Alternative Mechanisms to Generate Short CDP/Cux Isoforms in Cancer Cells

Brigitte Nicole Goulet

A thesis submitted to the Faculty of Graduate and Post-doctoral Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Biochemistry
McGill University
Montreal, Quebec, Canada
August 2003

Your file Votre référence
ISBN: 0-612-98263-7
Our file Notre référence
ISBN: 0-612-98263-7

Abstract

CDP/Cux is a transcription factor involved in the control of cellular differentiation and proliferation. It contains four DNA binding domains, three Cut Repeats (CR) and a Cut homeodomain (HD). The full-length p200 protein makes an unstable interaction with DNA via CR1CR2 and is responsible for the CCAAT displacement activity of CDP/Cux. A more stable DNA binding activity could also be detected. This particular activity increased as cells progress from G1 to S phase of the cell cycle following two events. One event is the dephosphorylation of CR3HD by the by the cdc25A phosphatase. The other event involves the proteolytic processing of the full length CDP/Cux protein into a p110 isoform. At the transcriptional level, the p110 protein is able to repress the *P21$^{waf1/Cip1}$* gene promoter and to activate the *DNA polymerase alpha* gene promoter as well as other S phase specific genes. In human uterine leiomyomas, an increase in the steady state levels of p110 was observed compared to the normal surrounding myometrium.

The goal of my project was to identify the protease involved in the proteolytic processing of CDP/Cux. I showed that the "lysosomal" cysteine protease cathepsin L translocates to the nucleus in a cell cycle dependent manner and was able to cleave p200 CDP/Cux into the p110 form. I also demonstrated that this protease was translated at downstream AUGs to produce truncated proteins devoid of signal peptide. These truncated cathepsin L species were unable to enter the endoplasmic reticulum and be routed to the lysosomes. In transformed cells, I found that CDP/Cux processing was increased and no longer regulated in a cell cycle dependent manner. This correlated with the enhanced expression and nuclear activity of cathepsin L. I also identified a novel shorter CDP/Cux isoform, which contains only CR3HD as DNA binding domains and is translated from an

mRNA initiated within the intron 20 of CUTL1. Interestingly, the intron20-initiated mRNA was distributed in a tissue specific manner, and was expressed only weakly or not at all in normal breast tissue and in human mammary epithelial cells. However, it was detected in many breast tumor cell lines and invasive lobular carcinomas. These results indicated that in cancer, several mechanisms could generate shorter active CDP/Cux isoforms, and that these amino-terminally truncated CDP/Cux proteins play a role in tumorigenesis.

Sommaire

Le facteur de transcription CDP/Cux est impliqué dans la differentiation et la prolifération cellulaire. Il possède quatre domaines de liaison à l'ADN; trois Répétitions Cut (RC) et un homéodomaine Cut (HD), L'intéraction de la protéine entière p200 avec l'ADN se fait de façon instable, via les domaines RC1RC2. Ceux-ci sont responsables de l'activité de déplacement CCAAT de CDP/Cux ("CCAAT displacement activity"). Toutefois, CDP/Cux peut également lier l'ADN de façon stable. Cette liaison stable apparaît lorsque les cellules passent de la phase G1 à la phase S du cycle cellulaire, et ce, suite à deux évènements. La région du RC3HD est déphosphorylée par la phosphatase Cdc25A, et la protéine p200 est clivée en une forme de 110 kDa (p110), qui ne contient pas la portion amino-terminale. La forme p110 peut réprimer la transcription du promoteur du gène de $p21^{waf1/Cip1}$ et activer, en autres, celle du promoteur du gène de la *polymérase de l'ADN alpha*. En outre, plusieurs leiomyomes utérins humains analysés contiennent plus de p110 que leurs myometrium normaux adjacents respectifs.

Le but de mon projet a été d'identifier la protéase responsable du clivage protéolytique de CDP/Cux. J'ai établi que la cathépsine L, une protéase à cystéine résidente du lysosome, est transférée dans le noyau, en fonction du cycle cellulaire, pour cliver p200 en sa forme de 110 kDa. La localisation nucléaire de la cathépsine L depend d'une traduction à partir d'AUGs situés en aval de l'AUG initiateur. Ces protéines plus courtes ne contiennent donc pas de peptide signal, qui leur permet habituellement d'être insérées dans la lumière du réticulum endoplasmique et d' être éventuellement ciblées vers les lysosomes. J'ai de plus montré que dans les cellules transformées, le clivage de CDP/Cux était d'une part augmenté, et que d'autre part, il n'était plus contrôlé en fonction

du cycle cellulaire. Cette observation concorde avec l'augmentation, dans ces cellules, de l'expression et de l'activité de la cathépsine L nucléaire. J'ai également identifié une nouvelle isoforme de CDP/Cux qui ne contient, comme domaines de liaison à l'ADN, que le RC3 et le HD. Cette protéine est traduite d'un ARNm qui est initié dans l'intron 20 du gène CUTL1. Curieusement, cet ARNm est spécifiquement distribué dans certains tissus, et aucune expression n'a été observée ni dans les glandes mammaires humaines, ni dans les cellules épithéliales humaines issues de glandes mammaires. En revanche, il a été détecté dans plusieurs lignées cellulaires de tumeurs du sein, ainsi que dans les carcinomes lobulaires invasifs. Ces résultats indiquent que dans les cellules cancereuses, plusieurs mécanismes peuvent générer des petites formes actives de CDP/Cux. Ces protéines tronquées en leur amino-terminal, semblent donc avoir un rôle lors de la tumorigénèse.

Acknowledgements

I would like to acknowledge my supervisor, Dr. Alain Nepveu for his intellectual input through these studies, and for his good methodology in scientific processes.

I also acknowledge the financial support of Canderel, FCAR, the faculty of medicine and the Royal Victoria Institute.

I wish to greatly thank the past and present members of Dr. Nepveu's lab for their friendship. Their support, encouragements, endless discussions and laughs during all these years were invaluable. Nathalie, Nam-Sung, Oriana, Peter, Madeleine, Lelia, Adela, Marianne, Mary, Laurent, Hélène, Ryoko, Brian, Mélanie, Ginette and Lam, I don't know how to thank you enough. You guys are wonderful. I wish to also extend my thanks to other members of MOG, especially people from Dr. Morag Park's lab, Christiane, Tanya, Louie, Hanane, Pascal, Lisa, Greg, Melanie, Jasmine, Sonia, Veena and Caroline for their friendship and encouragements. The diners on the terrace and the beers we had on Friday nights were always agreeable! Thank you all for the great time.

Caroline, merci pour tes judicieux conseils et pour ton amitié. Ginette, merci aussi pour ces heures passées à m'écouter et à m'encourager durant ces années. Ton aide au labo à également été grandement appréciée. Thank you Hashem, for encouraging me in continuing my thesis when I was very discouraged.

Monica, I really appreciated all your help in tissue culture. Your knowledge of cells is priceless. Thanks also Melanie for showing me a faster way of doing mutagenesis. I particularly appreciated your secret ingredient: Barry White! A big thank you to Raf, not only for his friendship but also for making sure that we all have clean glassware to do our experiments.

Lam, je ne sais pas comment te remercier pour ton aide quotidienne au labo. Tes connaissances autant scientifiques qu'humaines sont inestimables. J'ai évolué en tant qu'individu et en tant que scientifique, grâce à tes multiples conseils, qui me resteront.

I also wish to also express my deep gratitude to Mary, Brian and Gabriel for the editing of my thesis and useful suggestions. Your time and efforts were really appreciated. Thank you Maureen Caron, for everything you can do for us, students!

Merci la gang. Qu'aurais-je fais sans ces soirées mémorables, ces dïners de sacoches et notre chère amitié. Anne-Claude, Jeanne, et Madani, les Thompson House, les bouffes, les discussions, les encouragements et surtout les rigolades ont grandement contribués à ma santé mentale. Je vous en remercie chaudement.

