the identified targets of CDP/Cux are genes that are repressed in proliferating precursor cells and are turned on as cells become terminally differentiated and CDP/Cux activity ceases (4, 9, 28, 32, 33, 39, 46, 51, 54, 57). It was thus proposed that mammalian CDP/Cux proteins function as transcriptional repressors that inhibit terminally differentiated gene expression during early stages of differentiation (27, 39, 51). In addition, a role for CDP/Cux in cell cycle progression was suggested from the findings that CDP/Cux DNA binding activity oscillates during the cell cycle, reaching a maximum at the end of G1 and during the S phase (21). CDP/Cux was shown to repress the $p21^{WAF1/CIP1/SDI1}$ gene and also to bind to the promoters of various histone genes, which are regulated in a cell-cycle dependent manner (8, 9, 21, 23, 36, 55, 56). Intriguingly, binding of CDP/Cux to histone gene promoters has been associated with both activation

and repression of these genes, and it has been suggested that CDP/Cux activity could be modulated by Rb and Rb-related proteins (9, 23, 36, 55, 56).

Sequence homology between Drosophila and mammalian Cut proteins is limited to five evolutionarily conserved domains: a region predicted to form a coiled-coil structure (CC), three regions of ~70 amino-acids, the Cut repeats (CR1, CR2 and CR3), that share from 52 to 63% amino-acid identity with each other, and a Cut-type homeodomain (HD) (12, 46). The high degree of conservation of Cut repeats suggested that they may have an important biochemical function. Indeed, Cut repeats were found to function as specific DNA binding domains (3, 6, 24, 25). CDP/Cux proteins therefore are unique in that they contain four DNA binding domains: the Cut homeodomain and the three Cut repeats.

The available data suggest that CDP/Cux has the capability to bind to a wide range of DNA sequences. Reported binding sites included sequences related to either CCAAT, ATCGAT, Sp1 sites and AT-rich matrix attachment regions (MAR) (20, 22, 40, 57). On the other hand, PCR-mediated site selection with GST fusion proteins containing various CDP/Cux DNA binding domains led to the isolation of several types of sequences which could be aligned to either ATNNAT (mainly ATCGAT) or CCAAT. However, a sizable fraction of the selected sequences (~20%) diverged greatly from these consensus and yet represented excellent binding sites (3, 6, 24). These results therefore indicate that CDP/Cux can tolerate a certain degree of flexibility in their DNA targets. This property may not be unique to CDP/Cux proteins, as a similar relaxed sequence specificity was found for the GATA factors when submitted to PCR-mediated site selection (35, 45).

The evolutionary conservation of four DNA binding domains within the same protein strongly suggests that this peculiar organization permits the execution of

functions that could not otherwise be fulfilled. At this point, however, it is not well understood how CDP/Cux proteins interact with DNA. So far, DNA binding by Cut repeats has been investigated using glutathione S-transferase fusion proteins, which exist as dimers (3, 6, 24, 25). It has not been convincingly demonstrated whether each Cut repeat can bind DNA as a monomer or whether it requires cooperation with another Cut repeat or with the Cut homeodomain. As a monomer, Cut repeat 3 was shown to cooperate with the Cut homeodomain to bind to DNA with high affinity and specificity (24, 25). Yet, it is not known whether Cut repeats 1 and 2 can cooperate with each other or with the Cut homeodomain. If this was the case, we could extrapolate that various combinations of Cut repeats and the Cut homeodomain may confer to the protein the capacity to interact with a large spectrum of DNA sequences. Moreover, if CDP/Cux proteins were capable of binding simultaneously to multiple binding sites, then cooperativity should increase the overall affinity for a given DNA segment. Another issue that remains to be investigated is which of the four CDP/Cux DNA binding domains are responsible for the CCAAT-displacement activity, and what are the DNA sequence requirements for this activity to take place.

To begin to decipher the various modes of interaction of CDP/Cux with DNA, we have expressed each Cut repeat as a monomer, either alone or with another Cut repeat or the Cut homeodomain. As monomers, either two Cut repeats or one Cut repeat and the Cut homeodomain were required for efficient DNA binding. The DNA binding properties of the most efficient combinations, CR1+2 and CR3HD, were characterized and compared to that of the full length CDP/Cux protein. We then defined the DNA sequence requirement for the CCAAT-displacement activity and investigated the contribution of

each domain to this activity. In contrast to what was previously reported, our results demonstrate that the Cut repeats 1 and 2 are responsible for the CCAAT-displacement activity, without the participation of the Cut homeodomain. A model is presented to illustrate the known CDP/Cux DNA binding activities.

MATERIALS AND METHODS

Plasmid Construction. Plasmids for expression of histidine-tagged fusion proteins were prepared by inserting various fragments from the human CDP/Cux cDNA (GenBank, accession No. M74099) into the bacterial expression vectors pET-15b (Novagen) (46). CR1: nt 1605 - 2019 into the Xho I site of pET-15b. CR2: nt 2861-3153 into the Xho I site of pET-15b. CR1HD and CR2HD: a fragment encoding the homeodomain, nt 3737-3949, was placed in frame at the carboxy terminus of his-CR1 and his-CR2 in the vectors described above. CR1+2: nt 1694-3127 into the Bam HI site of pET-15b. CR2+3: nt 2861-3737 into the Bam HI site of pET-15b. For CR1-L-L-HD, three separate fragments, nt 1694-2853, nt 3071-3409 and nt 3673-3963 were placed in frame into the Xho I site of pET-15b. For CR2-L-HD, two fragments, nt 2853-3409 and nt 3673-3963, were placed in frame into the Bam HI site of pET-15b. The MBP-CR3 and MBP-CR3HD vectors have previously been described (24). MBP-CR1 was prepared by inserting nt 1605 - 2019 into the pMal-C2 plasmid (New England Biolabs). For expression of the full length CDP/Cux protein in SF9 insect cells, nt 27-5101 were inserted into the pBlueBac His2b (InVitrogen). The resulting plasmid was cotransfected with a helper plasmid to obtain Baculoviruses expressing CDP/Cux.

Expression and purification of CDP/Cux fusion proteins. pET-15b and pMal-C2 derived vector were introduced into the BL21(DE3) or DH5 strain of *E.coli* and induced with IPTG. SF9 insect cells were infected with Baculovirus encoding his-CDP/Cux and incubated for 3 days. The fusion proteins were purified by affinity chromatography using procedures provided by the suppliers.

Electro Mobility Shift Assay (EMSA). EMSA were performed with either 10 ng of purified fusion protein or 5 μg of nuclear extract from mammalian cells. The samples were incubated at room temperature for 5 min. in a final volume of 30 μl of 25 mM NaCl, 10 mM Tris, pH 7.5, 1 mM $MgCl_2$, 5 mM EDTA, pH 8.0, 5 % of glycerol, 1 mM of DTT, with 30 ng of poly dIdC and 3 μg of BSA as nonspecific competitors. End-labeled double stranded oligonucleotides (~10 pg) were added and further incubated for 15 min. at room temperature. Samples were loaded on a 5% or 4% polyacrylamide gel (30:1) and separated by electrophoresis at 8V/cm in 0.5 x Tris borate-EDTA. Gels were dried and visualized by autoradiography.

Oligonucleotides. The sequences of oligonucleotides used in this study are as follows: 29FP, TCGAGACGGTATCGATAAGCTTCTTTTC; "CGGT-ATCGAT", same as 29FP; "CGATATCGAT", TCGAGAAAAGAAGCTTATCGATATCGTC; "GG-GCATCGAT", TCGAGAAAAGAAGCTTATCGAT-GCCCT; N11, ACGCGATCGATATTTCGA-ATTCGCC; N15, CTAGATCGATATTTAGTTC-GAATTCG; N19, CTAGATCGATATTTAGTTCGT-TCGAATTCG; N21, CTAGATCGATATTTAGAGA-TCTCGTTCGAATTCG; "CGAATTCG", TCGACT-TGTCATTGGCGAATTAGAACACGCA; "CCAAT", TCGAGAAAAGAACAACCAATCACCCGTC

; "CCAAT--CGAT", TCGAGAAAAGAACAACCAA-TCACCGATC.

PCR-mediated random site selection. Binding site selections were performed essentially as described previously, using 50 ng of hisCR1+2 fusion protein in the first selection cycle, and then 10 ng in the subsequent cycles (10, 24). The sequence of the oligonucleotide used was 5'-AGACCTGCAGTC-TGC$(N)_{15}$CTGTCGTCTAGAGGA-

3". Protein-DNA complexes were separated from the free oligonucleotides by electrophoresis on a 5% polyacrylamide gel (30:1). After the fifth cycle, the PCR products were digested with PstI and Xba I and cloned into the plasmid pBlueScript pKS (Stratagene). Sequencing of the inserts were performed with the T7 polymerase sequencing kit (USB).

Calculation of the DNA binding affinity. To determine the dissociation constant (K_D), EMSA were performed essentially as described above, but using a fixed amount of DNA (≤10 pM) and a wide range of protein concentrations and with the following modifications: less than 10 pM of DNA was used, and protein and DNA were incubated for 15 min at room temperature. The binding affinity (K_D) was calculated using the method described by Janet Carey (17, 18). The amount of free and bound DNA was quantitated by scanning of the autoradiograms on a Phosphoimager (Fuji). Scintillation counting of the excised bands in one case gave similar results. The data was plotted as the fraction of free DNA Vs log of protein concentration. Since the protein concentrations did not take into account the fraction of inactive proteins, which were estimated in independent experiments to be less than 30% in each case, our data are referred to as the apparent dissociation constant (K_D(app.)).

On (k_{on}) and off (k_{off}) rate Determination To estimate the on rate, the reaction was performed at 4°C. 50 ng of purified protein was preincubated at 4°C for 30 min. in EMSA binding buffer prior to adding 1 ng of radiolabelled probe. Once the probe was added to the reaction mixture, an aliquot was taken and loaded on non denaturing polyacrylamide gel at different time points in the presence of electric current. To estimate the off rate, 50 ng of purified proteins was incubated with 1 ng of radiolabelled probe for

15 min. A 1000 fold excess of the unlabeled probe was added to the reaction mixture, and aliquots were taken and loaded on non denaturing polyacrylamide gel at different time points in the presence of electric current.

Preparation of nuclear cell extract. Monolayers of NIH3T3 cells were grown in Dulbecco modified Eagle supplemented with 10% fetal bovine serum. Nuclear extracts were prepared according to the procedure of Lee et al., except that nuclei were obtained by submitting cells to 3 freeze/thaw cycles in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT) (37).

Methylation interference assay. The assay was performed essentially as previously described (58) The probe was prepared from the pKS pBlueScript vector (Stratagene) whose polylinker includes an ATCGAT motif. To label the upper or lower DNA strand, the plasmid was digested with either Eco R1 or Xho 1, end-labeled with [g-^{32}P]ATP using T4 polynucleotide kinase and digested with the second enzyme, Xho 1 or Eco R1. The probes were purified from a non-denaturing acrylamide gel by passive elution. The labeled probes were partially methylated at purines using dimethylsulfate (DMS). 20 000 c.p.m. of partially methylated probe was used in EMSA as described above. The free and the retarded DNAs were visualized using autoradiography and purified by passive elution. The DNAs were cleaved at guanine and adenine bases using 0.2 M of NaOH. The digested products were analyzed on 12 % denaturing polyacrylamide.

RESULTS

Monomers of Cut repeats do not efficiently bind to DNA. The DNA binding properties of Cut repeats have previously been investigated using glutathione S-transferase (GST) fusion proteins (3, 7, 19, 24, 25). Since GST fusion proteins exist as dimers, these prior studies did not resolve whether individual Cut repeats can bind to DNA (1, 24, 50). To address this issue, we expressed each Cut repeat as a fusion protein with a histidine-tag (His) and assessed their DNA binding properties in electrophoretic mobility shift assays (EMSA) using 29FP oligonucleotides encoding a CDP/Cux consensus binding site, ATCGAT. Representative results using 10 ng of his-tagged fusion proteins are shown in Fig. 1B. Similar results were obtained when the maltose binding protein (MBP) was fused to either CR1, CR3 and CR3HD (data not shown). When expressed as monomeric fusion proteins, none of the Cut repeats nor the Cut homeodomain were able to bind efficiently to DNA (Fig. 1B, lanes 1-4). We then asked whether various combinations of Cut repeats and the Cut homeodomain could bind to DNA as monomers. The Cut repeats 1 and 2 (CR1+2), but not the Cut repeats 2 and 3 (CR2+3), were able to bind to DNA (Fig. 1B, lanes 5 and 6). In addition, each Cut repeat bound to DNA when expressed together with the Cut homeodomain: CR1HD, CR2HD and CR3HD (Fig. 1B, lanes 7-9). In the latter fusion proteins, the Cut repeat 3 and the Cut homeodomain were maintained in the arrangement in which they exist in the CDP/Cux protein. Thus, the result obtained with the CR3HD fusion protein most likely reflects the activity of these domains in the context of the CDP/Cux protein. However, the same cannot be said of CR1HD and CR2HD, because in these fusion proteins the

homeodomain has been brought in close proximity to Cut repeat 1 or 2, unlike in the original CDP/Cux protein. To verify whether the Cut repeats 1 and 2 could cooperate with the Cut homeodomain in the context of the entire CDP/Cux protein, we engineered larger fusion proteins in which the Cut repeats 2 and/or 3 had been precisely deleted, keeping intact the linker regions between Cut repeats: CR1-L-L-HD and CR2-L-HD (Fig. 1A). These two fusion proteins were able to bind to DNA, indicating that the Cut repeats 1 and 2 can cooperate with the Cut homeodomain even when positioned at a distance from each other in the primary sequence (Fig. 1B, lanes 10 and 11). Whether such cooperation can occur in the context of a full length CDP/Cux protein which contains CR3 could not be demonstrated, because any DNA sequence that is recognized by CR1HD or CR2HD is also bound by CR3HD (see discussion). In summary, when expressed as monomers individual Cut repeats cannot bind to DNA but can do so in conjunction with another Cut repeat or the Cut homeodomain. Therefore, at least two DNA binding domains are necessary for DNA binding, and the combinations that work are CR1+2, CR1HD, CR2HD and CR3HD.

CR1+2 binds to sequences containing two CA/GAT sites, organized either as direct or inverted repeats, and separated by a variable number of nucleotides. We have assessed the DNA binding specificity of CR1+2 by performing polymerase chain reaction-mediated random oligonucleotide selection using oligonucleotides containing random bases at 15 positions. Selected sequences were then analyzed in order to find the best alignment and derive a consensus binding site. Strikingly, all selected binding sites contained a CAAT or CGAT sequence almost invariably positioned at the beginning of the random sequence, indicating that CA/GAT is a preferred binding site but that the

flanking, non random, sequences also contribute to the binding site (Fig. 2A). For this reason, we have included the last three non random bases in the compilation of selected sequences. In addition, all selected binding sites contained a second CA/GAT sequence (although sometimes with one mismatch), either as a direct or inverted repeat and positioned at a variable distance from the first CA/GAT site. Consequently, the selected sequences have been organized in two groups on the basis of whether the CA/GAT sites occur as direct or inverted repeats. Thus, the derived consensus binding sites for CR1+2 include two half-sites organized as direct or inverted repeats: $cCA/GAT_{(n)}CA/GAT$.

The consensus binding site for CR1+2 is quite different from the ATCGAT consensus that was previously identified using GST fusion proteins (3, 7, 24). However, we noted that the juxtaposition of two CGAT sites, as in the case of sequences where N equals zero (N=0), reconstitutes the ATCGAT consensus binding site. Such a sequence would be expected to be recognized by CR3HD, whereas other sequences where N is superior to zero (N>0) would be bound by CR1+2 only. Indeed, CR1+2 bound well to oligonucleotides where the CA/GAT sites were separated by either 0, 5 or 6 bases, but CR3HD recognized only the $N_{(0)}$ oligonucleotide (Fig. 2B).

Monomers of CR1 or CR2 do not bind to a recognition site that is optimal for CR1+2. We showed in Fig. 1B that monomers of CR1 or CR2 could not bind efficiently to oligonucleotides containing the ATCGAT motif. This failure of Cut repeat monomers to bind DNA could have resulted from the fact that the ATCGAT site is not optimal for Cut repeats. To resolve this issue, we performed EMSA with oligonucleotides encoding a consensus binding site for CR1+2. Monomers of CR1 or CR2 were compared with the CR1+2 bipartite domain. While CR1+2 generated a strong

retarded complex in the nanomolar range, a monomer of CR1 produced only a faint retarded complex when present in the 10^{-7}M range, and the CR2 monomer did not exhibit any DNA binding (Fig. 2C). Altogether, results in Fig. 1B and 2C demonstrate that Cut repeats do not efficiently bind to DNA as monomers.

CR1+2 binds to the ATCGAT binding site more efficiently when a second CGAT motif is also present. In previous studies, the universal Cut repeat consensus binding site was defined as ATCGAT (3, 6, 24). Indeed, CR1+2 bound well to the 29FP oligonucleotide containing this motif (Fig. 1, lane 5). We reasoned that this oligonucleotide was well recognized by CR1+2 because it contained one perfect CGAT half-site (within the ATCGAT motif), and one imperfect inverted half-site, CGGT, immediately upstream (Fig. 3A). To test this hypothesis, we generated two sets of oligonucleotides which differed within the second half site. In one case, a point mutation generated a second, perfect, CGAT site. As shown in Fig. 3, CR1+2 bound more efficiently to this sequence (compare panels A and B). In the other case, two point mutations created a GGGC motif, which diverged completely from the CGAT consensus. CR1+2 bound the least efficiently to this sequence (Fig. 3, compare panel C with A and B). In contrast, the affinity of CR3HD for the three sets of oligonucleotides did not vary significantly (Fig. 3, compare panels D, E and F). These results demonstrate that the affinity of CR1+2 for a given DNA sequence depends on the presence of two half-sites that conform to the CA/GAT consensus.

CR1+2 and CR3HD exhibit similar DNA binding affinity but very different DNA binding kinetics. Since the Cut repeats and homeodomain exist in their normal configuration in the CR1+2 and CR3HD fusion proteins, we decided to characterize

further their DNA binding properties. The DNA binding affinity of each fusion protein for its consensus binding site was measured by calculating the apparent dissociation constant (K_D), as described in Materials and Methods (Fig. 4). The apparent dissociation constants were calculated by plotting the fraction of free DNA versus the log of protein concentration. CR3HD and CR1+2 exhibited similar apparent dissociation constant (K_D) for their respective consensus binding sites: 1.6 and 1.1 x 10^{-9} M (Fig. 4 and Table 1).

Whereas the off rate of CR1+2 was less than 1 min., the CR3HD/DNA complex was much more stable (Fig. 5 A and C). We considered the possibility that the low stability of the CR1+2/DNA complex was due to the fact that this binding site was not optimal. We thus recalculated the off rate, this time using the $N_{(5)}$ oligonucleotides containing the CCAAT---ATTG sequence (note that ATTG is the reverse complement of CAAT). Again, the retarded complex had completely disappeared 1 min. after addition of the specific competitor DNA (Fig. 5 B). We conclude that a fast off rate is an intrinsic property of CR1+2. The Cut homeodomain must be responsible for the increased stability of the protein/DNA complex since CR1HD and CR2HD exhibited slower off rates than CR1+2 (Fig. 5, D, E and A, respectively).

Another striking difference was observed with regards to the ON rate. Whereas an equilibrium was reached within the first minute in the reaction with CR1+2, the intensity of the retarded complex continued to increase for at least 15 min. in the case of CR3HD (Fig. 5 F and G). Thus, CR1+2 rapidly associates with DNA, while formation of CR3HD/DNA complexes takes place much more slowly. As a result, even though the CR1+2/DNA complex was very unstable, the rapidity with which CR1+2 bound to DNA helps explain why its apparent K_D was not higher than that of CR3HD. In summary,

CR1+2 and CR3HD exhibited strikingly different DNA binding kinetics: CR1+2 rapidly bound to DNA, but just as rapidly dissociated from it. In contrast, CR3HD slowly associated with DNA but remained bound to it for a much longer period of time.

CR3HD wraps around the DNA more than CR1+2. To understand the difference in DNA binding kinetics between CR1+2 and CR3HD, we performed DNA methylation interference assays. The most obvious differences were at the second and fifth positions of the ATCGAT core, as indicated by the stars in Fig. 6A. Methylation at these position interfered with CR3HD but not CR1+2 binding. In contrast, methylation of the adjacent guanine residue interfered with both CR1+2 and CR3HD binding. Since dimethylsulfate is known to methylate adenine at the N-3 position within the minor groove, and guanine at the N-7 position within the major groove, these results suggest that CR3HD makes contact within the minor and major groove at these positions, whereas CR1+2 makes contact within the major groove only.

To confirm that CR1+2 and CR3HD make contact within the major groove at the CG positions within the ATCGAT core, we performed EMSA using oligonucleotides in which the G:C base pairs were replaced for C:I. The structure of the C:I base pair differs from that of the C:G base pair only within the minor groove. Thus, these mutations should not affect DNA binding by CR1+2 and CR3HD if indeed these proteins bind within the major groove at these positions. This is exactly what was observed: CR1+2 and CR3HD efficiently bound to a DNA molecule containing the ATCIAT sequence in place of ATCGAT (Fig. 6B, compare lanes 1 with 3, and 4 with 6). In contrast, a C:I base pair differs from a T:A base pair only within the major groove. Thus, replacing the T:A and A:T base pairs with C:I and I:C at the second and fifth position of the ATCGAT

core, respectively, should abolish DNA binding if CR1+2 and CR3HD normally make contact within the major groove at these positions. Indeed, CR1+2 and CR3HD did not bind efficiently to a DNA molecule containing the ACCGIT sequence instead of ATCGAT (Fig. 6B, compare lane 1 with 2, and 4 with 5).

Altogether results from methylation interference assays and EMSA with mutated binding sites demonstrate that CR1+2 and CR3HD make contact within the major groove at positions 2, 3, 4 and 5 of ATCGAT, and within the minor groove at position 1 and 6 of ATCGAT (see diagram in Fig. 6C). The main difference was at the second and fifth positions of ATCGAT. CR3HD makes contact within the major and minor grooves at these positions whereas CR1+2 makes contacts only within the major groove. Thus, CR3HD wraps around the DNA, at least to some extent, at these positions, whereas CR1+2 interacts with only one side of the double helix. These differences in DNA binding contacts are likely to explain the higher stability of the CR3HD/DNA complex as compared to CR1+2.

CR1HD binds preferably to an ATCAAT site. DNA binding studies presented in Fig. 1 demonstrated that CR1 can cooperate with the Cut homeodomain to form a bipartite DNA binding domain. These results raised the possibility that the CDP/Cux protein may exist in alternative conformations, one of which would favor the association of the homeodomain with CR1 instead of CR3. To investigate this possibility, we would need first to identify sequences that are recognized by CR1HD but not by CR3HD. To this end, PCR-mediated site selection was performed with CR1HD. The derived consensus binding site was ATCAAT, although a few of the selected sequences included different bases at the two central positions of the ATNNAT core (data not shown). The

same sequences were previously reported to be selected at high frequency by CR3HD (24). Therefore, as of yet we do not know a DNA sequence that is recognized uniquely by CR1HD. As a result, it is not possible to verify whether CDP/Cux may exist in different conformational states.

The full length CDP/Cux exhibits limited cooperativity between CR1+2 and CR3HD. We next investigated the DNA binding properties of the full length CDP/Cux protein. To this end, the protein was purified as a histidine-tagged fusion protein using a baculovirus expression system. A coomassie stain of the purified protein is presented (Fig. 7A). Since CR1+2 and CR3HD can bind to distinct DNA sequences, we verified whether the full length protein would bind with higher affinity to a DNA molecule containing two binding sites, one for CR1+2 and one for CR3HD. EMSA were performed with a series of double-stranded oligonucleotides containing either one or two binding sites separated by varying distances. Oligonucleotides were designed such that their total lengths did not vary by more than just a few base pairs. Judging from the intensity of the retarded complexes, the full length CDP/Cux protein exhibited the highest affinity for a DNA molecule, N19, containing two binding sites separated by 19 base pairs (calculated from the center of each binding site) (Fig. 7B). The apparent dissociation constants indicated that the affinity was approximately 3-fold higher for N19 (K_D(app.) = 0.5 x 10^{-9} M), than for the single ATCGAT site (K_D(app.) = 1.6 x 10^{-9} M). In contrast, CR1+2 or CR3HD, when expressed separately, did not bind significantly better to the DNA molecule containing two sites (Table 1). These results demonstrate that the full length CDP/Cux protein is capable of cooperativity, albeit to a limited extent. The

possible reasons for the lack of extensive cooperativity will be addressed in the discussion.

The NF-Y (CCAAT) binding site is recognized by CR1+2, but not CR3HD, CR2HD nor CR23HD. We asked which of the CDP/Cux DNA binding domains were able to bind to a CCAAT site. First, EMSA were performed using CR1+2 or CR3HD fusion proteins and oligonucleotides encoding either the ATCGAT site (29FP), the CCAAT site, or both the CCAAT site and a CGAT site. The rationale for the design of the latter oligonucleotides was that the optimal binding sequence for CR1+2, as determined in Fig. 2, includes not just one but two CA/GAT sites. While both CR1+2 and CR3HD were able to bind to the ATCGAT sequence, only CR1+2 efficiently bound to the single CCAAT site (Fig. 8, panels A to F). Importantly, replacement of two bases to create a second CA/GAT site (CCAAT+CGAT) greatly increased the affinity of CR1+2 for the oligonucleotides (Fig. 8, panels B and C; see also Table 1). Thus, CR1+2 binds with higher affinity than CR3HD to a prototype CCAAT site, and the affinity of CR1+2 for this site is further increased when a second CA/GAT sites is present. These results are in agreement with the findings of several studies which reported the presence of more than one CCAAT-like sequence within promoters that are regulated by the CCAAT-displacement activity (2, 9, 23, 32, 33, 36, 43, 46, 48, 51, 53). Importantly, fusion proteins containing CR2 and the Cut homeodomain, CR2HD and CR23HD, did not efficiently bind to the CCAAT sites although they could bind to ATCGAT (Fig. 8, panels G to L). Thus, in contrast to previous claims, the CCAAT-displacement activity does not involve cooperation between the Cut repeat 2 and the Cut homeodomain (3). Altogether, these results suggest that the CCAAT-displacement activity of CDP/Cux is

provided by CR1+2 and is optimized by the proximity of another half-site, CAAT or CGAT.

The full length CDP/Cux displays DNA binding kinetics and specificity similar to that of CR1+2. To confirm that CR1+2 is active in the context of the full length CDP/Cux protein, we tested whether the protein could bind to the CCAAT and CCAAT+CGAT DNA sequences, two sites that are well recognized by CR1+2 but not by CR3HD. The full length CDP/Cux protein efficiently bound to these DNA sequences (Fig. 8, panels N and O; see also Table 1). Moreover, CDP/Cux displayed DNA binding kinetics essentially similar to that of CR1+2 (Fig. 9, A and B). It rapidly formed a complex with DNA and it rapidly dissociated from it. Altogether, these results suggest that the CR1+2 domains are active in the context of the full length protein and play an important role in determining the DNA binding kinetics and specificity of CDP/Cux.

CR1+2 is responsible for the CCAAT-displacement activity of CDP/Cux. The results accumulated so far suggested that the CCAAT-displacement activity would require the presence of another half site, CAAT or CGAT, and would involve competition by CR1+2. To confirm these hypotheses, we analyzed binding to the CCAAT site using nuclear extracts from NIH3T3 cells. With a probe that contained CCAAT plus an CGAT motif, two main retarded complexes were observed (Fig. 10A, lane 2). The fast migrating complex disappeared in the presence of anti-NF-Ya antibodies, while antibodies against CDP/Cux caused a supershift of the slower migrating, demonstrating that these two complexes contained NF-Y and CDP/Cux, respectively (Fig. 10A, lanes 3 and 4). The slow migrating CDP/Cux complex was not observed with a probe containing a simple CCAAT site, indicating that CDP/Cux does not efficiently bind

to a CCAAT site in the absence of a second half site (Fig. 10A, lanes 6-8). In contrast, with a CCAAT--CGAT probe, overexpression of CDP/Cux led to an increase in the intensity of the slow migrating complex with a corresponding decrease of the NF-Y complex (Fig. 10B, lanes 1 and 2). As expected, the CCAAT-displacement activity was abolished by adding an excess of cold oligonucleotides containing a CDP/Cux consensus binding site (Fig. 10, lanes 2 and 3). These results confirm that the CCAAT-displacement activity is strengthened by the presence of a CA/GAT motif at close proximity. We then verified whether the addition of purified CR1+2 or CR3HD to the nuclear extract would similarly compete with NF-Y. Purified CR3HD did not bind to the CCAAT--CGAT probe (Fig. 10C, lane 5), and addition of CR3HD to the nuclear extract did not significantly affect any of the retarded complexes (Fig. 10, lanes 6-8). In contrast, 1 ng of CR1+2 produced a strong retarded complex, and addition of the same amount of CR1+2 to the nuclear extract was sufficient to decrease both the CDP/Cux and NF-Y retarded complexes (Fig. 10C, lanes 2 and 3). Altogether, these results demonstrate that Cut repeats 1 and 2 are responsible for the CCAAT-displacement activity of CDP/Cux.

DISCUSSION

Cut proteins belong to a novel class of homeodomain proteins that exhibit the unique feature of containing multiple DNA binding domains. In previous studies, the DNA binding properties of various GST/Cut repeat fusion proteins were analyzed (3, 7, 19, 24, 25). However, GST fusion proteins were shown to exist as dimers, a property which may have affected their interaction with DNA (1, 24, 50). Moreover, how the native Cut protein interacts with DNA using four DNA binding domains, has not been thoroughly investigated. In the present study, we analyzed monomeric histidine-tagged fusion proteins containing one or two Cut DNA binding domains. We then investigated the properties of the full length CDP/Cux protein, either as a purified protein from a Baculovirus expression system, or as an endogenous protein present in nuclear extracts from NIH3T3 cells. These experiments have established that a single Cut repeat does not efficiently bind to DNA as a monomer (Fig. 1B and 2C). However, a Cut repeat can efficiently bind to DNA in cooperation with either another Cut repeat or with the Cut homeodomain. The combinations that worked best were CR1+2, CR1HD, CR2HD and CR3HD. Thus, not only can CR3 cooperate with the homeodomain, but also CR1 and CR2, even when positioned at a distance from the homeodomain. Obviously, the linker regions between the Cut repeats and the Cut homeodomain allow sufficient flexibility to permit interactions between different Cut repeats and the homeodomain. We can therefore envisage, at least in theory, multiple modes of DNA binding by CDP/Cux. However, the demonstration that CR1 or CR2 can interact with HD when CR3 is also present in the protein would require that we use a binding site that is recognized by CR1HD or CR2HD, but not by CR3HD. Unfortunately, such a binding site does not

exist because the DNA binding specificities of the various Cut repeats are not different enough.

While CR1 and CR2 were able of cooperative DNA binding, CR2 and CR3 were not. This was surprising, in light of the fact that CR2 was able to interact with either CR1 or HD, that CR3 could interact with HD, and that each Cut repeat can cooperate with itself when expressed as a GST homodimer (3, 6, 24). We have considered the possibility that the failure of CR2+3 to bind to DNA resulted from the fact that we did not use an optimal binding site for this particular combination. However, using PCR-mediated site selection, we have not been able to pull out high affinity binding sites for CR2+3, even after seven amplification/selection cycles. These results lead us to conclude that some intrinsic properties of CR2 and CR3 or the linker region between them do not permit a functional association between these two domains.

In contrast, the linker region between CR1 and CR2 appears to be highly flexible since the optimal binding sites as defined in PCR-mediated site selection contained two CA/GAT sites organized either as direct or inverted repeats. Moreover, the distance between these two sites could be varied (within the limits of the random sequence) without affecting the efficiency of DNA binding. Thus, CR1 and CR2 can exist in multiple configurations in relation to one another. Although we did not determine the upper limit of the distance between two CA/GAT sites, it is likely that CR1+2 are able to accommodate a fairly large distance since cooperative binding to CCAAT sites at a distance has been reported for the gp91phox gene promoter (43).

The kinetics of DNA binding by CR1+2 and heterodimers made of one Cut repeat and the Cut homeodomain were very different. CR1+2 displayed rapid on and off-rates,

whereas much slower on and off-rates were observed when the Cut homeodomain was tested in association with any of the Cut repeats. The results from methylation interference assays suggested that the differences in binding kinetics is due to the fact that CR1+2 binds the second and fifth positions of ATCGAT only within the major groove, while CR3HD is able to make contacts both within the major and minor grooves. This latter findings is in agreement with the crystal structure of the Pou domain and Pou homeodomain in association with their binding site (5, 34). In this complex, the Pou homeodomain was found to wrap around the double helix and make contact within the minor groove.

The results from PCR-mediated site selection indicated that CR1+2 binds with higher affinity to DNA sequences containing two CA/GAT sites in either orientations. Since the CAAT sequence is part of the consensus NF-Y binding site, our results suggested that CR1+2 would bind with high affinity to this site when a second CAAT or CGAT site was present. Indeed, our results confirmed this prediction. In contrast, CR3HD preferred ATCGAT or ATCAAT, and did not bind if the CAAT sequence was not preceded by AT (24). Similarly, CR2HD and CR23HD efficiently bound to ATCGAT but not to the NF-Y binding site. These results are in disagreement with previous claims that the CCAAT-displacement activity involves cooperation between the Cut repeat 2 and the Cut homeodomain (3). Our results clearly demonstrate that the CCAAT-displacement activity involves CR1+2 without the participation of the Cut homeodomain. It is important to note, however, that the efficiency of the CCAAT-displacement activity will not be the same on all promoters containing a CCAAT site, and will rely on the presence of a second CAAT or CGAT site. The strength of the CCAAT-

displacement activity will be maximal when two perfect half sites are present, and will decrease when one or both half sites diverge from the consensus binding site. In agreement with this prediction, we note that all promoters so far characterized as being regulated by a CCAAT displacement activity indeed contain two CA/GAT sites (see Table 2) (2, 9, 23, 32, 33, 36, 43, 46, 48, 51, 53). Moreover, in EMSA with DNA sequences from these promoters, detection of a CDP/Cux retarded complex required that a fairly long piece of DNA be used as a probe. In retrospect, we interpret these results to mean that a single CCAAT site was not sufficient for CDP/Cux binding.

Our results revealed that cooperation between various CDP/Cux DNA binding domains can generate at least two DNA binding activities with distinct binding kinetics and specificities (Fig. 11). On the one hand Cut repeats 1 and 2 bind rapidly but transiently to sequences containing two CA/GAT sites in either orientation. On the other hand, Cut repeat 3 and the Cut homeodomain can form a stable complex with the ATCGAT DNA sequence (Fig. 11). Moreover, we were able to show that the purified full length protein binds with higher affinity to oligonucleotides containing two binding sites (Fig. 7). Thus, the two bipartite domains CR1+2 and CR3HD can cooperate, albeit weakly, to bind to DNA. Surprisingly, however, *in vitro* the full length Cut protein bound with kinetics similar to that of CR1+2, with fast on and off rates (Fig. 9). Although a fast on rate was to be expected because of CR1+2, the simultaneous presence of CR3HD should have stabilized the protein on DNA. These results suggest the possibility that CR3HD is not very active in the context of the full length CDP/Cux protein. This would help explain the lack of extensive cooperativity as noted above. While such behavior by CDP/Cux in cells could be explained by invoking the

phosphorylation of the homeodomain during the G1 phase of the cell cycle, the same explanation cannot hold in the case of the purified full length protein which was dephosphorylated *in vitro* prior to DNA binding (21). We considered the possibility that CR1+2, in a manner analogous to HMG proteins, may impart a conformational change to DNA that would cause the quick release of CR3HD from its adjacent binding site (15, 47, 49). But when tested as individual proteins, CR1+2 and CR3HD were able to bind simultaneously to the N19 probe which contains binding sites for both proteins (data not shown). Therefore, CR1+2 does not prevent the stable binding of CR3HD to an adjacent site. Our results instead suggest that something else, perhaps some conformational constraint, prevents CR3HD from binding to DNA with high affinity in the context of the purified full length protein. Yet, the same protein when present within nuclear extracts in S phase can stably bind to DNA (21). The molecular basis for this discrepancy is not known, however, our results clearly point to a difference in the behavior of the CDP/Cux protein *in vitro* and in nuclear extracts.

ACKNOWLEDGMENTS

A.N. is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec. N.M.S. is the recipient of a fellowship from Medical Research Council of Canada. 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.

REFERENCES

1. **Aceto, A., P. Caccuri, T. Sacchetta, B. Bucciarelli, N. Dragani, G. Rosato, and C. Di Ilio.** 1992. Dissociation and unfolding of Pi-class glutathione transferase. Evidence for a monomeric inactive intermediate. Biochem J **285:**241-45.

2. **Ai, W., E. Toussaint, and A. Roman.** 1999. CCAAT displacement protein binds to and negatively regulates human papillomavirus type 6 E6, E7, and E1 promoters. J Virol **73:**4220-9.

3. **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:**245-57.

4. **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:**321-34.

5. **Assa-Munt, N., R. J. Mortishire-Smith, R. Aurora, W. Herr, and P. E. Wright.** 1993. The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage lambda repressor DNA-binding domain. Cell **73:**193-205.

6. **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:**7757-61.

7. **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:**7757-61.

8. **Aziz, F., A. J. Vanwijnen, P. S. Vaughan, S. J. Wu, A. R. Shakoori, J. B. Lian, K. J. Soprano, J. L. Stein, and G. S. Stein.** 1998. The integrated activities of Irf-2 (Hinf-M), Cdp/Cut (Hinf-D) and H4tf-2 (Hinf-P) regulate transcription of a cell cycle controlled human histone H4 gene - mechanistic differences between distinct H4 genes. Mol Biol Rep **25:**1-12.

9. **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:**347-59.

10. **Blackwell, T. K., and H. Weintraub.** 1990. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science **250:**1104-10.

11. **Blanc, R.** 1942. The production of wing scalloping in *Drosophila melanogaster*. Univ Calif Publ Zool **49**.

12. **Blochlinger, K., R. Bodmer, J. Jack, L. Y. Jan, and Y. N. Jan.** 1988. Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila. Nature **333:**629-35.

13. **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:**1124-35.

14. **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 D. melanogaster. Cell **51:**293-307.

15. **Boonyaratanakornkit, V., V. Melvin, P. Prendergast, M. Altmann, L. Ronfani, M. E. Bianchi, L. Taraseviciene, S. K. Nordeen, E. A. Allegretto, and D. P. Edwards.** 1998. High-mobility group chromatin proteins 1 and 2 functionally interact with steroid hormone receptors to enhance their DNA binding in vitro and transcriptional activity in mammalian cells. Mol Cell Biol **18:**4471-87.

16. **Braun, W.** 1942. The effect of puncture on the developing wing of several mutants of *Drosophila melanogaster.* J Exp Zool **84:**325-50.

17. **Carey, J.** 1991. Gel retardation. Methods Enzymol **208:**103-17.

18. **Carey, J.** 1988. Gel retardation at low pH resolves trp repressor-DNA complexes for quantitative study. Proc Natl Acad Sci USA **85:**975-9.

19. **Catt, D., W. Luo, and D. G. Skalnik.** 1999. DNA-binding properties of CCAAT displacement protein cut repeats [In Process Citation]. Cell Mol Biol (Noisy-Le-Grand) **45:**1149-60.

20. **Chattopadhyay, S., C. E. Whitehurst, and J. Chen.** 1998. A nuclear matrix attachment region upstream of the T cell receptor beta gene enhancer binds Cux/CDP and SATB1 and modulates enhancer-dependent reporter gene expression but not endogenous gene expression. J Biol Chem **273:**29838-46.

21. **Coqueret, O., G. Berube, and A. Nepveu.** 1998. The mammalian Cut homeodomain protein functions as a cell-cycle-dependent transcriptional repressor which downmodulates P21(Waf1/Cip1/Sdi1) In S phase. EMBO J **17:**4680-94.

22. **Dickinson, L. A., C. D. Dickinson, and T. Kohwi-Shigematsu.** 1997. An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. J Biol Chem **272:**11463-70.

23. **el-Hodiri, H. M., and M. Perry.** 1995. Interaction of the CCAAT displacement protein with shared regulatory elements required for transcription of paired histone genes. Mol Cell Biol **15:**3587-96.

24. **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:**129-40.

25. **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:**2062-7.

26. **Hertweck, H.** 1931. Anatomie und variabilitaet des nerven-systems und der sinneorgane von *Drosophila melanogaster.* J Exp Zool **139:**559-663.

27. **Heuvel, G. B. V., R. Bodmer, K. R. McConnell, G. T. Nagami, and P. Igarashi.** 1996. Expression of a Cut-related homeobox gene In developing and polycystic mouse kidney. Kidney International **50:**453-61.

28. **Higgy, N. A., H. A. Tarnasky, I. Valarche, A. Nepveu, and F. A. Vanderhoorn.** 1997. Cux/CDP homeodomain protein binds to an enhancer in the rat c-Mos locus and represses its activity. Biochimica et Biophysica Acta - Gene Structure & Expression **1351:**313-24.

29. **Jack, J., and Y. DeLotto.** 1992. Effect of wing scalloping mutations on cut expression and sense organ differentiation in the Drosophila wing margin. Genetics **131:**353-63.

30. **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:**735-47.

31. **Jack, J. W.** 1985. Molecular organization of the cut locus of Drosophila melanogaster. Cell **42:**869-76.

32. **Khanna-Gupta, A., T. Zibello, S. Kolla, E. J. Neufeld, and N. Berliner.** 1997. CCAAT displacement protein (CDP/cut) recognizes a silencer element within the lactoferrin gene promoter. Blood **90:**2784-95.

33. **Kim, E. C., J. S. Lau, S. Rawlings, and A. S. Lee.** 1997. Positive and Negative Regulation Of the Human Thymidine Kinase Promoter Mediated By CCAAT Binding Transcription Factors Nf-Y/Cbf, Dbpa, and Cdp/Cut. Cell Growth & Differentiation **8:**1329-38.

34. **Klemm, J. D., M. A. Rould, R. Aurora, W. Herr, and C. O. Pabo.** 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. Cell **77:**21-32.

35. **Ko, L. J., and J. D. Engel.** 1993. DNA-binding specificities of the GATA transcription factor family. Mol Cell Biol **13:**4011-22.

36. **Last, T. J., M. Birnbaum, A. J. van Wijnen, G. S. Stein, and J. L. Stein.** 1998. Repressor elements in the coding region of the human histone H4 gene interact with the transcription factor CDP/cut. Gene **221:**267-77.

37. **Lee, K. A. W., 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 Techn **5:**22-31.

38. **Li, S., L. Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, and M. J. Walsh.** 1999. Transcriptional repression of the cystic fibrosis transmembrane conductance regulator gene, mediated by CCAAT displacement protein/cut homolog, is associated with histone deacetylation. J Biol Chem **274:**7803-15.

39. **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:**12745-50.

40. **Liu, J., D. Bramblett, Q. Zhu, M. Lozano, R. Kobayashi, S. R. Ross, and J. P. Dudley.** 1997. The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. Mol Cell Biol **17:**5275-87.

41. **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:**133-43.

42. **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:**151-9.

43. **Luo, W., and D. G. Skalnik.** 1996. CCAAT displacement protein competes with multiple transcriptional activators for binding to four sites in the proximal gp91(phox) promoter. J Biol Chem **271:**18203-10.

44. **Mailly, F., G. Berube, R. Harada, P. L. Mao, S. Phillips, and A. Nepveu.** 1996. The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy. Mol Cell Biol **16:**5346-57.

45. **Merika, M., and S. H. Orkin.** 1993. DNA-binding specificity of GATA family transcription factors. Mol Cell Biol **13:**3999-4010.

46. **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. Nature Genet **1:**50-5.

47. **Onate, S. A., P. Prendergast, J. P. Wagner, M. Nissen, R. Reeves, D. E. Pettijohn, and D. P. Edwards.** 1994. The DNA-bending protein HMG-1 enhances progesterone receptor binding to its target DNA sequences. Mol Cell Biol **14:**3376-91.

48. **Pattison, S., D. G. Skalnik, and A. Roman.** 1997. CCAAT displacement protein, a regulator of differentiation-specific gene expression, binds a negative regulatory element within the 5' end of the human papillomavirus type 6 long control region. J Virol **71:**2013-22.

49. **Prendergast, P., S. A. Onate, K. Christensen, and D. P. Edwards.** 1994. Nuclear accessory factors enhance the binding of progesterone receptor to specific target DNA. J Steroid Biochem Mol Biol **48:**1-13.

50. **Sacchetta, P., A. Aceto, T. Bucciarelli, B. Dragani, S. Santarone, N. Allocati, and C. Di Ilio.** 1993. Multiphasic denaturation of glutathione transferase B1-1 by guanidium chloride. Eur J Biochem **215:**741-45.

51. **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:**16736-44.

52. **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 g-globulin gene. EMBO J **7:**3099-107.

53. **Teerawatanasuk, N., D. G. Skalnik, and L. G. Carr.** 1999. CCAAT displacement protein (CDP/cut) binds a negative regulatory element in the human tryptophan hydroxylase gene. J Neurochem **72:**29-39.

54. **Valarche, I., J. P. Tissier-Seta, M. R. Hirsch, S. Martinez, C. Goridis, and J. F. Brunet.** 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:**881-96.

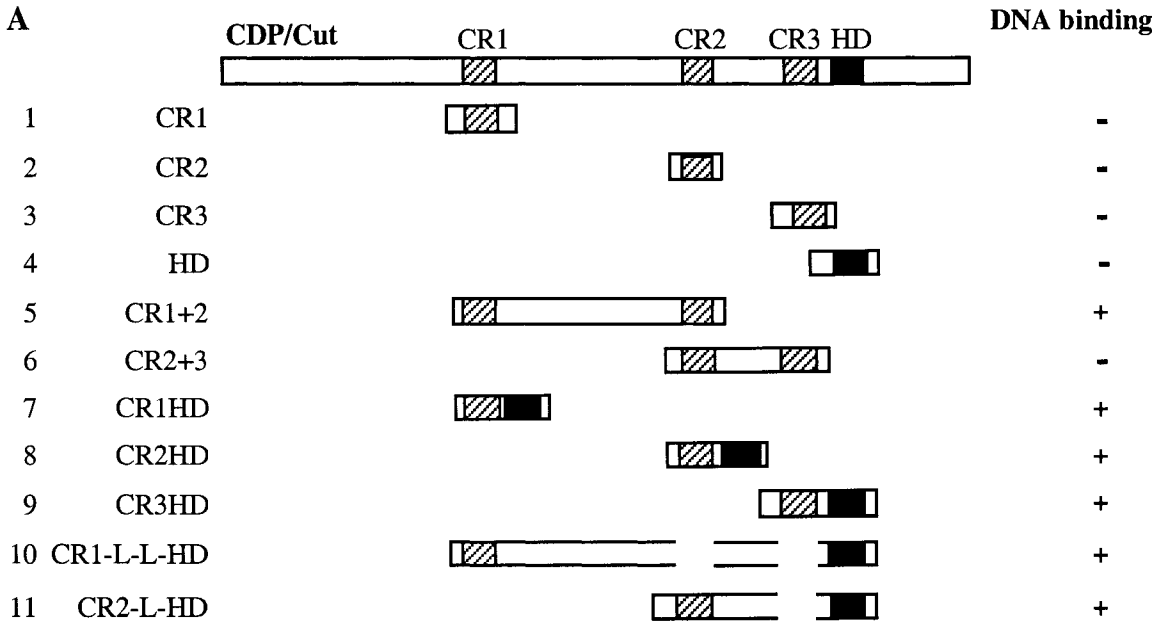
55. **van Wijnen, A. J., C. Cooper, P. Odgren, F. Aziz, A. De Luca, R. A. Shakoori, A. Giordano, P. J. Quesenberry, J. B. Lian, G. S. Stein, and J. L. Stein.** 1997. Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells. J Cell Biochem **66:**512-23.

56. **van Wijnen, A. J., M. F. van Gurp, M. C. de Ridder, C. Tufarelli, T. J. Last, M. Birnbaum, P. S. Vaughan, A. Giordano, W. Krek, E. J. Neufeld, J. L. Stein, and G. S. Stein.** 1996. CDP/cut is the DNA-binding subunit of histone gene transcription factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. Proc Natl Acad Sci USA **93:**11516-21.

57. **Wang, Z., A. Goldstein, R. T. Zong, D. Lin, E. J. Neufeld, R. H. Scheuermann, and P. W. Tucker.** 1999. Cux/CDP homeoprotein is a component of NF-muNR and represses the immunoglobulin heavy chain intronic enhancer by antagonizing the bright transcription activator. Mol Cell Biol **19:**284-95.

58. **Wissmann, A., and W. Hillen.** 1991. DNA contacts probed by modification protection and interference studies. Methods Enzymol **208:**365-79.

A


CDP/Cut
CR1
CR2
CR3
HD
DNA binding
1 CR1 -
2 CR2 -
3 CR3 -
4 HD -
5 CR1+2 +
6 CR2+3 -
7 CR1HD +
8 CR2HD +
9 CR3HD +
10 CR1-L-L-HD +
11 CR2-L-HD +

B

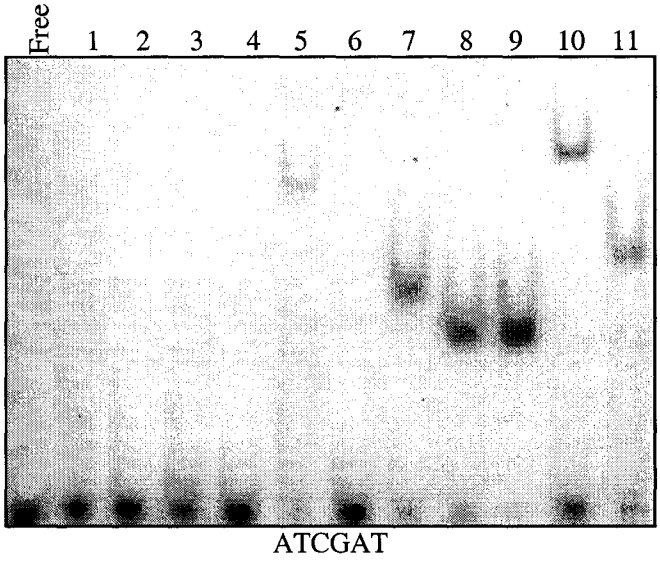

Free
1 2 3 4 5 6 7 8 9 10 11
ATCGAT

Figure 1. Electrophoretic mobility shift assay of CDP/Cut fusion proteins.

(A) Schematic representation of CDP/Cut fusion proteins. Fusion proteins were expressed in bacteria as His-tagged fusion proteins. (B) Radiolabeled 29FP oligonucleotides encoding the CDP/Cut consensus site, ATCGAT, were incubated either alone or with 10 ng of purified fusion proteins at room temperature for 15 min. and resolved on a non denaturing polyacrylamide gel. Note that lane numbers in (B) correspond to the fusion proteins shown in (A).

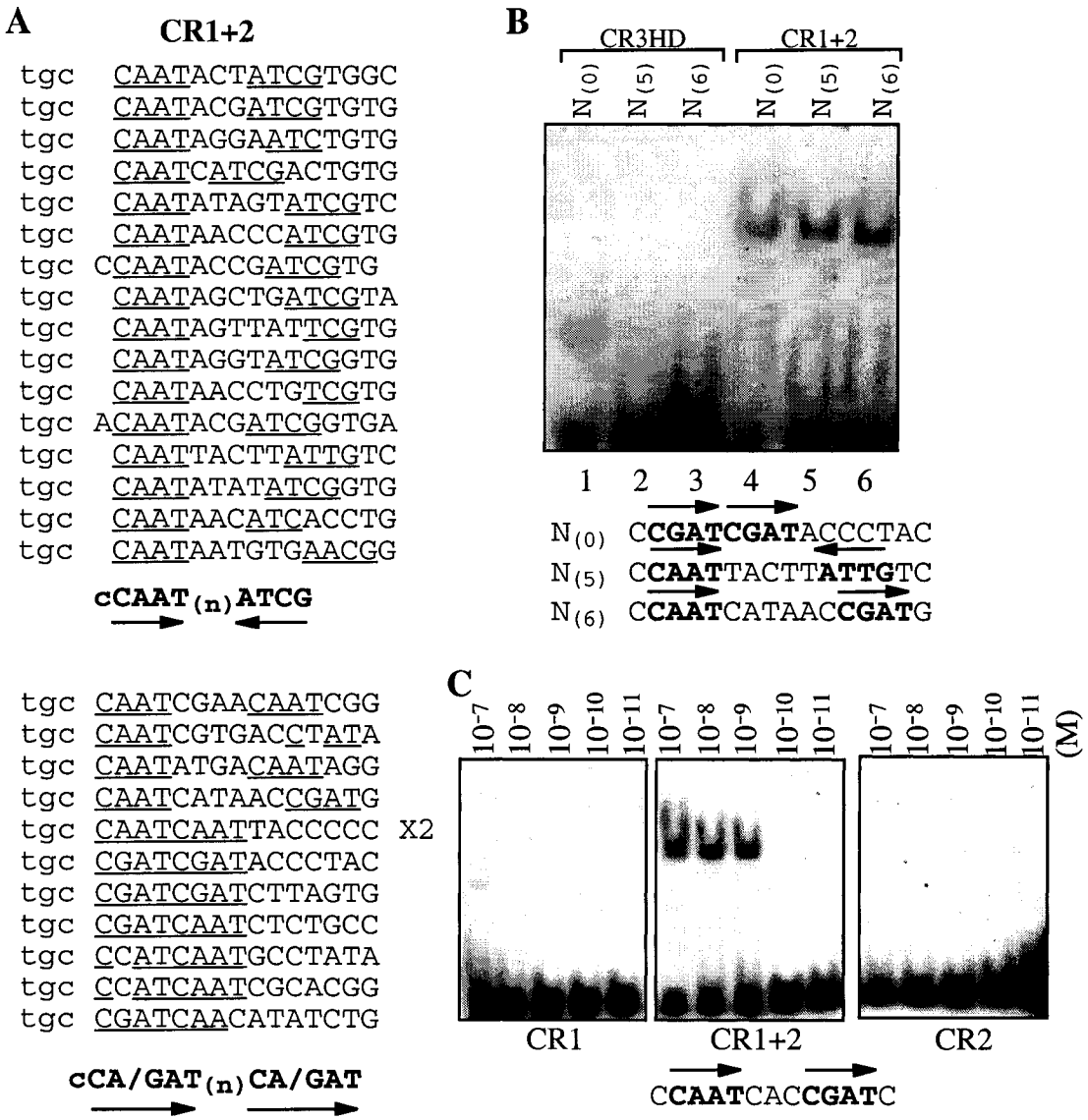

A
CR1+2
tgc CAATACTATCGTGGC
tgc CAATACGATCGTGTG
tgc CAATAGGAATCTGTG
tgc CAATCATCGACTGTG
tgc CAATATAGTATCGTC
tgc CAATAACCCATCGTG
tgc CCAATACCGATCGTG
tgc CAATAGCTGATCGTA
tgc CAATAGTTATTCGTG
tgc CAATAGGTATCGGTG
tgc CAATAACCTGTCGTG
tgc ACAATACGATCGGTGA
tgc CAATTACTTATTGTC
tgc CAATATATATCGGTG
tgc CAATAACATCACCTG
tgc CAATAATGTGAACGG
cCAAT(n)ATCG
B
CR3HD
CR1+2
N(0) N(5) N(6) N(0) N(5) N(6)
1 2 3 4 5 6
N(0) CCGATCGATACCCTAC
N(5) CCAATTACTTATTGTC
N(6) CCAATCATAACCGATG
tgc CAATCGAACAATCGG
tgc CAATCGTGACCTATA
tgc CAATATGACAATAGG
tgc CAATCATAACCGATG
tgc CAATCAATTACCCCC X2
tgc CGATCGATACCCTAC
tgc CGATCGATCTTAGTG
tgc CGATCAATCTCTGCC
tgc CCATCAATGCCTATA
tgc CCATCAATCGCACGG
tgc CGATCAACATATCTG
cCA/GAT(n)CA/GAT
C
10-7 10-8 10-9 10-10 10-11 10-7 10-8 10-9 10-10 10-11 10-7 10-8 10-9 10-10 10-11 (M)
CR1
CR1+2
CR2
CCAATCACCGATC

Figure 2. Compilation of sequences selected by CR1+2.

(A) DNA sequences were selected using purified his-tagged CR1+2 and PCR-generated random oligonucleotides, as described in Material and Methods. After five rounds of selection, selected oligonucleotides were cloned and analyzed by DNA sequencing. The DNA sequences were organized to fit the best possible alignment. On the basis of the frequency of each nucleotide at each position, a core consensus sequence was deduced.

(B) Radiolabeled oligonucleotides encoding sequences selected by CR1+2 were incubated with 100 ng of the indicated fusion protein at room temperature for 15 min. and resolved on a non denaturing polyacrylamide gel.

(C) Radiolabeled oligonucleotides encoding the CR1+2 consensus site (CAAT--CGAT) were incubated with decreasing concentrations of monomers of CR1, CR2 or CR1+2 at room temperature for 15 min. and resolved on a non denaturing polyacrylamide gel.

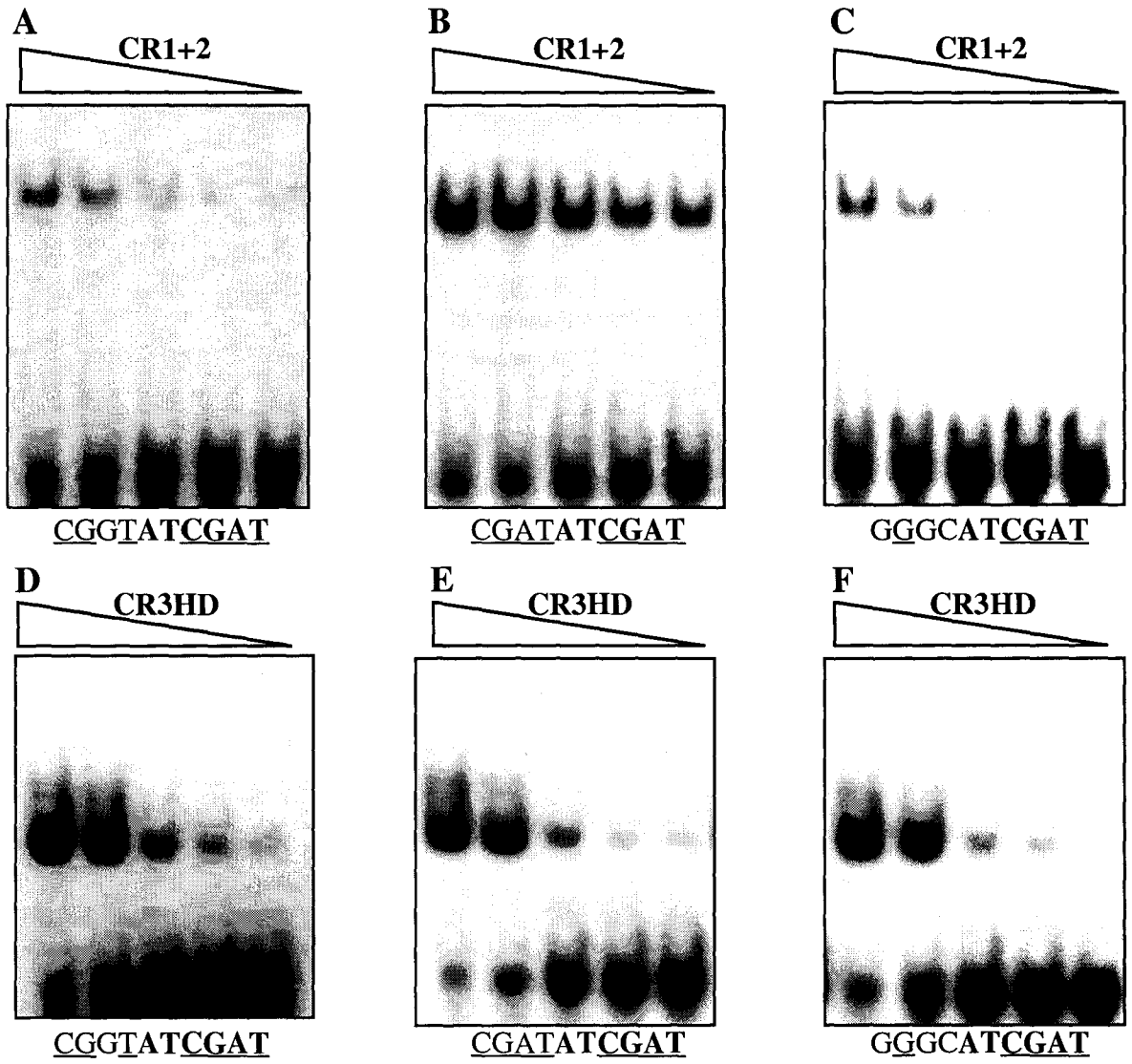

A
CR1+2
CGGTATCGAT
B
CR1+2
CGATATCGAT
C
CR1+2
GGGCATCGAT
D
CR3HD
CGGTATCGAT
E
CR3HD
CGATATCGAT
F
CR3HD
GGGCATCGAT

Figure 3. CR1+2, but not CR3HD, binds better to ATCGAT oligonucleotides that contain an additional CGAT motif.

Radiolabeled oligonucleotides encoding the ATCGAT site with various flanking sequences were incubated either alone or with 10 ng of purified CR1+2 fusion proteins at room temperature for 15 min. and resolved on a non denaturing polyacrylamide gel. On the left, the original 29FP oligonucleotides containing the ATCGAT site also includes the sequence CGGT, which closely resembles the CGAT consensus half-site for CR1+2. In the middle and on the right, point mutations were introduced in the second half-site either to produce a perfect CGAT (middle) or make a degenerate half-site, GGGC (right).

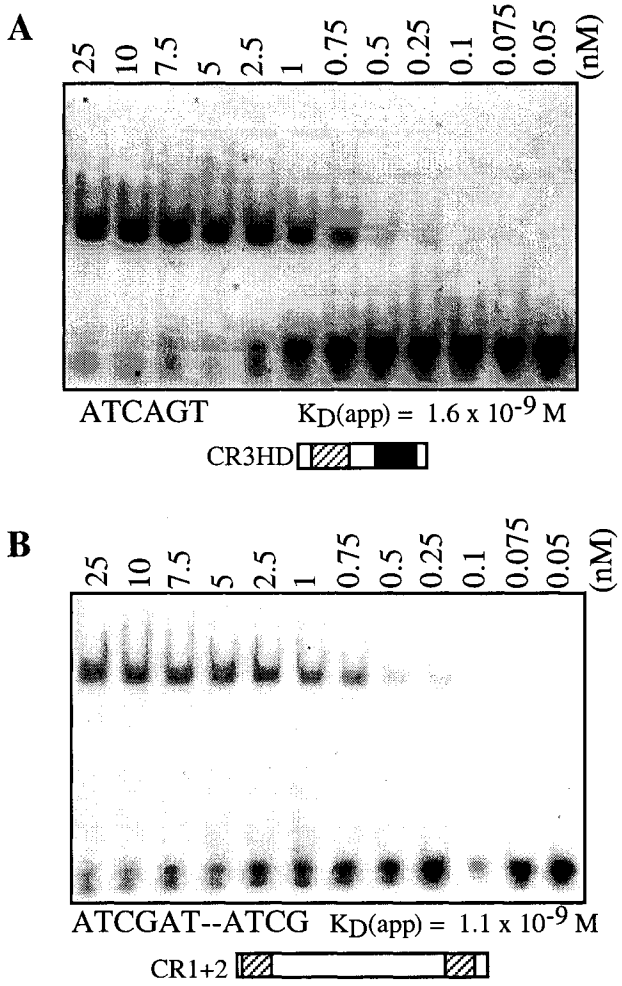

A
25
10
7.5
5
2.5
1
0.75
0.5
0.25
0.1
0.075
0.05
(nM)
ATCAGT
$K_{D(app)}$ = 1.6 x 10^{-9} M
CR3HD
B
25
10
7.5
5
2.5
1
0.75
0.5
0.25
0.1
0.075
0.05
(nM)
ATCGAT--ATCG
$K_{D(app)}$ = 1.1 x 10^{-9} M
CR1+2

Figure 4. Apparent dissociation constants (K_D) of CR1+2 and CR3HD for their respective binding sites.

(A) EMSA: ten pM of radiolabeled oligonucleotides encoding either the CR3HD or the CR1+2 consensus binding sites was incubated with decreasing concentrations of CR3HD or CR1+2 fusion proteins and analyzed in EMSA as in Fig. 1. The percentage of free DNA (relative to the amount in the lane with no protein added) was plotted against the log of protein concentration to obtain the apparent dissociation constant, $K_{D(app)}$.

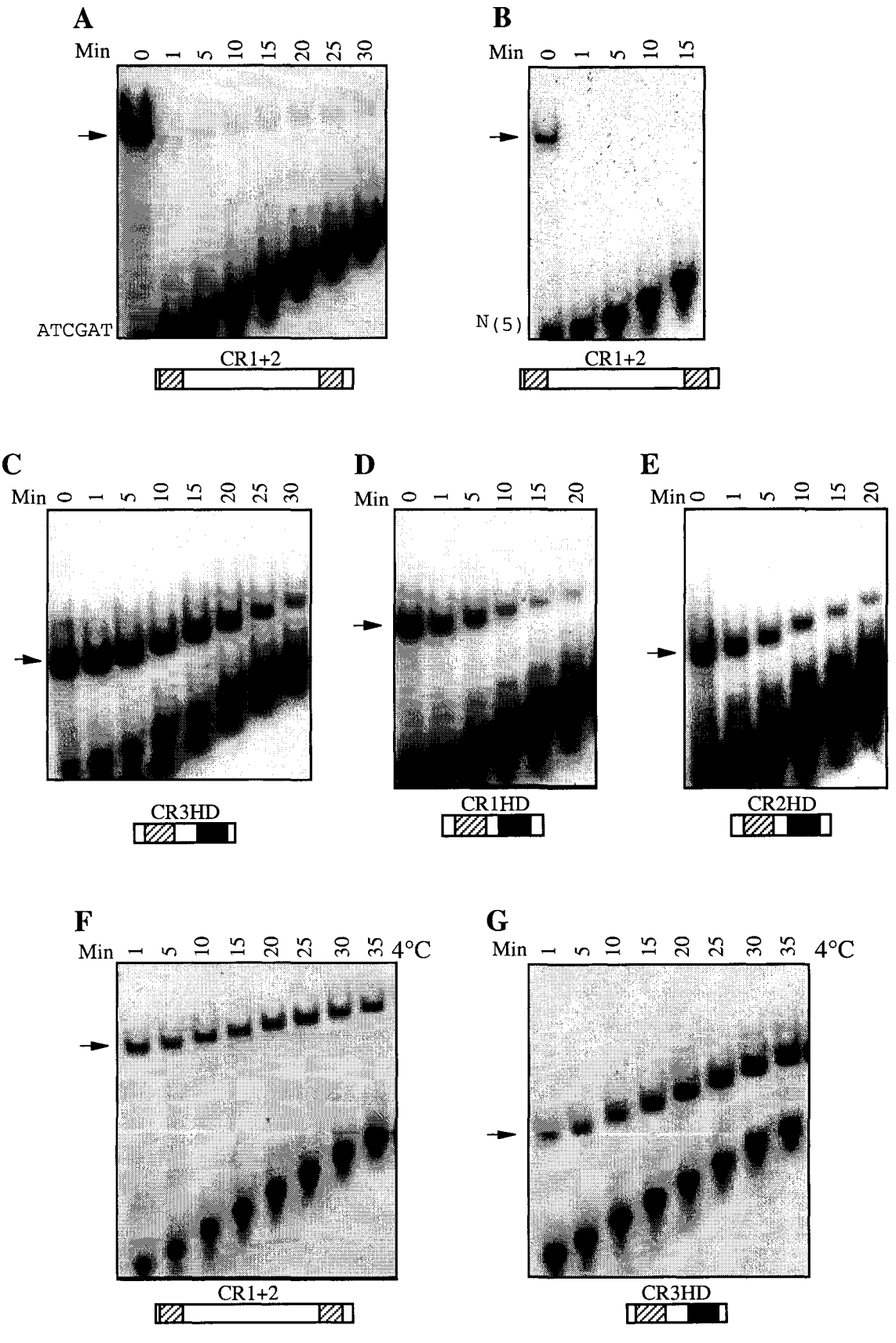

A
Min 0 1 5 10 15 20 25 30
ATCGAT
CR1+2
B
Min 0 1 5 10 15
N(5)
CR1+2
C
Min 0 1 5 10 15 20 25 30
CR3HD
D
Min 0 1 5 10 15 20
CR1HD
E
Min 0 1 5 10 15 20
CR2HD
F
Min 1 5 10 15 20 25 30 35 4°C
CR1+2
G
Min 1 5 10 15 20 25 30 35 4°C
CR3HD

Figure 5. Off and on rates of CR1+2, CR1HD, CR2HD and CR3HD.

(A-E) 100 ng of the indicated fusion protein was incubated with radiolabeled oligonucleotides at room temperature until the equilibrium was reached (15 min.). 1000 fold molar excess of unlabeled oligonucleotides was added, and at the indicated time points aliquots of the mixture were taken and analyzed in EMSA as in Fig. 1. Oligonucleotides encoding the ATCGAT site were used in panels A, C, D and E, while the N(5) site, CAAT--ATTG (see Fig. 2B), was used in panel B.

(F-G) 100 ng of either CR1+2 (A) or CR3HD (B) fusion protein was mixed with radiolabeled oligonucleotides encoding the ATCGAT site. The incubation took place at 4°C and at the indicated times aliquots of the mixture were taken and analyzed in EMSA as in Fig. 1.

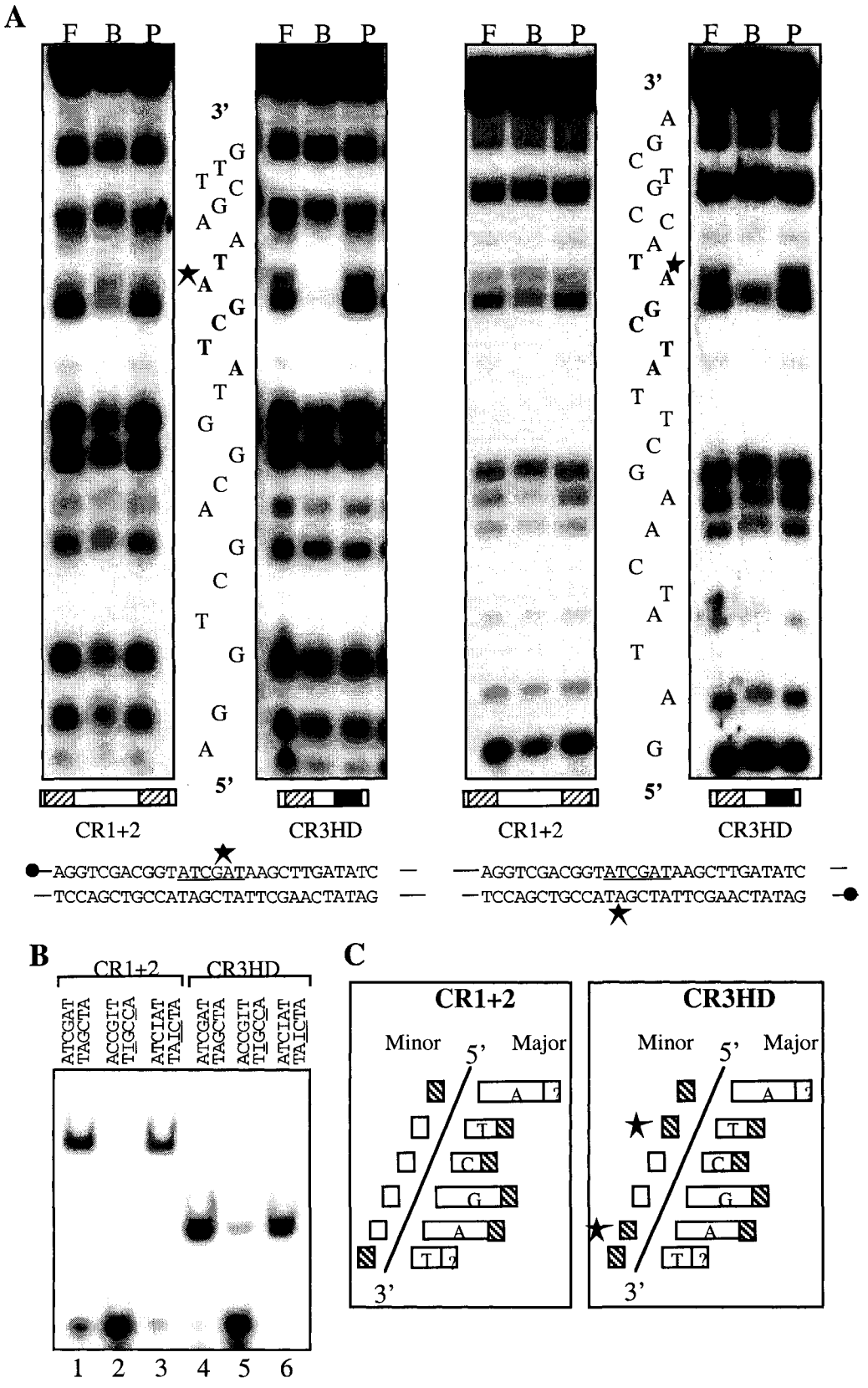

A
F B P
F B P
F B P
F B P
3'
5'
CR1+2
CR3HD
CR1+2
CR3HD
AGGTCGACGGTATCGATAAGCTTGATATC
TCCAGCTGCCATAGCTATTCGAACTATAG
B
CR1+2
CR3HD
ATCGAT TAGCTA
ACCGIT TIGCCA
ATCIAT TAICTA
ATCGAT TAGCTA
ACCGIT TIGCCA
ATCIAT TAICTA
1 2 3 4 5 6
C
CR1+2
CR3HD
Minor
5'
Major
3'

Figure 6. Methylation interference assay of CR1+2 and CR3HD.

(A) A DNA fragment containing the ATCGAT site was end labeled within the upper or lower strand, partially methylated and used in EMSA with either CR1+2 or CR3HD fusion proteins. The free and bound DNA molecules were separated by PAGE and purified. Free DNA (F), bound DNA (B) and the probe before EMSA (P) were digested with NaOH and resolved on a 12% denaturing polyacrylamide gel.

(B) The ATCGAT binding site was mutated as shown and the mutated oligonucleotides were tested in EMSA with CR1+2 (lanes 1, 2, and 3) and CR3HD (lanes 4,5, and 6).

(C) Diagram showing the contact points made by CR1+2 and CR3HD within the major and minor grooves of the ATCGAT core sequence. Note that, in contrast to CR1+2, CR3HD makes contact within the minor groove at the second and fifth position.

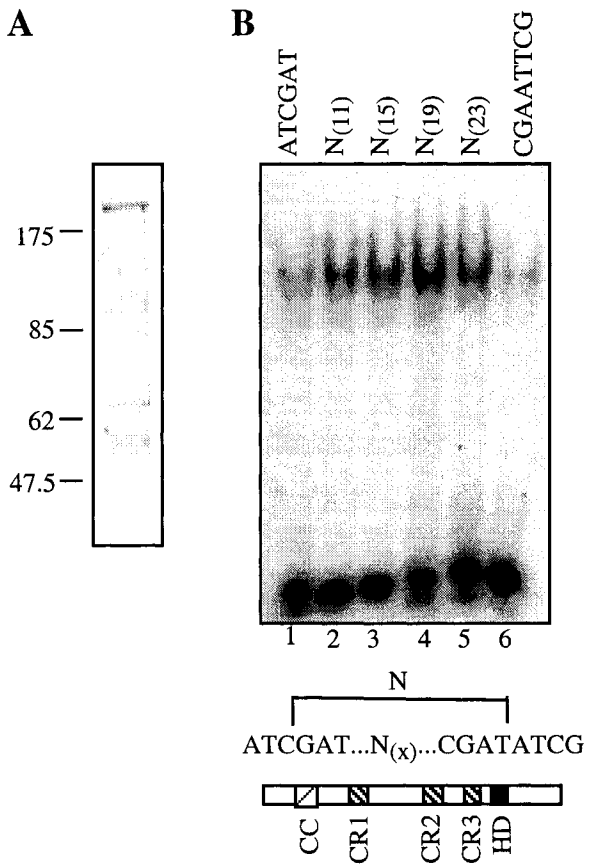


Figure 7. Cooperative DNA binding by the full length CDP/Cut protein.

(A) The his-tagged full length CDP/Cut protein was expressed and purified from SF9 insect cells. The purified protein was analyzed on a 8% polyacrylamide gel and stained with Coomassie blue. (B) Following treatment with calf intestinal phosphatase, 10 ng of purified CDP/Cut was tested in EMSA with radiolabeled oligonucleotides encoding a single binding site, either ATCGAT (lane 1) or CGATATCG (lane 6), or these two binding sites separated by varying distance (lanes 2 to 5). The number of base pairs from the middle of one binding site to the middle of the second site is indicated at the top of each lane.