Je tiens finalement à profondement remercier ma famille. Mes parents et Elodie ont été d'un soutien inconditionnel durant toutes ces années. Je ne vous remercirais jamais assez d'avoir été presents durant les moments difficiles que j'ai eu a traverser. Gabriel, I also deeply thank you for your patience, your understanding, your continuous encouragements and your love.

Preface

The guidelines concerning thesis preparation issued by the Faculty of Graduate and Postdoctoral Studies at McGill University reads as follows:

1. 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. (Reprints of published papers can be included in the appendices at the end of the thesis.)

2. 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 preceeding and following each manuscript are mandatory.

3. The thesis must conform to all other requirements of the "Guidelines for Thesis Preparation" in addition to the manuscripts.

The thesis must include the following:

1. a table of contents;
2. a brief abstract in both English and French;
3. an introduction which clearly states the rational and objectives of the research;
4. a comprehensive review of the literature (in addition to that covered in the introduction to each paper);
5. a final conclusion and summary;
6. a thorough bibliography;
7. Appendix containing an ethics certificate in the case of research involving human or animal subjects, microorganisms, living cells, other biohazards and/or radioactive material.

4. 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 judgement to be made of the importance and originality of the research reported in the thesis.

5. 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 defence. 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.

I have chosen to write my thesis according to these guidelines, with three published manuscripts and one submitted manuscript. The thesis is organized into seven chapters: (I) a general introduction and literature review, (II-V) manuscripts, each with their own abstract, introduction, materials and methods, results, discussion and references, (VI) a general discussion of all results with references, and (VII) claims to original research.

Publications arising from work of the thesis

First-author publications

1- **Goulet B.**, Watson P., Poirier M., LeDuy L., Bérubé G., Meterissian S., Jolicoeur P. and Nepveu A. (2002). Characterization of a Tissue-Specific CDP/Cux Isoform, p75, Activated in Breast Tumor Cells. *Cancer Research* **62**:6625-6633.

2- **Goulet B.***, Baruch A.*, Moon N.S., Poirier M., Erickson A., Bogyo M. and Nepveu A. A Cathepsin L Isoform that Is Devoid of a Signal Peptide Localizes to the Nucleus in S Phase and Processes the CDP/Cux Transcription Factor. *Submitted*

* These authors contributed equally

3- **Goulet B.**, Baruch A., LeDuy L., Bogyo M. and Nepveu A. Proteolytic Processing of the CDP/Cux Transcription Factor Is Activated in Transformed Cells and Correlates with Increased Cathepsin L Activity. *To Be Submitted*

Other publications:

1- Truscott M., Raynal L., Premdas P., **Goulet B.**, LeDuy L., Bérubé G. and Nepveu A. (2003) CDP/Cux Stimulates Transcription from the DNA Polymerase α Gene Promoter. *Molecular and Cellular Biology* **23** (8): 3013-3028

Contribution of Authors

Watson P.

In manuscript number 1, he provided us with human breast tumors samples and contributed the statistical analysis of the expression of intron 20 mRNA in these tumors (in figure 8).

Poirier M.

In manuscript number 1, she carried out the staining for the indirect immunofluorescence in figure 3D.

In manuscript number 2, she helped me in figure 5A by performing indirect immunofluorescence and counted cells to make the table of figure 5B. She also went to confocal microscopy and did figure 5C.

LeDuy L.

In manuscript number 1, he provided technical support where he did RT-PCRs in figures 1C, 2A, 6B, 6D and 8 and RNAse protection assay of figure 1B.

In manuscript number 3, he helped me with the centrifugal elutriation and Western blots presented in figure 5.

Bérubé G.

In manuscript number 1, she constructed the plasmid expressing the human intron20-mRNA and contributed the RNAse protection assay of figure 1A.

Jolicoeur P.

In manuscript number 1, he provided us with single ($CD4^+$ and $CD8^+$) and double positive ($CD4^+CD8^+$) mouse thymocytes used in figure 2B.

Baruch A.

In manuscript number 2, he performed figure 2 (*in vitro* processing assays and Edman degradation), the Western blot of figure 3B, and the cathepsin L activity tests of figure 4.

In manuscript number 3, he did the cathepsin L activity test of figure 6B.

Moon N.S.

In manuscript number 2, he carried out the Western blot and EMSA of figure 1A and 1B.

Erickson A.

In manuscript number 2, she provided us her polyclonal cathepsin L antibody.

In the paper of Truscott *et al.,* I performed Western blot analysis and EMSA of various short CDP/Cux isoforms.

Table of Contents

List of Figures

Chapter I Literature Review

Chapter II A Cathepsin L Isoform that Is Devoid of Signal Peptide Localizes to the Nucleus in S phase and Process the CDP/Cux Transcription Factor

Chapter III Elevated Cathepsin L Expression In Cancer Cells Correlates with Increased Proteolytic Processing of the CDP/Cux Transcription Factor

Chapter IV Characterization of a Tissue-Specific CDP/Cux Isoform, p75, Activated in Breast Tumor Cells

Chapter V General Discussion

Chapter I Introduction

The normal development of an organism requires that cell homeostasis be tightly regulated. Improper expression of particular genes or modification of other specific proteins can result in cellular transformation and cancer. These cancer genes, called oncogenes (deregulation thereof gives rise to a "gain of function") or tumor suppressor genes ("loss of function") are usually involved in regulating cellular proliferation, growth, differentiation and death. These key proteins are essential components of the multiple signaling cascades that bridge a given exterior stimulus to a cellular response. For example, a growth factor will bind to a specific cell surface receptor, which will then activate a cascade of signal transduction molecules. This results in a modification at the level of gene transcription or, in certain cases, protein translation. Once genes are transcribed and proteins are translated, the activity of a protein can also be regulated by post-translational modifications such as phosphorylation, glycosylation or acetylation. Another modification is site-specific proteolysis. Proteolysis is a regulated event that can serve multiple purposes. It can degrade a protein that is no longer necessary (for example via the proteasome pathways or the lysosomal pathways), or control the activity of another protein. In this latter case, proteolysis of a full-length protein can either generate a novel isoform with different functions (for example, proteolysis of the Notch cell surface receptor can generate a transcription factor), or can change its activity or its location (as is the case for the endoplasmic-reticulum-membrane bound transcription factor ATF6). Again, as for signaling molecules, transcription factors or translational regulators, deregulation of protease activity can also be detrimental to normal cellular function. In cancer, increased

protease activity has often been linked to degradation of the extracellular matrix and metastasis.

This literature review will first focus on the regulation of gene expression, particularly at the level of protein translation. Site-directed proteolysis and cathepsin L, a member of the family of cysteine proteases, will then be discussed. Cellular transformation and breast cancer will also be reviewed. Finally, the features and roles of CDP/Cux, a transcription factor involved in cellular differentiation in *Drosophila melanogaster* and, in cell cycle progression in mammalian cells, will be summarized.

1- Regulation of Gene Expression

In cells, "housekeeping genes" are constitutively expressed and encode structural proteins and metabolic enzymes. Other genes can be expressed in a cell type-specific manner, at certain stages of development or, under proliferative conditions. It is the regulation of these genes that must be tightly controlled.

1.1 Regulation by Gene Transcription

Interestingly, protein-coding sequences represent only 2% of the human genome. A large portion of the genomic sequences is thought to be used to regulate gene expression (199). In order to target gene expression at the right place and time, an amalgam of factors and *cis*-elements are needed to interact.

DNA is wrapped around histones forming the nucleosomes, which are considered as obstructions to the regulatory sequences. Following a specific signal issued from a transduction pathway, these nucleosomes have to be displaced for transcriptional activation

to occur. They are remodeled by chromatin modifying complexes. The core promoter of a gene is situated upstream of the transcription start site. It contains a TATA box, and/or an initiator element (INR) and/or a downstream promoter element (DPE), which recruits the TFIID subunit of TATA binding protein (TBP). At the 5' end of the core promoter are binding sites for factors that either activate or repress transcription of a given gene (307). These factors may alternatively act as "tethering elements" to recruit some distal elements to the core promoter (46). Enhancers are DNA sequences that can increase transcriptional activity. They can be located at the 5' or 3' end of the gene or within introns. They usually direct tissue or cell type-specific expression of the gene. Silenced domains are inactive regions of chromatin composed of hypo-acetylated histones and methylated histone H3 along with silencing factors (89). These regions preclude the spread of heterochromatin (250). Finally, insulators prevent regulation of neighboring genes by restricting enhancer and silencer functions (117).

Specificity and diversification of gene expression patterns is thus obtained by a complex interplay of many factors, depending upon cellular context, and type of signals i.e. combining regulation by initiation complexes with that of cofactors and chromatin remodeling/modifying complexes (199).

1.1.1 Regulation of Transcription Factor Activity

Transcription factors are DNA binding proteins that, with their co-factors can either activate or repress the transcription of a given gene. Interestingly, some transcriptional activators and repressors can be encoded by the same gene. This usually occurs via alternative transcription initiation sites, alternative splicing or alternative translational initiation (see figure 1). These events generate proteins of different lengths that can have

different properties. These novel isoforms may or may not possess a DNA binding domain, an activation domain, or a repression domain (80, 108).

Once synthesized, transcription factor activity can be regulated by post-translational modifications or by binding of a ligand. The former can modulate the cellular localization of the factor, its DNA binding activity, or its binding with other partners (47). One versatile modification is phosphorylation (147). For some proteins, phosphorylation close to or within a DNA binding domain modifies DNA binding activity (38, 147, 203). In other cases, phosphorylation can change protein conformation and either mask or unmask a nuclear localization signal (NLS), a DNA binding domain, or an inhibitory domain, as exemplified by NF-AT, SAF-1 and Elk-1, respectively (275, 377, 387). The phosphorylation status of a transcription factor can also regulate its transcriptional activity by modulating its interactions with different partners. Dephosphorylated NF-κB interacts with HDAC1 deacetylase, and when phosphorylated it binds to p300/CBP transcription coactivators (386). Phosphorylation can also regulate protein steady-state levels. For example, phosphorylation of E2F-1 by components of the RNA polymerase II holoenzyme can result in degradation via the ubiquitin-proteasome pathway (323, 350). Interestingly, ubiquitylation and transcription are linked in two manners. First, ubiquitin ligases are enzymes that usually add a chain of ubiquitin residues to a target protein as a signal to send it for degradation via the proteasome pathway (63, 66, 141)(and see section 3). Second, ubiquitylation can occur within an activation domain, enhancing its activity (233, 285). For example, the ubiquitin-ligase skp2 can stimulate transcription induced by c-myc as well as its degradation (170, 357). It is of interest to note that the acetylation of transcription factors by co-factors, such as P/CAF or p300/CBP can produce effects similar to those described for phosphorylation (118, 125, 265, 274).

1.2 Regulation by Protein Translation

Following the initiation of transcription, a given eukaryotic RNA transcript is first synthesized as a precursor heterogeneous nuclear RNA (or hnRNA), which possess introns and exons. Shortly after the beginning of transcription, the 5' end of the precursor RNA is capped (m^7GpppN); with a methyl moiety on the G nucleotide (297). This cap structure helps stabilize the new RNA, and is important for the initation of translation (182). A poly(A) tail is then added at the 3'end of the transcript. This tail is important for the stability of the mRNA, as well as for the nuclear export of the RNA, and, via poly(A) binding protein (PABP), for the initation of translation (325, 326). Before nuclear export, the hnRNA is spliced to remove non-coding introns and is now termed the messenger RNA (or mRNA). Each of these steps is intricately controlled. Once in the cytosol, translation of that specific mRNA is tightly regulated and allows for a more rapid response following a given signal than starting directly from transcription (123). As mentioned with gene transcription, translation can also generate multiple proteins with different functions by the use of alternative translation initiations sites (figure 1) (80).

1.2.1 Overview of Translation Initiation-

Protein synthesis can be controlled at the level of translation initiation. This depends on the availability of the various translation initiation factors (eIFs), which are involved in recruiting the ribosome to the mRNA, helping the recognition and selection of the initiator AUG codon (scanning), facilitating the binding of Met-$tRNA_i^{Met}$ to the small ribosomal 40S subunit, and promoting elongation (10, 88, 178, 259). Thirteen different initiation factors have been identified so far in eukaryotes. The recognition of the mRNA by Met-$tRNA_i^{Met}$-40S-initiation factors (43S pre-initiation complexes) is mediated by eIF3

and eIF4. The eIF4F 5'cap binding complex contains the cap binding protein eIF4E, the scaffold adaptor protein eIF4G, and the ATPase/RNA helicase eIF4A that unwinds secondary structures in the 5'UTR (218). This process is believed to facilitate ribosomal recruitment and formation of the 43S complex (257). The 43S ribosome complex then scans the mRNA until it reaches the first initiation codon. Once the initiating AUG is recognized, some initiation factors dissociate from the ribosome and the 60S large ribosomal subunit joins the complex, forming a functional 80S ribosome that can start the elongation phase of translation (88, 260).

1.2.2 Regulation of Protein Translation

The 3'-untranslated region (3'UTR) of mRNA is important in translational control as it contains various *cis*-regulatory elements on which specific translation factors can bind to repress or activate translation. This is the case for CPEB (cytoplasmic polyadenylation element binding protein) an RNA binding protein involved in the repression or activation of translation, depending on its binding partners (231). The interaction between proteins from the 5'-end (eIF-4G) and 3'end (PABP) of the mRNA (making a circular mRNA) is also believed to stimulate translational initiation by a currently unknown mechanism (267, 268).

In 1991, Marilyn Kozak proposed that the initiation of translation in higher eukaryotes was modulated by 5 aspects of the mRNA structure: the 5'cap structure, the sequence surrounding the AUG codon, the position of the AUG (whether it is the first AUG or not), the secondary structure of the mRNA and the length of the 5'end (182). Her studies using specific mutagenesis revealed that the optimal context for initiation of translation in tissue culture cells is the sequence GCC(A/G)CCAUGG (176, 180). It is still

not known how the nucleotides surrounding the first AUG interact with ribosomes and are involved in the recognition of this initiator AUG. In the scanning model, the ribosome migrates linearly in a 5' to 3' direction and stops at the first AUG.

The 5'-UTR sometimes contains stable secondary structures that affect translation initiation efficiency. GC-rich regions form stem-loops that can prevent the binding of the ribosome, especially if they are close to the 5'end of the mRNA. Other genes may have small upstream open reading frames (uORFs) in their 5' end that will be translated by the ribosomes first (111). The main ORF is then translated by re-initiation, a mechanism that greatly limits protein production. The generation of alternative transcripts that eliminate secondary structures and uORFs can therefore increase protein synthesis (177). In some cases, when the context surrounding the initial AUG is not optimal, internal AUG codons can be translated by a process called leaky scanning. This mechanism allows the generation of shorter isoforms of a protein that usually have different functions (48, 181, 317, 382). Another mechanism of translational control used to generate shorter isoforms of a protein is cap-independent translation initiation, mediated by internal ribosomal entry sites (IRES). IRES were initially identified in viruses, but are now being discovered in some mammalian mRNAs that either have a highly structured 5'UTR or are translated when cap-dependent translation is inhibited (258, 344). It must be noted that the existence, and significance of cellular IRES is still controversial (179, 293).