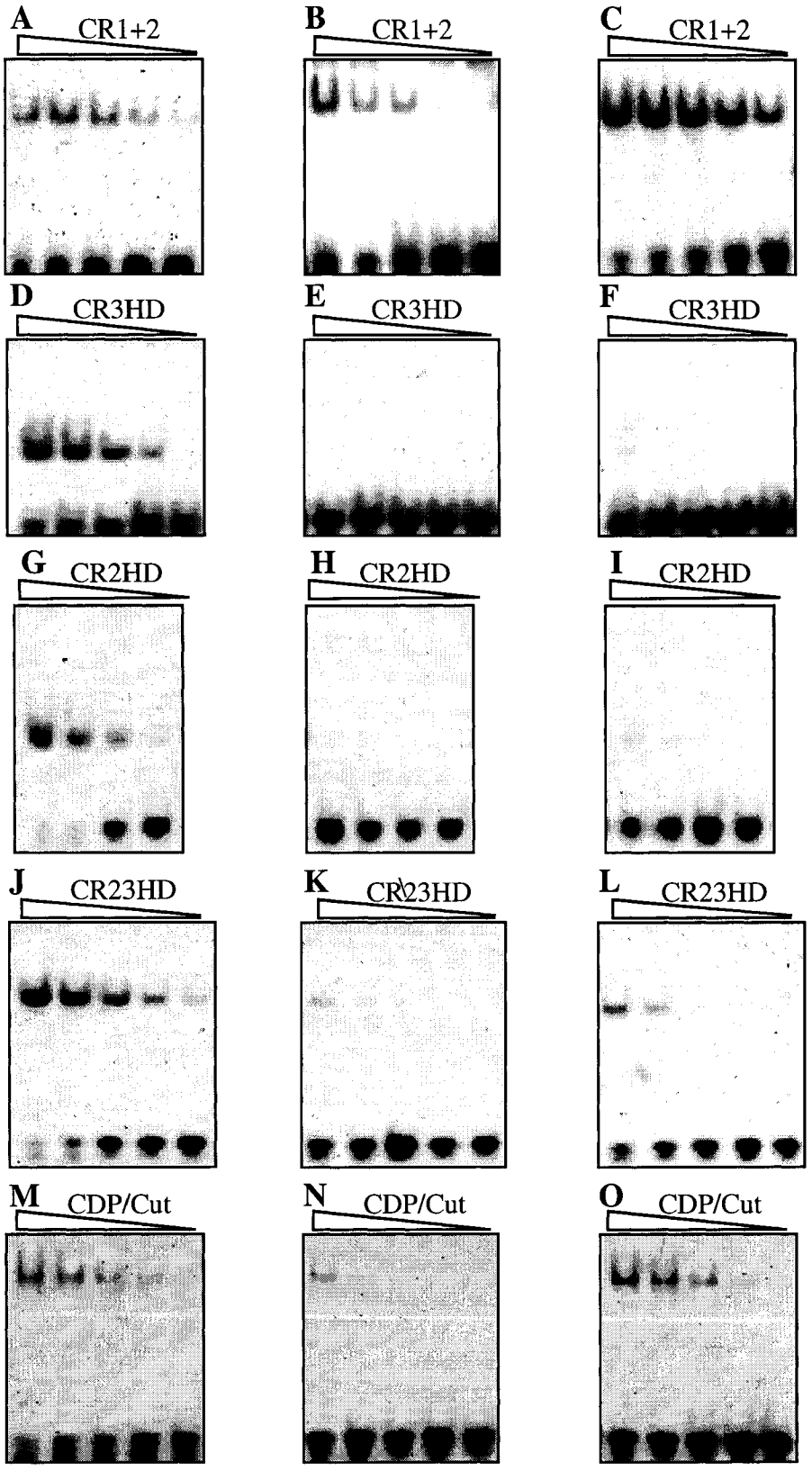

A CR1+2
B CR1+2
C CR1+2
D CR3HD
E CR3HD
F CR3HD
G CR2HD
H CR2HD
I CR2HD
J CR23HD
K CR23HD
L CR23HD
M CDP/Cut
N CDP/Cut
O CDP/Cut
ATCGAT
GACGGTATCGATAAGC
NF-Y
CAACCAATCACCCGTC
NF-Y--CGAT
CAACCAATCACCGATC

Figure 8. The NF-Y (CCAAT) binding site is recognized by CR1+2, but not CR3HD, CR2HD nor CR23HD.

Radiolabeled oligonucleotides encoding the ATCGAT site (29FP) or the NF-Y consensus binding site (CCAAT) flanked or not with a CGAT site were incubated with decreasing concentrations of various CDP/Cut fusion proteins at room temperature for 15 min. and resolved on a non denaturing polyacrylamide gel.

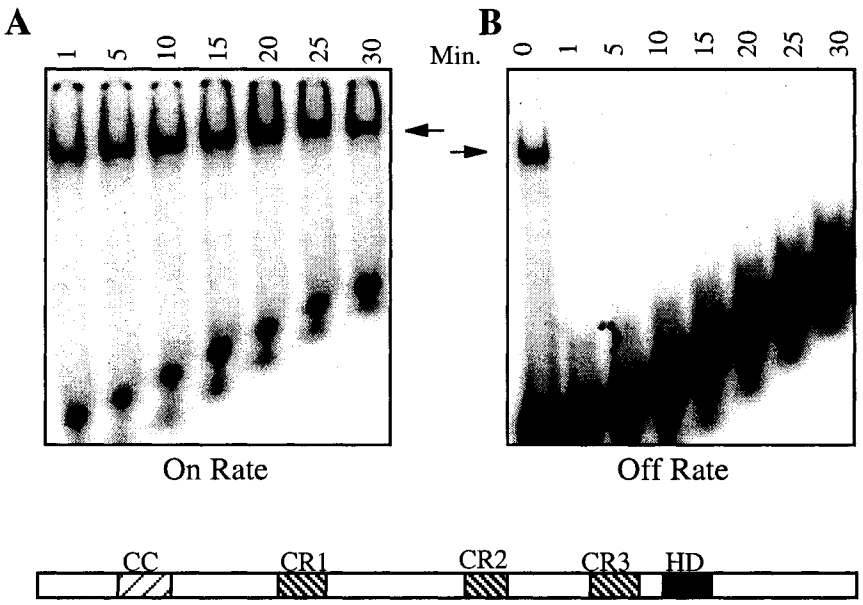
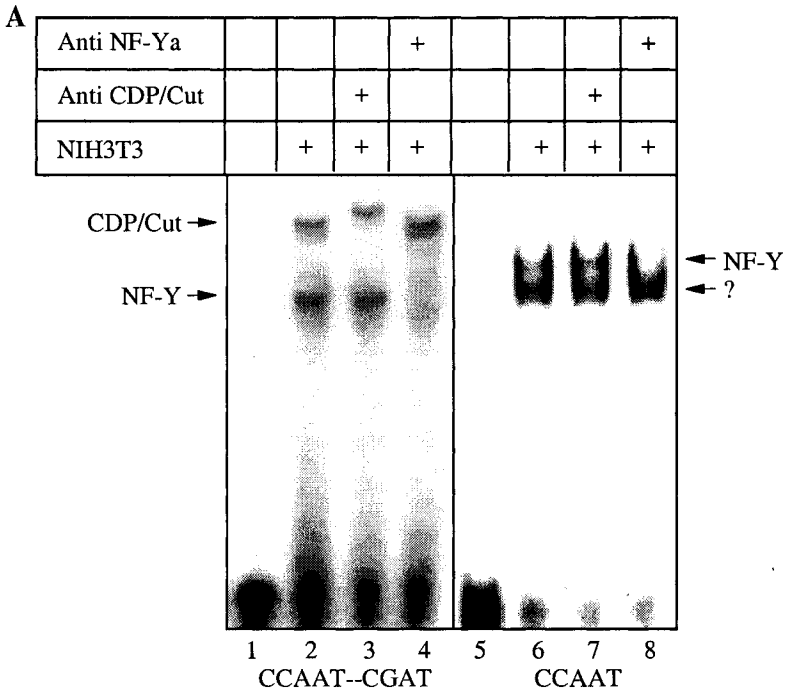


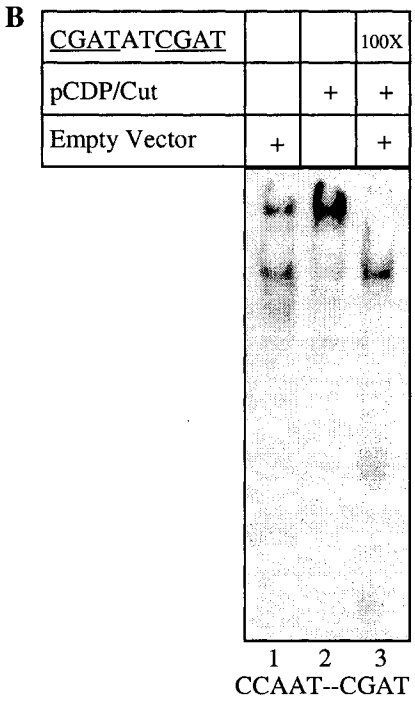
Figure 9. On and off rates of full length CDP/Cut.

100 ng of dephosphorylated full length CDP/Cut was incubated with radiolabeled 29FP oligonucleotides encoding the ATCGAT site.

(A) The incubation took place at 4°C and at the indicated times aliquots of the mixture were taken and analyzed in EMSA as in Fig. 1.

(B) The incubation took place at room temperature until the equilibrium was reached (15 min.). 1000 fold molar excess of unlabeled oligonucleotides was added, and at the indicated time points aliquots of the mixture were taken and analyzed in EMSA as in Fig. 1.


A
Anti NF-Ya
Anti CDP/Cut
NIH3T3
CDP/Cut
NF-Y
NF-Y
?
1 2 3 4 5 6 7 8
CCAAT--CGAT
CCAAT


B
CGATATCGAT
pCDP/Cut
Empty Vector
100X
1 2 3
CCAAT--CGAT

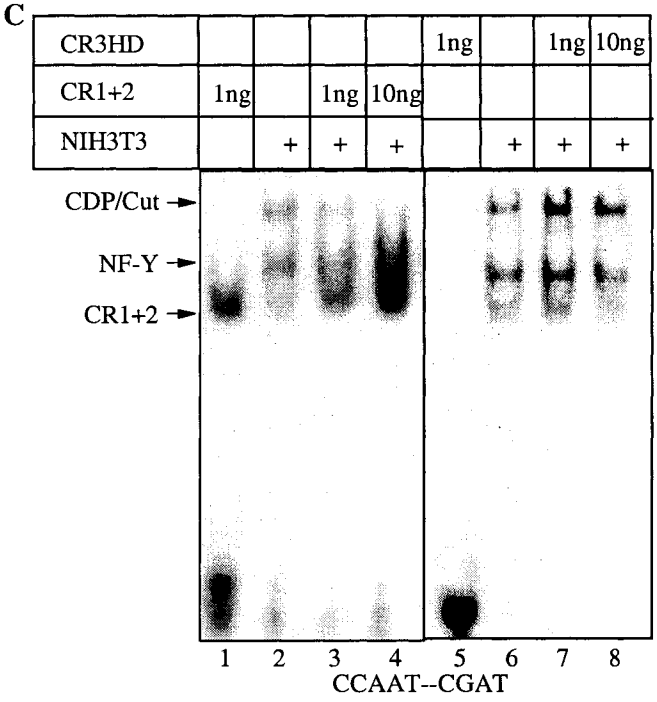

C
CR3HD
CR1+2
NIH3T3
1ng
1ng
10ng
1ng
1ng
10ng
CDP/Cut
NF-Y
CR1+2
1 2 3 4 5 6 7 8
CCAAT--CGAT

Fig. 10 CR1+2 is responsible for the CCAAT-displacement activity of CDP/Cut.

(A) Nuclear extracts from NIH3T3 cells were tested in EMSA with oligonucleotides encoding the NF-Y consensus binding site, CCAAT, either alone or flanked with a CGAT motif. The DNA and proteins were incubated with either no antibody (lanes 1, 2, 5 and 6), anti-CDP/Cut antibodies (lanes 3 and 7) or anti-NF-Ya antibodies (lanes 4 and 8). The positions of known retarded complexes are indicated by the arrows.

(B) NIH3T3 cells were transfected with a vector encoding either nothing or the CDP/Cut protein. Whole cell extracts were prepared 48 hours later and tested in EMSA using oligonucleotides encoding both the CCAAT and CGAT motifs. In lane 3, the incubation took place in the presence of 100-fold excess of oligonucleotides encoding both ATCGAT and CGAT motifs.

(C) EMSA were performed as in (A) except that purified fusion proteins were added to the reaction as indicated.

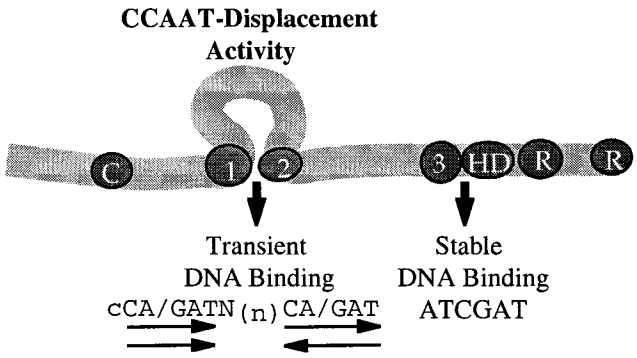


Figure 11. Modes of DNA binding by CDP/Cut.

Schematic diagrams of the two modes of DNA binding by CDP/Cut. The CCAAT-displacement activity involves CR1+2, whereas stable DNA binding to ATCGAT involves CR3HD.

Table 1
Apparent KD of CDP/Cut DNA binding domains for various DNA sequences

	29FP	N19	NF-Y+CGAT	NF-Y
	CGGT-ATCGAT	ATCGAT-CGAT-CGAT	CCAAT-CGAT	CCAAT
CR1+2		0.9×10^{-9}	1.1×10^{-9}	6.7×10^{-9}
CR3HD	1.6×10^{-9}	1.8×10^{-9}		
CDP/Cut	1.6×10^{-9}	0.5×10^{-9}	1.9×10^{-9}	8.2×10^{-9}

Table 2

CDP Binding Sites Reported in the Literature	
GP91phox	CAGTTCCAATGATTATTAGCCAATTTCTGATAAAAGA
NCAM	AAAATCGAACCGAATCTAAAATT
hTPH	CAGGTCATTGTGTCGATAATAGGCGTTATCT
H4	GTTTTCAATCTGGTCCGATACTCTT
H2B-1 (Sea Urchin)	TGATTGGTAGATTATCATTCGCGCTGAT
TK	CCGGGCGCTGATTGGCCCCATGGCGGC CGGCTCGTGATTGGCCAGCA
H2a H2B (Xenopus laevis)	GGGTGCGAGGAAGCCTAACCAATCGGCA GACACTGATTCCCATTGGCTGGG
γ-Globin	CTTGACCAATAGCCTTGACAAGGCAAACTTGACCAATA G

Chapter III S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by the CDP/Cux Homeodomain Protein.

Preface

Since full length CDP/Cux protein contains four DNA binding domains, CR1, CR2, CR3, and HD, it is expected to display the DNA binding characteristic of both CR1+2 and CR3HD. In the previous chapter, full length CDP/Cux was found to posses DNA characteristic similar to that of CR1+2 *in vitro*. However, the off-rate experiment indicated that DNA binding activity of CR3HD, stable binding to DNA, is near inactive. This was surprising finding considering the fact that strong DNA binding to ATCGAT, a consensus binding site of CR3HD, by CDP/Cux is readily detected from a cellular extract. Interestingly, the previous study also demonstrated that this stable DNA binding activity of CR3HD is regulated in a cell cycle dependent manner and strongest in S-phase. In this manuscript, one of the key steps to expose the stable DNA binding activity by CDP/Cux in a cell cycle dependent manner is discussed.

ABSTRACT

The CCAAT-displacement protein (CDP), the homologue of the *Drosophila melanogaster* Cut protein, contains four DNA binding domains that function in pairs. Cooperation between the Cut repeat 3 and the Cut homeodomain (CR3HD) allows stable DNA binding to the ATCGAT motif, an activity previously shown to be upregulated in S phase. Here, we showed that the full length CDP/Cut protein is incapable of stable DNA binding, and that the ATCGAT binding activity present in cells involves a 110 kDa carboxy-terminal peptide of CDP/Cut. A vector expressing CDP/Cut with Myc and HA epitope tags at either end generated N- and C-terminal products of 90 and 110 kDa, suggesting that proteolytic cleavage was involved. *In vivo* pulse/chase labeling experiments confirmed that the 110-kDa protein was derived from the full length CDP/Cut protein. Proteolytic processing was weak or not detectable in G0 and G1, but increased in populations of cells enriched in S phase, and the appearance of the 110-kDa protein coincided with the increase in ATCGAT DNA binding. Interestingly, the amino-truncated and the full-length CDP/Cut isoforms exhibited different transcriptional properties in a reporter assay. We conclude that proteolytic processing of CDP/Cut at the G1/S transition generates a CDP/Cut isoform with distinct DNA binding and transcriptional activities. These findings, together with the cleavage of the Scc1 protein at mitosis, suggest that site-specific proteolysis may play an important role in the regulation of cell cycle progression.

INTRODUCTION

Genetic studies in *Drosophila* indicated that *cut* plays an important role in determining cell-type specificity in several tissues (reviewed in (35). Defects caused by *Cut* mutations appear to result from the fact that some cells have enrolled in the wrong developmental program (6, 8-10, 12, 13, 24, 29-31). In higher vertebrates, Cut proteins were originally characterized as CCAAT-displacement proteins (CDP), and have been termed CDP (CCAAT displacement protein), Clox (Cut-like homeobox), Cux-1 (Cut homeobox), and CDP-2 (2, 36, 43, 50).

CDP/Cut proteins are unique in that they contain four DNA binding domains: the Cut homeodomain (HD) and the three Cut repeats (CR1, CR2 and CR3), three regions of ~70 amino acids that share from 52 to 63% amino-acid identity with each other (1, 3, 7, 20, 21, 36). Using a panel of fusion proteins, we have shown that a Cut repeat cannot bind to DNA as a monomer, but that certain combinations of domains exhibit high DNA binding affinity: CR1+2, CR3HD, CR1HD and CR2HD (33). CR1+2 displayed rapid on and off-rates, and bound preferably to two CAAT or CGAT sites, organized as direct or inverted repeats. Accordingly, only CR1+2 was able to bind to the CCAAT sequence, and its affinity was increased by the presence of a CA/GAT site at close proximity. Moreover, CR1+2, but not CR3HD, was able to displace the NF-Y factor. Thus, the CCAAT-displacement activity involves Cut repeats 1 and 2, but not the Cut homeodomain. In contrast, various combinations involving the Cut homeodomain and one Cut repeat were found to make a stable interaction with ATNNAT motifs. In particular, CR3HD stably bound to DNA, preferably with the ATCGAT sequence. CDP/Cut proteins were generally found to function as transcriptional repressors (2, 17, 27, 28, 32, 38, 40, 43). In addition,

CDP/Cut may also be able to participate in gene activation on specific promoters (45, 50). Indeed, CDP/Cut was found to be a component of the promoter complex HiNF-D, which is believed to contribute to the transcriptional induction of several histone genes at the G1/S phase transition of the cell cycle (5, 44, 45). In this instance, transcriptional activation could not be demonstrated in co-transfection assays, but it was proposed that the regulatory effect of CDP/Cut on transcription might vary depending on the proteins with which it interacts (45). In particular, supershift assays using antibodies against pRb-related proteins suggested that CDP/Cut might alternatively interact with pRb or p107 on different promoters (45).

We have recently obtained evidence that CDP/Cut may play an important role in cell cycle progression (16). Using the ATCGAT site as a probe, little CDP/Cut DNA binding was detected in G0 and early G1, unless cell extracts were previously treated with alkaline phosphatase. In contrast, strong DNA binding was observed in S phase. This was shown to result, at least in part, from dephosphorylation of the Cut homeodomain by the Cdc25A phosphatase. Thus, apart for cdk2, CDP/Cut is the only known substrate for Cdc25A, a phosphatase that is required for G1/S transition, and whose overexpression triggers S phase entrance (11, 25, 46). The increase in CDP/Cut activity coincided with a decrease in $p21^{WAF1/CIP1/SDI1}$ mRNAs. In co-transfection experiments, CDP/Cut repressed a reporter controlled by the p21 promoter, whereas an antisense CDP/Cut vector restored p21 expression in S phase. Moreover, p21 expression was repressed equally well by either Cdc25A or CDP/Cut. Altogether, these results led us to propose a model whereby Cdc25A activates the CDP/Cut repressor that subsequently down-regulates transcription of p21 in S phase.

Site-specific proteolysis has emerged as an important regulatory mechanism that plays a role in a number of cellular processes including transduction within the Notch and Hedgehog signaling pathways, sister-chromatid separation at anaphase, and generation of the amyloidogenic peptide in Alzheimer's disease (4, 15, 19, 42). In the field of transcription, proteolytic processing of transcription factors was shown to redirect the localization of these proteins or to generate specific isoforms with distinct biochemical properties (4, 14, 18, 22, 47, 49).

The present study was triggered by a surprising result. We found that a purified CDP/Cut protein exhibited DNA binding properties similar to that of CR1CR2, suggesting that CR3HD was weakly or not active in the context of the full-length protein. Moreover, in mammalian cells the ATCGAT binding activity that is upregulated in S phase could be supershifted with C-terminal, but not with N-terminal-specific anti-CDP/Cut antibodies. Further investigation led to the discovery that the full length CDP/Cut protein is proteolytically cleaved to generate N-terminal and C-terminal peptides of 90 and 110 kDa, and that processing is regulated during cell cycle progression. Interestingly, the full-length and 110 kDa CDP/Cut isoforms displayed different effects in a reporter assay suggesting that processing of CDP/Cut serves to generate an isoform with distinct transcriptional properties.

MATERIALS AND METHODS

Plasmid Construction. For expression of the full length CDP/Cut protein in SF9 insect cells, nt 27-5101 from the human CDP/Cut cDNA (GenBank, accession No. M74099) were inserted into the pBlueBac His2b (InVitrogen)^{Tm}. The resulting plasmid was cotransfected with a helper plasmid to obtain Baculoviruses expressing CDP/Cut with a histidine tag at the amino-terminus. The DNA polymerase alpha/luciferase reporter construct was prepared as follows. PCR-amplification was performed to obtain a fragment of genomic DNA containing nt 56 to 1657 of the human DNA polymerase alpha gene 5' end (GenBank # M64481). The primer at the 5' and 3' ends included EcoRI and BamH1 sites, respectively, to allow cloning into the pBluescript KS vector (Stratagene). An EcoRV-NcoI fragment, including nt 56 to 1621 of the DNA polymerase alpha gene, was then subcloned into the corresponding sites of the luciferase reporter vector, pGL3 (Promega). The resulting plasmid contains sequences from -1517 to +49 relative to the transcription initiation site of the DNA polymerase alpha gene.

Expression and purification of CDP/Cut fusion proteins. The his-tagged CR1CR2, CR3HD and full-length CDP/Cut have been described previously (33). ET-15b derived vector were introduced into the BL21(DE3) of *E.coli* and induced with IPTG. SF9 insect cells were infected with Baculovirus encoding his-CDP/Cut and incubated for 3 days. The fusion proteins were purified by affinity chromatography using procedures provided by the suppliers.

Cell culture and synchronization.

NIH3T3, HeLa cells and HS578T cells were grown in DMEM medium supplemented with 10% fetal bovine serum (FBS) and 293 cells, in DMEM with 10% horse serum. To obtain cells in G0 using the serum starvation/stimulation method, NIH3T3 cells were maintained for 3 days in DMEM plus 0.4% FBS (Serum 0h). The medium was then changed for DMEM plus 10% FBS and cells were harvested either, 3 hours later to obtain cells in early G1 (Serum 3h), 10 hours for mid-G1 or 18 hours of S phase. Synchronization in G1/S was performed using the double thymidine procedure (39). Cells were cultured overnight in DMEM plus 10% FBS supplemented with 2 mM thymidine, washed the next day, cultured for 10 hours in DMEM plus 10% FBS and finally further incubated overnight in the presence of 2 mM thymidine (Thymidine 0h). To allow cells to progress in the cell cycle, the medium was replaced with DMEM plus 10% FBS and cells were harvested three hours later (Thymidine 3h). FACS analysis was performed as previously described (16)

Preparation of nuclear extracts. Nuclear extracts were prepared according to the procedure of Lee et al., except that nuclei were obtained by submitting cells to 3 freeze/thaw cycles in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT) (26). Nuclei were then resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM $MgCl_2$, 420 mM $NaCl_2$, 0.2 mM EDTA.) and incubated at 4 °C for 30 min. After 15 min. of centrifugation, the supernatant was collected. Buffers A and C were supplemented with protease inhibitor mix tablet purchased from Roche. Total extracts were prepared by applying buffer X (50 mM Hepes, pH 7.9, 0.4 M KCl, 4 mM NaF, 4 mM Na_3VO_4, 0.2 mM EGTA, 0.2 mM EDTA, 0.1% NP-40, 10 % glycerol, 0.5 mM DTT, Protease inhibitor mix

tablet from Roche) to a monolayer plate. After 10 min. incubation on ice, the resulting slurry was centrifuged for 15 min at 4 °C and the supernatant was collected.

Electrophoretic mobility shift assay (EMSA). EMSA were performed with 2 to 5 µg of nuclear extract from mammalian cells. The samples were incubated at room temperature for 5 min. in a final volume of 30 µl of 25 mM NaCl, 10 mM Tris, pH 7.5, 1 mM $MgCl_2$, 5 mM EDTA, pH 8.0, 5 % of glycerol, 1 mM of DTT, with 30 ng of poly dIdC and 30 µg of BSA as nonspecific competitors. End-labeled double stranded oligonucleotides (~10 pg) were added and further incubated for 15 min. at room temperature. Samples were loaded on a 5% or 4% polyacrylamide gel (30:1) and separated by electrophoresis at 8V/cm in 0.5 x Tris borate-EDTA. Gels were dried and visualized by autoradiography.

Oligonucleotides. The sequences of oligonucleotides used in this study are as follows: Universal CDP/Cut consensus binding site, CGATATCGAT, TCGAGACGAT**ATCGAT**AAGCTTCTTTTC. ATCGAT, TCGAGACGGTATCGATAAGCTTCTTTTC.

CDP/Cut antibodies and Western blot analysis. To generate polyclonal antibodies against various regions of CDP/Cut (1505 a.a.), rabbits were injected with 500 µg of purified bacterial fusion protein containing various regions of CDP/Cut in Freund's complete adjuvant. The αN was raised against amino acids (a.a.) 227 to 1003; α861, a.a. 861 to 936. The animals were boosted twice with 250 µg of protein, and serum was collected 10 days after the last boost. Polyclonal antibodies were purified by affinity chromatography. The

serum was passed through two GST affinity columns and the flow-through was then applied to a GST-CDP/Cut affinity column to isolate antibodies against CDP/Cut. To generate N-term antibodies, the N serum was immunodepleted by passing through a column containing GST-CDP/Cut 861-936 and GST-CDP/Cut 940-1036. For Western blot analysis, protein extracts were recovered as described above and separated by electrophoresis on 8% (Fig. 2) or 6% polyacrylamide gels. Western blot analysis with αCut 861 and αCut N-term was performed as previously described (16). Western blot analysis for Myc and HA epitopes were done with minor modification. For the primary antibody incubation, αMyc or αHA antibodies were incubated with the membrane in TBST (10 mM Tris pH8, 150 mM NaCl , 0.1% Tween) for 1 hour at room temperature. After four 10 min. washes with TBST, secondary antibodies were added to membrane in TBST and incubated for 45 min. at room temperature. Following four 10 min. washes with TBST, proteins were visualized with the ECL system from Amersham.

DNA affinity chromatography. A 5' biotinylated oligonucleotide containing an ATCGAT site was annealed with its unlabeled complementary oligonucleotide and bound to streptavidine agarose beads for one hour in EMSA reaction buffer. 100 μg of HeLa nuclear extract were precleared with streptavidine agarose beads for one hour and added to streptavidine-biotin-DNA beads and incubated for one hour in EMSA buffer. The DNA-protein-beads complex was washed three times for 5 min. with EMSA buffer at 4C°, resolved by SDS-PAGE and visualized by western blot analysis.

Luciferase assays

HS578T cells were plated at $1x10^5$ cells per 22.1 well. The next day, a total amount of 1.5 μg of DNA was transfected including 0.5 μg of reporter DNA and increasing amount of effector DNA using ExGen500 (MBI, Fermentas). Each transfection was carried in 6 separate wells. Cells were harvested 40 hours later: cells from 3 wells served to prepare separate cytoplasmic extracts for luciferase assay, and cells from 3 other wells serve to prepare total extracts for Western blot analysis. Luciferase assays were performed as previously described with minor modifications (16). Because the internal control plasmid is itself often repressed by CDP/Cut, as a control for transfection efficiency the purified β-galactosidase protein (Sigma) was included in the transfection mix, as previously described (23). The luciferase activity was then normalized based on β-galactosidase activity.

RESULTS

The full-length CDP/Cut protein exhibits DNA binding kinetics similar to that of CR1CR2. The CDP/Cut protein contains four DNA binding domains: CR1, CR2, CR3 and HD. We have previously shown that two DNA binding domains are required for efficient DNA binding, and that the two most efficient combinations, CR1CR2 and CR3HD, exhibit different DNA binding specificities and kinetics (33). Whereas CR1CR2 only transiently bound to sequences containing direct or inverted repeats of the CGAT or CAAT motifs, CR3HD made a stable interaction with the ATCGAT sequence. To analyze DNA binding by the full length CDP/Cut protein, we designed oligonucleotides that can be recognized both by CR1CR2 and CR3HD. Oligonucleotides containing the CGATATCGAT sequence were recognized efficiently by CR1CR2 since two CGAT direct repeats are present (CGATATCGAT), and by CR3HD since the ATCGAT sequence is present (CGATATCGAT) (Fig. 1A). These oligonucleotides thus can serve as universal CDP/Cut consensus binding sites.

To investigate the DNA binding properties of the full length CDP/Cut protein, the protein was purified as a histidine-tagged fusion protein using a baculovirus expression system. When incubated with the CGATATCGAT probe, CDP/Cut generated one retarded complex (Fig. 1A). We next investigated the DNA binding kinetics of the full length CDP/Cut protein. Like CR1CR2, but in contrast to CR3HD, CDP/Cut bound to DNA with a fast "off" rate (Fig. 1B). Thus, the full length CDP/Cut protein binds only transiently to DNA, although it contains CR3HD that, on its own, makes a stable interaction with DNA. Altogether, these results suggest that the CR1CR2 bipartite DNA binding do-

main is active in the context of the full length CDP/Cut protein, whereas CR3HD is weakly or not active.

An amino-truncated 110 kDa CDP/Cut is responsible for the lower of two retarded complexes observed with the universal CDP/Cut consensus-binding site. Electrophoretic mobility shift assays were performed using the universal CDP/Cut consensus binding site, CGATATCGAT, and nuclear extracts from NIH3T3 cells that had been transfected with a vector expressing the full length CDP/Cut cDNA. Two main retarded complexes were observed (Fig. 2A, lane 1). Antibodies raised against different regions of the CDP/Cut protein were used in supershifting experiments in order to assess the nature of the protein(s) responsible for the retarded complexes. While the 861 antibodies supershifted both retarded complexes, the N-term antibodies supershifted only the highest complex, suggesting that the higher complex involved the full length CDP/Cut protein and the lower complex, a CDP/Cut protein that is truncated at its amino-terminal end (Fig. 2A, lanes 2 and 3). In Western blot analyses with the N-term and 861 antibodies, several lower molecular weight species were observed in addition to the full length 200 kDa CDP/Cut protein (Fig. 2C, lanes 1 to 4)). Interestingly, a 110-kDa protein detected with the 861 antibody was also present in untransfected 293 cells, indicating that this smaller protein is normally expressed in untransfected human cells (Fig. 2C, lanes 3-4).

To confirm and extend the above results, we prepared a vector, called MCH (Myc-Cut-HA), expressing the full length CDP/Cut protein with Myc and HA epitope tags at its amino- and carboxy-termini, respectively. NIH3T3 cells were transfected with the MCH vector and tested in EMSA with the universal CDP/Cut consensus-binding site. Again, two

main retarded complexes were observed (Fig. 2B, lane 1). The Myc antibody supershifted only the highest retarded complex, the HA antibody supershifted both complexes, whereas an unrelated antibody had no effect (Fig. 2B, lanes 2 to 4). These results indicate that the protein responsible for the highest complex contains both epitope tags and therefore must be full length. In contrast, the protein generating the lower retarded complex contains an HA tag but not a Myc tag, and must be truncated at its amino-terminal end. In Western blot analysis, the HA antibody revealed two main proteins: one with an apparent M.W. of 200 kDa corresponding to the full length CDP/Cut, and one of approximately 110 kDa. The Myc antibody also revealed two main proteins: one of 200 kDa and one of approximately 90 kDa. These results demonstrate that shorter CDP/Cut proteins can be generated from the full-length CDP/Cut coding sequences, and that an amino-truncated CDP/Cut protein binds to the universal CDP/Cut binding site.

Two mechanisms could account for the generation of shorter CDP/Cut proteins: A proteolytic processing event may at once produce the 110 kDa protein seen with the HA antibody, and the 90 kDa protein seen with the Myc antibody. Alternatively, the amino-truncated 110-kDa protein may be generated by translation at an internal start codon, while the carboxy-truncated 90-kDa protein would be generated by premature translation termination or proteolytic cleavage of the full-length protein. Experiments described below will address this issue. It is important to stress that several experiments were performed to ensure that shorter CDP/Cut isoforms were not simply generated as a result of proteolytic cleavage occurring post lysis. First, our lysis buffer contained an extensive cocktail of protease inhibitors (see Materials and Methods). Secondly, the shorter CDP/Cut isoforms were observed with a series of lysis buffers containing various concentrations of ionic and

(data not shown). Thirdly, treatment of cells with certain protease inhibitors for 4 hours caused a reduction in the amount of short CDP/Cut isoforms (Fig. 8, and data not shown). However, no reduction was observed when the same protease inhibitors were applied to cells only minutes prior to cell lysis (data not shown).

DNA affinity chromatography leads to the specific enrichment of a 110-kDa CDP/Cut protein. The results presented in Fig. 2 showed that the most abundant CDP/Cut protein species was the 200-kDa full-length protein. However, in EMSA the most abundant retarded complex was the lowest one. Assuming that the lowest retarded complex involved a shorter CDP/Cut protein, these results could only be reconciled if the putative shorter species exhibited a higher DNA binding affinity or was able to bind more stably to DNA. To investigate this possibility, we performed DNA affinity chromatography, and assessed the relative amounts of CDP/Cut protein species before and after the procedure (see Materials and Methods). In unfractionated nuclear extracts from HeLa cells, the 110 kDa species was clearly a minor species as compared to the 200-kDa species (Fig. 3, lanes 1 and 6). However, DNA affinity chromatography led to the specific enrichment of the 110-kDa protein (Fig. 3, lanes 5 and 7). The 110-kDa protein was recognized by the 861 and 1300, but not by the N-term CDP/Cut antibodies (see Fig. 2 for the position of their respective epitopes). These results clearly demonstrate the N-terminally truncated 110 kDa CDP/Cut protein can bind to DNA with high affinity.

S phase-specific proteolytic cleavage of CDP/Cut. We have previously shown that CDP/Cut is regulated in a cell cycle-dependent manner and that DNA binding to the

ATCGAT sequence increases as cells progress into S phase (16). We therefore verified whether expression of the shorter 110-kDa protein was also regulated during the cell cycle. NIH3T3 cells were transfected with the MCH vector, and then synchronized in G0, early G1, mid-G1 and S phase using serum starvation/restimulation or thymidine block (see Materials and Methods). Total extracts were prepared and analyzed in Western blot using the HA and Myc antibodies (Fig. 4A). The 110 kDa protein was not detected in the population of cells enriched in G0, it was barely visible in early G1 and mid-G1, but it was strongly expressed in S phase using either method of synchronization. Interestingly, with the Myc antibody a protein of 90 kDa was detected in the two populations of cells enriched in S phase, suggesting that the 110 kDa HA-tagged protein and the 90 kDa Myc-tagged proteins were generated at the same time, most likely by proteolytic cleavage. This hypothesis will be tested in pulse-chase labeling experiments (see Fig. 6). Interestingly, the Myc antibody also revealed a carboxy-terminally truncated protein that migrated with an apparent M.W. of 180 kDa. These findings raise the possibility that another processing event also takes place within the carboxy-terminal domain of the protein. This possibility is currently under investigation. EMSA were performed with the same extracts and oligonucleotides containing the ATCGAT consensus (Fig. 4B). These oligonucleotides were previously shown to be well recognized by CR3HD but weakly or not by CR1CR2 (33). Indeed, only one retarded complex was detected, and its intensity was increased in S phase. Altogether these results demonstrate that the increase in ATCGAT binding activity at the G1/S transition coincides with the production of the N-terminally truncated 110 kDa CDP/Cut protein.

We then verified whether expression of the endogenous CDP/Cut protein was also regulated in a cell cycle-dependent manner. NIH3T3 cells were synchronized by serum starvation/restimulation and analyzed in Western blot using the 861 antibodies, and in EMSA using the ATCGAT binding site. Interestingly, a doublet of bands at 110 kDa was observed at 18 hours following serum stimulation in the population of cells enriched in S phase (Fig. 5A). Whether there are two processing events, or one processing event followed by some post-translational modification is not known at this point. It should be noted that only one 110-kDa band was detected in 293 cells using the same antibodies (see Fig. 2). Thus, there could be a difference in this respect between human and murine cells. Importantly, these novel bands correlate with the increase in the retarded complex seen in EMSA (Fig. 5B). In conclusion, amino-truncated CDP/Cut proteins are expressed predominantly in S phase, and the increase in their expression correlates with the increase in DNA binding to the ATCGAT sequence.

Pulse-chase labeling *in vivo* indicates that the 110-kDa protein derives from the 200-kDa full length CDP/Cut protein. The fact that the 110 kDa N-terminally truncated CDP/Cut protein was expressed concomitantly with a 90 kDa C-terminally truncated protein suggested that a proteolytic processing event was involved in the generation of the 110-kDa protein. To verify this hypothesis, we performed ^{35}S pulse-chase labeling *in vivo* and followed expression of the 200 and 110 kDa CDP/Cut proteins. NIH3T3 cells were transfected with a vector expressing MCH and synchronized in G0 by serum starvation for 72 hours. After serum starvation, cells were stimulated to re-enter the cell cycle by replacing the medium for complete medium with 10% FBS. 12 hours later, cells were incubated

with complete labeling medium for 30 minutes. Total extracts were prepared from one plate (Fig. 6A, sample 1), while cells in the second plate were incubated with normal medium for an additional 4 hours at which time total extracts were prepared (Fig. 6A, sample 2). This scheme was chosen because previous cell synchronization experiments revealed that the 110 kDa protein was weakly or not expressed in G1, but appeared later as cells progressed into S phase (Fig. 4 and 5). Following immunoprecipitation with the 861 antibody, the 110 kDa protein was not detected in sample 1, indicating that it is not synthesized in mid-G1 (Fig. 6A, lane 1). Yet, the 110-kDa protein was clearly visible in cells that had been incubated an additional 4 hours in the presence of "cold" medium (Fig. 6A, lane 2). Since radiolabeling took place from 12h. to 12h30 and only full length CDP/Cut was visible at 12h30, the presence of the 110-kDa protein at 16h30 could only be accounted for by supposing that it was generated by cleavage of the full length CDP/Cut protein. Moreover, as both samples were prepared in the same manner, the presence of the 110-kDa protein in one sample but not in the other further demonstrates that the proteolytic cleavage did not occur post-lysis.