Another mechanism of regulating protein translation involves initiation factors. One example is eIF4F. The eIF4E-binding protein (4E-BP) reduces the formation of eIF4F by competing with eIF4G for a common binding site on eIF4E. This binding is dependent upon the phosphorylation status of 4E-BP. When 4E-BP is hypophosphorylated, it binds to eIF4E, inhibiting translation, but when hyperphosphorylated, it can not bind and translation

proceeds (122). Various stimuli, such as growth factors, hormones or nutrients, induce 4E-BP phosphorylation, whereas nutrient deprivation or stress conditions reduce it. A known inhibitor of cap-dependent translation is rapamycin, an immunosuppressant drug that prevents 4E-BP phosphorylation (23, 123). Rapamycin, as an immunosuppressant, is used in organ transplantation to prevent allograft rejections, and as a cell cycle inhibitor, and is also currently in clinical trial as an anti-tumor growth drug (79). The gene encoding eIF4E is considered to be an oncogene. Its overexpression causes malignant transformation by increasing translation of mRNAs that contain extensive secondary structures in their 5' UTR (48, 175). Important genes coding for regulatory proteins are transcribed in mRNAs with highly structured 5'UTR that are more resistant to translation initiation (177).

2- Proteolysis

Although proteolysis was often associated with terminal protein degradation (as a "housekeeping" function), it is now becoming clear that it is also an important way to regulate protein activity. Contrary to other post-translational modifications such as phosphorylation or acetylation, proteolysis is irreversible.

2.1 Proteolysis and the Cell Cycle

Cell cycle progression depends on the spatio-temporal expression of specific regulators. The various cyclins and their relative kinases, the cdks, are key mediators of the different cell cycle transitions. The steady state levels of these proteins and their specific inhibitors vary in a cell cycle-dependent manner. The ubiquitin/proteasome pathway allows each cell cycle transition by removing the mediators of the previous phase.

The proteasome 26S complex is located both in the cytosol and in the nucleus. It is composed of a barrel-shaped proteolytic core unit (20S) capped by two regulatory subunits, one of which contains ATPase activity and the other binds the ubiquitin chains. These regulatory units unfold ubiquitylated proteins and translocate them into the barrel for degradation (356). A protein targeted for degradation will first be flagged by the attachment of a polyubiquitin chain. The formation of polyubiquitin chains requires three components, a ubiquitin activating enzyme (E1), a ubiquitin-conjuging enzyme (E2) and a substrate recognition factor (E3). Two E3 complexes were shown to regulate the cell cycle, the SCF (skp-cullin-F-box complex) and the APC (anaphase-promoting complex). The APC complex is involved in mitosis, by targeting mitotic cyclins and spindle proteins. The SCF complex is the key regulator of the G1/S transition, marking (usually in a phosphorylation dependent manner), the G1 cyclins, the cdk inhibitors and some transcription factors for degradation (174).

An important regulator of G1 and S phase progression is the cdk inhibitor (CKI) p27 (139). Variations in the levels of p27 are crucial for normal cell cycle progression and deregulation is associated with cancer (32). Translation and two proteolytic events regulate the amount of p27 available (140). p27 is initially highly expressed in G0, where it binds tightly to cyclin E/cdk2 and inhibits its activity (139). Following a mitogenic signal, p27 synthesis is reduced and a fraction of p27 is degraded by the proteasome, independently of the phosphorylation of Thr 187 (32, 217). p27 is phosphorylated on Ser10 by hKIS and is actively exported to the cytoplasm where it is degraded by the proteasome pathway (34, 67, 150). Most cyclin E/cdk2 is then active and can phosphorylate p27 on Thr 187. This generates a docking site for the SCF-E3 ligase skp2, which ubiquitylates p27 targeting it for degradation via the proteasome pathway, thus liberating all cyclin E/cdk2 (335).

2.2 Site-Specific Proteolysis

Contrary to degradation, partial proteolysis regulates proteins by altering their activity, stability, and/or their cellular localization.

2.2.1 Change in Cellular Localization Following Proteolysis

Membrane Proteins

Interestingly, the proteasome was shown to not only degrade proteins, but also to have endoproteolytic activity (207). Biologically active proteins can therefore be generated from precursors by proteasomal processing. This is exemplified by the transcription factor NF-κB and its yeast homologues SPT23 and MGA2 (273). The transcription factors SPT23 and MGA2 are synthesized as inactive endoplasmic reticulum (ER) membrane-tethered precursors involved in the adjustment of the unsaturated fatty acid pool in yeast by activating the Δ9 fatty acid desaturase (146). Following mono-ubiquitination by RSP5, SPT23 dimerizes and is processed into p90 and p120 isoforms by the proteasome, with the help of the chaperone $CDC48^{UFD1/NPL4}$. $CDC48^{UFD1/NPL4}$ then removes p90 from its partner, allowing it to enter to the nucleus (273).

In mammalian cells, this regulated intramembrane proteolysis (RIP) therefore allows for a controlled release of active proteins and various examples of substrates and proteases can be cited (44, 362). These intramembrane cleaving proteases (I-CliPs) process directly within the lipid bilayers of the membranes. The first I-CliP was the metalloprotease S2P (site-2-protease), one of the proteases involved in the proteolytic processing of the transcription factor SREBP (sterol regulatory element binding proteins).

SREBP controls the transcription of genes involved in the biosynthesis of cholesterol and fatty acids (43). S2P is also responsible for the cleavage of ATF-6, another ER-membrane-bound precursor that becomes activated in response to the detection of unfolded proteins (378, 379). In normal conditions, the ER chaperone BiP binds to the luminal domain of the inactive ATF-6. When misfolded proteins accumulate, they sequester BiP, which liberates ATF-6 for trafficking in the Golgi, where it is then cleaved by S2P and S1P, another protease (58, 299). Once cleaved, ATF-6 translocates to the nucleus and activates stress genes. Other I-CliPs include the presenilins PS-1 and PS-2, which make up the catalytic subunit of the γ-secretase complex. They are well studied because of their involvement in the cleavage of the β-amyloid precursor protein (β-APP) in Alzheimer's disease. The presenilins also process the Notch receptor and ErbB4 receptors, liberating an intracellular domain, which then translocates to the nucleus to act as a transcription factor (362).

Cytoplasmic Proteins

Hedgehog (Hh) is a secreted signaling molecule in *Drosophila melanogaster* that is involved in tissue development and Cubitus interruptus (Ci) is one of its downstream effectors (232). Ci is a transcription factor that is mainly retained in the cytoplasm. In the absence of Hh signaling, Ci is cleaved and translocates to the nucleus. Contrary to the full-length protein, this truncated isoform is a transcriptional repressor (12). It contains the DNA binding domain but lacks the transactivation domain and the nuclear export signal (57). Regulation of proteolysis of the full length Ci depends on multiple phosphorylation events by PKA (protein kinase A), followed by GSK3 (glycogen synthase kinase 3) and CKI (casein kinase I) (59, 160, 269, 270). Phosphorylated Ci is then believed to be a target

of the F-box protein Slimb (161, 270). Interestingly, Ci proteolysis is proteasome dependent, but ubiquitylation has not been yet detected (57).

2.2.2 Change in Activity

A change in protein activity following proteolytic processing was observed for many transcription factors and precursor proteins.

The easiest example is the proteases themselves, which are synthesized as inactive precursors. Cleavage of the prodomain allows for enzymatic activity (77, 295).

The activity of some transcription factors changes following proteolytic cleavage. IRF-2 (interferon response factor 2) was first identified as a repressor of the β-interferon promoter, by displacing the activator IRF-1 (129, 130). Following viral infections, IRF-2 is cleaved into an amino-terminally truncated isoform with different DNA binding activities from the full-length protein. Whereas the full-length isoform binds only transiently to DNA, the truncated isoform makes a more stable interaction and is a stronger repressor (254, 368). The protease involved in this processing is likely a member of the ICE family of cysteine proteases (367).

Host-cell factor (HCF-1) is a chromatin-associated transcription factor and coactivator. It was initially identified as a cofactor for herpes simplex virus (HSV) VP16 immediate-early gene transcription (184). HCF-1 is proteolytically processed at a series of six 26 amino-acid repeats located towards the center of the precursor (371). The cleavage occurs in an autocatalytic manner to generate a series of polypeptides that remain tightly associated together (355). HCF-1 was shown to be involved in two stages of the cell cycle; passage through the G1 phase and exit from mitosis. Proteolytic processing is essential to separate these two functions of HCF-1. The amino-terminus of HCF-1 promotes cell

proliferation, whereas the carboxy terminus region is involved in proper exit of mitosis (165). Interestingly, HCF-1 was also shown to regulate transcription in a positive or negative manner. The amino terminus contains a transcriptional activation domain, and the carboxy terminus contains both activation and repression activities (212, 262, 355, 374).

The Signal Transducer and Activator of Transcription (STAT) is involved in cytokine signaling. Some of the members (STAT 3 and 5) were also shown to be proteolytically processed at the carboxy terminus. The truncated version of STAT 5 lacks a transactivation domain and thus functions as a dominant negative. This processing was found to be important in myeloid maturation (261). A nuclear serine protease and calpain were shown to mediate this cleavage (103, 249).

These are only few examples underlying the importance of proteolytic processing in the generation of different biological outputs from the same precursor.

3- Cysteine Proteases

Proteolytic enzymes are classed depending on the amino acid present in their catalytic site. These are the cysteine, serine or aspartic acid proteases. Alternatively, some proteases, called metalloproteases possess a zinc residue in their active site.

3.1 Family of Cysteine Proteases

Various cysteine protease or peptidase families have emerged from different evolutionary origins and can be separated into different clans. It is believed that the origin of the "peptidase domain", which is responsible for the peptide-bond-hydrolysis, arose from an ancestor of both bacteria and archea. Cysteine proteases have therefore been

evolving for at least three billion years, and have diverged into different families with unique specificities and structures (21). Cysteine proteases can be grouped into two superfamilies: the interleukin 1β converting enzymes (ICE) and the papains. Although both families have a cysteine in their active site, they do not share any sequence similarities (24).

The papain (from *Carica papaya)* superfamily of cysteine proteases (clan CA, family C1) can be divided into calpains, bleomycin hydrolases and the papain group (21). All of these enzymes can be aligned against each other in the region of the active site, which is composed of a catalytic dyads containing cysteine and histidine residues for their activity (24). Within the papain group are the mammalian lysosomal enzymes known as cathepsins. It must be noted that not all cathepsins are cysteine proteases, since cathepsin D and E are aspartyl proteases, cathepsin III is a metalloprotease, and cathepsins A and G are serine proteases (227). Although initially thought to be involved in terminal protein degradation, the availability of new technologies is helping to elucidate new physiological roles for cathepsins. Most cathepsins such as cathepsins L, B, H (I), C (dipeptidyl peptidase I or cathepsin J), O, X and F are expressed ubiquitously. Others are expressed in a tissue-specific manner, and are found to have specific roles in these tissues (339). For example, Cathepsin K (O2) is highly expressed in osteoclasts and was shown to play a role in bone resorption and remodeling (358). Cathepsin K appears to be upregulated at sites of inflammation (53). A study has also reported expression of cathepsin K in human breast carcinomas (206). Inactivation of this enzyme is associated with the bone disorder pycnodysostosis, underling its important role in extracellular matrix remodeling (116). Cathepsin V (L2, U) is mainly expressed in thymus and testis, whereas cathepsin W is found in CD8+ T lymphocytes (205, 288). Finally, cathepsin S is expressed in the spleen

and in the peripheral antigen presenting cells (APCs, including lymphocytes, macrophages, dendritic cells) (300, 301). Cathepsin S is involved the degradation of the invariant chain li of MHC II and in peptide loading for antigen presentation in the cells where it is expressed (241, 278, 354). Recent studies described a role for cathepsin S in angiogenesis as well as in atherogenesis (320). Cathepsins Q, M, and P were found to be specifically expressed in rat and mouse placenta (312-314).

Most cathepsins are endopeptidases except Cathepsin B, which also has a dipeptidyl carboxypeptidase activity, Cathepsin H, which is an aminopeptidase and Cathepsin C, which only has dipeptidyl aminopeptidase activity. On the other hand, cathepsin X exhibits carboxypeptidase or dipeptidase activity depending on the substrate (339). The overall structure, mechanisms of action and the regulation of activity of the lysosomal cysteine proteases will be discussed for cathepsin L in the following section.

3.2 Cathepsin L

Cathepsin L is a ubiquitously expressed lysosomal cysteine protease that has now also been shown to have other important physiological roles. It is notable that various specific cathepsin L inhibitors are currently being designed for anti-cancer therapies (41, 62, 167).

3.2.1 Gene and Protein Structure

Gene Structure

Cathepsin L was identified about 25 years ago, but the cDNA was only cloned ten years later (112, 153, 164, 171). Human Cathepsin L is localized on chromosome 9q21-22, while mouse cathepsin L is on chromosome 13 (56). In parallel, cathepsin L was also

cloned as the major excreted protein (MEP) in transformed cells (331). Human and mouse cathepsin L genes are composed of 8 exons and 7 introns and span approximately 5100bp and 7400bp, respectively. The intron-exon junctions are well conserved between human and mouse. Exon 1 is larger in human than in mouse but exons 2 to 7 are the same size. However, variations are observed in the structure of their introns (56). In both species, the AUG translation initiation site is located within exon 2. One major transcription initiation site is situated at –290 on the human cDNA sequence (or at –1489 on the gene) from the starting AUG (9, 15, 56, 157, 164). This mRNA corresponds to hCATL A. Another mRNA, hCATL B, was also described and encodes the same cathepsin L protein (56, 164). hCATL B differs from hCATL A in its 5' untranslated region. Three spliced variants of hCATL A were then identified; hCATL AI, hCATL AII, hCATL AIII (9, 15, 157, 277). These spliced forms lack 27 nt, 90 nt and 144 nt from the 3' end of exon 1, respectively, compared to hCATL A, leading to mRNA species that differ in their 5' untranslated ends, but that are translated into identical proteins. The proximal promoter of human cathepsin L is a TATA-less promoter containing one CCAAT motif and two GC boxes. NF-Y binds to the CCAAT sequence whereas Sp1 and Sp3 bind the GC regions. Mutation of these three binding sites decreases the expression of a cathepsin L gene promoter-luciferase construct (15, 157).

Protein Synthesis, Modification and Routing

At the protein level, like all lysosomal enzymes, cathepsin L is translated as an inactive pre-pro-enzyme (see figure 2A). The preregion, located at the amino terminus of the protein, is a 17 amino-acid signal peptide (or signal sequence). This hydrophobic stretch, which is also present in secreted and membrane proteins, directs the ribosome to the endoplasmic reticulum (ER) and recognizes the signal recognition particle (SRP). After

crossing the ER membrane, the signal peptide is cleaved after the Ala-Leu-Ala sequence by a signal peptidase, and the protein continues to be translated in the lumen of the ER. While being translated, human cathepsin L is N-glycosylated at Asn 204 (Asn 221for mouse cathepsin L). This glycosylation event is not required for proper folding of the protein, nor for its enzymatic activity or stability, but is important for lysosomal targeting (166, 309). The glycosylation/phosphorylation of cathepsin L first starts by the transfer of a complex oligosaccharide on the Asn of the Asn-Asp-Thr sequence (consensus sequence being Asn-X-Ser/Thr) in the ER. In the golgi vesicles, the oligosaccharide is then trimmed and further obtains high mannose carbohydrates. At the level of the *cis* golgi saccules, some of the mannose residues are then phosphorylated forming the mannose-6-phosphate (M6P) lysosomal recognition marker. Once the protein reaches the *trans* golgi network, this M6P sorting signal is recognized by the M6P receptor (MPR) allowing the enzyme to be routed to the lysosomes. The receptor/ligand complex is first sorted to a prelysosome (late acidic endosome) compartment, where the receptor dissociates from the enzyme and is recycled back to the *trans* golgi. These late endosomes then mature into lysosomes.