Another mechanism for the production of the N-terminally truncated 110-kDa protein would be translation initiation at an internal start site. To investigate this possibility, we modified the MCH vector by inserting a stop codon in the coiled-coil region upstream of CR1. The presence of a stop codon in the 5' coding region should not affect expression of the 110-kDa protein if translation at a downstream start site is involved. The MCH-STOP vector produced a short, truncated, protein that was easily detected with the Myc antibody, indicating that the stop codon was functional (Fig. 6B, lane 1). However, the 110 kDa protein could not be detected with the HA antibody (Fig. 6B, lane 2). We conclude

that expression of the 110-kDa protein requires the prior production of the full length 200 kDa CDP/Cut protein.

Altogether, these results exclude translation at an internal initiation site as the mechanism leading to the expression of the 110-kDa protein, and are most consistent with the notion that this shorter protein is generated from the 200-kDa protein by proteolytic processing. Thus, translation in mid-G1 leads to the synthesis of a 200 kDa full length CDP/Cut protein that is later processed into an amino-terminally truncated protein of 110 kDa with DNA binding properties similar to that of CR3HD.

Mapping of the sequences required for processing. To map the sequences required for processing, we generated a series of vectors expressing amino-truncated Myc-Cut-HA proteins (Fig. 7A). The name of each vector corresponds to the first amino acid encoded by this vector. Only the full length MCH and the 659 vectors produced the 110-kDa proteins (Fig. 7A). The 878, 958 and 1029 vector generated shorter proteins. Thus, the site of cleavage is located between CR1 and CR2, more precisely between amino acids 659 and 878. In EMSA with the ATCGAT sequence, the 878 and 958 vectors generated retarded complexes that migrated faster than those produced by the full length MCH vector. In contrast, 659 produced two retarded complexes, one comigrating with the MCH complex, and one more intense that migrated more slowly. These results suggest that the full-length 659 protein, in contrast to the full length MCH protein, is able to bind efficiently to the ATCGAT sequence. Since the full-length 659 protein is more abundant than its processed form (Fig. 7A, lane 2), it gives rise to the most abundant retarded complex in EMSA. We deduce from these results that the ATCGAT DNA binding activity is inhibited in the

context of the full-length protein, and that this inhibition can be released following amino-terminal truncation of the protein.

We then generated a series of MCH vectors with progressive carboxy-terminal deletions. Since processing was observed with each construct, we conclude that the carboxy-terminal domain, the Cut homeodomain and the Cut repeat 3 can be removed without affecting processing (Fig. 7C). An internal deletion from a.a. 659 to 878 completely abolished processing (Fig. 7D), whereas a construct encoding only the Cut repeats 1 and 2 and the linker between them generated a shorter protein of the expected M.W. (Fig. 7D). Altogether, our mapping data indicate that proteolytic cleavage occurs between a.a. 659 and 878 and does not require domains of the protein that are located far away from the cleavage site. Whether sequences between a.a. 659 and 878 are sufficient for cleavage is currently under investigation.

Processing changes the transcriptional properties of CDP/Cut. Since the short isoform is capable of stable DNA binding, whereas the full-length protein only transiently binds to DNA, we envisaged that the two isoforms could exhibit different transcriptional properties. CDP/Cux has been characterized by us and other groups as a transcriptional repressor (2, 17, 27, 28, 32, 38, 40, 43). However, some results in the literature suggest that it may also participate in transcriptional activation (45, 50). This is also suggested by our own studies using a reporter construct containing the DNA polymerase alpha gene reporter (Truscott et al., in preparation). To compare the effect of the two CDP/Cut isoforms in a transient reporter assay, we needed to express the full-length protein in the absence of the processed isoform. In parallel studies, we found that little or no processing of CDP/Cut

takes place in confluent cells. Therefore, in this experiment, cells were plated at a higher cell density and, following transfection, were allowed to reach near confluence before harvesting them. As can be seen in Fig. 8B, no processed isoform was observed in cells transfected with the full length Myc-Cut-Ha construct. In the reporter assay, we found that a recombinant protein corresponding to the p110 CDP/Cut isoform was able to stimulate expression from a DNA polymerase alpha gene reporter (Fig. 8A). In contrast, expression was decreased in the presence of full-length CDP/Cut (Fig. 8A). These results do not demonstrate that p110 directly activates the DNA polymerase alpha reporter. The results are also consistent with a model whereby the short CDP/Cut isoform would repress expression of a protein that, in turn, acts as a down-modulator of the DNA polymerase alpha gene promoter. Notwithstanding the exact mechanism leading to an increase in reporter gene expression, these results suggest that an important consequence of CDP/Cut processing is to generate an isoform with distinct transcriptional properties.

DISCUSSION

The CDP/Cut transcription factor contains two bipartite DNA binding domains, CR1CR2 and CR3HD, that exhibit distinct DNA binding activities. When expressed as fusion proteins, CR1CR2 bound with rapid "on" and "off" rates to DNA sequences containing direct or inverted repeats of the CAAT or CGAT motifs, whereas CR3HD made a stable interaction preferably with the ATCGAT sequence (33). To our surprise, we found that the full length CDP/Cut protein could only make a transient interaction with DNA (Fig. 1). In cells, the CDP/Cut protein that binds to the ATCGAT site was found to be an amino-terminally truncated 110-kDa protein (Fig. 2 and 3). Production of this protein was shown to involve the synthesis of the full-length protein followed by an S phase-specific proteolytic cleavage (Fig. 4 to 7). In reporter assays, an amino-terminally truncated CDP/Cut protein caused an increase whereas the full-length CDP/Cut caused a decrease in the expression from the DNA polymerase alpha gene promoter (Fig. 8). Altogether our results demonstrated that proteolytic processing of CDP/Cut leads to the production of an isoform that displays distinct DNA binding and transcriptional activities (Fig. 9).

The proteolytic processing event leading to the production of the 110 kDa CDP/Cut protein was found to occur at a much higher rate in populations of cells enriched in S phase, either by serum starvation/stimulation or thymidine block, than in unsynchronized cells or in populations enriched in G0 or G1 (Fig. 4 and 5). This suggests that CDP/Cut processing begins as cells progress into S phase. It is not clear whether the triggering event at the end of G1 involves the protease or CDP/Cut. Two scenarios can be envisaged: The protease itself may be expressed or activated at the end of G1. Alternatively, we can envisage that

the protease acts in a constitutive manner, but that CDP/Cut needs to be post-translationally modified in order to become a substrate of this protease.

Interestingly, S phase-specific proteolytic processing of the p27 cdk inhibitor has previously been reported (37). Processing was assumed to cause the inactivation of p27 by removing the cyclin interacting domain, CY. The cleavage of p27 and CDP/Cut suggests that proteolytic processing may be an important regulatory mechanism in the control of the G1/S transition. These results, along with the findings that sister-chromatid separation at anaphase is promoted by cleavage of the cohesin subunit Scc1, indicate that proteolytic processing may be involved in different phases of the cell cycle (42). Thus, two sorts of proteolysis appear to be needed to ensure cell cycle progression. The first is the destruction of several cell cycle regulators following ubiquitination by the Skp1–Cdc53/cullin–F-box protein (SCF) complexes or the anaphase-promoting complex/cyclosome (APC/C) (reviewed in (41)). The second is the precise cleavage of certain proteins to generate peptides with properties distinct from that of their precursors.

In most cases thus far reported, the consequence of proteolytic processing was to redirect the localization of these proteins, in particular from a cytoplasmic site to the nucleus. For example, cleavage of SREBP and ATF6 in response to sterol-deprivation and ER-stress caused the release of their cleaved products from the ER membrane and their movement to the nucleus (22, 48). Cleavage of Cubitus interruptis in the absence of Hedgehog signaling enabled the translocation of the Ci 75-kDa isoform from microtubules in the cytoplasm to the nucleus where it functioned as a repressor (4). Processing of NFκB p105 into NFκB p50 was proposed to contribute to its nuclear translocation by separating the P50 N-terminal peptide from the ankyrin repeats which would retain it to a cytoplasmic

anchor (18). In a few cases, specific proteolytic cleavage serves to generate novel isoforms with altered biochemical properties. Cleavage of full-length 38-kDa C/EBPβ leads to the production of a dominant negative C/EBPβ isoform of 21-kDa, LIP (liver-enriched transcriptional inhibitory protein), that can bind to DNA but is devoid of transactivation potential (49). Autocatalytic processing of the C1 factor was found to generate a number of polypeptides that remained tightly associated together (47). In this case, processing resulted in the production of a protein complex with different properties, but whether processing also affected its localization remained unclear. In the case of CDP/Cut, processing did not seem to alter sub-cellular localization, as antibodies raised against N- or C-terminal peptides generated a strong signal primarily, if not exclusively, in the nucleus (data not shown). We cannot of course exclude the possibility that processing affects localization in a more subtle way within the nucleus itself. Clearly, however, processing of CDP/Cut generated an isoform with different DNA binding properties.

An amino-terminally truncated CDP/Cut protein, but not the full length CDP/Cut protein, was able to stimulate expression from a reporter construct containing the DNA polymerase alpha gene promoter (Fig. 8). We do not know the mechanism of action of p110: it may directly activate transcription of the reporter; alternatively, it may repress expression of another repressor that binds to the DNA polymerase alpha promoter. Whatever its mechanism of action, we can speculate that the distinct transcriptional activities of the short and full-length CDP/Cut proteins may result from their differences in DNA binding activities. Whereas the full length CDP/Cut protein could only make transient interaction with oligonucleotides containing the ATCGAT motif, shorter proteins containing CR3HD or CR2CR3HD were capable of stable DNA binding (Fig. 1 and (33)). It is gener-

ally assumed that transcriptional activation requires stable interaction with the promoter. On the other hand, we and others have shown that CDP/Cut can repress by two mechanisms: active repression and competition for binding site occupancy (27, 28, 32). While transient DNA binding could cause repression by preventing the binding of an activator to an overlapping binding site, active repression would be expected to require stable DNA binding to the promoter in order to recruit a histone deacetylase. Future experiments should test the notion that stable DNA binding is required for active repression and transcriptional activation. Moreover, our results raise the interesting possibility that CDP/Cut isoforms may exhibit distinct biochemical activities on different promoters.

An important component of the transcriptional regulation at the G1/S transition involves the phosphorylation of Rb by G1 cyclin/CDKs and the concomitant release of E2F, which can then activate the transcription of genes whose products are required for DNA replication (reviewed in (34)). However, some of the genes previously reported to be activated by E2F in co-transfection assays, including the DNA polymerase alpha gene, were recently found not to be primary targets of E2F, implying the involvement of other effectors acting downstream of, or in parallel with, E2F (46). Our results suggest that proteolytic cleavage of CDP/Cut in part may serve to activate the transactivation function of CDP/Cut (Fig. 8). These findings raise the possibility that the processed CDP/Cut isoform represents one of the G1/S effectors. Future experiments should aim to verify whether CDP/Cut function is required for the induction of S phase by E2F.

ACKNOWLEDGMENTS

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

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:**245-57.

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:**321-34.

3. **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:**7757-61.

4. **Aza-Blanc, P., F. A. Ramirez-Weber, M. P. Laget, C. Schwartz, and T. B. Kornberg.** 1997. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. Cell **89:**1043-53.

5. **Aziz, F., A. J. Vanwijnen, P. S. Vaughan, S. J. Wu, A. R. Shakoori, J. B. Lian, K. J. Soprano, J. L. Stein, and G. S. Stein.** 1998. The integrated activities of Irf-2 (Hinf-M), Cdp/Cut (Hinf-D) and H4tf-2 (Hinf-P) regulate transcription of a cell cycle controlled human histone H4 gene - mechanistic differences between distinct H4 genes. Mol Biol Rep **25:**1-12.

6. **Blanc, R.** 1942. The production of wing scalloping in *Drosophila melanogaster*. Univ Calif Publ Zool **49**.

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 identity in Drosophila. Nature **333:**629-35.

8. **Blochlinger, K., R. Bodmer, 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:**1322-31.

9. **Blochlinger, K., L. Y. Jan, and Y. N. Jan.** 1993. Postembryonic patterns of expression of cut, a locus regulating sensory organ identity in Drosophila. Development **117:**441-50.

10. **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:**1124-35.

11. **Blomberg, I., and I. Hoffmann.** 1999. Ectopic expression of Cdc25A accelerates the G(1)/S transition and leads to premature activation of cyclin E- and cyclin A-dependent kinases. Mol Cell Biol **19:**6183-94.

12. **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 D. melanogaster. Cell **51:**293-307.

13. **Braun, W.** 1942. The effect of puncture on the developing wing of several mutants of *Drosophila melanogaster*. J Exp Zool **84:**325-50.

14. **Brown, M. S., and J. L. Goldstein.** 1997. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell **89:**331-40.

15. **Chan, Y. M., and Y. N. Jan.** 1998. Roles for proteolysis and trafficking in notch maturation and signal transduction. Cell **94:**423-6.

16. **Coqueret, O., G. Berube, and A. Nepveu.** 1998. The mammalian Cut homeodomain protein functions as a cell-cycle-dependent transcriptional repressor which downmodulates P21(Waf1/Cip1/Sdi1) In S phase. EMBO J **17:**4680-94.

17. **Dufort, D., and A. Nepveu.** 1994. The human cut homeodomain protein represses transcription from the c-myc promoter. Mol Cell Biol **14:**4251-7.

18. **Fan, C. M., and T. Maniatis.** 1991. Generation of p50 subunit of NF-kappa B by processing of p105 through an ATP-dependent pathway. Nature **354:**395-98.

19. **Gervais, F. G., D. Xu, G. S. Robertson, J. P. Vaillancourt, Y. Zhu, J. Huang, A. LeBlanc, D. Smith, M. Rigby, M. S. Shearman, E. E. Clarke, H. Zheng, L. H. Van Der Ploeg, S. C. Ruffolo, N. A. Thornberry, S. Xanthoudakis, R. J. Zamboni, S. Roy, and D. W. Nicholson.** 1999. Involvement of caspases in proteolytic cleavage of Alzheimer's amyloid- beta precursor protein and amyloidogenic A beta peptide formation. Cell **97:**395-406.

20. **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:**129-40.

21. **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:**2062-7.

22. **Haze, K., H. Yoshida, H. Yanagi, T. Yura, and K. Mori.** 1999. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. Mol Cell Biol **10:**3787-99.

23. **Howcroft, T. K., S. L. Kirshner, and D. S. Singer.** 1997. Measure of transient transfection efficiency using beta-galactosidase protein. Anal Biochem **244:**22-7.

24. **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:**735-47.

25. **Jinno, S., K. Suto, A. Nagata, M. Igarashi, Y. Kanaoka, H. Nojima, and H. Okayama.** 1994. Cdc25A is a novel phosphatase functioning early in the cell cycle. EMBO J **13:**1549-56.

26. **Lee, K. A. W., 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 Techn **5:**22-31.

27. **Li, S., L. Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, and M. J. Walsh.** 1999. Transcriptional repression of the cystic fibrosis transmembrane conductance regulator gene, mediated by CCAAT displacement protein/cut homolog, is associated with histone deacetylation. J Biol Chem **274:**7803-15.

28. **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:**12745-50.

29. **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:**133-43.

30. **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:**151-9.

31. **Ludlow, C., R. Choy, and K. Blochlinger.** 1996. Functional analysis of Drosophila and mammalian Cut proteins In flies. Dev Biol **178:**149-59.

32. **Mailly, F., G. Berube, R. Harada, P. L. Mao, S. Phillips, and A. Nepveu.** 1996. The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy. Mol Cell Biol **16:**5346-57.

33. **Moon, N. S., G. Berube, and A. Nepveu.** 2000. CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain. J Biol Chem **275:**31325-34.

34. **Muller, H., and K. Helin.** 2000. The E2F transcription factors: key regulators of cell proliferation. Biochim Biophys Acta **1470:**M1-12.

35. **Nepveu, A.** 2001. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. Gene **270:**1-15.

36. **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. Nature Genet **1:**50-5.

37. **Shirane, M., Y. Harumiya, N. Ishida, A. Hirai, C. Miyamoto, S. Hatakeyama, K. Nakayama, and M. Kitagawa.** 1999. Down-regulation of p27(Kip1) by two mechanisms, ubiquitin-mediated degradation and proteolytic processing. J Biol Chem **274:**13886-93.

38. **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:**16736-44.

39. **Stein, G. S., and T. W. Borun.** 1972. The synthesis of acidic chromosomal proteins during the cell cycle of HeLa S-3 cells. I. The accelerated accumulation of acidic residual nuclear protein before the initiation of DNA replication. J Cell Biol **52:**292-307.

40. **Superti-Furga, G., A. Barberis, E. Schreiber, and M. Busslinger.** 1989. The protein CDP, but not CP1, Footprints on the CCAAT region of the g-globulin gene in unfractionated B-cell extracts. Biochim Biophys Acta **1007:**237-42.

41. **Tyers, M., and P. Jorgensen.** 2000. Proteolysis and the cell cycle: with this RING I do thee destroy. Curr Op Genet Dev **10:**54-64.

42. **Uhlmann, F., F. Lottspeich, and K. Nasmyth.** 1999. Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1 [see comments]. Nature **400:**37-42.

43. **Valarche, I., J. P. Tissier-Seta, M. R. Hirsch, S. Martinez, C. Goridis, and J. F. Brunet.** 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:**881-96.

44. **van Wijnen, A. J., C. Cooper, P. Odgren, F. Aziz, A. De Luca, R. A. Shakoori, A. Giordano, P. J. Quesenberry, J. B. Lian, G. S. Stein, and J. L. Stein.** 1997. Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells. J Cell Biochem **66:**512-23.

45. **van Wijnen, A. J., M. F. van Gurp, M. C. de Ridder, C. Tufarelli, T. J. Last, M. Birnbaum, P. S. Vaughan, A. Giordano, W. Krek, E. J. Neufeld, J. L. Stein, and G. S. Stein.** 1996. CDP/cut is the DNA-binding subunit of histone gene transcription

factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. Proc Natl Acad Sci USA **93:**11516-21.

46. **Vigo, E., H. Muller, E. Prosperini, G. Hateboer, P. Cartwright, M. C. Moroni, and K. Helin.** 1999. CDC25A phosphatase is a target of E2F and is required for efficient E2F-induced S phase. Mol Cell Biol **19:**6379-95.

47. **Vogel, J. L., and T. M. Kristie.** 2000. Autocatalytic proteolysis of the transcription factor-coactivator C1 (HCF): a potential role for proteolytic regulation of coactivator function. Proc Natl Acad Sci USA **97:**9425-30.

48. **Wang, X., R. Sato, M. S. Brown, X. Hua, and J. L. Goldstein.** 1994. SREBP-1, a membrane-bound transcription factor released by sterol- regulated proteolysis [see comments]. Cell **77:**53-62.

49. **Welm, A. L., N. A. Timchenko, and G. J. Darlington.** 1999. C/EBPalpha regulates generation of C/EBPbeta isoforms through activation of specific proteolytic cleavage. Mol Cell Biol **19:**1695-704.

50. **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:**18453-62.

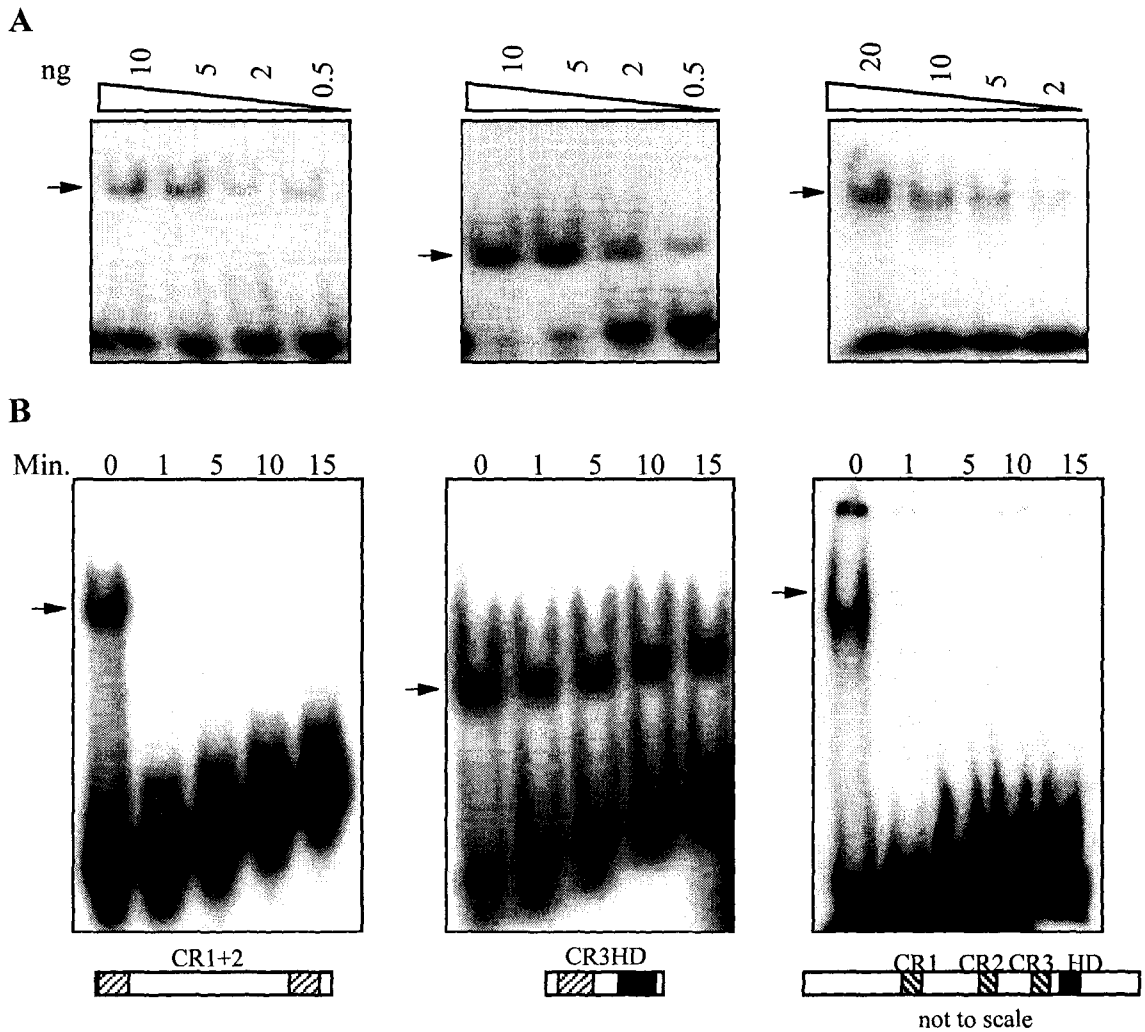

A
ng
10
5
2
0.5
10
5
2
0.5
20
10
5
2
B
Min.
0 1 5 10 15
0 1 5 10 15
0 1 5 10 15
CR1+2
CR3HD
CR1 CR2 CR3 HD
not to scale

Figure 1. The full length CDP/Cut protein exhibits DNA binding kinetics similar to that of CR1CR2.

(A) DNA binding specificity of CR1CR2, CR3HD and CDP/Cut for CGATATCGAT. Decreasing amounts of the indicated fusion proteins were mixed with radiolabeled oligonucleotides containing the CGATATCGAT, at room temperature until the equilibrium was reached. DNA-protein complexes were resolved on a nondenaturing polyacrylamide gel that was dried and autoradiographed. Note that oligonucleotides containing the CGATATCGAT sequence are recognized efficiently by CR1CR2 since two CGAT direct repeats are present (CGATATCGAT), and by CR3HD since the ATCGAT sequence is present (CGATATCGAT).

(B) DNA binding kinetics of CR1CR2, CR3HD and CDP/Cut. 100 ng of the indicated fusion protein was incubated with radiolabeled oligonucleotides, containing the CGATATCGAT sequence, at room temperature until the equilibrium was reached (15 min.); 1000 fold molar excess of unlabeled oligonucleotides was added, and at the indicated time points aliquots of the mixture were taken and analyzed in EMSA.

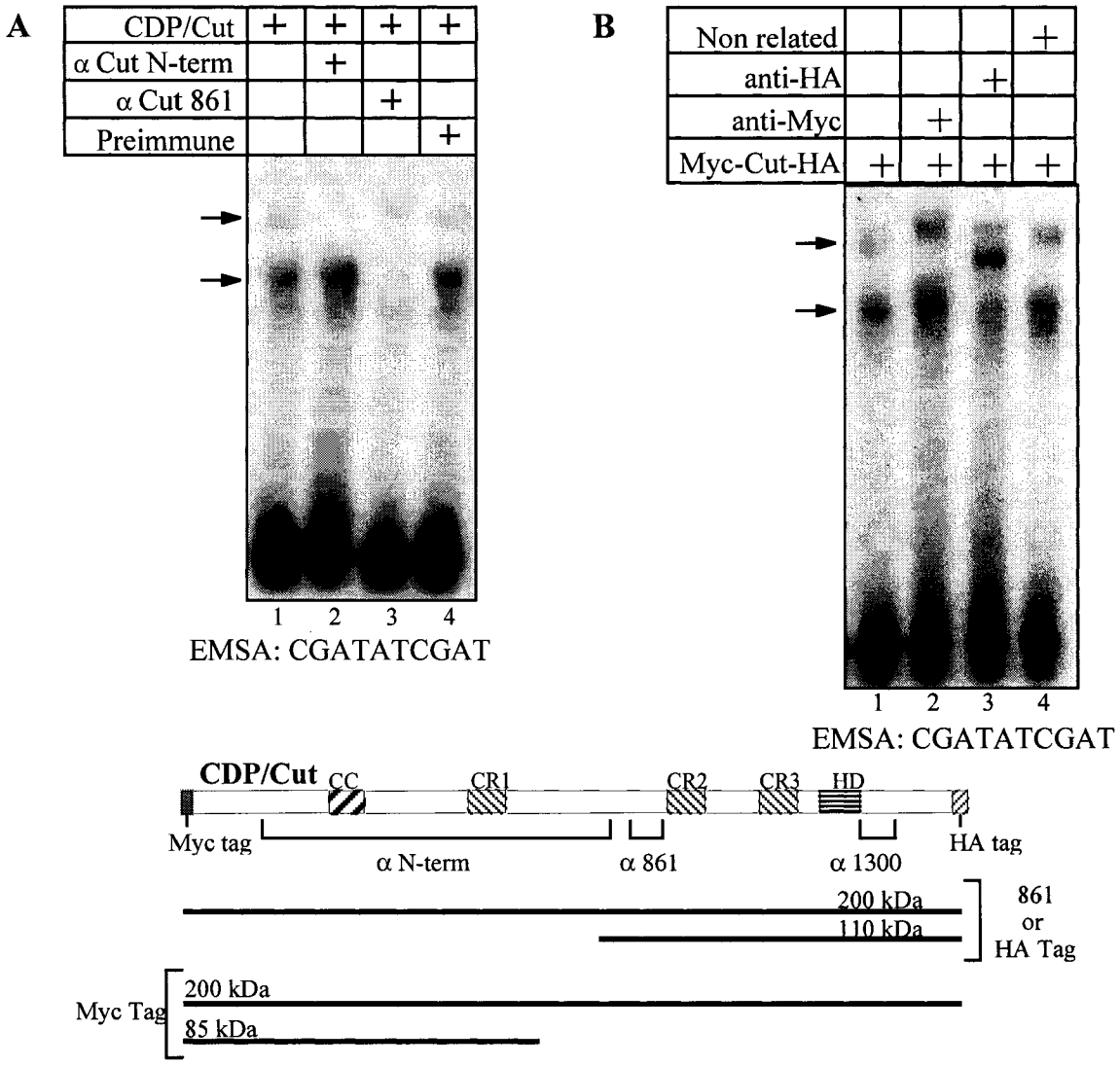
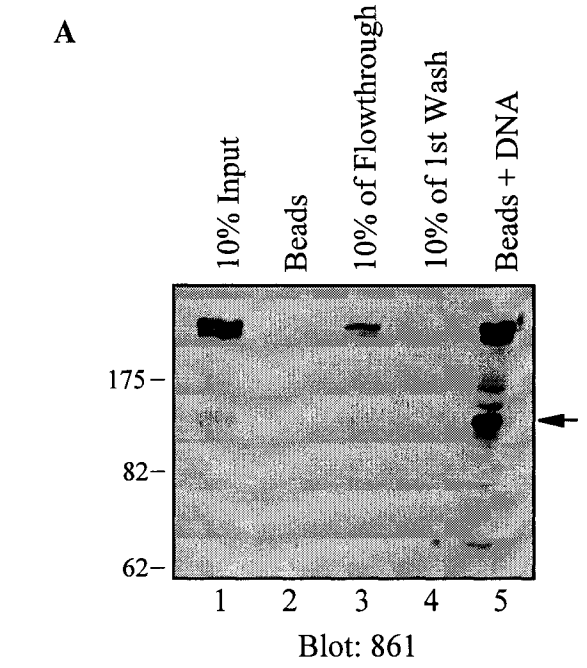
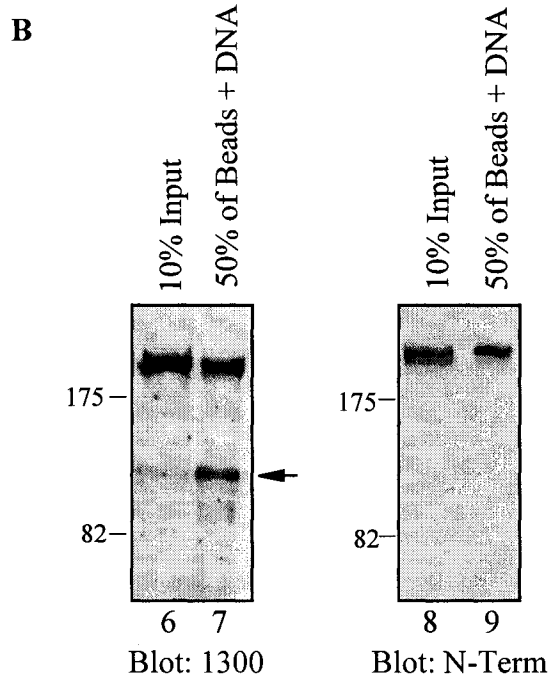

A
CDP/Cut + + + +
α Cut N-term +
α Cut 861 +
Preimmune +
1 2 3 4
EMSA: CGATATCGAT
B
Non related +
anti-HA +
anti-Myc +
Myc-Cut-HA + + + +
1 2 3 4
EMSA: CGATATCGAT
CDP/Cut
CC
CR1
CR2
CR3
HD
Myc tag
α N-term
α 861
α 1300
HA tag
200 kDa
110 kDa
861 or HA Tag
Myc Tag
200 kDa
85 kDa
C
293
CDP/Cut
kDa
175
83
1 2
Blot: N-term
293
CDP/Cut
kDa
175
83
3 4
Blot: 861
MCH
Vector
kDa
175
83
5 6
Blot: HA
MCH
Vector
kDa
175
83
7 8
Blot: Myc

Figure 2. Two distinct retarded complexes are generated by CDP/Cut proteins, only one of which is supershifted by antibodies raised against the n-terminal region of the protein.

(A) and (B) NIH3T3 cells were transfected with a vector expressing the CDP/Cut protein (A), or Myc-Cut-HA, a CDP/Cut protein with Myc and HA epitope tags at its amino- and carboxy-terminus, respectively (B). Nuclear extracts were prepared and analyzed in EMSA with oligonucleotides containing a universal CDP/Cut consensus-binding site (CGATATCGAT). The DNA and proteins were incubated with either no antibody (lanes 1), or the indicated antibodies (lanes 2 to 4). The preimmune serum is that for anti-861, the non-related antibody is a goat anti-rabbit IgG. The arrows indicate the positions of the retarded complexes. Note that the lower retarded complex can be supershifted with anti-861 (A) and anti-HA (B) antibodies, but not with anti-N-term (A) not anti-Myc (B) antibodies. A diagram of the protein is shown indicating the evolutionarily conserved domains and the regions recognized by the respective antibodies (CC, coiled-coil; CR1, CR2 and CR3: Cut repeats 1, 2 and 3; HD, homeodomain).

(C) Nuclear extracts from untransfected 293 cells and transfected NIH3T3 cells (CDP/Cut, MCH, Vector) were separated by electrophoresis on 8% polyacrylamide gels and analyzed in Western blots with CDP/Cut N-term, CDP/Cut 861, HA and Myc antibodies. Note that the relative abundance of the full length CDP/Cut protein appears lower than in subsequent blots using 6% polyacrylamide (Fig. 3, 4 and 5). This is because transfer of the 200-kDa protein to the membrane is less efficient in 8% polyacrylamide.


A
10% Input
Beads
10% of Flowthrough
10% of 1st Wash
Beads + DNA
175
82
62
1 2 3 4 5
Blot: 861


B
10% Input
50% of Beads + DNA
175
82
6 7
Blot: 1300
10% Input
50% of Beads + DNA
175
82
8 9
Blot: N-Term

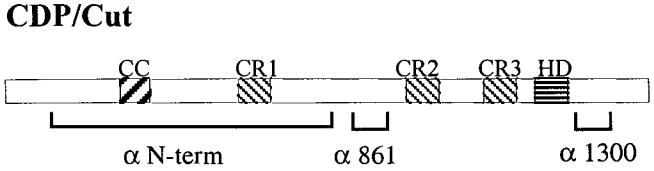
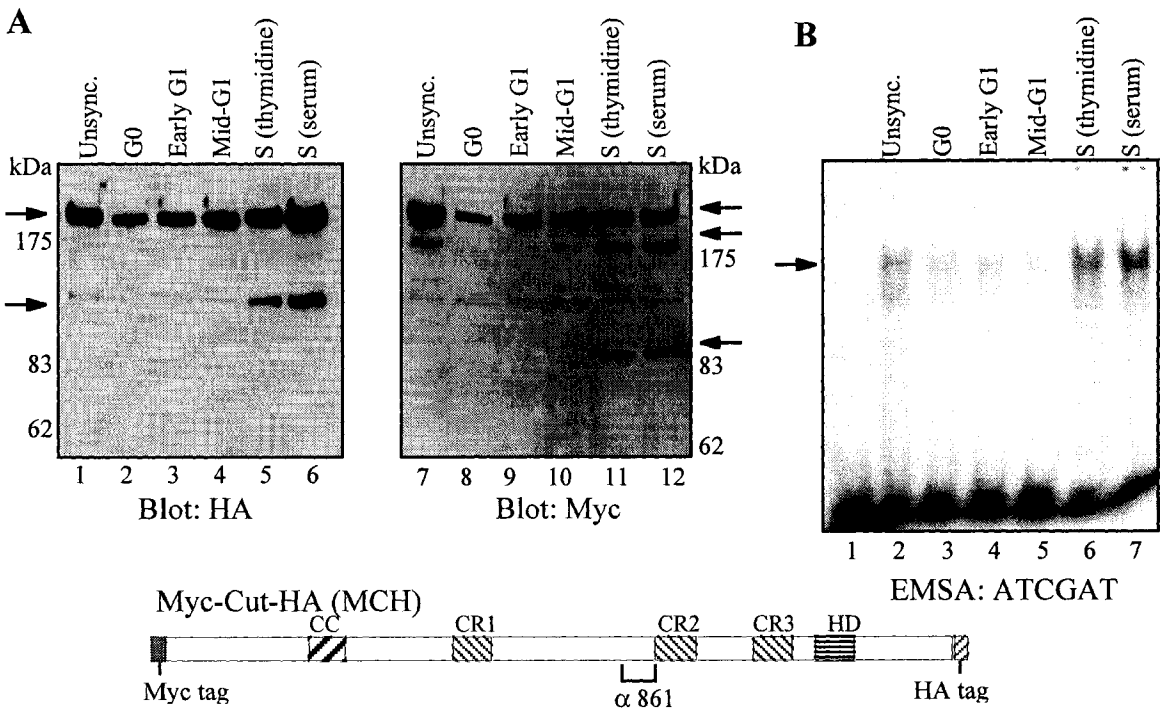

CDP/Cut
CC
CR1
CR2
CR3
HD
α N-term
α 861
α 1300

Figure 3. DNA affinity chromatography with the ATCGAT sequence leads to the specific enrichment of a 110 kDa CDP/Cut protein

(A) Nuclear extracts from HeLa cells were subjected to affinity chromatography using as a bait biotinylated oligonucleotides containing the ATCGAT sequence. Samples were then separated by electrophoresis on 6% polyacrylamide gels and analyzed by Western blot analysis with the anti-CDP/Cut 861 (lanes 1-5), 1300 (lanes 6-7) and N-term (lanes 8-9) antibodies.


A
Unsync.
G0
Early G1
Mid-G1
S (thymidine)
S (serum)
kDa
175
83
62
1 2 3 4 5 6
Blot: HA
7 8 9 10 11 12
Blot: Myc
B
1 2 3 4 5 6 7
EMSA: ATCGAT
Myc-Cut-HA (MCH)
CC
CR1
CR2
CR3
HD
Myc tag
α 861
HA tag

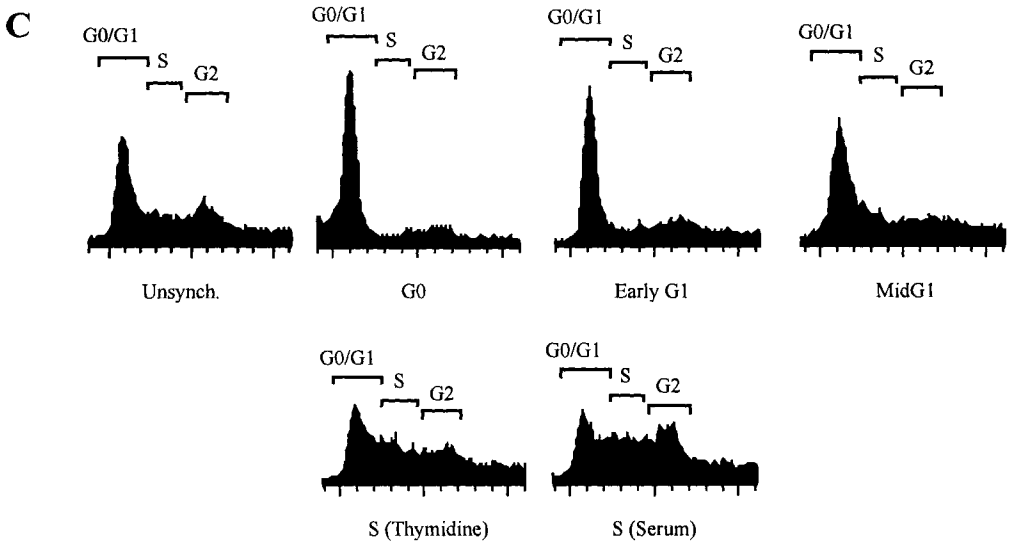

C
G0/G1
S
G2
Unsynch.
G0
Early G1
MidG1
S (Thymidine)
S (Serum)

Figure 4. S phase-specific proteolytic cleavage of CDP/Cut

A cDNA encoding Myc-Cut-HA generates short proteins specifically in S phase. NIH3T3 cells were transfected with a vector expressing Myc-Cut-HA, a CDP/Cut protein with Myc and HA epitope tags at its amino- and carboxy-termini, respectively. Cells were synchronized either by serum starvation/stimulation or thymidine block, as described in Materials and Methods. Total extracts were prepared and analyzed in Western blots with anti-Myc and anti-HA antibodies (A), and in EMSA with an ATCGAT probe (B). Note that only a lower retarded complex is visible in EMSA with total extracts, whereas two complexes were observed with nuclear extracts (see Fig. 1). Cell cycle distribution was monitored by fluorescence-activated cell sorting (FACS) analysis after staining of the DNA with propidium iodide (C).