In certain cases, especially in transformed cells, secretion of procathepsin L (MEP) is observed (124). Various mechanisms have been proposed to explain this secretion. Initially, this secretion was thought to result from missorting of the enzyme within the golgi apparatus, resulting in secretion via the default constitutive pathway (282). Other mechanisms were also proposed, including low affinity of the protease for the MPR (96), saturation of the MPR or altered glycosylation of the protein (1, 190). It was then suggested that the secretion of Cathepsin L could be independent of a MPR pathway. In this situation, procathepsin L would self-associate into aggregates and be secreted via the formation of some dense vesicles (380) (see figure 2B)

Role of the Proregion

The prodomain is a chaperone. All cathepsins have differences in the length, sequence and structure of their propiece, although some residues are very well conserved. It first plays an important role in the folding of the enzyme, and then in inhibiting its catalytic activity (77, 324). Tao and colleagues have shown, using chimeric and deletion constructs, that the proregion of cathepsin L was required for ER trafficking, stability and mannose phosphorylation (324). It must be noted that in their studies, all of the prodomain was removed or exchanged with the one from aleurain (another cysteine protease). An earlier study demonstrated that the proregion was needed for the folding of a denatured bacterial cathepsin L (308). The crystal structure of procathepsin L was resolved (75). Structurally, like most enzymes from the papain family, the mature cathepsin L consist of two globular regions, the R-domain (right) and L- domain (left) (340). These two domains are organized to form an open V-shaped active site cleft. The propiece of cathepsin L can also be separated into two regions. The amino-terminal part, which consists of the first 60 amino acids, is important for proper folding and glycosylation of procathepsin L (53). Some of those residues interact with ER membrane proteins called lysosomal proenzyme receptor (LPR) while others contact the mature enzyme on its "proregion binding loop" or PBL (77, 228, 229). The carboxy-terminal part of the prodomain is responsible for the inhibitory role of the propiece, by preventing the entry of the substrate into the active cleft. In order to do so, while the amino-terminus binds to the PBL, the carboxy-terminus of the proregion bends over the groove of the active site in the opposite direction that the substrate would have been (75).

3.2.2 Regulation of the Protease Activity

Activation of the Zymogen

In order for the enzyme to become active, the proregion must be removed. This regulated proteolytic step is therefore crucial for controlling the activity of the enzyme in a location-dependent manner. The removal of the propiece occurs via an intra and/or inter molecular mechanism as the zymogen reaches the acidic environment of late endosomes or lysosomes (152, 247, 248, 286). Many studies suggest that the drop of pH in the acidic compartments weakens the interaction between the propeptide and the catalytic part (113, 159, 230, 286). The proregion binds less tightly to the active site producing a change in the propeptide conformation (159). Even if details of the initiation of the propeptide cleavage cascade are still unknown, some groups have shown that the proenzyme has a low, but detectable level of catalytic activity (221, 230). One possible explanation for this would be that this residual activity may be enough to initiate cleavage of some proenzymes (223). The small amount of catalytically active enzyme produced could then, via an intermolecular mechanism, initiate the activation cascade. It is believed that as for other cathepsins, the proregion is degraded after its removal (225, 283). Glycosaminoglycans found in the lysosomes (or extracellular matrix) can also stimulate autoactivation of procathepsin L by loosening up the interaction between the prodomain and the enzyme (151, 223). Studies have shown that other lysosomal enzymes, such as cathepsin D can also process procathepsin L and be involved in the initial steps of activation (247). In some cell types, the mature single chain from of cathepsin L is further processed into a two chains form –heavy chain and light chain- by cleavage of the carboxy-terminus (104, 113, 222).

Activity of Cathepsin L

The optimal activity of mature cathepsin L requires a slightly acidic pH and a reducing environment (so the active cysteine can be oxidized). The maximal activity of cathepsin L, using small synthetic peptides as substrates is at pH 5.5. The enzyme is most stable between pH 4.5 and 5.5. Like all cysteine proteases, the active site of cathepsin L is composed of a reactive cysteine (Cys 25), and a histidine (His 159). In the active form, both residues are charged, forming a thiolate-imidazolium ion pair (227). The general mechanism by which the enzymatic/proteolytic cleavage occurs is that the thiolate anion attacks the carbonyl-carbon of the amide bond of the substrate. The covalent bond is then disrupted following an acid-base hydrolytic process (53, 266). However, in order for the cleavage to occur, the substrate has to first enter into the active cleft. The substrate enters in an extended conformation so that the targeted bonds are well oriented towards the active nucleophile (341). Specific residues within the substrate (P3, P2, P1 and P1') will bind to other key residues in the active site of the enzyme (S sites), allowing the important embedding of the substrate into the cleft. Cathepsin L prefers a hydrophobic residue (mainly L/I) in the P2 position (cleavage occurs between residues P1 and P1') (53).

Inactivation of Cathepsin L

Because mature cathepsin L is a potent protease, its activity has to be tightly regulated. As mentioned above, cathepsin L is synthesized as an inactive precursor, which controls the timing of activation. Other factors can eventually play a role, such as the pH of the environment and the existence of cellular protease inhibitors. At neutral or alkaline pH, mature cathepsin L was found to be the most unstable lysosomal cysteine protease (20, 337, 338). In support of this data it was observed that above pH 7.0, the activity drops quickly

(222, 338). This inactivation of the enzyme is due to the irreversible disruption in the secondary structures, mainly the α-helices (100, 338). Similarly, under very acidic conditions (pH 3.8), such as in the mature lysosomes, cathepsin L is also inactivated. It is first irreversibly denatured and then degraded by cathepsin D (337). At this low pH, substrates and inhibitors can no longer bind the enzyme (222, 337). Turk and colleagues therefore proposed that cathepsin L was active only during certain stages of the life of a lysosome (337). Interestingly, MEP (the secreted procathepsin L), which is also active at pH 5.5 is more stable in neutral conditions (221). This suggests that the prodomain protects the enzyme from inactivation, which is in accordance with its role in protein stability.

Cysteine proteases have various endogenous inhibitors, such as the cystatins, the thyropins and the general protease inhibitor a2-macroglobulin. The cystatins include the stefins, the cystatins and the kininogenes and are less specific in inhibiting their target enzymes than the thyropins (thyroglobulin type-1 domain inhibitor) (195, 196). The stefins (type I, or cystatin A and B) are intracellular inhibitors, whereas cystatin C (type II) is an extracellular protease inhibitor, and the kininogenes (type III) are circulating inhibitors (340, 342). Cathepsin L was shown to be efficiently inhibited by Stefin A and B, the serpin squamous cell carcinoma antigen, the Li p41 fragment and the cytotoxic T-lymphocyte antigen 2B (25, 86, 321). Recently, a novel member of the Serpin family, Hurpin, was found to inhibit cytoplasmic Cathepsin L (365). In addition, another serpin, MENT (Myeloid and Erythroid Nuclear Termination stage-specific protein) inhibited nuclear protease activity (148). Many synthetic inhibitors are currently being generated to target individual Cathepsins with high specificity (41, 62, 167).

3.2.3 Different Roles for Cathepsin L: Examples from Cathepsin $L^{-/-}$ Mice

Cathepsin L is a lysosomal enzyme thought to be only involved in terminal protein degradation. However, knockout mice showed that terminal protein degradation was not the work of a single cathepsin as none of these mice had defects in protein degradation. The various phenotypes of these mice rather, pointed towards evidence that this protease has other specific biological roles.

One knockout was generated, and two mice with natural mutations within the cathepsin L gene were also identified, *Furless and Nackt*. The *furless* mice possess a G-to-A mutation, which substitute an arginine for a glycine at position 149 of the cathepsin L protein sequence, resulting in an inactive enzyme (281). The *nackt* mice display a deletion in the *cathepsin L* gene, preventing the generation of any functional protein (22). These three animal phenotypes have revealed that cathepsin L plays a role in epidermal homeostasis, in the regulation of the hair cycle, in the MHC class-II mediated antigen presentation and in positive selection of $CD4^+$ T cells in cortical epithelial cells of the thymus (22, 240, 281). Recently, cathepsin L expression in thymocytes was also shown to be essential for natural killer cell development (145). In general, pups lacking cathepsin L also have a slightly higher mortality upon weaning than their littermates (276).

3.2.4. Cathepsin L and Cancer

As previously mentioned, cathepsin L was initially identified as the major excreted protein (MEP) secreted from transformed fibroblastic cells (124, 331, 333). Oncogenic signals such as Ras, Raf, v-Src, fos, SV40 Large T as well as tumor promoters like phorbol ester can induce MEP expression and secretion (137, 163, 193, 322, 332). Moreover,

cathepsin L secretion correlated with the metastatic potential of transformed cell lines (50, 87). Increased cathepsin L activity and secretion was also observed in many human cancers (55, 94, 136, 169, 197, 246, 256, 263, 298, 303, 305, 360). Various reports also suggested that cathepsin L levels could be used as a potential indicator of tumor aggressiveness (metastasis) (256, 329). In breast cancer, numerous studies indicate that secreted cathepsin L could be a strong and independent prognostic factor, with a prognostic impact similar to that of lymph node status and grading (49, 99, 107, 132, 133, 198, 328, 329). One study has found no correlation between secreted cathepsin L activity and tumor aggressiveness (186). Interestingly, one group has found increased intracellular cathepsin L activity in cancer cells (375). This finding may develop into an important feature of the involvement of cathepsin L in cancer.

Therefore, although the involvement of cathepsin L in cancer is well established, its specific roles are not yet fully elucidated.

4- Cellular Transformation and Cancer

Cancer is a multistep disease that occurs as a result of the accumulation of many changes. It was suggested that most cancers acquire essentially the same six alterations: self-sufficiency in growth signals, insensitivity to growth inhibitory signals, evasion of apoptosis (programmed cell death), limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (128).

4.1 Events Triggering Cellular Transformation

Various genetic and epigenetic events can promote cellular transformation by activating oncogenes and/or inactivating tumor suppressor genes.

4.1.1 Genetic Events

Normal cells proliferate, differentiate and eventually die in a very controlled manner. Cancer is a disease that results when the normal behavior of a cell is deregulated. As first mentioned by Rudolph Virchow in 1858, this abnormal phenotype is heritable from cell to cell (27). During the late 18th century epidemiologists had already made the observation that external carcinogens could cause cancer. However, it was not until the 20th century that researchers came to understand that these external causes might have an "internal" link: DNA. This conclusion arose from cytogenetic studies that revealed the importance of genes and chromosomes in cancer and from the discovery of tumor viruses. Peyton Rous continued on the first discovery of Vilmen Ellerman and Oluf Bang, who had found that an infectious extract from blood cells could transmit leukemia from one chicken to another. Rous isolated a virus that could cause tumors (sarcomas) in chicken, the Rous sarcoma virus. It is only decades later that the transforming agent was identified as a normal gene captured from the retrovirus from a host cell. This transduced gene was called v-*src* and was different than its normal cellular counterpart c-*src*. The former was shown to be constitutively active. This gene was called an oncogene, and its normal counterpart a proto-oncogene. From then, it became clear that genetic changes were responsible for neoplastic transformation (27). Oncogenes were identified as genes involved in activating cellular proliferation, growth and survival whereas tumor suppressor genes were those inhibiting normal proliferation. Oncogenes are activated by dominant mutations, or "gain of function" of a proto-oncogene, while tumor suppressor genes are inhibited by recessive mutations, or "loss of function".

Many genetic events are involved in tumorigenesis. These include, (but are by no means all-encompassing), viral insertions/infections, chromosomal rearrangements, gene amplification, or point mutations.

RNA tumor viruses can also induce cellular transformation, independently of the presence of a transduced oncogene in their genome. Following infection, the provirus can integrate either within a proto-oncogene, or within its regulatory sequences, thereby either creating chimeric proteins, or deregulating the gene's expression (185, 243). The discovery of viral integration was useful for the identification of many oncogenes. DNA tumor viruses cause transformation of non-permissive cells. Stable integration of viral DNA and expression of early viral genes that interact with key cellular proteins can induce cellular neoplasia. For example, the oncogenic viral proteins E1A (adenovirus), Large T antigen (SV40), E6 and E7 (papillomaviruses) can target and inactivate the tumor suppressor genes RB and p53 (101, 189, 289, 366, 369). Interestingly, very few cancers are induced by viral infections in humans. In some cases, the hepatitis B virus (liver carcinoma), herpesvirus (African Burkitt's lymphoma), the papillomaviruses (cervical carcinoma) and the retrovirus HTLV (T-cell leukemia) are all associated with human cancers (84, 114, 264, 388).

Chromosomal translocations occur when two different chromosomal regions are joined together, which can result in the formation of chimeric genes. In many cases, the generated fusion proteins are oncogenic. Two well-studied examples are the translocation of the *myc* oncogene in Burkitt's lymphomas and the Philadephia chromosome, found in 90% of patients with chronic myelogenous leukemia (CML). In Burkitt's lymphomas, *myc* expression is deregulated, whereas in CML, the chimeric gene encodes a constitutively active oncogene, BCR-ABL (2, 52, 78, 82, 220, 302).

A number of oncogenes are activated following DNA amplification, which results in an increasing number of gene copies per cell (42). This phenomenon was first observed as a mechanism by which cells acquired resistance to chemotherapeutic agents, but has now been detected in many cancers and transformed cells (4, 318). *c-myc* was also the first oncogene found to be amplified in neoplasias (3, 65, 219). The mechanisms underyling gene amplification are not fully understood. Extra copies of amplified DNA are visible in cytogenetic studies as double minute chromosomes (DM), which, when incorporated into chromosomal regions, are called homogeneously stained regions (HSR). Following double-strand breaks within chromosomal fragile sites (CFSs), breakage-fusion-bridge (BFB) cycles can result in an increase in the number of a gene copy after multiple rounds of cell division (68, 69, 138).

Lastly, point mutations induced by chemical carcinogens can activate oncogenes and inactivate tumor suppressor genes. Known examples are the *ras* family members, where mutations have been found in 50% of human cancers (see section 4.3 for more details on Ras) and the p53 tumor suppressor gene, which is also mutated in 50% of human cancers (36, 296).

4.1.2 Epigenetic Events

Hypermethylation of certain promoters is a novel mechanism associated with tumorigenesis (162). DNA methylation on CpG dinucleotides is involved in specific cellular memory function during development (26). However, some stretches of GC rich regions called 'CpG islands' are usually unmethylated in normal cells (7). These CpG islands are short sequences located close to proximal promoter regions. In cancer, they often are hypermethylated, which results in silencing of key genes such as repair genes or

tumor suppressor genes (106, 238, 239). Functional consequences include an increase in point mutations and genomic instability, which could be early events in tumor progression. The causes of *de novo* methylation in cancer are still not fully understood, and may be linked with cellular proliferation (352). Histone modifications (deacetylation, methylation, phosphorylation) are a prerequisite for DNA methylation, giving another potential level of dysregulation in cancer cells (16, 158).

4.2 Oncogenes and Tumor Suppressor Genes

Oncogenes and tumor suppressor genes are responsible for cellular transformation. They represent any key protein involved in promoting or inhibiting cellular proliferation, respectively.