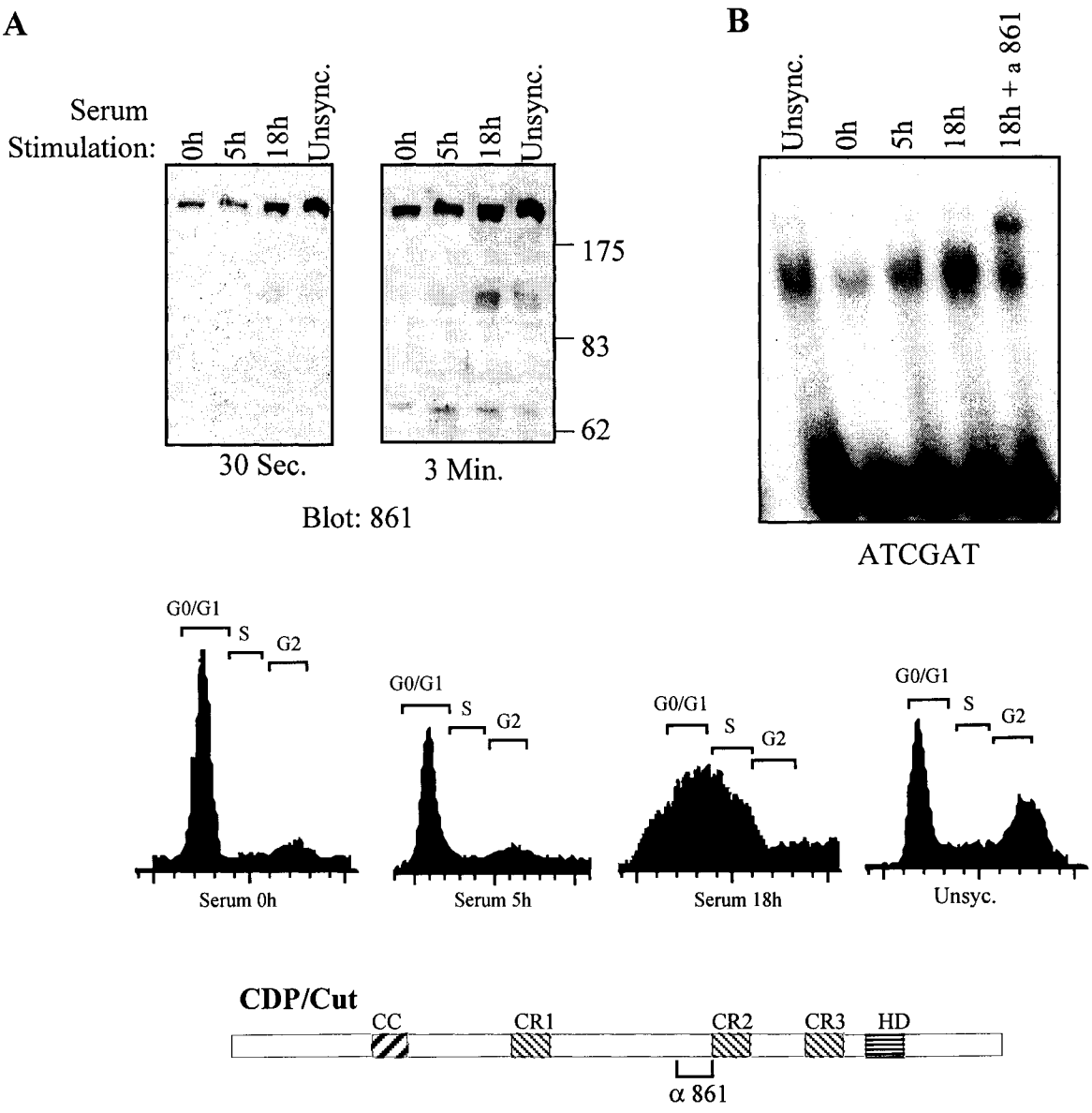

A
Serum Stimulation:
0h
5h
18h
Unsync.
0h
5h
18h
Unsync.
175
83
62
30 Sec.
3 Min.
Blot: 861
B
Unsync.
0h
5h
18h
18h + a 861
ATCGAT
G0/G1
S
G2
Serum 0h
Serum 5h
Serum 18h
Unsyc.
CDP/Cut
CC
CR1
CR2
CR3
HD
α 861

Figure 5. 110 kDa CDP/Cut proteins are expressed in S phase

NIH 3T3 cells were synchronized in G0 by serum starvation for 72 h and then stimulated with fresh DMEM plus 10 % FBS. The indicated time correspond to the time elapsed since release from the G0 block. Nuclear extracts were prepared and analyzed in Western blots with anti-Cut 861 antibodies (A), and in EMSA with the ATCGAT probe (B). Cell cycle distribution was monitored by fluorescence-activated cell sorting (FACS) analysis after staining of the DNA with propidium iodide.

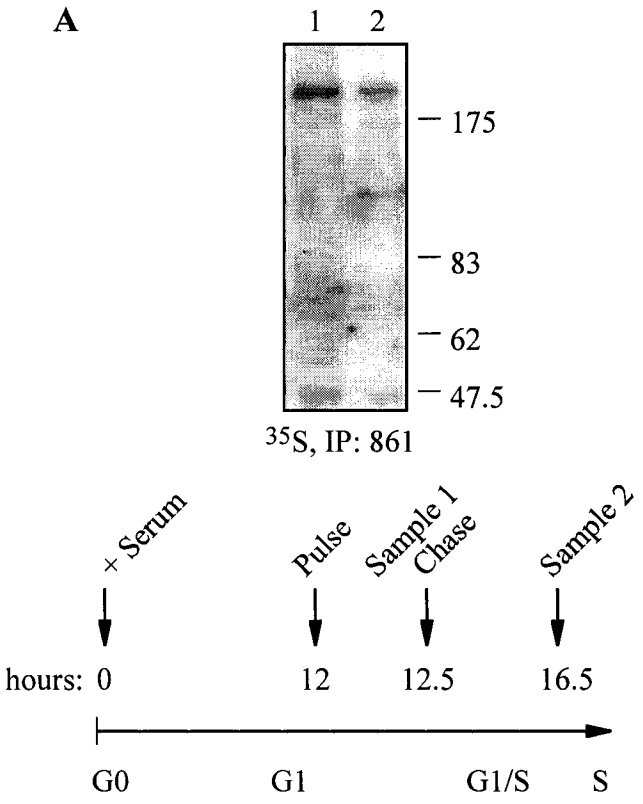

A
1 2
175
83
62
47.5
35S, IP: 861
+ Serum
Pulse
Sample 1
Chase
Sample 2
hours: 0
12
12.5
16.5
G0
G1
G1/S
S
B
MCH Stop
MCH Stop
MCH
175
83
62
47.5
1
Blot:Myc
2 3
Blot:HA
MCH
CC
CR1
CR2
CR3 HD
Myc tag
HA tag
MCH STOP
CC
CR1
CR2
CR3 HD
Myc tag
UAA
HA tag

Figure 6. ^{35}S Pulse chase labeling demonstrates that the 110 kDa amino-truncated protein derives from the 200 kDa full length CDP/Cut protein. (A) NIH3T3 cells were transfected with cDNA expressing Myc-Cut-HA in two 150 mm-plates and synchronized in G0 by serum starvation for 72 h and then stimulated with fresh DMEM plus 10 % FBS. The indicated time correspond to the time elapsed since release from the G0 block. At 12 hours, the medium was replaced for complete labeling medium containing ^{35}S-labeled methionine and cysteine. At 12 hours and 30 minutes, total extracts were prepared from one plate of cells, while the medium was replaced for complete medium with cold methionine and cysteine in the second plate. At 16 hours and 30 minutes, total extracts were prepared from the second plate. Samples 1 and 2 were immunoprecipitated with 861 antibody, resolved by PAGE and revealed by autoradiography.

(B) A cDNA expressing Myc-Cut-HA and Myc-Cut-HA with a stop codon inserted in the coiled-coil region were transfected in NIH3T3. Nuclear extracts were prepared and analyzed by Western blot with Myc and HA anti body.

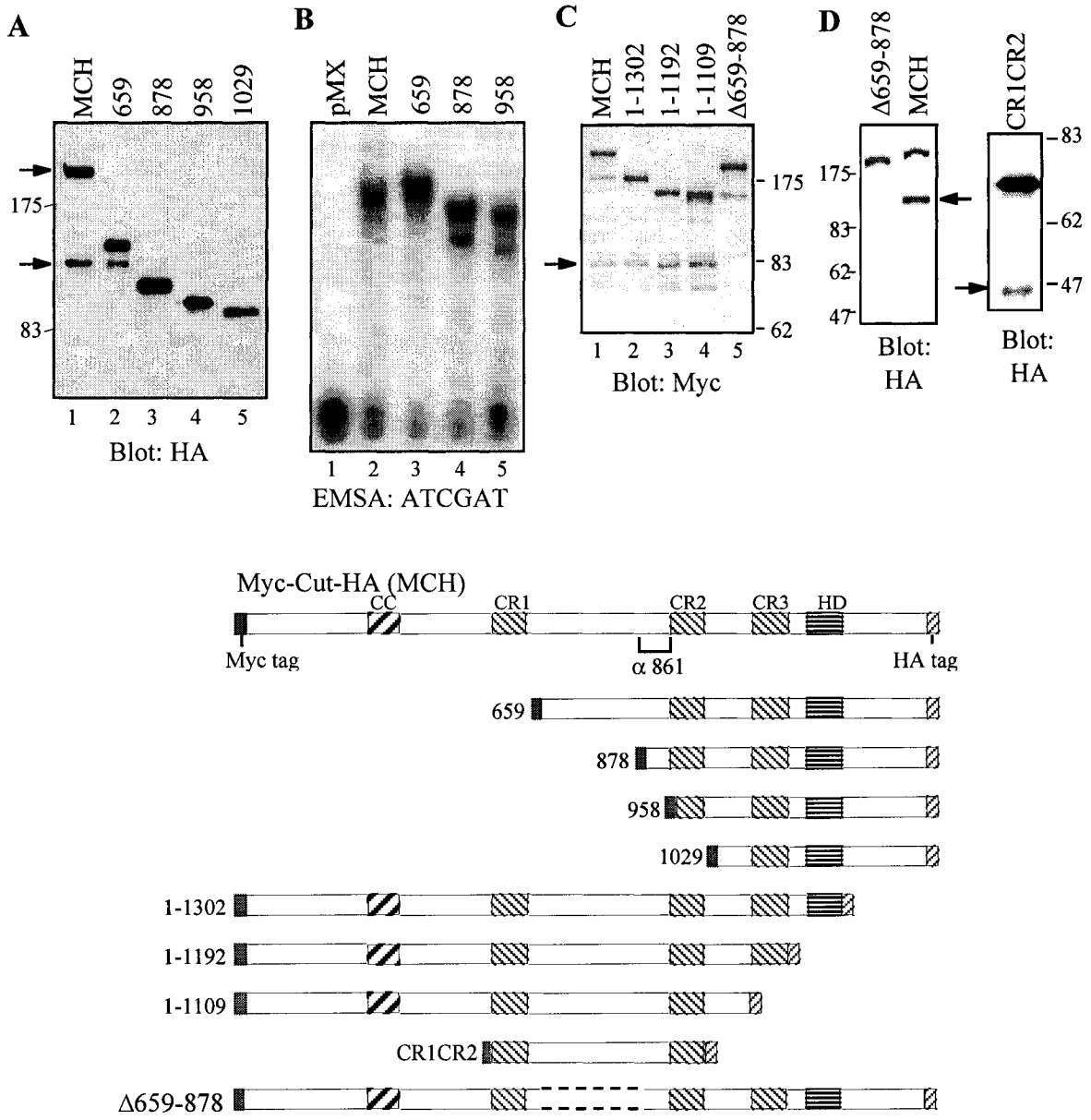

A
MCH
659
878
958
1029
175
83
1 2 3 4 5
Blot: HA
B
pMX
MCH
659
878
958
1 2 3 4 5
EMSA: ATCGAT
C
MCH
1-1302
1-1192
1-1109
Δ659-878
175
83
62
1 2 3 4 5
Blot: Myc
D
Δ659-878
MCH
175
83
62
47
Blot:
HA
CR1CR2
83
62
47
Blot:
HA
Myc-Cut-HA (MCH)
CC
CR1
CR2
CR3
HD
Myc tag
α 861
HA tag
659
878
958
1029
1-1302
1-1192
1-1109
CR1CR2
Δ659-878

Figure 7. Proteolytic cleavage occurs downstream of CR1, between a.a. 659 and 878. NIH3T3 cells were transfected with a series of vectors expressing amino- and carboxy-terminally truncated Myc-Cut-HA (MCH) proteins. Nuclear extracts were prepared and analyzed in Western blots with anti-HA or anti-Myc antibodies, as indicated, and in EMSA with the ATCGAT probe (B). A diagram of the proteins encoded by each construct is shown at the bottom.

Note that only the full length MCH and the 659 are processed into a shorter form. In EMSA, 659 generates two main retarded complexes, the higher of which migrating more slowly than what is seen with MCH. This is because the non-processed 659 is more abundant than its processed derivative, and is capable of stable binding to the ATCGAT probe.

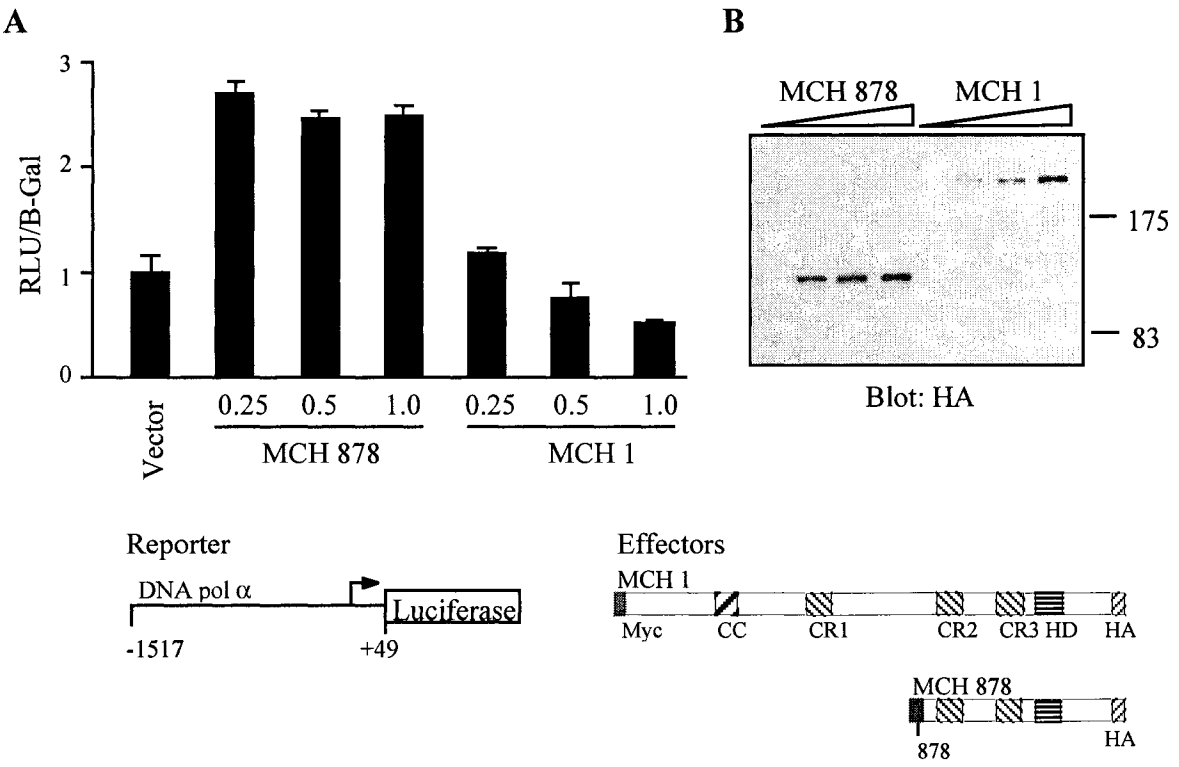


Fig. 8 An amino-terminally truncated CDP/Cut isoform is capable of transcriptional activation

HS578T cells were cotransfected with the DNA polymerase alpha/luciferase reporter construct and a vector expressing either nothing, full-length Myc-Cut-HA (MCH 1) or a Myc-Cut-HA protein starting at a.a. 878 (MCH 878). (A) Cytoplasmic extracts were prepared and processed to measure luciferase activity. Mean of 6 transfections are shown and the results are expressed as relative light units (RLU) normalized to b-galactosidase activity from an internal control. (B) Total extracts were prepared in parallel and analyzed in Western blots with anti-HA antibodies.

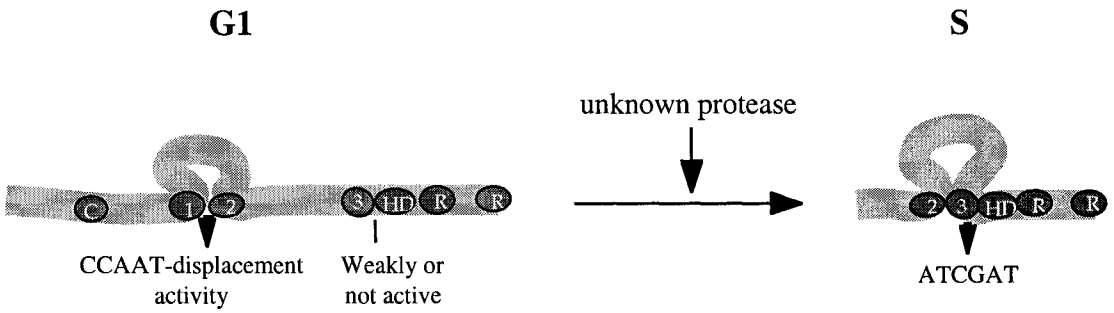


Fig. 9 Modulation of CDP/Cut at the G1/S transition

In early G1, CDP/Cut exists as a full-length protein that cannot stably bind to the ATCGAT sequence. As cells reach S phase, CDP/Cut is proteolytically processed into an amino-truncated 110-kDa protein in which the CR3HD bipartite DNA binding domain is fully active.

Chapter IV Expression of N-Terminally Truncated Isoforms of CDP/Cux Is Increased in Human Uterine Leiomyomas.

Preface

The previous chapter described a cell cycle dependent site specific proteolytic event that produces a N-terminally truncated isoform of CDP/Cux. p110 CDP/Cux. p110 CDP/Cux is different from full length CDP/Cux, p200 CDP/Cux, that it is capable of inter-acting DNA more stably and has a different effect on transcription of a gene. In particular, p110 CDP/Cux was able to activate the transcription of DNA polymerase α subunit while p200 can not, suggesting that p110 CDP/Cux might contribute to the G1 to S phase transition. The activity of many genes that promote the G1 to S phase transition, such as c-Myc and Cyclin D, are often upregulated in human cancer. The following chapter describes an observation that was made from human uterine Leiomyomas. It demonstrates that the cells that over-expresses N-terminally truncated isoforms of CDP/Cux are selected for during tumorigenesis.

Abstract

Genetic analyses and mRNA expression studies have implicated CUTL1 as a candidate tumor suppressor gene in uterine leiomyomas and breast cancers. However, the modulation of CDP/Cux, the protein encoded by CUTL1, does not agree with this notion. The activity of CDP/Cux, which is the DNA binding subunit of the histone nuclear factor D (HiNF-D), was shown to be upregulated as normal cells progress into S phase and constitutively elevated in several tumor cell lines. Recently, activation of CDP/Cux at the G1/S transition was found to involve the proteolytic processing of the protein to generate a shorter isoform. Uterine leiomyomas represent a unique reagent for molecular analysis because they are resected as homogenous tumor tissue together with the adjacent normal myometrium and they are often very large. In the present study, proteins were isolated from 16 pairs of matched tumors and adjacent myometrium and analyzed by Western blot analysis and electrophoretic mobility shift assays. Strikingly, in 11 out of 16 tumors, the steady-state level of small CDP/Cux isoforms was increased as compared to the normal control tissue. Where tested, a corresponding increase in CDP/Cux stable DNA binding activity was observed. DNA sequencing analysis of CUTL1 cDNAs from 6 leiomyomas, including 4 with LOH of CUTL1, did not reveal any gross rearrangement nor point mutations. Altogether these findings suggest that CUTL1 is probably not the tumor suppressor on 7q22. Moreover, the frequent increase in smaller CDP/Cux isoforms indicates that molecular events associated with the truncation of CDP/Cux proteins may be selected in uterine leiomyomas.

Introduction

Uterine leiomyomas are benign tumors that arise from smooth muscle cells of the myometrium [1, 2]. It is estimated that between 20 to 30% of women over the age of 30 will develop a leiomyoma [1, 3]. Approximately half of the uterine leiomyomas exhibit cytogenetic abnormalities. The most frequent genetic alterations are translocations involving the 12q15 region and deletions within the large arm of chromosome 7[1, 4]. The 12q15 chromosome target was identified as the high mobility group protein gene HMGI-C [5, 6]. Following translocation, the HMG1-C gene is truncated and fused to various ectopic DNA sequences to create a fusion gene [5]. Deletions on chromosome 7q, del(7q), were found in approximately 35% of studied cases with cytogenetic abnormalities (128/366), and the smallest commonly deleted region of 7q was mapped to band 7q22 [1, 7-10]. The high proportion of cytogenetically detectable deletions of 7q22 in leiomyomas and other types of cancers suggests that a tumor suppressor gene may be located within this chromosomal region [1, 9]. We have previously shown that CUTL1 is present in a commonly deleted region in 7 out of 50 uterine leiomyomas samples examined [9]. Furthermore, Northern blot analysis revealed that CUTL1 mRNA levels were reduced in eight tumors out of thirteen [9]. Loss-of-heterozygosity of CUTL1 was also reported in 12 out of 66 sporadic breast cancers [11]. Altogether these results raised the possibility that CUTL1 may be the tumor suppressor gene on 7q22. However, this notion is in contradiction with DNA binding studies in tissue culture systems. CUTL1 encodes for the CDP/Cux protein, which was characterized independently as the CCAAT-displacement protein (CDP) and the DNA binding sub-unit of the histone nuclear factor D (HiNF-D) [12-14]. The DNA binding activity of HiNF-D was shown to be upregulated in S phase in normal cells, but to be constitutively activated in

tumor cell lines [15-18]. CDP/Cux is a transcription factor that contains four DNA binding domains and whose expression has been associated with cellular proliferation, the repression of genes that are turned on in terminally differentiated cells, and the regulation of matrix attachment regions (MARs) [19]. In addition, a number of evidence points to a role of CDP/Cux in cell cycle progression. CDP/Cux was found to be a component of the histone nuclear factor D (HiNF-D), whose presence on various histone promoters coincides with their upregulation in the cell cycle. [12, 13, 15, 20, 21]. Moreover, one of CDP/Cux DNA binding activities was shown to be upregulated as normal cells progress into S phase, and inhibited in G2 following the phosphorylation of the Cut homeodomain by cyclin A/CDK1 [15, 22, 23]. More recently, the upregulation of CDP/Cux DNA binding activity at the G1/S transition was found to involve the proteolytic processing of p200 CDP/Cux into an N-terminally truncated isoform of 110 kDa [24].

From the point of view of the experimentalist, uterine leiomyomas represent an excellent tumor system for molecular investigation. These tumors occur very frequently and are always resected together with abundant adjacent normal myometrium which can be used as control tissues in molecular studies. Moreover, leiomyomas present themselves as homogenous tumor tissue, with little or no infiltration by normal cells such as macrophages. This is best illustrated in LOH studies: in contrast to what is observed with other types of tumors, those leiomyomas with LOH of a particular marker show no or very weak signal for the lost allele [9]. Finally, leiomyomas are sometimes very large, a feature that makes it possible to isolate RNA and proteins from the same tumor. The expression pattern of candidate tumor suppressor genes can therefore be investigated directly in the tumor and the adjacent normal tissue. In the present study, we took advantage of these

features to investigate CDP/Cux protein expression and activity in uterine leiomyomas, including 2 with LOH of CUTL1. In parallel, we cloned CUTL1 cDNAs from leiomyomas with and without LOH and investigated whether these transcripts carried short deletions or point mutations.

Materials and Methods

Preparation of Total Protein Extracts from Tumors

Uterine leiomyomas and matched normal tissue samples were obtained at the time of hysterectomy in the Royal Victoria Hospital, Montreal, Canada, according to the guidelines of the Canadian Health Research Institute. All patients have signed an inform consent approved by the local Ethics Committees. Specimens were grossly and then microscopically examined by a pathologist and were classified as benign leiomyomas. No cellular atypia, increased mitotic activity, or necrosis was identified in any of the leiomyomas. Tissue samples measuring 0.5 cubic cm were collected from each uterus: one from the central area of the largest leiomyoma and the second from the grossly unaffected myometrium situated more than 2 cm away from the leiomyoma capsule. Tissue samples were snap-frozen in liquid nitrogen and were kept at –80 °C until protein extraction. To make protein extracts, tissues were grinded to small powder using mortar and pestle in the presence of liquid nitrogen. Grinded tissues were resuspended in buffer X (50 mM Hepes, pH 7.9, 0.4 M KCl, 4 mM NaF, 4 mM Na_3VO_4, 0.2 mM EGTA, 0.2 mM EDTA, 0.1% NP-40, 10 % glycerol, 0.5 mM DTT, Protease inhibitor mix tablet from Roche) and incubated at 4 °C for 30 min. The resulting slurry was centrifuged for 15 min at 4 °C and the supernatant was collected.

Western Blot Analysis

For each sample, 100 μg of protein was loaded on an 8% polyacrylamide-SDS gel. Following electrophoresis, CDP/Cux isoforms were visualized by immunobloting using α861 antibody as previously described [24]. To verify equal loading of protein extracts, 10 μg

of each protein extracts were analyzed by Western blotting using smooth muscle specific actin antibody (Neomarker, cat. #: MS-113; clone 1A4).

Electrophoretic Mobility Shift Assays

For electrophoretic mobility shift assays (EMSA), nuclear extracts from selected samples of pairs of uterine leiomyomas and matched normal myometrium were prepared as follows. Grinded tissue powders were submitted to 3 freeze/thaw cycles in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT, Protease inhibitor mix tablet from Roche). After a brief spin, nuclei were resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM $MgCl_2$, 420 mM $NaCl_2$, 0.2 mM EDTA, Protease inhibitor mix tablet from Roche) and incubated at 4 °C for 30 min. After 15 min. of centrifugation at maximum speed, the supernatant was collected. 10 μg of nuclear extracts were analyzed in EMSA with oligonucleotides containing a CDP/Cux consensus-binding site (CGATATCGAT). For supershift experiments, the DNA and protein mixtures of one selected sample was incubated with the indicated antibodies.

RNA isolation, cDNA cloning and sequencing

RNA was prepared using the procedure described by Chomczynski and Sacchi (1987) [25]. 5 μg of RNA was used in reverse transcription with Superscript reverse transcriptase (InVitrogen) using as primers either oligo dT or nt 3097-3080 of CDP (NCBI accession number M74099). The reaction was stopped by boiling for 10 minutes. An aliquot of 2 μl was used as template for polymerase chain reaction (PCR) using the Elongase Enzyme mix (InVitrogen) and primers F1 and B1 or F2 and B2. F1, nt 24-49; B1, 2805-2788; F2, 2695-2713; B2, 4567-4547. PCR was performed in a final volume of 100 μl, containing approximately 1 ng cDNA, 1.5 mM $MgCl_2$, 10 μl standard 10x PCR buffer (200 mM Tris-

HCl pH 8.4, 500 mM KCl), 0.45 μM of each primer, 0.2 mM dNTPs, 5% DMSO, and 2.5 units of Elongase Enzyme mix (InVitrogen). An initial step of 4 minutes at 95°C was followed by 30 cycles of 45s of denaturation at 95°C, 45s of annealing at 53°C, and 3 minutes of extension at 68°C, followed by a final extension step of 7 minutes at 68°C. The amplified fragments were purified on agarose gel . An aliquot was saved for cloning into the pCRII plasmid (InVitrogen). For each ligation, three independent clones were expanded and analyzed by DNA sequencing. In parallel, the rest of the amplified fragment was used for direct DNA sequencing using oligonucleotide primers situated approximately 400 nt apart from each other.

Results

Total protein extracts from 16 pairs of uterine leiomyomas and matched normal myometrium were examined by Western blot analysis using α861 anti-CDP/Cux antibodies (Fig. 1; see Fig. 2 for a diagram of the proteins and antibodies). These included two leiomyomas with LOH of CUTL1, #68 and #98. To control for equal loading, the samples were analyzed in parallel for expression of the smooth muscle specific actin. The first observation that can be made is that CDP/Cux proteins were expressed in all tumor samples. In most samples, three bands corresponding to three isoforms could be distinguished: p200, p110 and p100. The p200 and p110 have previously been characterized in various cell lines [19, 24]. However, an isoform with an apparent molecular weight of 100 kDa has not yet been reported. The p100 protein is an amino-terminally truncated isoform of CDP/Cux since it is recognized by the 861 and 1300 but not by the N-term antibodies (Fig. 1 and data not shown). Since p100 has not been observed elsewhere, we consider the possibility that it represents a tissue or cell type-specific isoform of CDP/Cux. Since we have not found an alternatively spliced transcript that could account for its production, p100 is probably generated by proteolytic cleavage of p200, like we have shown for p110 [24]. We are currently characterizing this isoform to determine the exact positions of its amino- and carboxy-termini. Comparison between the normal and tumor sample within each pair of samples revealed that the steady-state level of the full length CDP/Cux isoform, p200, was reduced in 6, equal in 1 and increased in 9 out of 16 samples. Among the two samples with LOH, 68 and 98, p200 was increased in one and reduced in the other. More strikingly, the shorter isoforms, p110 and p100, were present at higher level in 11 out of 16 samples. These comparisons are presented in a tabulated form in

Tables 1 and 2. We conclude that there is no obvious selection against CDP/Cux protein expression in uterine leiomyomas. In fact, in most patients there is an increase in the steady-state of CDP/Cux proteins, in particular the processed isoforms of 110 and 100 kDa.

Electrophoretic mobility shift assays (EMSA) were performed using oligonucleotides containing a CDP/Cux consensus binding site to verify whether DNA binding activity could be correlated with CDP/Cux steady-state levels (Fig. 2) . DNA binding assays could not successfully be performed with total protein extracts. Therefore, nuclear extracts were prepared from a number of sample pairs for which more material was available. To confirm the specificity of retarded complexes, the reaction was incubated in the presence of various CDP/Cux antibodies as previously described [22, 24]. The region recognized by each antibody is indicated on the diagram (Fig. 2). As previously shown, the antibody raised against the amino-terminal portion of the protein, N-term, supershifted only the higher retarded complex since this antibody recognizes the full length protein but not the processed isoforms (Fig. 2, lane 11)[24]. In contrast, antibodies raised against more carboxy-terminally located epitopes, 861, 1300 and C-term, were able to supershift both the higher and lower retarded complexes, albeit to a variable extent (Fig. 2, lane 12 to 14). These results confirm that the retarded complexes observed with the CDP/Cux consensus binding site involve both the full length and processed CDP/Cux isoforms. In EMSA with 4 pairs of samples, the intensity of the higher and lower retarded complexes correlated perfectly with the variations in CDP/Cux steady-state levels. The tumor with reduced CDP/Cux proteins, 204, also displayed less intense retarded complexes (Fig. 2, lane 6 and 7), whereas more intense retarded complexes were observed with tumors expressing higher levels of CDP/Cux proteins: 401, 59 and 402 (Fig. 2, lane 2 to 5, 8 and 9). We conclude

that the steady-state level of CDP/Cux proteins as judged from Western blot analysis provides a correct estimate of the protein activity in tumor samples.

CUTL1 mRNA sequences from 6 pairs of samples were analyzed for the presence of mutations: 66, 67, 68, 98, 100 and 402. The first four had LOH of CUTL1, whereas the last two had no LOH. Patients 98, 100 and 402 showed higher, whereas patient 68 had lower, CDP/Cux protein expression in the tumor than in the adjacent myometrium (Fig. 1). The status of CDP/Cux protein expression in patients 66 and 67 is not known because the small size of their tumors did not allow us to prepare protein extracts. The cDNAs were cloned by reverse-transcriptase polymerase chain reaction (RT-PCR) using the primers depicted in Fig. 3, and the DNA sequence was analyzed by two methods: either PCR-sequencing or isolation of individual cDNA clones followed by DNA sequencing of 3 individual clones for each leiomyoma. All amplified cDNA fragments were of the expected molecular weight, indicating that gross molecular rearrangements had not taken place. Moreover, no point mutation or short deletion was identified in any tumor. These results are in agreement with those from a similar study in ovarian cancers [26].

Discussion

Overall, our results indicate that CDP/Cux proteins are expressed in 16 uterine leiomyomas, including 2 tumors with LOH of the CUTL1 gene. From our sequencing analysis of 6 leiomyomas, including 4 tumors with LOH, it appears that the coding region of the CUTL1 gene does not carry mutations. We cannot exclude the possibility that mutations are present within non-coding regions at the 5' or 3' end of the mRNA, or within regulatory sequences that control the expression of the gene. However, if these mutations exist, they do not affect the overall level of CDP/Cux protein expression. In fact, one surprising findings of this study was that a majority of uterine leiomyomas were found to express a higher level of CDP/Cux proteins. In particular, the proteolytically processed isoforms of CDP/Cux were found to be expressed at higher level in 11 out of 16 leiomyomas. Since proteolytic processing of CDP/Cux has been shown to be activated as normal fibroblastic cells progress into the S phase of the cell cycle, it is tempting to speculate that the elevated levels of CDP/Cux processed isoforms in uterine leiomyomas is linked to the high proliferation index of these tumors. Whether activated CDP/Cux processing in uterine leiomyomas is only the consequence of their proliferative capacity or whether it plays a causative role in this process remains to be investigated. In conclusion, our results suggested that CUTL1 is not the tumor suppressor gene on 7q22. In this regards, it should be noted that CUTL1 covers a very large distance, over 470 Kbp, and that several putative genes have already been identified within this chromosomal region [27]. Our results also revealed that molecular events leading to the activation of CDP/Cux processing may be selected in uterine leiomyomas.

Acknowledgments

A.N. is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec. N.M.S. is the recipient of a fellowship from the Canadian Institute for Health Research. This research was supported by a grant from the Canadian Breast Cancer Research Initiative to A.N.

References

1. Ozisik YY, Meloni AM, Surti U, Sandberg AA. Deletion 7q22 in uterine leiomyoma. A cytogenetic review. *Cancer Genetics & Cytogenetics* 1993;71(1):1-6.
2. Andersen J. Growth factors and cytokines in uterine leiomyomas. *Semin Reprod Endocrinol* 1996;14(3):269-82.
3. Wilcox LS, Koonin LM, Pokras R, Strauss LT, Xia Z, Peterson HB. Hysterectomy in the United States, 1988-1990. *Obstetrics & Gynecology* 1994;83(4):549-55.
4. Dei Tos AP, Dal Cin P. The role of cytogenetics in the classification of soft tissue tumours. *Virchows Arch* 1997;431(2):83-94.
5. Hennig Y, Wanschura S, Deichert U, Bartnitzke S, Bullerdiek J. Rearrangements of the high mobility group protein family genes and the molecular genetic origin of uterine leiomyomas and endometrial polyps. *Mol Hum Reprod* 1996;2(4):277-83.
6. Schoenmakers EF, Wanschura S, Mols R, Bullerdiek J, Van den Berghe H, Van de Ven WJ. Recurrent rearrangements in the high mobility group protein gene, HMGI-C, in benign mesenchymal tumours. *Nat Genet* 1995;10(4):436-44.
7. Ishwad CS, Ferrell RE, Davare J, Meloni AM, Sandberg AA, Surti U. Molecular and cytogenetic analysis of chromosome 7 in uterine leiomyomas. *Genes Chromosomes Cancer* 1995;14(1):51-5.
8. Ishwad CS, Ferrell RE, Hanley K, Davare J, Meloni AM, Sandberg AA, et al. Two discrete regions of deletion at 7q in uterine leiomyomas. *Genes, Chromosomes & Cancer* 1997;19(3):156-60.

9. Zeng WR, Scherer SW, Koutsilieris M, Huizenga JJ, Filteau F, Tsui LC, et al. Loss of heterozygosity and reduced expression of the Cutl1 gene In uterine leiomyomas. *Oncogene* 1997;14(19):2355-65.

10. Sargent MS, Weremowicz S, Rein MS, Morton CC. Translocations in 7q22 define a critical region in uterine leiomyomata. *Cancer Genetics & Cytogenetics* 1994;77(1):65-8.

11. Zeng WR, Watson P, Lin J, Jothy S, Lidereau R, Park M, et al. Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene. *Oncogene* 1999;18(11):2015-21.

12. van Wijnen AJ, Wright KL, Lian JB, Stein JL, Stein GS. Human H4 histone gene transcription requires the proliferation-specific nuclear factor HiNF-D. Auxiliary roles for HiNF-C (Sp1-like) and HiNF-A (high mobility group-like). *J Biol Chem* 1989;264(25):15034-42.

13. van Wijnen AJ, van Gurp MF, de Ridder MC, Tufarelli C, Last TJ, Birnbaum M, et al. CDP/cut is the DNA-binding subunit of histone gene transcription factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93(21):11516-21.

14. Neufeld EJ, Skalnik DG, Lievens PM, Orkin SH. Human CCAAT displacement protein is homologous to the Drosophila homeoprotein, cut. *Nature Genetics* 1992;1(1):50-5.

15. Holthuis J, Owen TA, van Wijnen AJ, Wright KL, Ramsey-Ewing A, Kennedy MB, et al. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. *Science* 1990;247(4949 Pt 1):1454-7.

16. Wright KL, Dell'Orco RT, van Wijnen AJ, Stein JL, Stein GS. Multiple mechanisms regulate the proliferation-specific histone gene transcription factor HiNF-D in normal human diploid fibroblasts. *Biochemistry* 1992;31(10):2812-8.

17. van Wijnen AJ, Cooper C, Odgren P, Aziz F, De Luca A, Shakoori RA, et al. Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells. *Journal of Cellular Biochemistry* 1997;66(4):512-23.

18. van Wijnen AJ, Choi TK, Owen TA, Wright KL, Lian JB, Jaenisch R, et al. Involvement of the cell cycle-regulated nuclear factor HiNF-D in cell growth control of a human H4 histone gene during hepatic development in transgenic mice. *Proc Natl Acad Sci U S A* 1991;88(6):2573-7.

19. Nepveu A. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. *Gene* 2001;270(1-2):1-15.

20. el-Hodiri HM, Perry M. Interaction of the CCAAT displacement protein with shared regulatory elements required for transcription of paired histone genes. *Molecular & Cellular Biology* 1995;15(7):3587-96.

21. van den Ent FM, van Wijnen AJ, Lian JB, Stein JL, Stein GS. Cell cycle controlled histone H1, H3, and H4 genes share unusual arrangements of recognition motifs for

HiNF-D supporting a coordinate promoter binding mechanism. *J Cell Physiol* 1994;159(3):515-30.

22. Santaguida M, Ding Q, Berube G, Truscott M, Whyte P, Nepveu A. Phosphorylation of the CCAAT Displacement Protein (CDP)/Cux Transcription Factor by Cyclin A-Cdk1 Modulates Its DNA Binding Activity in G2. *J Biol Chem* 2001;276(49):45780-90.

23. Coqueret O, Berube G, Nepveu A. The mammalian Cut homeodomain protein functions as a cell-cycle dependent transcriptional repressor which downmodulates p21WAF1/CIP1/SDI1 in S phase. *EMBO Journal* 1998;17(16):4680-94.

24. Moon NS, Premdas P, Truscott M, Leduy L, Berube G, Nepveu A. S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by the CDP/Cut Homeodomain Protein. *Molecular & Cellular Biology* 2001;21:6332-45.

25. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analyt Biochem* 1987;162:156-59.

26. Neville PJ, Thomas N, Campbell IG. Loss of heterozygosity at 7q22 and mutation analysis of the CDP gene in human epithelial ovarian tumors. *Int J Cancer* 2001;91(3):345-9.

27. Glockner G, Scherer S, Schattevoy R, Boright A, Weber J, Tsui LC, et al. Large-scale sequencing of two regions in human chromosome 7q22: analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci reveals 17 genes. *Genome Res* 1998;8(10):1060-73.

28. Zeng RW, Soucie E, Sung Moon N, Martin-Soudant N, Bérubé G, Leduy L, et al. Exon/intron structure and alternative transcripts of the CUTL1 gene. *Gene* 2000;241(1):75-85.

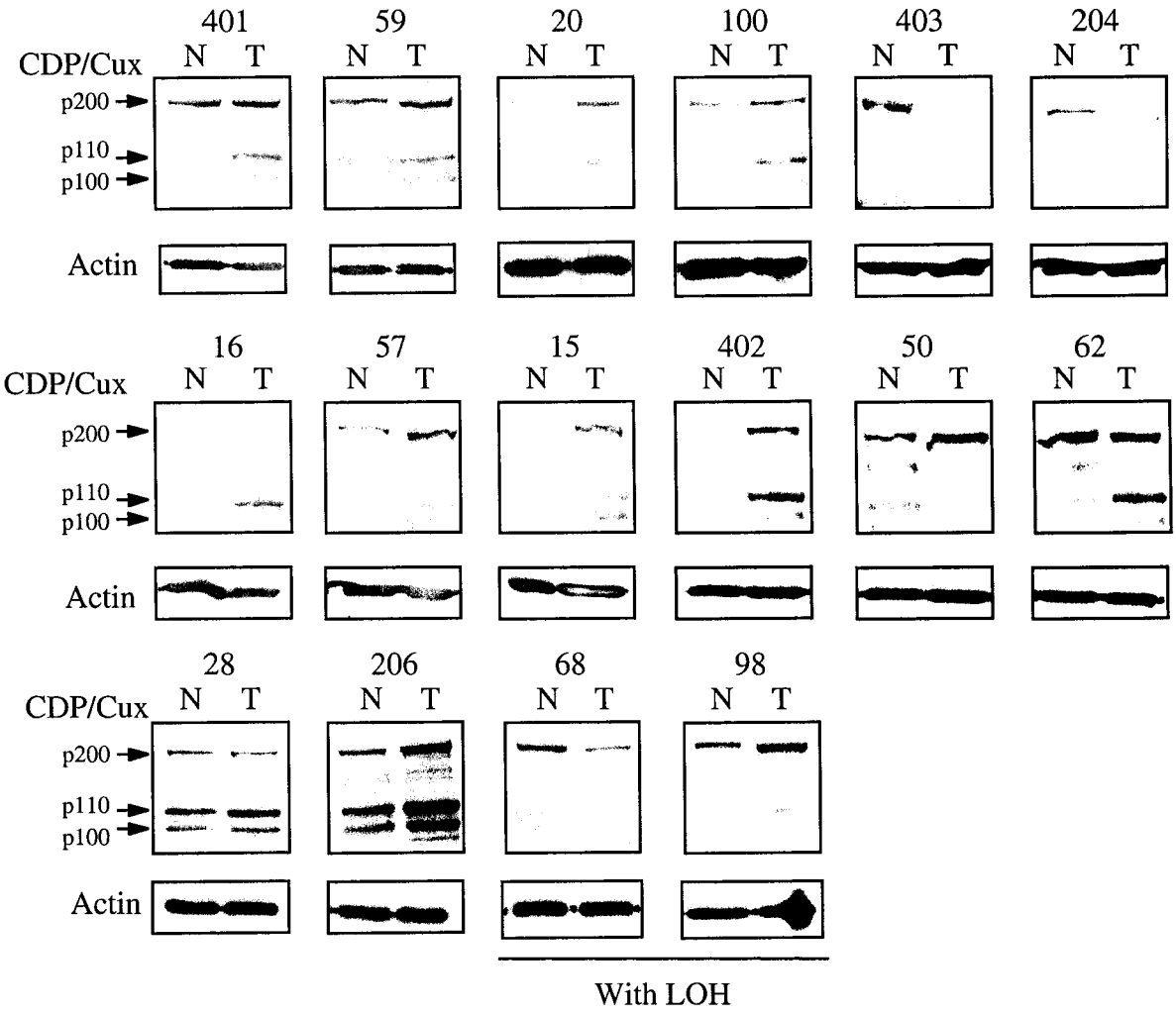


Figure1 Expression of CDP/Cux proteins in uterine leiomyomas and matched normal myometrium.

Total protein extracts from 16 pairs of uterine leiomyomas and matched normal myometrium were prepared and analyzed by Western blot analysis using α861 anti-CDP/Cux antibodies. Leiomyomas #68 and #98 have previously been shown to exhibit LOH of CUTL1 [9].

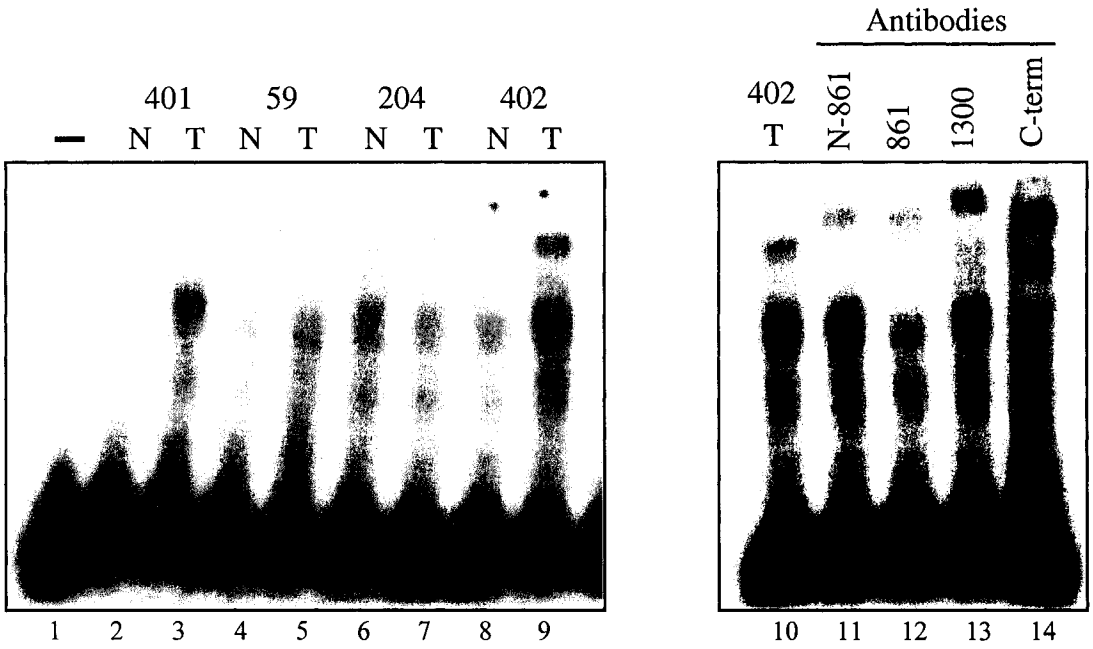


Fig. 2 EMSA of in uterine leiomyomas and matched normal myometrium using a CDP/Cux consensus binding site.

Nuclear extracts from selected samples of pairs of uterine leiomyomas and matched normal myometrium were prepared as described in Materials and Methods and analyzed in EMSA with oligonucleotides containing a CDP/Cut consensus-binding site (CGATATCGAT). The DNA and proteins were incubated either with no antibody (lanes 1-10), or with the indicated antibodies (lanes 11 to 4). The diagram shows a schematic representation of the CDP/Cux proteins with their conserved domains and the regions recognized by the respective antibodies [22, 24]. CC, coiled-coil; CR1, CR2 and CR3: Cut repeats 1, 2 and 3; HD, homeodomain; R, repression domain. The amino- and carboxy-termini of the p100 isoform remain to be precisely defined.

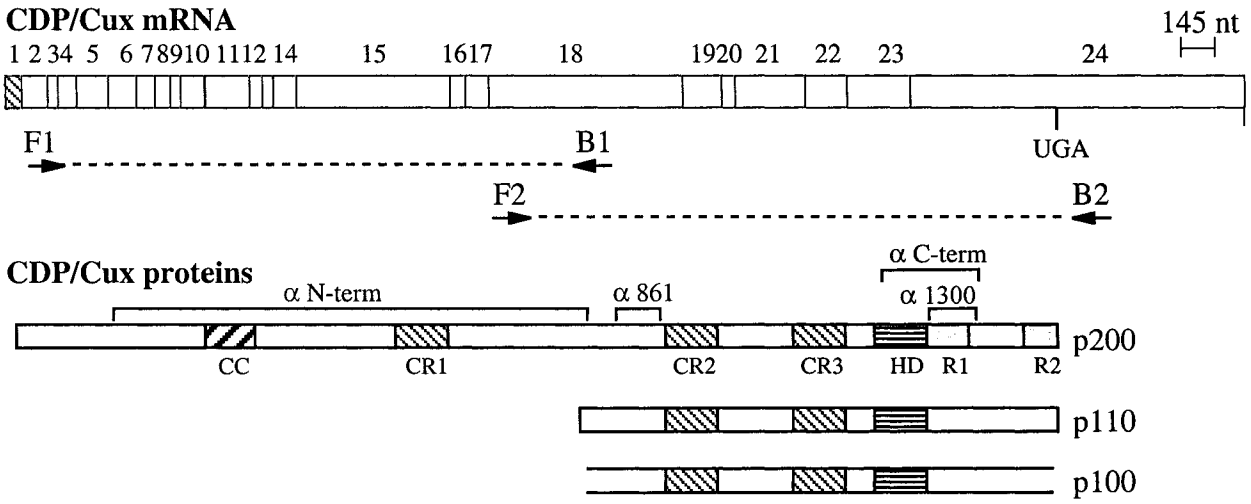


Fig. 3 DNA sequence analysis of CDP/Cux cDNA isolated from uterine leiomyomas with CUTL1 LOH

The diagram shows a schematic representation of the CDP/Cux mRNA with its 24 exons[28], and the primers that were used for cDNA cloning. CDP/Cux cDNA fragments were generated by RT-PCR using primers F1-B1 and F2-B2, as indicated. The cDNA fragments were either cloned into the pCRII plasmid or used as substrates in PCR-sequencing using a second set of primers. For each cDNA fragment cloned, three individual clones were analyzed by DNA sequencing. Similar results were obtained from PCR-sequencing and cDNA cloning and sequencing: no mutation was found.

Table 1 CDP/Cux protein expression and processing

The expression of each CDP/Cux isoform has been measured semi-quantitatively using a densitometer. For each pair of leiomyoma and matched normal tissue, we have indicated the ratio of CDP/Cux proteins in tumor versus the normal sample. This comparison was done either for all CDP/Cux isoforms (p200 + p110 + p100), the full length isoform only (p200), or the processed isoforms (p110 + p100).

Patient	Ratio of CDP/Cux proteins (Tumor/Normal)		
	p200 + p110 + p100	**p200**	**p100 + p110**
401	2.1	1.1	>10
59	2.1	1.4	6.9
20	5.8	4.4	>10
100	3.2	1.9	6.6
403	0.1	0.1	0.2
204	0.6	0.3	0.9
16	2.6	0.2	4.4
57	2.0	3.0	1.4
15	>10	>10	>10
402	4.3	7.7	3.4
50	1.0	1.4	0.5
62	1.1	0.7	1.6
28	1.1	0.7	1.2
206	2.2	2.5	2.2
68	0.4	0.4	0.3
98	3.1	2.5	>10

Table 2 Comparative CDP/Cux protein expression in uterine leiomyomas and matched myometrium

CDP/Cux Proteins	Tumors with higher expression	Tumors with equal expression	Tumors with lower expression
p200	9	1	6
p110 + p100	11	2	3

Chapter V Phosphorylation of CDP/Cux by Cyclin D/Cdk4 in G1 Inhibits Its Proteolytic Processing and DNA binding Activity.

Preface

In the chapter III, the cell cycle dependent site-specific proteolysis that cleaves full length CDP/Cux into p110 CDP/Cux has been described. So far, there are two events identified that regulate the stable DNA binding activity of CDP/Cux in a cell cycle dependent manner. One is the previously described phosphorylation in the region of the Cut homeodomain and the other is the processing event. Both events control the activity of CR3HD in a manner that it is only active in S phase. Even though, the consequence of the processing is clear, the identity of the protease and the mechanism regulating the proteolytic event have not been identified. One way to regulate the cell cycle dependent processing is to regulate the activity of the proteolysis. Alternatively, a cell cycle dependent modification on the substrate, CDP/Cux, could regulate the processing event. In this chapter, the phosphorylation by cyclin D/Cdk4 that seems to control both cell cycle dependent DNA binding activity of CR3HD and the proteolysis that produce p110 CDP/Cux is described.

ABSTRACT

Two regulatory events contribute to stimulate a stable DNA binding activity of the CDP/Cux transcription factor at the G1/S transition of the cell cycle: dephosphorylation of the Cut homeodomain by the Cdc25A phosphatase and proteolytic cleavage of CDP/Cux p200 to generate the N-terminally truncated CDP/Cux p110 isoform. Inhibition of DNA binding in G1 was previously shown to involve the phosphorylation of the Cut homeodomain. Here, we present evidence showing that cyclin D/Cdk4 interacts with CDP/Cux and phosphorylates it on serine residues. Phosphorylation of two of those, serines 1237 and 1270, was sufficient to inhibit DNA binding by a recombinant CDP/Cux p110 protein. In addition, over-expression of cyclin D/Cdk4 also inhibited proteolytic processing of CDP/Cux p200. This inhibition was observed only with CDP/Cux substrates that contained the Cut homeodomain, which is necessary for the interaction with cyclin D/Cdk4. CDP/Cux mutants were generated in which serines were replaced with alanine. Cyclin D/CDK4 inhibited the processing of most mutants except for one that contained 8 mutations. Whereas wild type CDP/Cux was not processed until S phase, this mutant was processed earlier, in G1. Altogether these results demonstrate that CDP/Cux is a target of cyclin D/Cdk4, which acts to delay its processing and inhibit its stable DNA binding activity until the end of G1. These findings introduce a novel concept in cell cycle regulation: cyclin-dependent kinases not only serve to stimulate the transition from one phase to another, but also must ensure that some events do not take place prematurely.

INTRODUCTION

The CDP/Cux/Cut transcription factors are a family of evolutionarily conserved homeodomain proteins that contain four DNA binding domains: three regions called Cut repeat 1, 2 and 3, and the Cut homeodomain (reviewed in (34)). In *Drosophila melanogaster*, *cut* has been implicated in the determination of cell-type identity in several tissues (6, 21, 26). One mutation preventing the function of a distant wing specific enhancer caused the formation of truncated or cut wings (7, 20). This is the phenotype that gave its name to the locus, *cut*. In vertebrates, the CCAAT-displacement activity was first characterized in sea urchin, early B cells and myeloid precursor cells (3, 39). Purification of this activity from HeLa cells and subsequent cloning of the corresponding cDNA revealed that the CCAAT-displacement protein, CDP, was the human ortholog of Cut (5, 35). Other cDNAs have been isolated from various species. In particular, the mouse cDNA was called Cux (Cut homeobox) in accordance with the nomenclature rules (41). The term CDP/Cux will be used in the remainder of this manuscript to designate the mammalian protein.

In the mouse, replacement of the Cut repeat 3 and/or the Cut homeodomain by gene targeting has revealed several phenotypes including partial neonatal lethality, growth retardation, delayed differentiation of lung epithelia, altered hair follicle morphogenesis, perturbed B and T lymphopoiesis and myeloid hyperplasia (15, 28, 38). In tissue culture, the expression and activity of CDP/Cux has been associated with cellular proliferation, the repression of genes that are turned on in terminally differentiated cells, the up- and down-regulation of genes encoding histones and the thymidine kinase, and the regulation of matrix attachment regions (MARs) (34). CDP/Cux proteins were found generally to function as transcriptional repressors (1, 10, 24, 25, 29, 39-41). However, certain results suggest, but

do not prove, that CDP/Cux may also be able to participate in gene activation (45, 47). Indeed, Cut was found to be a component of the promoter complex HiNF-D, which is believed to contribute to the transcriptional induction of several histone genes at the G1/S phase transition of the cell cycle (2, 44, 45). More recently, one isoform of CDP/Cux was found to stimulate expression of a reporter plasmid carrying the DNA pol α gene promoter (32). Whether this stimulation involved direct transcriptional activation by CDP/Cux or an indirect mechanism, via the down-modulation of a DNA pol α repressor, is currently under investigation.

A significant amount of evidence supports a role for CDP/Cux in cell cycle progression. Interaction of CDP/Cux with a consensus binding site or with histone H4 gene promoter sequences (as part of HiNF-D) was found to be upregulated as cells progress from G1 to S phase (8, 17). CDP/Cux was shown to bind to the core promoter of the $p21^{WAF1/CIP1}$ gene and in transient transfection assays, CDP/Cux repressed a $p21^{WAF1/CIP1}$ reporter whereas an antisense CDP/Cux construct was able to restore $p21^{WAF1/CIP1}$ expression levels in S phase (8). Increase in DNA binding at the G1/S transition involved two regulatory events: dephosphorylation of the Cut homeodomain by the Cdc25A phosphatase, and specific proteolytic cleavage of CDP/Cux between CR1 and CR2 to generate an amino-terminally truncated protein of 110 kDa (8, 32). Interestingly, CDP/Cux p200 was found to be incapable of stable interaction with DNA, whereas p110 could bind stably to DNA (31, 32). Thus, cell cycle-dependent processing of CDP/Cux serves to generate a p110 isoform with distinct DNA binding and transcriptional properties. More recently, a cyclin-Cdk complex that is present at the end of S phase and in G2, cyclin A/Cdk1, was shown to interact with p110 CDP/Cux, phosphorylate the Cut homeodomain and inhibit its

DNA binding activity (37). Overall, the existing data suggest that the activity of CDP/Cux is regulated throughout the cell cycle following post-translational modifications by kinases, phosphatases and a site-specific protease.

Cell cycle progression is controlled by a series of cyclin-dependent kinases (Cdk), phosphatases and Cdk inhibitors. Each phase of the cell cycle, and the transition between any two phases, is controlled by certain cyclin/Cdk complexes whose activation requires several steps (reviewed in (33)). While the mechanisms of Cdk regulation have been deciphered to some extent, less is known about the targets of cyclin/Cdks and their functions. Only a few targets have been identified, and for most cyclin/Cdks we still do not know what targets are necessary and sufficient for progression into the next phase of the cell cycle. For example, G1 Cdks, the cyclin D/Cdk4 and 6 and cyclin E/Cdk2, have been shown to phosphorylate Rb, triggering its release from the E2F transcription factor (11, 22). Yet, cyclin E was also found to induce the G1/S transition in an Rb-independent manner, suggesting that at least one other target is essential for progression into S phase (27). Cyclin D has been found to interact with a number of transcription factors including the estrogen receptor, DMP1, v-myb, MyoD and Stat3 (4, 16, 18, 19, 48, 49). However, so far only pRb has been shown to be phosphorylated by cyclin D/Cdk4 (reviewed in (12, 13)). Despite the abundant literature regarding the regulation of pRb during cell cycle progression in normal and cancer cells, whether or not the role cyclin D/Cdk complexes is limited to the phosphorylation of pRb is currently unknown.

The present study was inspired by our previous findings regarding the phosphorylation of a recombinant CDP/Cux protein containing the Cut repeat 3 and the Cut homeodomain (CR3HD) during the G1 phase of the cell cycle (8). In light of the accumulating

evidence linking the modulation of CDP/Cux with cell cycle progression, we considered that a G1-specific Cdk complex might be responsible for modulating the DNA binding activity of CDP/Cux during this period of the cell cycle. Indeed, we obtained results confirming that cyclin D/Cdk4 can interact with CDP/Cux and inhibit DNA binding of a recombinant protein containing CR3HD. However, we found that the effect of cyclin D/Cdk4 on CDP/Cux was more complex than originally anticipated. Cyclin D/Cdk4 was able to phosphorylate CDP/Cux at several positions upstream of CR3HD, and phosphorylation of CDP/Cux was associated with the lack of proteolytic processing. These results suggests a novel role for cyclin D/Cdk4 in ensuring that some events that take place at the onset of S phase do not occur prematurely during the G1 phase.

MATERIALS AND METHODS

Cell Culture and Synchronization.

NIH3T3 cells were grown in DMEM medium supplemented with 10% fetal bovine serum (FBS). To obtain cells in G0 using the serum starvation/stimulation method, NIH3T3 cells were maintained for 2 to 3 days in DMEM plus 0.4% FBS. The medium was then changed for DMEM plus 10% FBS and cells were harvested either 9 hours later for mid-G1 or 16 hours for S phase. However, it should be noted that the time required to reach S phase varies with the length of time cells were maintained in medium with low serum. The longer this incubation, the longer it takes for cells to reach S phase upon re-addition of serum. In practice, cell cycle progression must be monitored by FACS analysis for each experiment. Synchronization in G1/S was performed using the single thymidine block procedure. Cells were cultured overnight in DMEM plus 10% FBS supplemented with 2 mM thymidine. FACS analysis was performed as previously described (8)

Preparation of Nuclear extract

The nuclei were obtained by submitting cells to 3 freeze/thaw cycles in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM DTT). After quick centrifugation, resulting nuclei were then resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM $MgCl_2$, 420 mM NaCl, 0.2 mM EDTA.) and incubated at 4 °C for 30 min. After 15 min. of centrifugation, the supernatant was collected. Buffers A and C were supplemented with protease inhibitor mix tablets purchased from Roche and with E-64 from Calbiochem.

Plasmid Construction

Plasmids for expression of CDP/Cux proteins in both bacteria and mammalians cells were prepared by inserting various fragments from human *CDP/Cux* (Genbank accession number NM 001913) cDNA into the specified vectors. GST-CDP/Cux constructs and the mammalian expression vector Myc-CDP/Cux-HA (MCH) were described previously (37). DNA sequence coding for amino acid 612-1328 and 522-1027 of CDP/Cux were cloned into a 5' 6XHis expression vector TriEx 2.1 (Novagen) with an HA tag at the carboxy terminus. The same TriEx vector were used for expression in mammalian and bacterial cells. Serine to alanine mutations were made at following amino acid position of CDP/Cux by site directed mutagenesis according to the method of Deng and Nickoloff (9): mutant A, S749; mutant B, S887, S909 and S914; mutant C, S1054 and S1059; mutant D, S1237 and S1270; mutant ABCD, S749, S887, S909, S914, S1054, S1059, S1237 and S1270. Dr. Bob Fisher kindly provided Baculovirus expressing Cyclin D and Cdk 4 and the Mammalian expression vectors for Cyclin D, Cdk 4. The dominant negative form of Cdk 4 and GST-C-term pRb were generous gift from Dr. Sander Van Den Heuvel and Dr. Nicholas Dyson, respectively.

Expression and Purification of Fusion Protein

PTriEx-CDP/Cux 612-1328 was introduced into the BL21(DE3) strain of *E. coli* and induced with 1 mM IPTG for 1.5 hours. Proteins were purified on Ni-NTA agarose (Qiagen) according to the manufacturer's instructions. Vectors encoding the indicated region of GST-CDP/Cux fusion proteins were introduced in the DH5 strain of *E. coli* and protein expression was induced with 1 mM IPTG for 1.5 hours. Proteins were purified on glutathione sepharose (Amersham Pharmacia Biotech) according to the manufacturer's instructions.

Antibodies

Antibodies against CDP/Cux were previously described (8). Anti -HA antibody was purchased from Covance (cat #139021002). Anti-Cyclin D1 and anti-Cdk 4 were purchased from Neo Markers (Cat. #MS-210 and Cat. #MS-229).

Immunoprecipitation

Rabbit or mouse antibodies were incubated with protein A sepharose beads (Gibco BRL) or protein G agarose beads (Gibco BRL), respectively, for 2 hours at 4°C. Beads were washed with lysis buffer 3 times. Cell lysates were then incubated with antibody-bound beads for 1 hour at 4°C followed by 5 washes in lysis buffer. Immunoprecipitated products were analyzed by SDS-PAGE. For co-immunoprecipitation, cells were lysed in buffer X (50 mM Hepes, pH 7.9, 0.4 M NaCl, 4 mM NaF, 4 mM $NaVO_3$, 0.2 mM EDTA, 0.2 mM EGTA, 0.1% NP-40, 10% glycerol, and protease inhibitor mix tablets (Roche)).

Electro Mobility Shift Assay (EMSA)

EMSA were performed with 1 μg of nuclear extract from transfected mammalian cells. The samples were incubated at room temperature for 5 min. in a final volume of 30 μl in 25 mM NaCl, 10 mM Tris, pH 7.5, 1 mM $MgCl_2$, 5 mM EDTA pH 8.0, 5 % of glycerol, 1 mM of DTT, with 60 ng of poly dIdC and 10 μg of BSA as nonspecific competitors. End-labeled double stranded oligonucleotides (~10 pg) were added and further incubated for 15 min. at room temperature. Samples were loaded on a 5% polyacrylamide (29:1),1X Tris/Glycine (100 mM Tris/ 575 mM Glycine/ 1.32 mM EDTA, pH 8.5) gel and separated by electrophoresis at 8V/cm in 0.5X Tris /Glycine. Gels were dried and visualized by autoradiography

GST Pull-down

Lysates from baculovirus Sf9 cells previously infected with baculovirus encoding cyclin D and Cdk 4 were mixed with glutathione beads bound with 1 μg of GST-CDP/Cux fusion proteins at 4°C for 1 hour in binding buffer (20 mM Tris pH 7.5, 0.05% NP40, 137 mM NaCl), washed 3 times in binding buffer and resolved by SDS-PAGE on a 10% acrylamide gel. Proteins were analyzed by western blotting with a monoclonal antibody for cyclin D or Cdk 4 (E2 3, Neo Markers) and a monoclonal antibody for HA (11, Covance).

Western blot

Western blot analysis for HA epitope, Cyclin D, Cdk 4 and CDP/Cux were done in an identical manner. Protein extracts were resolved by SDS-PAGE electrophoresis and transferred to PVDF membrane. The PVDF membrane was incubated with blocking solution (10 mM Tris pH8, 150 mM NaCl, 5% powder milk, 2% BSA) for 2 hours at room temperature. Primary antibodies were incubated with the membrane in TBST (10 mM Tris pH8, 150 mM NaCl , 0.1% Tween) for 1 hour at room temperature. After four 10 min. washes with TBST, secondary antibodies were added and incubated in TBST for 45 min. at room temperature. Following four 10 min. washes with TBST, proteins were visualized with the ECL system from Amersham.

***In vitro* Kinase Assay**

The method used was adapted from that of Meyerson and Harlow (30). Briefly, baculovirus-infected SF9 insect cells were lysed in kinase buffer and sonicated 3 times for 10 seconds. Cyclin D/Cdk 4 kinase complexes were isolated by immunoprecipitation using a Cdk4 monoclonal antibody (Ab-1, Neo Markers) for1 hour at room temperature. The precipitate

was washed with kinase buffer three times at 4°C. and 1 μg of proteins of interest was incubated with Cyclin D/Cdk4 in presence of 5 μCi ^{32}P γ-ATP (6000 Ci/ mmol) (Amersham Pharmacia Biotech) in kinase buffer (50 mM Hepes pH 7.5, 10 mM $MgCl_2$, 1 mM DTT, 1 mM NaF). Reactions were allowed to proceed for 10 minutes at 24°C and were terminated by adding 6 μl of SDS loading buffer and boiling for 5 minutes. The reaction mixtures were resolved by were resolved by SDS-PAGE electrophoresis on 8% acrylamide gel followed by coomassie staining. Phosphorylated proteins were visualized by autoradiography.

Phospho-amino acid analysis. 6His 612-1328 was phosphorylated by *in vitro* kinase assay as described. Phosphorylated 6His 612-1328 proteins were migrated on a 10% SDS-PAGE and electrophoretically transferred to PVDF membrane. The membrane was washed in a large volume of water, dried and rewetted with methanol and water and dried again. Membrane was exposed to X-ray film to visualize the phosphorylated species and cut out of membrane. Phospho-amino acid analysis was performed as previously described (8).

RESULTS

CDP/Cux interacts with cyclin D. We have previously shown that the DNA binding activity of a recombinant CDP/Cux protein containing CR3HD was inhibited during the G1 phase of the cell cycle (8). This inhibition correlated with serine phosphorylation within a 12 kDa segment containing the Cut homeodomain. Since the phosphorylation pattern of CDP/Cux displayed cell cycle dependence, we considered the possibility that the kinase phosphorylating CDP/Cux in G1 was in fact the cyclin-dependent kinase complex, cyclin D/Cdk4. As a first step to test this hypothesis, we performed co-immunoprecipitation assays using antibodies against cyclin D and CDP/Cux. Remarkably, an interaction between the two proteins was detected without having to over-express any of the proteins. Endogenous CDP/Cux was precipitated with cyclin D antibodies (Fig. 1A, lane 2). Reciprocally, endogenous cyclin D was brought down with CDP/Cux antibodies (Fig. 1A, lane 3). To determine the region of CDP/Cux that associates with cyclin D and Cdk4, pull-down assays were performed using a panel of GST-CDP/Cux fusion proteins and total extracts from Sf9 cells that had been infected with baculovirus expressing cyclin D and Cdk4. Both cyclin D and Cdk4 were pulled down by GST-CR3HD (Fig. 1B, lane 6). In addition, a weak interaction was observed with GST-CTD. It should be noted that these two fusion proteins share a region of overlap between a.a. 1299 and 1313 and that a cyclin interaction motif (CY), RRELF, has previously been mapped from a.a. 1299 to 1304 (37). This would explain why both fusion proteins interacted with cyclin D. The binding to GST-CR3HD was stronger, suggesting that the overall interaction was stabilized by the presence of the Cut homeodomain as was previously found for cyclin A/Cdk1 (37).

Phosphorylation of CDP/Cux by cyclin D/Cdk4 inhibits the DNA binding activity of CR3HD. Cyclin D/Cdk4 was found to phosphorylate the GST-CR3HD fusion protein *in vitro* (data not shown; see also Fig. 5). Phosphorylation within this region by cyclin A/Cdk1 was previously shown to inhibit DNA binding (37). We tested whether DNA binding by a recombinant protein corresponding to the proteolytically processed CDP/Cux isoform, p110, would be similarly inhibited upon co-expression with cyclin D/Cdk4. NIH3T3 cells were transfected with a vector expressing MCH878, a recombinant CDP/Cux protein containing amino acids 878-1505 and an influenza virus hematoglutinin (HA) tag at its carboxy-terminus (described in (32)). MCH878 was expressed either alone or with cyclin D and Cdk4 expressing vectors. Whereas the steady-state level of MCH878 was not affected (Fig. 2 A, left panel), its DNA binding activity was reduced in the presence of cyclin D/Cdk4 (Fig. 2 A, right panel). In contrast, no reduction in DNA binding was observed when the same experiment was repeated with a mutant, MCH878$^{S/A}$, in which serines 1237 and 1270 were replaced with alanines (Fig. 2B). These results indicate that cyclin D/Cdk4 can phosphorylate CDP/Cux in the region of the Cut homeodomain and inhibit its DNA binding activity.

Co-expression with cyclin D/Cdk 4 inhibits proteolytic processing of CDP/Cux. Proteolytic processing of CDP/Cux was previously shown to occur at the G1/S transition of the cell cycle (32). Thus, the CDP/Cux isoform expressed when cyclin D/Cdk4 is active in G1 is the full length CDP/Cux p200. We therefore tested the effect of cyclin D/Cdk4 on this CDP/Cux isoform. As expected, we observed a reduction in DNA binding (Fig. 3A, lanes 1-3). However, to our surprise, Western blot analysis revealed that the level of p110 was reduced in the presence of cyclin D/Cdk4 (Fig. 3A, lanes 4 and 5).

The reduction in p110 may have occurred for one of two reasons. Proteolytic processing of CDP/Cux p200 may have been inhibited, or the stability of CDP/Cux p110 may be reduced in cells over-expressing cyclin D/Cdk4. Since the steady-state level of the recombinant MCH878 protein was not affected by cyclin D/Cdk4 in earlier experiments (see Fig. 2A, lanes 1 and 2), the former possibility appears to be more likely. This notion is further confirmed by the finding that p90, the carboxy-terminal peptide generated as the result of CDP/Cux processing (32), is also less abundant in cells expressing cyclin D/Cdk4 (Fig. 3A, lanes 6 and 7). We then tested whether co-expression of CDP/Cux with a dominant negative mutant of Cdk4, $CDK4^{DN}$, would have an opposite effect on the processing of CDP/Cux from that of cyclin D/Cdk4. We observed no significant difference in CDP/Cux processing when this experiment was performed in NIH3T3 cells (data not shown; see also Fig. 4C, lanes 1 and 2). However, co-expression of $CDK4^{DN}$ in 293 cells lead to an increase in the level of the processed isoform, p110 (Fig. 3B, lanes 1 and 2). One possible explanation for the discrepancy between the results in NIH3T3 and 293 cells was provided from the Western blot analysis of cyclin D1 in each cell line. As shown in Fig. 3C, the steady-state level of cyclin D1 was much higher in NIH3T3 than in 293 cells. Since $CDK4^{DN}$ exerts its dominant negative effect by binding to cyclin D and displacing the endogenous Cdk4, it is possible that this effect is more easily observed in cells that express a lower level of cyclin D, as in 293 cells here. Altogether, results from Fig. 2 and 3 suggest that processing of CDP/Cux may be inhibited by cyclin D/Cdk4.

The CR3HD region is required for the inhibition of CDP/Cux processing by cyclin D/Cdk4. The potential mechanisms by which cyclin D/Cdk4 inhibits CDP/Cux processing can be broadly divided into two categories. Cyclin D/Cdk4 may inhibit the pro-

tease or alternatively, it may modify the CDP/Cux substrate in a manner that renders it refractory to processing. In either case, the effect may be direct or may involve several intermediary steps. One approach to distinguish between these alternative mechanisms is to investigate the processing of a CDP/Cux substrate that does not contain the sequences necessary for its interaction with cyclin D/Cdk4. A recombinant protein containing CR1, CR2 and the linker in between them, CDP/Cux 522-1027, was previously shown to be proteolytically processed in cells (32). We reasoned that since this protein does not contain the CR3HD region, it would not interact with cyclin D/Cdk4. If processing of this protein was still inhibited in the presence of cyclin D/Cdk4, it would suggest that the protease is targeted. On the contrary, no reduction in processing would suggest that the inhibition requires that cyclin D/Cdk4 interacts with CDP/Cux. NIH3T3 cells were transfected with a vector expressing CDP/Cux 522-1027, in the presence or absence of vectors expressing cyclin D/Cdk4. In Western blot analysis, the level of the processed product was not significantly affected by the presence of cyclin D/Cdk4 (Fig. 4A, lanes 1 and 2). We then tested a CDP/Cux substrate, CDP/Cux 612-1328, whose carboxy-terminus extends past the CR3HD. Importantly, this protein includes the CY motif present at a.a. 1299 to 1304. Processing of CDP/Cux 612-1328 protein was not as efficient in the presence of cyclin D and Cdk4 (Fig. 4B, lanes 1 and 2). These results indicate that sequences far downstream from the site of processing are required for the inhibition of processing by cyclin D/Cdk4. Moreover, the difference in electrophoretic mobility of the two recombinant proteins in the presence of cyclin D/Cdk 4 suggested that phosphorylation of CDP/Cux may be involved in the inhibition of processing. Indeed, the migration of CDP/Cux 522-1027 was not altered, whereas an additional band of slower mobility was observed when CDP/Cux 612-1328 was

expressed together with cyclin D/Cdk 4 (Fig. 4A and B, compare lanes 1 and 2). Importantly, both cyclin D and Cdk4 were required, as processing of CDP/Cux 612-1328 was not significantly affected when only cyclin D or Cdk4 were present (Fig. 4B, lanes 3 and 4). To obtain additional evidence that inhibition of processing was associated with the phosphorylation of CDP/Cux, we repeated the same experiment using either wild type Cdk4 or a catalytically inactive mutant of Cdk4, $Cdk4^{DN}$. Equal expression level of the two Cdk4 proteins was verified by Western blot analysis (data not shown). Processing of CDP/Cux 612-1328 was inhibited by cyclin D/Cdk4 but not by cyclin D/$Cdk4^{DN}$ (Fig. 4C, lanes 1-3). Moreover, the electrophoretic mobility of CDP/Cux 612-1328 was slightly retarded in the presence of cyclin D/Cdk4. Altogether, these results show that the inhibition of CDP/Cux processing by cyclin D/Cdk4 requires the presence of CR3HD, a region of CDP/Cux that interacts with cyclin D/Cdk4, and is potentially associated with the phosphorylation of CDP/Cux.

Cyclin D/Cdk 4 phosphorylates several serine residues of CDP/Cux *in vitro*. We investigated whether cyclin D/Cdk4 was able to phosphorylate CDP/Cux 612-1328 *in vitro*. Cyclin D/Cdk4 complexes isolated from baculovirus-infected Sf9 cells (see Material and Methods) were incubated together with a bacterially expressed CDP/Cux 612-1328 protein and radioactively labeled ATP. As controls for the efficiency and specificity of the kinase complex, other reactions were run in parallel using as substrates either BSA or a GST fusion protein carrying the carboxy-terminal domain of the retinoblastoma protein, GST/pRb-CTD (36). While BSA was not phosphorylated at all, CDP/Cux 612-1328 was phosphorylated to a higher level than that of GST/pRb-CTD (Fig. 5A, lanes 4, 5 and 6). Phospho-amino acid analysis revealed that phosphorylation of CDP/Cux 612-1328 oc-

curred on serine residues (Fig. 5B). To confirm that co-expression with cyclin D/Cdk4 can lead to an increase in the phosphorylation of CDP/Cux *in vivo*, transfected 293 cells were metabolically labeled with ^{32}P-orthophosphate, and the recombinant CDP/Cux 612-1328 was purified by affinity chromatography and revealed by coomassie blue staining or autoradiography (Fig. 5C, compare lanes 1 and 2). Examination of the amino acid sequence of CDP/Cux 612-1328 revealed the presence of 8 serine-proline (SP) residues that could serve as Cdk phosphorylation sites. We engineered 5 mutants in which specific serine residues were replaced with alanine either individually or as a group. The names and positions of the mutations are shown in a diagram (Fig. 5D). The mutant ABCD carried a mutation at each of the 8 putative phosphorylation sites. Kinase reactions were carried as previously with wild type CDP/Cux 612-1328 and mutants A, B, C, D and ABCD (Fig. 5D). The level of phosphorylation was significantly reduced in the case of mutants B, D and ABCD. These results indicate that cyclin D/Cdk4 can phosphorylate CDP/Cux at several positions. There was still detectable phosphorylation of mutant ABCD even though all putative Cdk phosphorylation sites containing the SP sequence had been removed. A number explanations could be evoked for this phenomenon. It is possible that another kinase was brought down when cyclin D/Cdk4 complexes were precipitated from insect cells with anti-cyclin D antibodies. However, we favor the possibility that cyclin D/Cdk4 displayed promiscuity in the *in vitro* kinase reaction, especially that only one potential substrate was available. In particular, we note that the sequence TP is present at five positions within CDP/Cux 612-1328. Although phospho-amino acid analysis of wild type CDP/Cux 612-1328 did not reveal the presence of phosphorylated threonine residues, it is possible that some of these sites may be targeted once all the serine phosphorylation sites have been eliminated.

Serine to alanine mutations within CDP/Cux prevent the inhibition of processing by cyclin D/ Cdk 4. We tested the effect of cyclin D/Cdk4 co-expression on the processing of CDP/Cux 612-1328 mutants. Cyclin D/Cdk4 inhibited the processing of wild type CDP/Cux as well as mutants A, B and CD (Fig. 6A, lanes 1-8). However, the processing of mutant ABCD was not affected in the presence of cyclin D/Cdk4 (Fig. 6A, lanes 9-10). To confirm that mutations ABCD were sufficient to prevent the effect of cyclin D/Cdk4 on processing, all mutations were transferred into a construct expressing the full length CDP/Cux. Again, cyclin D/Cdk4 inhibited the processing of most mutants except for mutant ABCD (Fig. 6B, lanes 1-10). Two conclusions can be drawn from these results. First, the mechanism by which cyclin D/Cdk4 inhibits the processing of CDP/Cux involves the phosphorylation of CDP/Cux. Secondly, the mutation of several serine residues is necessary to prevent the inhibition of processing by cyclin D/Cdk4. These results suggest that the phosphorylation of CDP/Cux at several alternative positions has an inhibitory effect on processing.

Serine to alanine mutations within CDP/Cux de-regulate cell cycle-dependent processing. Processing of CDP/Cux was previously shown to be tightly regulated during the cell cycle as no processing was observed before the G1/S transition (32). We considered the possibility that the molecular basis for this cell cycle regulation may involve the phosphorylation of CDP/Cux by cyclin D/Cdk4 in G1. If this hypothesis was correct, we would predict that the ABCD mutant should be able to be processed in G1. To test this hypothesis, NIH3T3 were transfected with vectors expressing either the wild type or mutated full length CDP/Cux protein, and were then synchronized by serum starvation and re-stimulation. Cells were harvested 9h and 16h following serum addition and analyzed

for CDP/Cux processing and cell cycle distribution. The FACS analysis indicated that most cells were still in G1 at the 9h time point, whereas a large proportion of cells had progressed past the G1/S transition at 16 hours (Fig. 7). As expected, the wild type CDP/Cux protein was processed at 16 hours but little or no processing was observed at 9 hours (Fig. 7, left panel). In contrast, processing of the ABCD mutant was essentially similar at the two time points (Fig. 7, right panel). These results indicate that the mutations of Cdk phosphorylation sites within CDP/Cux prevented the inhibition of processing in the G1 phase of the cell cycle.

Evidence from indirect immunofluorescence that mutations of Cdk phosphorylation sites alter the regulation of CDP/Cux processing in cells. Evidence that CDP/Cux is proteolytically processed in cells can be obtained from indirect immunofluorescence using antibodies against amino- and carboxy-terminal epitopes. Typically, following transfection of CDP/Cux we can observe two distinct staining patterns on the same cover slip: type 1 and type 2 (Fig. 8). In type 1, staining is identical with both antibodies and is limited to the nucleus. In type 2, however, the two antibodies generate different staining patterns. Whereas the antibody against the carboxy-terminal antibody stains only the nucleus, the antibody against the amino-terminal peptide stains the entire cells. The type 2 staining pattern can most easily be explained by assuming that the two epitopes are present on different molecules as would be expected if the protein was cleaved and the amino-terminal peptide, which does not contain a nuclear localization signal, diffused out of the nucleus. That a CDP/Cux protein truncated upstream of CR3HD localizes to the cytoplasm has previously been reported (15, 28). As can be seen in the tabulated data, when the wild type CDP/Cux protein is introduced into continuously growing cells, approximately

82% of the cells exhibit the type 1, and 18% the type 2 staining pattern (Fig. 8). In a population of cells enriched in S phase using the thymidine block procedure, the proportion of cells exhibiting type 2 staining pattern was increased to 35%. These results clearly support the notion that proteolytic processing of CDP/Cux is regulated in a cell cycle dependent manner. If, as shown from the Western blot analysis of synchronized cells, mutant ABCD can be processed in G1, we would predict that a higher fraction of continuously growing cells expressing this mutant would display the type 2 staining pattern. Indeed this is what we observed, as the type 2 staining was observed among 35% of cells transfected with this construct. As expected, synchronization using the thymidine block procedure did not further increase this proportion. These results confirm by yet another approach that the mutation of Cdk phosphorylation sites alters the cell cycle dependent regulation of CDP/Cux processing in cells.

DISCUSSION

DNA binding by the CDP/Cux protein was previously shown to be upregulated at the G1/S transition as a result of dephosphorylation by the Cdc25A phosphatase (8). CDP/Cux was later found to be proteolytically processed at the G1/S transition to generate an isoform, p110, which is capable of stable DNA binding (32). More recently, cyclin A/Cdk1 was shown to interact with CDP/Cux, phosphorylate the Cut homeodomain and inhibit DNA binding (37). In the present study, we have shown that a G1 cyclin/Cdk complex, cyclin D/Cdk4, can interact with CDP/Cux, phosphorylate it and inhibit both proteolytic processing and DNA binding. Altogether, these results point to a very tight regulation of CDP/Cux during the cell cycle, such that its stable DNA binding activity is maximal during S phase but is controlled before and after, in G1 and in G2. This type of regulation suggests that the p110 CDP/Cux isoform plays a role specifically in S phase. We speculate that CDP/Cux regulates the expression of a number of genes whose products are involved in DNA replication. Interactions between CDP/Cux and regulatory elements of several histone genes have been documented by several groups (3, 14, 17, 42, 43, 45, 46). Interestingly, the histone nuclear factor D (HiNF-D) complex had been characterized for a number of years before it was realized that CDP/Cux was its DNA binding subunit (45, 46). Whether CDP/Cux plays a role as an activator or a repressor of these genes has not yet resolved. Our group has obtained evidence that CDP/Cux would stimulate expression of the DNA polymerase alpha gene in transient transfection assays (37). However, whether this effect is direct or indirect is currently being investigated. CDP/Cux was also suggested to play both a positive and negative regulatory role in the case of the thymidine kinase gene (23).

The processing of CDP/Cux and its stable DNA binding activity were inhibited in cells over-expressing cyclin D/Cdk4. While we have shown that cyclin D/Cdk4 did not affect a mutant CDP/Cux protein in which 8 serines had been mutated to alanine, we have been unable to identify a single Cdk phosphorylation site that would be sufficient to mediate the inhibition of processing. Four other mutants, carrying either one, two or three mutations were assayed. The processing of each of these mutants was inhibited in the presence of cyclin D/Cdk4. The mutant CD, in which serines 1054 and 1059, 1237 and 1270 were mutated, seemed less affected than the wild type, as a processed isoform was detected even in cells over-expressing cyclin D/Cdk4 (Fig. 6). Yet, cyclin D/Cdk4 was still partially able to inhibit processing. These results strongly suggest that the inhibition of processing results from the phosphorylation of several residues over a rather large region of the protein. We speculate that the accumulation of negative charges in CDP/Cux leads to a conformational change such that the linker region between CR1 and CR2 is not as readily accessible to the protease. Alternatively, it is possible that processing takes place in a specific subcellular compartment and that phosphorylation of CDP/Cux precludes its trafficking to this location.

If phosphorylation of CDP/Cux by cyclin D/Cdk4 serves to delay its processing until S phase, then a mutated CDP/Cux protein in which all Cdk phosphorylation sites have been removed should be processed earlier in the cell cycle. This is what we observed (Fig. 7). In addition, we predicted that a larger fraction of cells in a population of unsynchronized cells should display evidence of processing. To test this notion, we performed indirect immunofluorescence using antibodies against an amino- and a carboxy-terminal epitope. Since the nuclear localization signal within CDP/Cux is located in the CR3HD region,

we expected that the amino-terminal peptide generated as a result of processing would localize to the cytoplasm, whereas the carboxy-terminal peptide containing CR3HD would stay in the nucleus. Indeed, following transfection of a wild type recombinant CDP/Cux protein, we observed two types of staining patterns. In one type, type 1, the amino- and carboxy-terminal antibodies both stained only the nucleus. In type 2, the amino-terminal antibody stained the entire cell. Although we cannot exclude that processing occurs in the cytoplasm, we favor the notion that the full length CDP/Cux protein localizes to and is cleaved in the nucleus, and that the amino-terminal peptide then diffuses out of the nucleus. Only 18% of the cells displayed the type 2 staining pattern, a proportion that is lower than that of the cells having progressed beyond the G1/S transition. Our current explanation for this discrepancy is that the amino-terminal peptide is unstable and that the rapid disappearance of this peptide causes the loss of a cytoplasm signal with the amino-terminal antibody in many cells that have progressed past the G1 phase. In support for this hypothesis, we note that we have not been able to detect the amino-terminal peptide generated from the processing of the endogenous CDP/Cux protein. In contrast to p110, detection of the N-terminal peptide requires the over-expression of CDP/Cux.

Our findings have uncovered a new target for cyclin D/Cdk4. In addition, the findings that cyclin D/Cdk4 acts to delay the activation of CDP/Cux introduces a novel concept in the field of cell cycle regulation. The role of cyclin/Cdk complexes is generally perceived to consist primarily in promoting the transition from one phase of the cell cycle to the next. Thus, the phosphorylation of pRb by cyclin D/Cdk4 contributes to trigger the dissociation of pRb from the E2F/DP complex, which can then activates the transcription of genes involved in DNA replication. In contrast, the effect of cyclin D/Cdk4 on CDP/Cux is to de-

lay the time at which the protein is proteolytically processed and its stable DNA binding activity is unleashed. This suggests that the role of cyclin D/Cdk4 is not limited to the preparation of the G1/S transition but also to ensure that some events meant to take place at the onset of S phase would not occur prematurely during the G1 phase. What happens then in cancer cells that over-express cyclin D/Cdk4? Is processing of CDP/Cux inhibited? This question merits investigation, but at this point we would speculate that the effect of cyclin D/Cdk4 may be counter-acted by the action of a phosphatase, perhaps Cdc25A, that is activated at the end of G1. This hypothesis should be investigated in future work.

ACKNOWLEDGMENTS

We are grateful to Dr. R.P. Fisher and Dr. S. Larochelle for the generous gift of Cyclin D and Cdk 4 Baculovirus constructs, Dr. S. van den Heuvel for the mammalian expression vector for wild type and mutant Cdk4, and Dr. Nicholas Dyson for the vector expressing GST/pRb-CTD. A.N. is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec. N.M.S. is the recipient of a fellowship from Medical Research Council of Canada. This research was supported by grant #3497 from the National Cancer Institute of Canada and grant #MT-11590 from the Canadian Institute of Health Research of Canada to A.N.

REFERENCES

1. **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:**321-34.

2. **Aziz, F., A. J. Vanwijnen, P. S. Vaughan, S. J. Wu, A. R. Shakoori, J. B. Lian, K. J. Soprano, J. L. Stein, and G. S. Stein.** 1998. The integrated activities of Irf-2 (Hinf-M), Cdp/Cut (Hinf-D) and H4tf-2 (Hinf-P) regulate transcription of a cell cycle controlled human histone H4 gene - mechanistic differences between distinct H4 genes. Molecular Biology Reports **25:**1-12.

3. **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:**347-59.

4. **Bienvenu, F., H. Gascan, and O. Coqueret.** 2001. Cyclin D1 represses STAT3 activation through a Cdk4-independent mechanism. J Biol Chem **276:**16840-7.

5. **Blochlinger, K., R. Bodmer, J. Jack, L. Y. Jan, and Y. N. Jan.** 1988. Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila. Nature **333:**629-35.

6. **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 D. melanogaster. Cell **51:**293-307.

7. **Braun, W.** 1940. The effect of puncture on the developing wing of several mutants of *Drosophila melanogaster*. J. Exp. Zool. **84:**325-350.

8. **Coqueret, O., G. Berube, and A. Nepveu.** 1998. The mammalian Cut homeodomain protein functions as a cell-cycle dependent transcriptional repressor which down-modulates p21WAF1/CIP1/SDI1 in S phase. EMBO J **17:**4680-4694.

9. **Deng, W. P., and J. A. Nickoloff.** 1992. Site-directed mutagenesis of virtually any plasmid by eliminating a unique site. Anal Biochem **200:**81-8.

10. **Dufort, D., and A. Nepveu.** 1994. The human cut homeodomain protein represses transcription from the c-myc promoter. Mol Cell Biol **14:**4251-7.

11. **Dynlacht, B. D., O. Flores, J. A. Lees, and E. Harlow.** 1994. Differential regulation of E2F transactivation by cyclin/cdk2 complexes. Genes Dev **8:**1772-86.

12. **Dyson, N.** 1998. The regulation of E2F by pRB-family proteins. Genes Dev **12:**2245-62.

13. **Ekholm, S. V., and S. I. Reed.** 2000. Regulation of G(1) cyclin-dependent kinases in the mammalian cell cycle. Curr Opin Cell Biol **12:**676-84.

14. **el-Hodiri, H. M., and M. Perry.** 1995. Interaction of the CCAAT displacement protein with shared regulatory elements required for transcription of paired histone genes. Mol Cell Biol **15:**3587-96.

15. **Ellis, T., L. Gambardella, M. Horcher, S. Tschanz, J. Capol, P. Bertram, W. Jochum, Y. Barrandon, and M. Busslinger.** 2001. The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev **15:**2307-19.

16. **Hirai, H., and C. J. Sherr.** 1996. Interaction of D-type cyclins with a novel myb-like transcription factor, DMP1. Mol Cell Biol **16:**6457-67.

17. **Holthuis, J., T. A. Owen, A. J. van Wijnen, K. L. Wright, A. Ramsey-Ewing, M. B. Kennedy, R. Carter, S. C. Cosenza, K. J. Soprano, J. B. Lian, and et al.** 1990. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science **247:**1454-7.

18. **Inoue, K., M. F. Roussel, and C. J. Sherr.** 1999. Induction of ARF tumor suppressor gene expression and cell cycle arrest by transcription factor DMP1. Proc Natl Acad Sci U S A **96:**3993-8.

19. **Inoue, K., and C. J. Sherr.** 1998. Gene expression and cell cycle arrest mediated by transcription factor DMP1 is antagonized by D-type cyclins through a cyclin-dependent-kinase- independent mechanism. Mol Cell Biol **18:**1590-600.

20. **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:**735-47.

21. **Jack, J. W.** 1985. Molecular organization of the cut locus of Drosophila melanogaster. Cell **42:**869-76.

22. **Kato, J., H. Matsushime, S. W. Hiebert, M. E. Ewen, and C. J. Sherr.** 1993. Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. Genes Dev **7:**331-42.

23. **Kim, E. C., J. S. Lau, S. Rawlings, and A. S. Lee.** 1997. Positive and Negative Regulation Of the Human Thymidine Kinase Promoter Mediated By CCAAT Binding Transcription Factors Nf-Y/Cbf, Dbpa, and Cdp/Cut. Cell Growth & Differentiation **8:**1329-1338.

24. **Li, S., L. Moy, N. Pittman, G. Shue, B. Aufiero, E. J. Neufeld, N. S. LeLeiko, and M. J. Walsh.** 1999. Transcriptional repression of the cystic fibrosis transmembrane

conductance regulator gene, mediated by CCAAT displacement protein/cut homolog, is associated with histone deacetylation. J Biol Chem **274:**7803-15.

25. **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.

26. **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:**151-9.

27. **Lukas, J., T. Herzinger, K. Hansen, M. C. Moroni, D. Resnitzky, K. Helin, S. I. Reed, and J. Bartek.** 1997. Cyclin E-induced S phase without activation of the pRb/E2F pathway. Genes Dev **11:**1479-92.

28. **Luong, M. X., C. M. van der Meijden, D. Xing, R. Hesselton, E. S. Monuki, S. N. Jones, J. B. Lian, J. L. Stein, G. S. Stein, E. J. Neufeld, and A. J. van Wijnen.** 2002. Genetic ablation of the CDP/Cux protein C terminus results in hair cycle defects and reduced male fertility. Mol Cell Biol **22:**1424-37.

29. **Mailly, F., G. Berube, R. Harada, P. L. Mao, S. Phillips, and A. Nepveu.** 1996. The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy. Mol Cell Biol **16:**5346-57.

30. **Meyerson, M., and E. Harlow.** 1994. Identification of G1 kinase activity for cdk6, a novel cyclin D partner. Mol Cell Biol **14:**2077-86.

31. **Moon, N. S., G. Berube, and A. Nepveu.** 2000. CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain. J Biol Chem **275:**31325-31334.

32. **Moon, N. S., P. Premdas, M. Truscott, L. Leduy, G. Berube, and A. Nepveu.** 2001. S Phase-Specific Proteolytic Cleavage Is Required to Activate Stable DNA Binding by the CDP/Cut Homeodomain Protein. Mol Cell Biol **21:**6332-6345.

33. **Morgan, D. O.** 1997. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol **13:**261-91.

34. **Nepveu, A.** 2001. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. Gene **270:**1-15.

35. **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. Nature Genetics **1:**50-5.

36. **Qin, X. Q., T. Chittenden, D. M. Livingston, and W. G. Kaelin, Jr.** 1992. Identification of a growth suppression domain within the retinoblastoma gene product. Genes Dev **6:**953-64.

37. **Santaguida, M., Q. Ding, G. Berube, M. Truscott, P. Whyte, and A. Nepveu.** 2001. Phosphorylation of the CCAAT Displacement Protein (CDP)/Cux Transcription Factor by Cyclin A-Cdk1 Modulates Its DNA Binding Activity in G2. J Biol Chem **276:**45780-45790.

38. **Sinclair, A. M., J. A. Lee, A. Goldstein, D. Xing, S. Liu, R. Ju, P. W. Tucker, E. J. Neufeld, and R. H. Scheuermann.** 2001. Lymphoid apoptosis and myeloid hyperplasia in CCAAT displacement protein mutant mice. Blood **98:**3658-3667.

39. **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.

40. **Superti-Furga, G., A. Barberis, E. Schreiber, and M. Busslinger.** 1989. The protein CDP, but not CP1, Footprints on the CCAAT region of the g-globulin gene in unfractionated B-cell extracts. Biochimical Biophysical Acta **1007:**237-242.

41. **Valarche, I., J. P. Tissier-Seta, M. R. Hirsch, S. Martinez, C. Goridis, and J. F. Brunet.** 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:**881-96.

42. **van den Ent, F. M., A. J. van Wijnen, J. B. Lian, J. L. Stein, and G. S. Stein.** 1994. Cell cycle controlled histone H1, H3, and H4 genes share unusual arrangements of recognition motifs for HiNF-D supporting a coordinate promoter binding mechanism. J Cell Physiol **159:**515-30.

43. **van Wijnen, A. J., T. K. Choi, T. A. Owen, K. L. Wright, J. B. Lian, R. Jaenisch, J. L. Stein, and G. S. Stein.** 1991. Involvement of the cell cycle-regulated nuclear factor HiNF-D in cell growth control of a human H4 histone gene during hepatic development in transgenic mice. Proc Natl Acad Sci U S A **88:**2573-7.

44. **van Wijnen, A. J., C. Cooper, P. Odgren, F. Aziz, A. De Luca, R. A. Shakoori, A. Giordano, P. J. Quesenberry, J. B. Lian, G. S. Stein, and J. L. Stein.** 1997. Cell cycle-dependent modifications in activities of pRb-related tumor suppressors and proliferation-specific CDP/cut homeodomain factors in murine hematopoietic progenitor cells. J Biol Chem **66:**512-23.

45. **van Wijnen, A. J., M. F. van Gurp, M. C. de Ridder, C. Tufarelli, T. J. Last, M. Birnbaum, P. S. Vaughan, A. Giordano, W. Krek, E. J. Neufeld, J. L. Stein, and G. S. Stein.** 1996. CDP/cut is the DNA-binding subunit of histone gene transcription

factor HiNF-D: a mechanism for gene regulation at the G1/S phase cell cycle transition point independent of transcription factor E2F. Proc Natl Acad Sci U S A **93:**11516-21.

46. **van Wijnen, A. J., K. L. Wright, J. B. Lian, J. L. Stein, and G. S. Stein.** 1989. Human H4 histone gene transcription requires the proliferation-specific nuclear factor HiNF-D. Auxiliary roles for HiNF-C (Sp1-like) and HiNF-A (high mobility group-like). J Biol Chem **264:**15034-42.

47. **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:**18453-62.

48. **Zwijsen, R. M., R. S. Buckle, E. M. Hijmans, C. J. Loomans, and R. Bernards.** 1998. Ligand-independent recruitment of steroid receptor coactivators to estrogen receptor by cyclin D1. Genes Dev **12:**3488-98.

49. **Zwijsen, R. M., E. Wientjens, R. Klompmaker, J. van der Sman, R. Bernards, and R. J. Michalides.** 1997. CDK-independent activation of estrogen receptor by cyclin D1. Cell **88:**405-15.

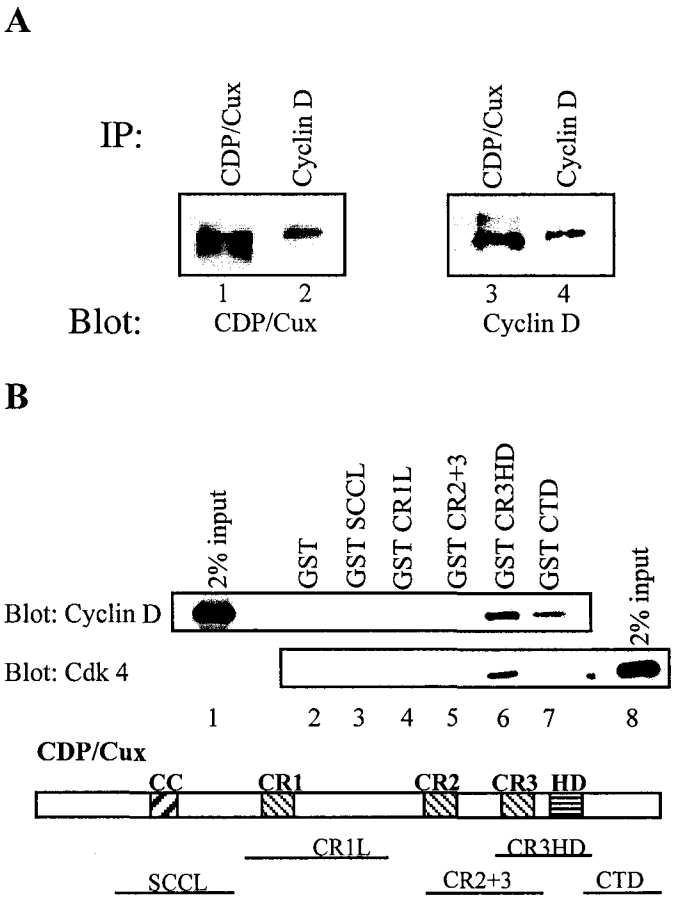
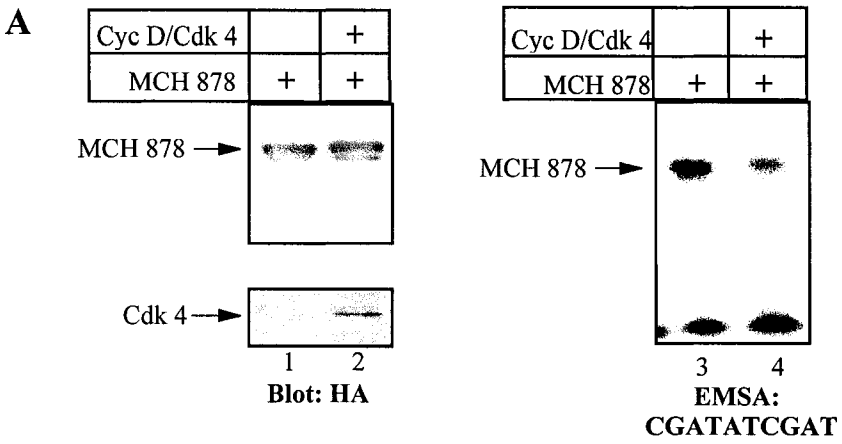
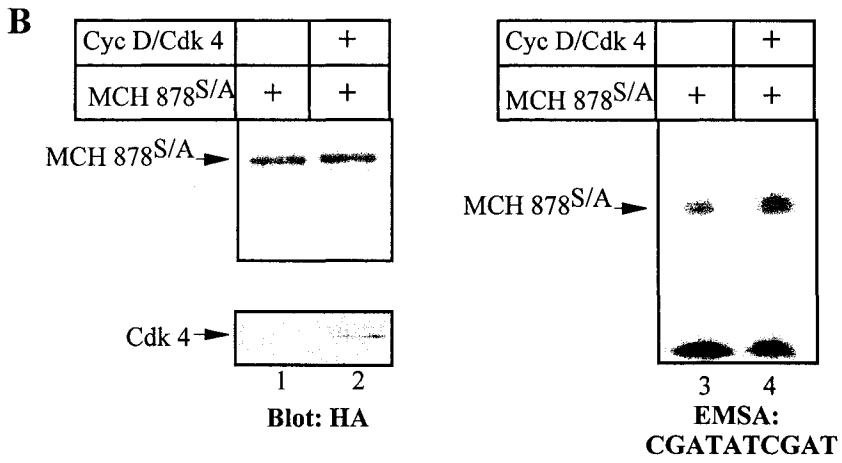


Fig.1 Interaction between Cyclin D and CDP/Cux. (A) Whole cell extracts from 293 cells were submitted to immunoprecipitation and immunoblotting using antibodies against CDP/Cux or cyclin D, as indicated. (B) Pull down assays were performed using GST-CDP/Cux fusion proteins and whole cell extracts from SF9 cells previously infected with baculovirus encoding cyclin D and Cdk 4. Cyclin D and Cdk 4 were visualized by immunoblotting. Below is a schematic representation of the CDP/Cux regions present within the GST-CDP/Cux fusion proteins.


A
Cyc D/Cdk 4
MCH 878
+
+
+
MCH 878
Cdk 4
1
2
Blot: HA
Cyc D/Cdk 4
MCH 878
+
+
+
MCH 878
3
4
EMSA:
CGATATCGAT


B
Cyc D/Cdk 4
MCH 878S/A
+
+
+
MCH 878S/A
Cdk 4
1
2
Blot: HA
Cyc D/Cdk 4
MCH 878S/A
+
+
+
MCH 878S/A
3
4
EMSA:
CGATATCGAT

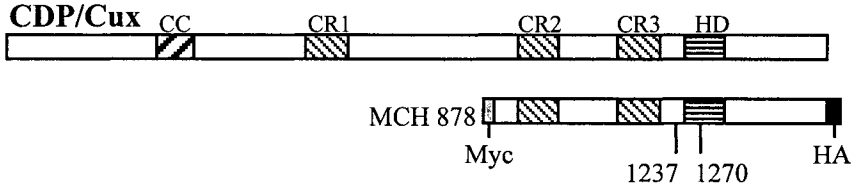
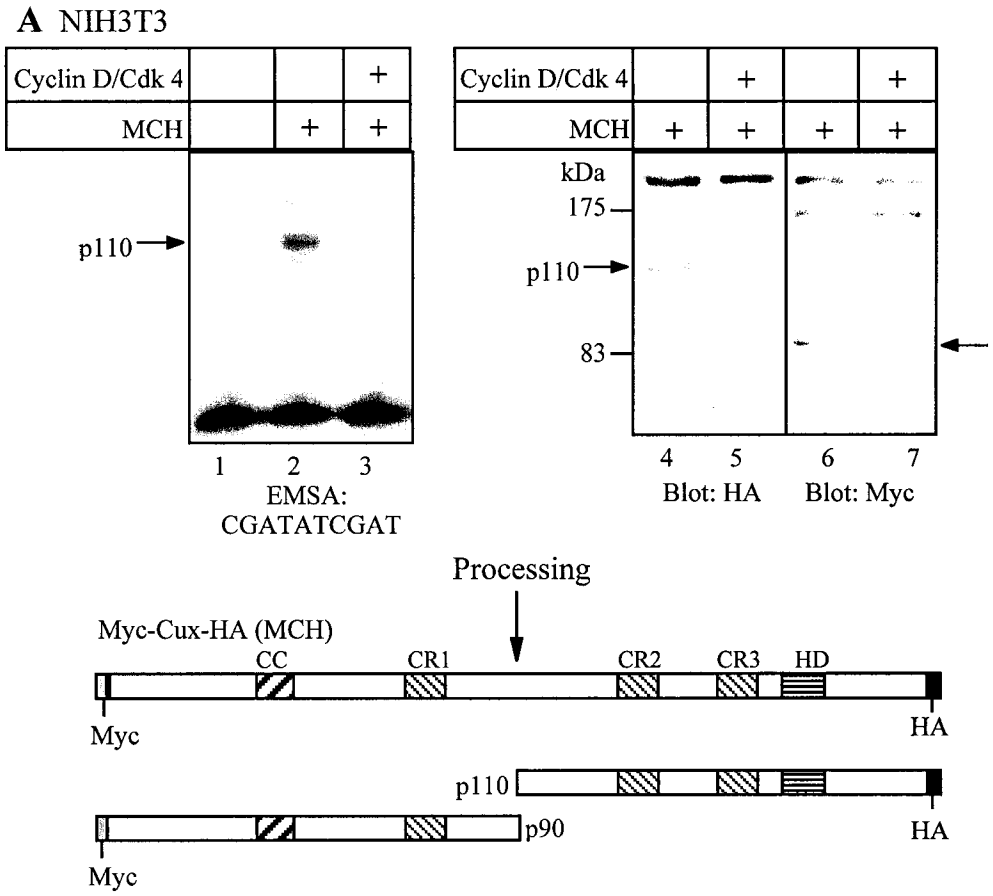

CDP/Cux
CC
CR1
CR2
CR3
HD
MCH 878
Myc
1237
1270
HA

Fig. 2 Phosphorylation of CR3HD by Cyclin D/Cdk 4 inhibits DNA binding. (A) NIH3T3 cells were transfected with a vector expressing a CDP/Cux recombinant protein, MCH878, containing amino acids 878-1505 and an HA epitope at its carboxy-terminus. The CDP/Cux vector was introduced either alone (lanes 1 and 3) or with vectors expressing cyclin D and Cdk 4 (lanes 2 and 4) as indicated. Nuclear extracts were analyzed in Western blot with anti HA antibody (left panel) and in EMSA with oligonucleotides containing a CDP/Cux consensus binding site (CGATATCGAT). (B) The same experiment was repeated with a the mutant $MCH878^{S/A}$ in which serines 1237 and 1270 were replaced with alanines. The positions of the mutations are indicated in the diagram.


A NIH3T3
Cyclin D/Cdk 4
MCH
p110
1 2 3
EMSA:
CGATATCGAT
kDa
175
p110
83
4 5 6 7
Blot: HA
Blot: Myc
Processing
Myc-Cux-HA (MCH)
CC
CR1
CR2
CR3
HD
Myc
HA
p110
p90
HA
Myc

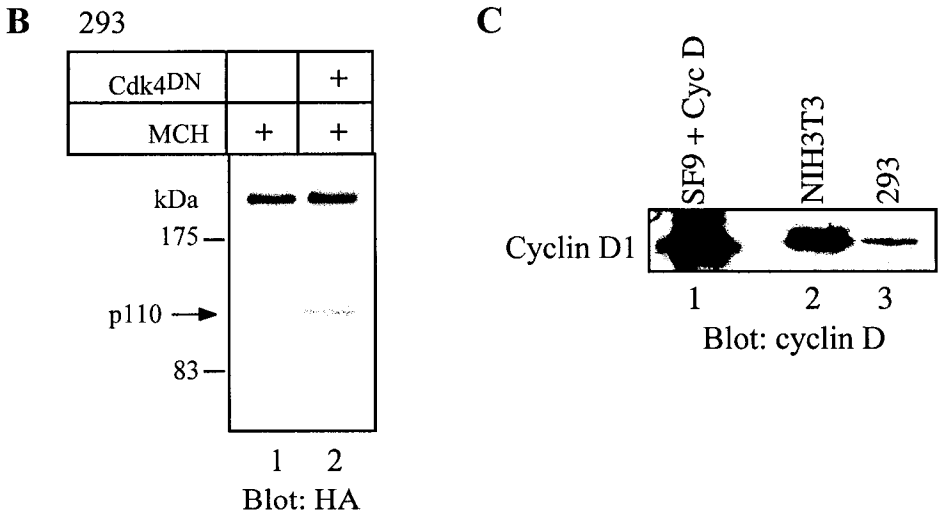

B 293
Cdk4DN
MCH
kDa
175
p110
83
1 2
Blot: HA
C
SF9 + Cyc D
NIH3T3
293
Cyclin D1
1 2 3
Blot: cyclin D

Fig. 3 Co-expression with cyclin D/Cdk 4 inhibits proteolytic processing of CDP/Cux.

(A) NIH3T3 cells were transfected with a vector expressing Myc-Cux-HA (MCH), a CDP/Cux protein with Myc and HA epitope tags, either alone or with vectors expressing Cyclin D and Cdk 4, as indicated. Nuclear extracts were analyzed in Western blot with an anti-HA antibody and in EMSA with oligonucleotides containing a CDP/Cux consensus binding site.

(B) 293 cells were transfected with the MCH vector either alone (lane1) or with a vector expressing a dominant negative mutant of Cdk 4 (lane 2). Nuclear extracts were prepared and analyzed in Western blot with HA antibody.

(C) Western blot analysis using antibodies against cyclin D1 and total extracts from baculovirus-infected Sf9 cells, NIH3T3 and 293 cells

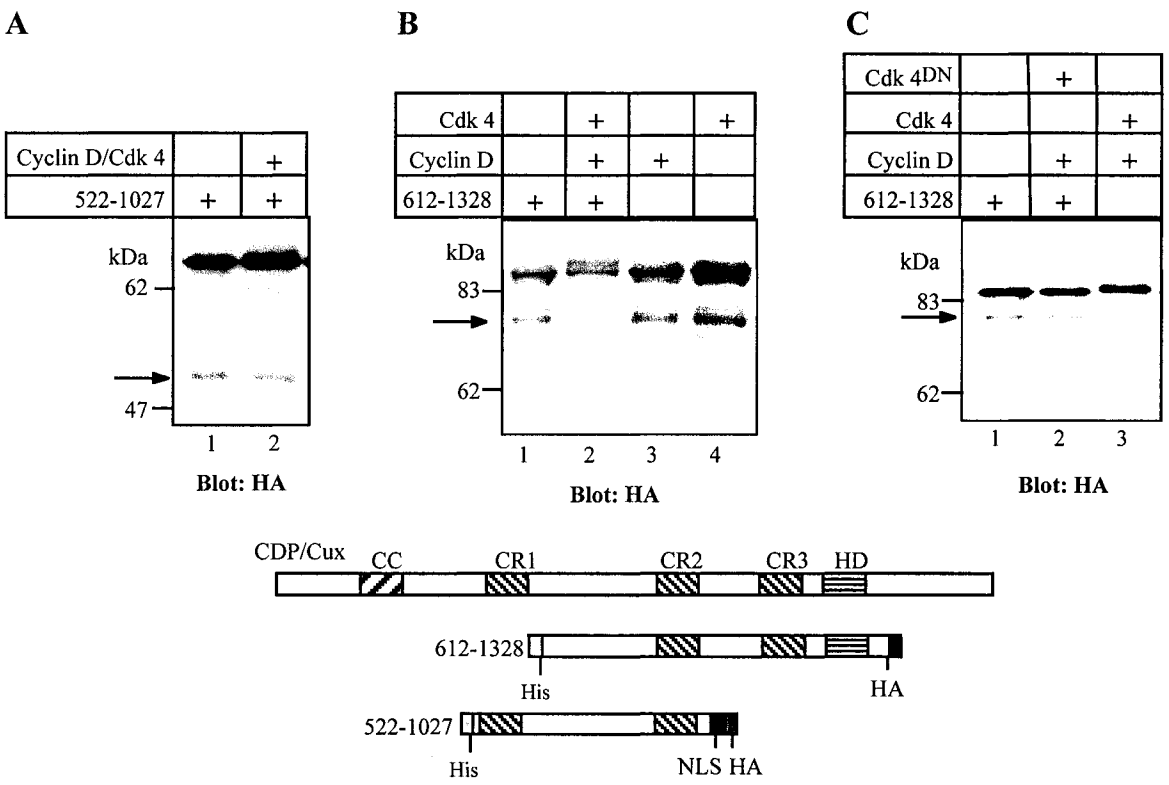


Fig. 4 Inhibition of processing by cyclin D/Cdk4 requires the presence of the CR3HD region and is associated with a reduction in the electrophoretic mobility of CDP/Cux. NIH3T3 cells were transfected with a vector expressing CDP/Cux recombinant proteins containing amino acids 522-1027 (A) or 612-1328 (B and C) and an HA epitope at its carboxy-terminus. The CDP/Cux vectors were introduced either alone (lanes 1) or with vectors expressing cyclin D with Cdk 4 or dominant negative form of Cdk4 as indicated. Nuclear extracts were analyzed in Western blot with anti HA antibody. The arrow indicates in each case the position of the processed CDP/Cux protein.

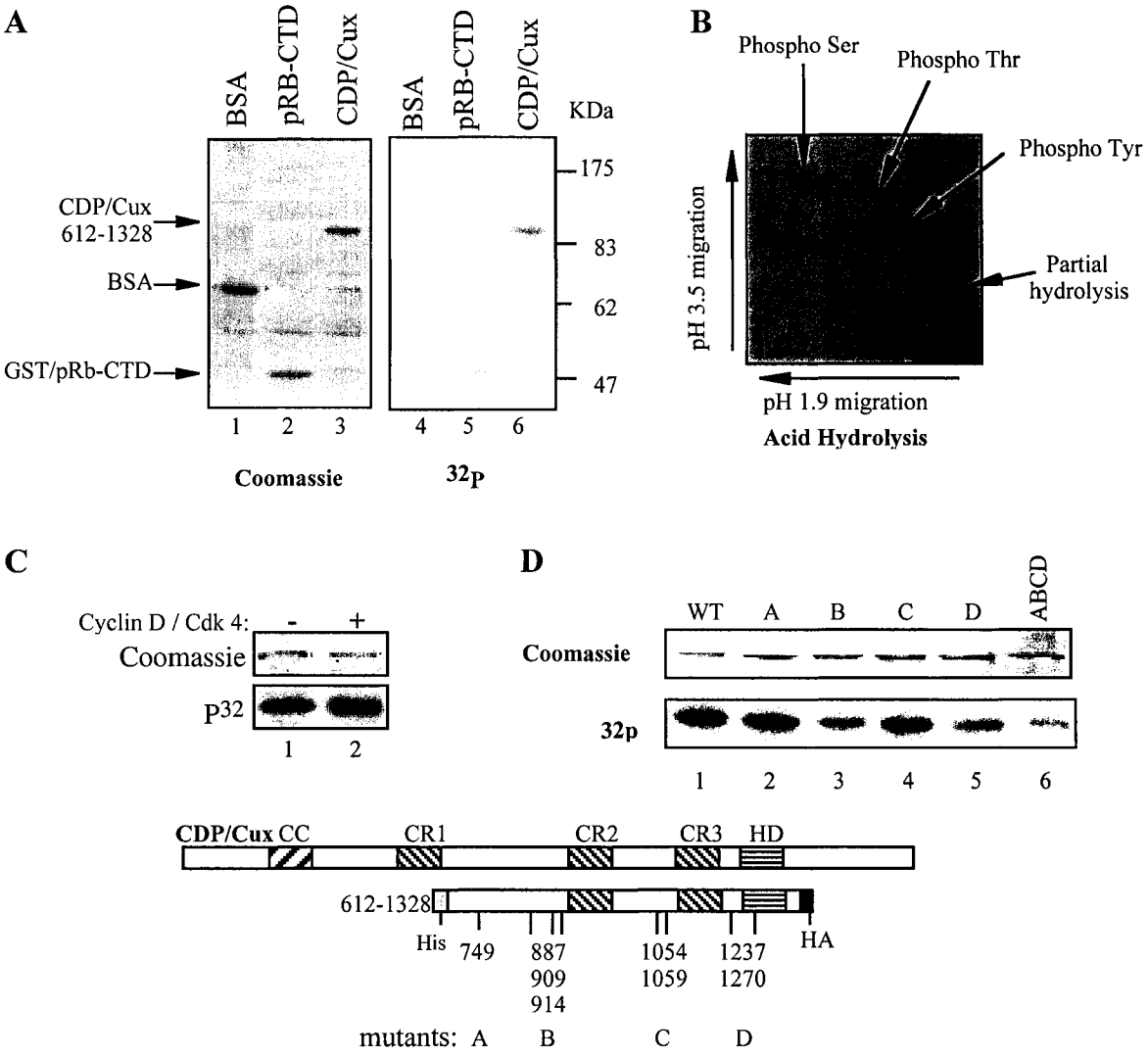

A
BSA
pRB-CTD
CDP/Cux
BSA
pRB-CTD
CDP/Cux
KDa
175
83
62
47
CDP/Cux
612-1328
BSA
GST/pRb-CTD
1 2 3 4 5 6
Coomassie
32p
B
Phospho Ser
Phospho Thr
Phospho Tyr
Partial
hydrolysis
pH 3.5 migration
pH 1.9 migration
Acid Hydrolysis
C
Cyclin D / Cdk 4: - +
Coomassie
P32
1 2
D
WT A B C D ABCD
Coomassie
32p
1 2 3 4 5 6
CDP/Cux CC CR1 CR2 CR3 HD
612-1328
His 749 887 909 914 1054 1059 1237 1270 HA
mutants: A B C D

Fig. 5 Cyclin D/Cdk 4 phosphorylates several serine residues of CDP/Cux *in vitro*.

(A) Cyclin D/Cdk4 complexes from Sf9 cells were immunoprecipitated using an anti-Cdk4 antibody, and were incubated with 500 ng of the indicated proteins for 20 min. in the presence of 5 μCi ^{32}P g-ATP. Proteins in each sample were separated by electrophoresis on a 12% SDS-polyacrylamide gel and revealed by staining with Coomassie-blue and autoradiography.

(B) *In vitro* phosphorylated CDP/Cux 612-1328 was hydrolyzed with acid. The hydrolysate was applied to TLC plates and subjected to two-dimensional electrophoresis in pH 1.9 and pH 3.5 buffers. The mobilities of phosphoserine, phosphothreonine and phosphotyrosine are indicated according to the migration of a phosphoaminoacid standard mixture.

(C) For the purpose of metabolic labeling, 293 cells were transfected with constructs expressing CDP/Cux 612-1328 with or without cyclin D and Cdk4 vectors as indicated. Cells were metabolically labeled with {^{32}P} orthophosphate. Recombinant CDP/Cux proteins were purified by affinity chromatography on nickel column and visualized by Coomassie blue staining and autoradiography. Note that this purification scheme allows the purification of only the full length 612-1328 protein, not the processed isoform.

(D) *In vitro* kinase assays were performed as described in A with either wild type (lane 1) or mutant CDP/Cux 612-1328 (lane 2 to 6) in which serine residues were replaced for alanine at position 749 (A), 887, 909 and 914 (B), 1054 and 1059 (C), and 1237 and 1270 (D). The positions of mutated serine residues are shown in the diagram. Proteins in each sample were separated by electrophoresis on a 12% SDS-polyacrylamide gel and revealed by staining with Coomassie-blue and autoradiography.

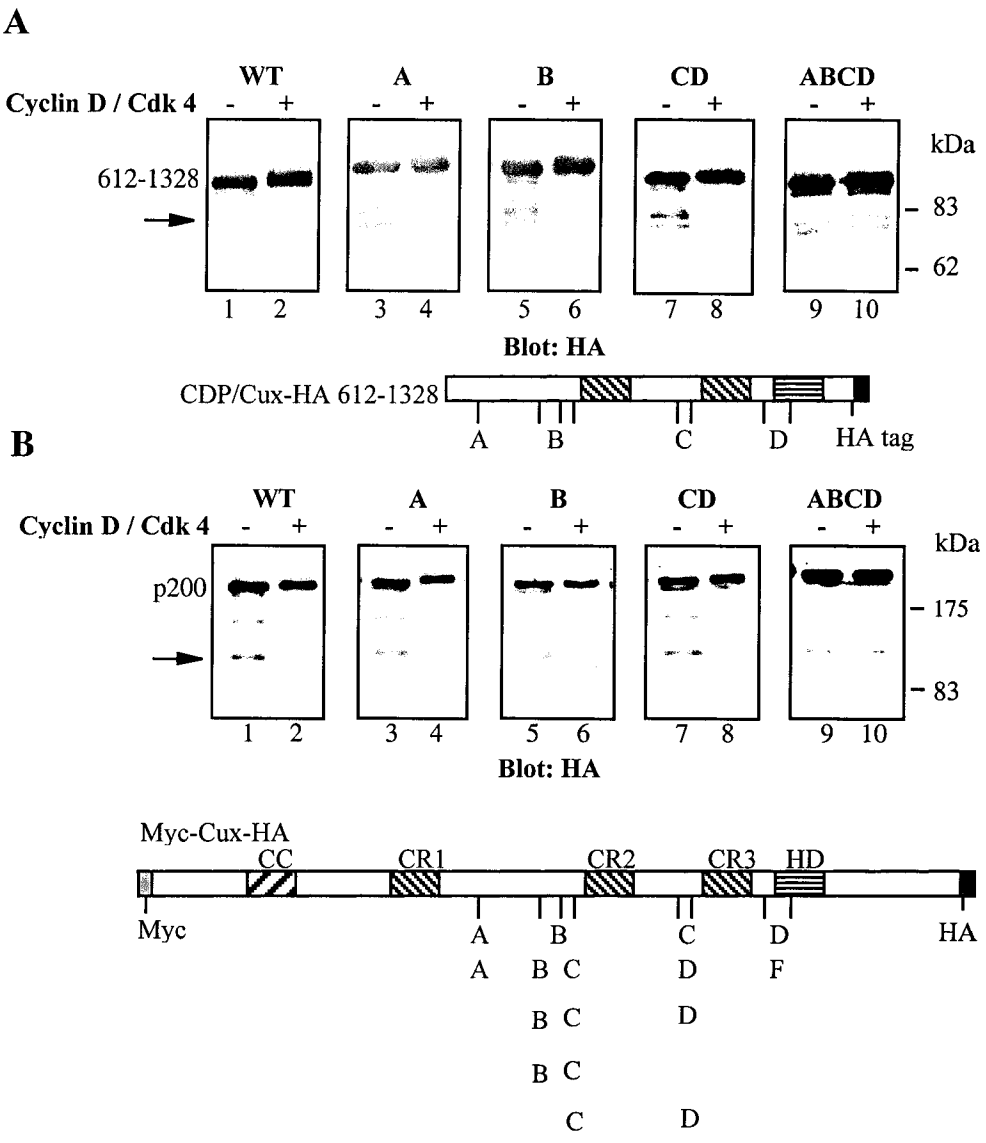


Fig. 6 Serine to alanine mutations within CDP/Cux prevent the effect of Cyclin D/ Cdk 4 on processing. NIH3T3 cells were transfected with a vector expressing a wild type or mutated versions of CDP/Cux 612-1328 (A) or full length CDP/Cux (B). The CDP/Cux vector was introduced either alone or with vectors expressing cyclin D and Cdk 4 as indicated. Nuclear extracts were analyzed in Western blot with anti HA antibody.

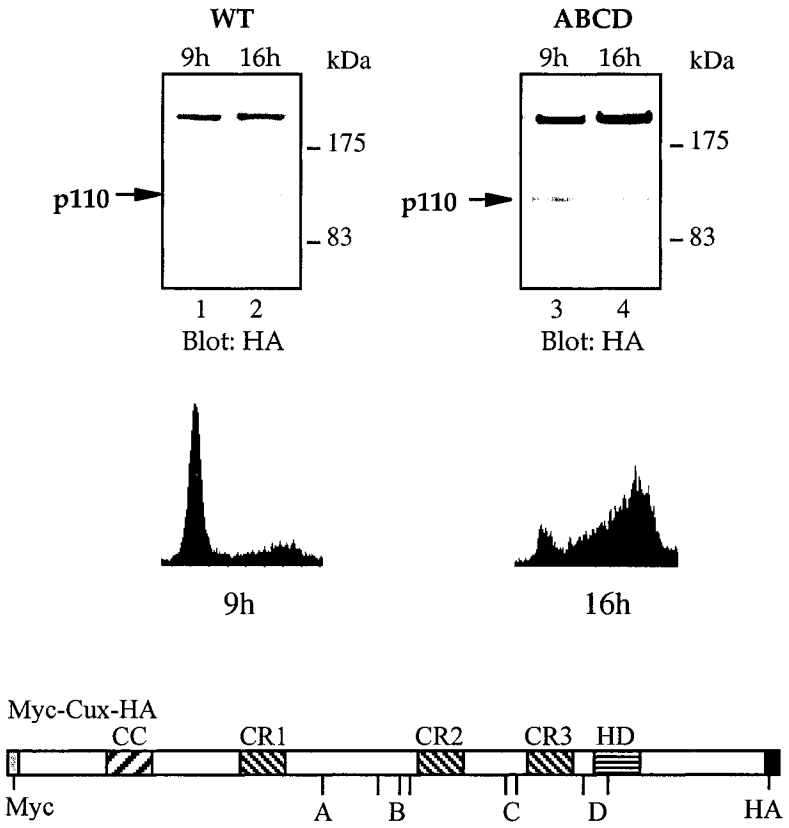
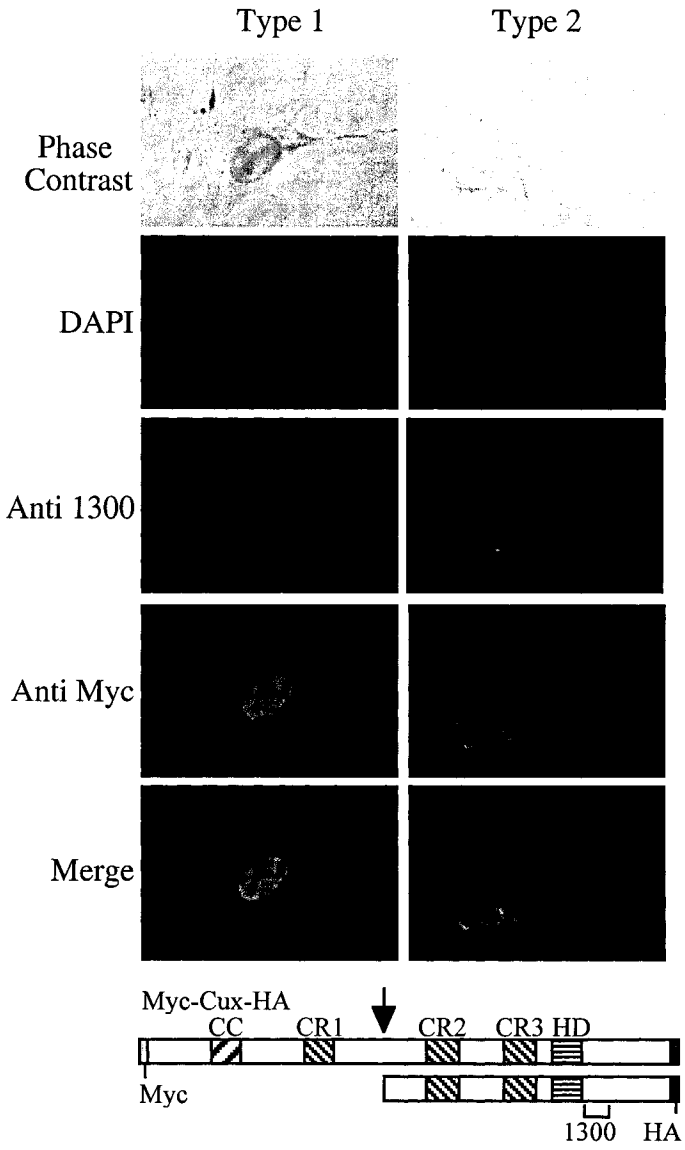


Fig. 7 Serines to alanine mutations within CDP/Cux allows processing to occur in G1. NIH3T3 cells were transfected with a vector expressing a wild type or a mutated version of full length CDP/Cux, mutant ABCD, in which serines 749, 887, 909, 914, 1054, 1059, 1237 and 1270 were replaced with alanine. Cells were then synchronized by serum starvation followed by serum addition for 9h (lane 1 and 3) or 16h (lane 2 and 4) respectively. Nuclear extracts were analyzed in Western blots with anti-HA antibody. Cell cycle distribution was monitored by fluorescence-activated cell sorting (FACS) analysis after staining of the DNA with propidium iodide.



Culture conditions		CDP/Cux	Frequency of Type 1	Frequency of Type 2
Unsyn-chronized	G0/G1 53% S 18% G2 29%	WT	82 %	18 %
		ABCD	65 %	35 %
Thymidine block	G0/G1 36% S 41% G2 23%	WT	65 %	35 %
		ABCD	72 %	28 %

Fig. 8 Indirect immunofluorescence of cells expressing wild type or mutant ABCD CDP/Cux proteins. NIH 3T3 cells on glass coverslips were transfected with a vector expressing wild type or mutant ABCD CDP/Cux (see diagram in Fig. 7). The recombinant CDP/Cux protein included a Myc epitope tag at the amino-terminus. One day later, indirect immunofluorescence was performed with the indicated antibodies as described in material and methods. The coverslips were mounted on slides and cells were visualized with incident light fluorescence. Two staining patterns are presented, type 1 and type 2. The table gives the proportion of cells that display type 1 and type 2 staining pattern in the indicated situation.

GENERAL DISCUSSION

In this thesis, I have described different modes of DNA binding by CDP/Cux, a transcription factor with four DNA binding domains, and some of the mechanisms that control it. Recent studies from other groups and ours have generated evidence that CDP/Cux may play an important role in the G1 to S transition. Figure 1 summarizes the finding from this thesis. Briefly, CDP/Cux exists mainly as full-length form in G1 that exhibits DNA binding characteristics similar to CR1+2, the domain responsible for the CCAAT-displacement activity. The DNA binding activity of CR3HD is inactive in G1 since the phosphorylation by cyclin D/Cdk4 inhibits the processing and DNA binding by CR3HD. As cells enter S phase, CDP/Cux is cleaved between CR1 and CR2. The N-terminally truncated p110 CDP/Cux stably binds to DNA and regulates S phase specific genes. In addition, an increase in the expression level of N-terminally truncated CDP/Cux isoforms is observed in human uterine leiomyomas. In this section, I will discuss the significance of new findings in understanding the function of CDP/Cux and questions that remain to be answered.

CCAAT Displacement Activity

The ability to prevent a transcriptional activator such as NF-Y to bind to the CCAAT element of a promoter is unique to p200 (full-length) CDP/Cux since the isoform contains CR1CR2. The identification of CR1CR2 as a bipartite DNA binding domain responsible for CCAAT displacement activity permitted us to study the mechanism of action *in vitro*. Since both CDP/Cux and NF-Y can bind to the CCAAT element, one is lead to ask which CCAAT sites are bound by CDP/Cux at a given time. Depending on the

factor occupying the CCAAT site of a gene promoter, transcription could be turned on or off. Since CDP/Cux needs at least two CAAT or CGAT like site in order to bind to DNA with high affinity, we can, to some extent, predict which CCAAT sites of a promoter are preferred by CDP/Cux. However, any DNA sequence with CCAAT consensus sequence that is longer than 128 bp (4X2X4X4) probably contains an additional C(A/G)AT sequence. Furthermore, since CR1CR2 can tolerate some variation in the recognition sequence, the chance of finding such a sequence increases. As mentioned in chapter II, this raises the issue of whether CDP/Cux binds to any DNA sequence long enough to contain two C(A/G)AT like sequence. It would be important to determine the maximum tolerable distance between two C(A/G)AT site recognized by CR1CR2. This is not easy to study *in vitro* as sequence specificity for the second recognition site by CR1CR2 is somewhat promiscuous and the difference in the affinity among those sequences may not be significant enough. The result of previously performed *in vitro* DNA binding assays with CDP/Cux should be reexamined. Another issue here is whether DNA looping may allow CR1CR2 to bind to two CGAT sites situated at a distance from one another. One could envision a situation whereby CR1CR2 binds only to those promoters where the second CGAT sites is located at less than, for example, 15 bp or more than 90 bp. The latter number representing the minimal distance at which DNA looping is possible, whereas the former number would represent the maximal distance between two closely spaced sites. Using technique such as ChIP assay, it will be important to investigate if the targets of CDP/Cux *in vivo* include genes that were previously shown to be bound by CDP/Cux *in vitro*.

The first attempt to create a *cux-1* knockout mouse resulted in a mouse expressing Cux-1 protein with deletion at CR1 (CuxΔCR1) due to the production of an alternatively

spliced mRNA (15). Since CR1CR2 is responsible for the CCAAT-displacement activity, the CuxΔCR1 protein should not display any CCAAT displacement activity. However, the only visible phenotype observed from the CuxΔCR1 mouse was a mild defect in hair morphology. Compared to what was observed in the *cut* mutants in *Drosophila*, the defect observed from this mouse was minor. It is possible that the resulting CuxΔCR1 protein, a protein with CR2, CR3, and the Cut HD, can still display some CCAAT-displacement activity, indicating that CR1 may be dispensable for this activity. Alternatively, the CCAAT-displacement activity may not be the major function of the Cux-1 protein. Recently, a *cux-1* knockout mouse with non-functional protein was generated and displayed more significant defects in the hair follicle, the lung development and B and T cell production (5, 14). However, it is difficult to determine how much of the phenotype observed in the total *cux-1* knockout is due to the absence of the CCAAT-displacement activity since the mouse does express neither p200 nor p110 CDP/Cux protein.

Stable DNA binding by CDP/Cux

CR3HD is a previously identified bipartite DNA binding domain of CDP/Cux (8, 9). Clearly, the DNA binding characteristics of CR3HD are different from those observed with CR1CR2. The p110 CDP/Cux, which contains CR2, CR3, and the Cut HD, was able to bind to DNA stably, in a manner similar to that of CR3HD. The absence of stable DNA binding activity in the context of the full-length CDP/Cux protein indicates that the DNA binding activity of CR3HD is specific to the p110 CDP/Cux protein. A proteolytic event seems to be required for this activity. No other modification has been identified to activate stable DNA binding activity in the context of p200 CDP/Cux. It appears that the presence of the N-terminal portion of CDP/Cux suppress the DNA binding activity of CR3HD. This

suggests the presence of an inhibitory domain in the N-terminal portion of p200 CDP/Cux. Ongoing experiments performed by a postdoctoral fellow in the laboratory have localized this localization of this domain to the first 300 amino acids of the protein (Raynal et al., manuscript in preparation). Interestingly, a new transcript initiating within intron 20 of CUTL1 has been cloned (Goulet et al., manuscript submitted for publication). It contains CR3HD as the DNA binding domains, an alternative way to produce a CDP/Cux protein with stable DNA binding activity.

The DNA binding specificity of CR3HD is different from that of CR1CR2. However, many of the CR3HD binding sites could be recognized by CR1CR2 since the consensus recognition sequence of CR3HD, ATCGAT, contains the consensus binding site of CR1CR2 (ATCGAT). If we presume that the p110 CDP/Cux proteins have a similar DNA binding specificity to CR3HD, it is tempting to speculate that some of the sequences recognized by p200 CDP/Cux might also be competitively bound by p110 CDP/Cux if any one of two C(A/G)AT site resembles ATCGAT. Since p200 and p110 CDP/Cux were shown to have an opposite effect on transcription, the relative expression levels of the two isoforms might be important for transcriptional regulation for genes with such sequences in their promoter. For genes shown to be regulated in a cell cycle dependant manner by CDP/Cux, the increased expression level of p110 CDP/Cux in G1/S might be a crucial event for their transcriptional regulation. There are many examples where both transactivation and repression of a gene is regulated by the same regulatory element within its promoter. Transcriptional regulation by Myc/Max/Mad (1) and E2F/DP/RB (4) transcription factors are such examples. It will be interesting to determine the specific targets of p200 and p110 CDP/Cux. However, identifying the specific *in vivo* targets of p200 and p110 CDP/Cux using technique such as ChIP assay is a difficult task since an

antibody that recognize p110 CDP/Cux would also recognize p200 CDP/Cux. An antibody specific to the proteolytically processed form of a protein is clearly necessary for this experiment.

Modulation of DNA binding Activity.

Phosphorylation on number of sites in the CDP/Cux has been shown to modulate the DNA binding activity. Phosphorylation by CKII, PKC, and Cdks occur either within a DNA binding domain or in close proximity to it (2, 3, 13). CKII and PKC phosphorylate conserved motifs in all three CRs and decrease their affinity for DNA. As a result, the DNA binding activities of both p200 and p110 CDP/Cux should be inhibited. It is still unclear in which physiological situation CKII and PKC down-regulate the CDP/Cux DNA binding activity. In contrast, modification by Cdks seems to regulate primarily the stable DNA binding activities of CDP/Cux (CR2CR3HD and CR3HD). Cyclin D/Cdk4 phosphorylates CDP/Cux to inhibit the stable DNA binding by CR3HD in G1 and the proteolytic cleavage of CDP/Cux, preventing the production of the p110 CDP/Cux protein. These results, together with the previous findings that Cdc25A stimulates DNA binding by CR3HD in S phase and that cyclin A/Cdk1 down-modulates p110 DNA binding in G2, indicate that the stable DNA binding activity of CDP/Cux is tightly regulated in a cell cycle dependent manner, such that it is highest in S-phase. In contrast, the CCAAT-displacement activity has not been found to be regulated in a cell cycle dependent manner. It is tempting to speculate that the first class of CDP/Cux target genes that are shown to be repressed in proliferating precursor cells are primarily regulated by the CCAAT-displacement activity since Cdks should not have much effect on CR1CR2 DNA binding activity in proliferating cells. p110 CDP/Cux is probably responsible for regulating the second class of CDP/Cux targets genes that are regulated in a cell cycle dependent manner. (See introduction)

Proteolysis of CDP/Cux

Contrary to several examples mentioned in the introduction, the site-specific proteolysis of CDP/Cux did not change the intracellular localization of the protein. Instead, it changed its DNA binding activity and effect on transcription. It is the first case to my knowledge in which the activity of a transcription factor is regulated by a site-specific protease in a cell cycle dependent manner. It is possible that the site-specific proteolysis is more commonly used post-translational modification for many different proteins than what one might suspect. However, there are several reasons why a proteolytically created isoform might have been ignored in different studies. First, the cleaved product might not contain the necessary epitope for the detection by Western blot or immunoflorescence. Second, it is hard to distinguish a product of site-specific proteolysis from a degradation product. Third, there is always a possibility that the proteolysis could have occurred after cells have been lysed. In the case of CDP/Cux, the immunoflorescence data provided an important evidence that the proteolysis occurred *in vivo*. Due to the clear differences in intracellular localization, it has been easier to demonstrate *in vivo* proteolysis of proteins such as Notch and SREBP (10, 17).

We still do not know the identity of the protease that cleaves CDP/Cux. So far, studies of cell cycle dependent proteolysis have mainly been focused on ubiquitin-dependent proteolysis by the proteasome (16). It is likely that the protease that cleaves CDP/Cux targets other proteins that might be important for the G1 to S transition. It would be interesting to identify the protease and other targets that are cleaved by it. Such a finding would provide new insight into the regulation of the G1 to S phase transition. Indeed, there has been a report that a cysteine protease inhibitor can block the G1 to S progression while lactacystin, a specific inhibitor of proteasome, could not (11).

CDP/Cux in cell cycle

Biochemical evidence clearly suggests that the stable DNA binding activity of CDP/Cux is activated in S phase, suggesting that p110 CDP/Cux has an important role in cell cycle progression. However, the *cux*-1 knockout mouse that does express neither p200 nor p110 CDP/Cux did not display any clear developmental defects that could result from abnormal cell cycle progression (5). Recently, the *cux-2* gene, another mouse homologue of *cut,* was cloned (12). It is possible that Cux-2 can compensate for the loss of Cux-1 in the *cux-1* knockout mouse. In an early study the expression pattern of cux-2 was shown to be restricted to the nervous tissue. However, a more recent study in chicken embryo demonstrated that cux-2 can be expressed in tissues other than the nervous tissue. Obviously, more work is required to resolve the tissue-specific expression of cux-2 and determine whether it may be expressed in additional tissues in the context of a cux-1 knockout. A double knockout of Cux-1 and Cux-2 should be generated to address this issue. However, there is an alternative explanation for the lack of cell cycle defects in the *cux-1* knockout mouse. The knockout mouse generated from this study does not express functional p200 and p110 CDP/Cux. Because both isoforms share some similarities in their target sequence specificity and are suggested to have an opposite effect on transcription, it is possible that p200 and p110 CDP/Cux antagonize each other's activity during the cell cycle. Since the expression of p110 CDP/Cux is S phase specific, a subset of CDP/Cux target genes could be repressed by p200 CDP/Cux in G1, and activated in S phase when the expression level of the p110 CDP/Cux is increased. Elimination of the functional p200 and p110 CDP/Cux protein could result in a milder phenotype since both negative and positive regulatory factors are lost for such genes. If this is true, then most of the phenotypes observed in the $cux^{-/-}$ mouse could be the result of the deregulation of genes regulated by

p200 CDP/Cux, and not by p110 CDP/Cux. The compensatory effect of losing two antagonizing factors has been suggested in the study of *Drosophila* E2F (6). It would be interesting to design a mouse model or a tissue culture system where the activity of a specific isoform of CDP/Cux is altered with out affecting the activity of the other.

Given that p110 CDP/Cux might activate transcription directly, the proteolysis event that creates p110 CDP/Cux seems to be a critical step for CDP/Cux to behave as an activator of transcription. It is still unclear how p110 CDP/Cux can activate transcription while p200 CDP/Cux cannot. One of the possible mechanisms how CDP/Cux is converted from a repressor to an activator of transcription is that the proteolysis uncovers motifs that are usually hidden in the context of p200 CDP/Cux. One can imagine that the truncation of the N-terminal half of CDP/Cux could change the overall structure of the protein or some covalent modifications in a specific CDP/Cux isoform could be used as interaction motifs. One important aspect to point out is that the extreme C-terminus of CDP/Cux contains two active repression domains. This means that somehow in the context of p110 CDP/Cux, the repression domain has to be inactivated. One interesting observation is that in S phase, when one detects the proteolysis between CR1 and CR2, another proteolysis event occurs that cleaves between the Cut HD and the C-terminal repression domain. It is possible that the true isoform of CDP/Cux that can activate transcription might be truncated both at the amino- and carboxy- terminus.

CDP/Cux in cancer.

12 out of 16 uterine Leiomyomas samples showed increases in the expression level of N-terminally truncated CDP/Cux isoforms as compared to the normal tissues. Clearly, tumor cells that express higher level of the shorter CDP/Cux protein have been selected for during tumor development. It is still unclear whether the higher expression of p110

CDP/Cux is the consequence of increased level of proliferation in tumor cells or it actually contributes to the higher level of proliferation in these cells. Given the fact that CDP/Cux participates in the regulation of genes necessary for S phase, it is probable that p110 CDP/Cux plays a more active role in tumors. It is the first evidence that suggests CDP/Cux might be an oncogene. It would be interesting to see if the expression of p110 is also elevated in other human cancers. However, it is hard to obtain enough pure protein extracts from most human tumors for Western blot analysis. An alternative way to investigate the protein expression pattern in tissues is to use an immuno-histological approach. Since all the antibodies available that recognize the p110 CDP/Cux protein also recognize full-length form, it will be important to generate an antibody specific to p110 CDP/Cux.

One possible mechanism by which the expression level of p110 CDP/Cux is increased in human uterine leiomyomas is that the regulatory modification on CDP/Cux that controls its proteolysis may become deregulated in tumors. For example, the level of phosphorylation on CDP/Cux by cyclin D/Cdk4 could be decreased in tumors. The absence of the kinase activity or an increased phosphatase activity in cancer could potentially results in over-expression of p110 CDP/Cux. However, cyclin D/Cdk4 is probably not responsible for the increased expression of p110 CDP/Cux in cancer since its activity is often upregulated in human cancers (7). In contrast, the Cdc25A phosphatase has been shown to be overexpressed in certain cancers. Alternatively, the protease that cleaves CDP/Cux could be upregulated in cancer, suggesting that an increase of an intra-cellular protease activity might contribute to tumor development. As mentioned previously, the targets of the protease could be important for G1 to S progression and increased protease activity might contribute directly to tumorigenesis. If this is true, it is possible to envisage the protease as a target for a chemotherapeutic reagent that could be

used in the treatment of cancer. Interestingly, there is evidence from our laboratory that the protease activity that cleaves p200 CDP/Cux is increased in some transformed tissue culture cells, and that the regulation of the cell cycle dependent expression of p110 CDP/Cux is lost.

REFERENCE

1. **Bernards, R.** 1995. Transcriptional regulation. Flipping the Myc switch. Curr Biol **5:**859-61.

2. **Coqueret, O., G. Berube, and A. Nepveu.** 1996. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Protein Kinase C. Journal of Biological Chemistry **271:**24862-68.

3. **Coqueret, O., N. Martin, G. Berube, M. Rabbat, D. W. Litchfield, and A. Nepveu.** 1998. DNA Binding By Cut Homeodomain Proteins Is Down-Modulated By Casein Kinase II. Journal of Biological Chemistry **273:**2561-66.

4. **Dyson, N.** 1998. The regulation of E2F by pRB-family proteins. Genes Dev **12:**2245-62.

5. **Ellis, T., L. Gambardella, M. Horcher, S. Tschanz, J. Capol, P. Bertram, W. Jochum, Y. Barrandon, and M. Busslinger.** 2001. The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev **15:**2307-19.

6. **Frolov, M. V., D. S. Huen, O. Stevaux, D. Dimova, K. Balczarek-Strang, M. Elsdon, and N. J. Dyson.** 2001. Functional antagonism between E2F family members. Genes Dev **15:**2146-60.

7. **Gillett, C., V. Fantl, R. Smith, C. Fisher, J. Bartek, C. Dickson, D. Barnes, and G. Peters.** 1994. Amplification and overexpression of cyclin D1 in breast cancer detected by immunohistochemical staining. Cancer Res **54:**1812-7.

8. **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. Molecular & Cellular Biology **15:**129-40.

9. **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. Journal of Biological Chemistry **269:**2062-7.

10. **Kopan, R., E. H. Schroeter, H. Weintraub, and J. S. Nye.** 1996. Signal transduction by activated mNotch: importance of proteolytic processing and its regulation by the extracellular domain. Proc Natl Acad Sci U S A **93:**1683-8.

11. **Mellgren, R. L.** 1997. Evidence for participation of a calpain-like cysteine protease in cell cycle progression through late G1 phase. Biochem Biophys Res Commun **236:**555-8.

12. **Quaggin, S. E., G. B. Vandenheuvel, K. Golden, R. Bodmer, and P. Igarashi.** 1996. Primary Structure, Neural-Specific Expression, and Chromosomal Localization Of Cux-

2, a Second Murine Homeobox Gene Related to Drosophila Cut. Journal of Biological Chemistry **271:**22624-34.

13. **Santaguida, M., Q. Ding, G. Berube, M. Truscott, P. Whyte, and A. Nepveu.** 2001. Phosphorylation of the CDP/Cux transcription factor by cyclin A/Cdk1 modulates its DNA binding activity in G2. J Biol Chem **2:**2.

14. **Sinclair, A. M., J. A. Lee, A. Goldstein, D. Xing, S. Liu, R. Ju, P. W. Tucker, E. J. Neufeld, and R. H. Scheuermann.** 2001. Lymphoid apoptosis and myeloid hyperplasia in CCAAT displacement protein mutant mice. Blood **98:**3658-67.

15. **Tufarelli, C., Y. Fujiwara, D. C. Zappulla, and E. J. Neufeld.** 1998. Hair defects and pup loss in mice with targeted deletion of the first cut repeat domain of the Cux/CDP homeoprotein gene. Dev Biol **200:**69-81.

16. **Tyers, M., and P. Jorgensen.** 2000. Proteolysis and the cell cycle: with this RING I do thee destroy. Curr Opin Genet Dev **10:**54-64.

17. **Wang, X., R. Sato, M. S. Brown, X. Hua, and J. L. Goldstein.** 1994. SREBP-1, a membrane-bound transcription factor released by sterol- regulated proteolysis. Cell **77:**53-62.

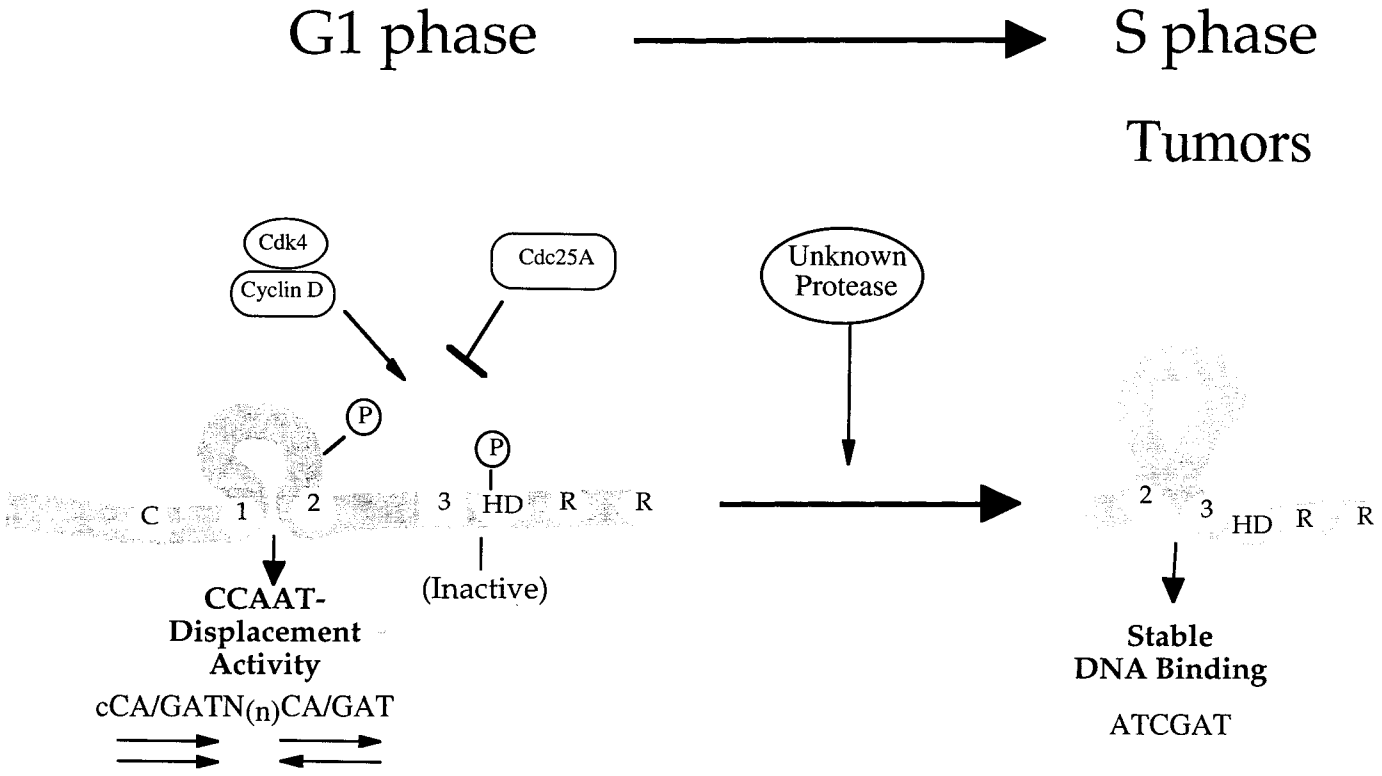


Figure 1. Alternative modes and cell cycle dependent regulation of CDP/Cux DNA binding .

Contribution to original research

1 Identification and characterization of CR1+2 as the bipartite DNA binding domain responsible for the CCAAT-displacement activity by CDP/Cux. This finding provided crucial insight on how CDP/Cux can recognize a wide range of DNA sequence and exert the CCAAT-displacement activity.

2 Identification and characterization of a new CDP/Cux isoform, p110 CDP/Cux, which is a product of S phase site-specific proteolysis. This isoform contains DNA binding characteristics and effects on transcription that differ from that of full length CDP/Cux, p200 CDP/Cux. This finding suggests that site specific proteolysis of CDP/Cux might be an important component of the G1 to S phase transition in mammalian cells.

3 The observation that the N-terminally truncated CDP/Cux protein is expressed at a higher level in human Uterine Leiomyomas than in normal tissue may be the first evidence that CDP/Cux might behave as an Oncogene.

4 Demonstration that the cell cycle dependant DNA binding activity and the S phase site-specific processing of CDP/Cux are both regulated by cyclin D/Cdk 4 phosphorylation. This suggests that CDP/Cux is a novel target of cyclin D/Cdk4 kinase complex.