4.2.1 Oncogenes

As previously mentioned, "gain of function" mutations activate proto-oncogenes, and from multiple genetic or epigenetic events result the acquired neoplastic phenotypes. One oncogene is able to transform immortalized rodent cells line, but not primary cells, where the activation of two cooperating oncogenes is necessary (188, 284). Proto-oncogenes are important components of the signaling cascades involved in cellular growth, proliferation, differentiation or death. They include the extracellular growth factors and cytokines, their transmembrane receptor tyrosine kinases, cytoplasmic signal transduction proteins and transcription factors.

Growth factors (mitogens) are polypeptides that control cellular proliferation by binding to their respective receptors and initiating a cascade of signal transduction (156, 187). Culture of fibroblasts requires the presence of exogenous growth factors present in

serum. In the absence of these mitogens, cell proliferation ceases. Following a period of growth factor depletion (or serum starvation) the addition of growth factors is necessary for the exit from quiescence (G0 phase) and entry into the G1 phase of the cell cycle. Many transformed cells are able to escape the necessity of growth factors (143). This can be achieved in part by their endogenous production of growth factors, thereby creating an autocrine stimulation of cell proliferation (316).

Another mechanism by which transformed cells can escape the requirement for growth factors or cytokines is via the constitutive activity of their transmembrane receptors. Upon ligand binding, the receptors polymerize and become activated by autophosphorylation. Phosphorylation of tyrosines in the intracellular domains creates docking sites for multiple signaling molecules. Constitutive activation of receptor tyrosine kinases occurs by different mechanisms. One of these is chromosomal rearrangement, which can generate constitutively dimerized (or oligomerized) chimeric receptors. This is observed with the TprMet oncogene, where the leucine zipper Tpr is fused to the intracellular domain of the Met receptor (279). The EGF receptor (EGFR or erbB) and family members are also notable examples. Amplification of EGFR or of another family member c-erbB2/neu/HER2 is detected in many cancers (19, 91, 204, 306, 343, 376). Alternatively, a truncated version of the EGFR or an erbB2 harboring point mutations in its transmembrane domain generates constitutive receptor activity (168, 224, 364).

Following receptor activation, various signaling pathways are activated, such as the PI3K pathway, PLC pathway, SRC pathway, and/or RAS pathway. Similarly to RTKs, mutations or upregulation of any key members of these pathways have oncogenic potential.

Ultimately, the cellular response is controlled by the pattern of gene expression. Various transcription factors with oncogenic properties have been identified. As

mentioned, genetic events such as gene amplification and chromosomal translocations (as for c-myc) can generate constitutively active proteins (4, 73).

4.2.2 The Ras Oncogene

Ras (*Rat sarcoma*) was initially identified, like many oncogenes, by viral transduction (135, 172). Ras membrane-associated proteins control several signaling pathways that regulate normal cell behavior. However, 20-30% of human cancers express a Ras protein activated by point mutation (36). The three cellular *ras* genes code for four homologous proteins of 21 kDa: N-Ras, H-Ras, K-Ras4A and K-Ras4B. Ras proteins have intrinsic GTPase activity and are turned "on" and "off" depending if they are GTP- or GDP- bound, respectively (37). Ras is initially synthesized as a cytoplasmic protein, which becomes membrane-associated via the addition of a farnesyl group (215). Proper localization to specific plasma membrane microdomains (lipid rafts) is then necessary for effective signaling (271). Recent papers are indicating that Ras is also active from intracellular endoplasmic and golgi membranes, where it would have specific biological roles (28, 61, 90). Membrane-GTP-bound Ras interacts with, and activates, various effector molecules. One of them is the serine/threonine kinase Raf that initiates the mitogen-activated protein (MAP) kinase pathway, which stimulates cell cycle progression by regulating specific transcription factors (370). Ras also activates phosphatidylinositol 3-kinase (PI3K) that regulates cell survival via the AKT/PKB pathway, protein translation via PDK1 (and AKT) and actin cytoskeleton remodeling via Rac (81, 183, 280). The Ras-related guanine nucleotide exchange factors (RalGEFs) are the third most studied Ras effectors and are involved in cell cycle progression as well as in vesicular transport (119) (35, 242). Activated Ras can therefore induce cellular transformation by multiple

mechanisms and targeting Ras signaling pathways in cancer therapy has proven quite complex (97). Interestingly, the types of tumors linked to oncogenic Ras are different between human and mouse, which lead to the observation that different downstream effectors of Ras were necessary to transform mouse and human cells (127). These results consequently challenged the validity of transposing results obtained in mouse models to transformation of human cells (127, 142).

4.2.3 Tumor Suppressor Genes

The concept of the tumor suppressor gene arose in the late 1960s, when Henry Harris and his colleagues found that a hybrid cell issued from the fusion of a cancer cell with a normal cell grew normally (134). The tumorigenesity was lost, implying that the normal cell had a gene acting as a "negative regulator", or a "break" of transformation. Knudson proposed the "two-hit" or "two-mutations" hypothesis following the study of inherited and sporadic retinoblastomas. In sporadic cancers, the two mutations occur on somatic cells, whereas in inherited cancer, one mutation was inherited in germline cells, and the other acquired somatically (173). A common genetic abnormality in many cancers is the loss of heterozygosity (LOH). LOH occurs when one allele of a tumor suppressor gene is missing or doesn't produce a functional protein, and the remaining normal allele is lost. LOH analysis has uncovered novel tumor suppressor genes in many cancers. *Rb1* was the first tumor suppressor gene discovered, and was later found to be a key player in cell cycle progression (45, 85, 109, 192). Rb mainly represses transcription of cell cycle genes by sequestering E2F, thereby preventing E2F-mediated activation, and by recruiting co-repressors (92). Mitogen stimulation is necessary for the cells to enter the G1 phase of the cell cycle until the first restriction point (255). It is believed that many of these signals

converge towards the regulation of Rb (363). Cell cycle regulation of Rb activity occurs via phosphorylation. In G0, hypophosphorylated Rb binds to E2F and inhibits its activity. As cells progress through G1, cyclin D/cdk4/6 followed by cyclin E/cdk2 phosphorylate Rb, which releases E2F, allowing it to activate S phase genes (92). Inhibitors of cell cycle progression, such as the cdk inhibitors, are also potential tumor suppressor genes.

4.3 Breast Cancer

Breast cancer is the most common malignancy affecting women. One in eight Caucasian women will develop breast cancer in their lifetime, and the recorded incidences are increasing. Breast cancer is a heterogeneous disease, and the appropriate therapy for each patient is usually selected depending on the stage of the disease and on the estimated risk/benefit between disease recurrence and toxicity of the therapy.

Breast cancers are diagnosed based on TNM (tumor-node-metastases) stages, axillary node status, tumor size, the presence of micrometastases ($\leq$ 2 mm) in sentinel nodes, hormone receptor and HER-2 status. Pathologic analysis of breast biopsies can reveal differences between the benign and malignant lesions. DCIS (Ductal Carcinoma *In Situ*) is a non-invasive atypical proliferation of ductal epithelial cells that eventually fills and plugs the duct. DCIS lesions can recur as invasive ductal carcinomas. In invasive DCIS, cells infiltrate the surrounding stroma and morphologically change into larger cells with large irregular hyperchromatic nuclei. LCIS (Lobular Carcinoma *In Situ*) is an acinus lesion, where slightly larger cells are proliferating. Following diagnosis of LCIS, 30% of women will develop invasive carcinoma in either breast. LCIS is a strong risk factor for breast cancer. Invasive lobular carcinomas represent only 5-10% of breast carcinomas but are quite particular. They tend to be bilateral and multicentric within the same breast.

Histologically, they consist of infiltrating small cells dispersed through the stroma one behind the other. Often, infiltrating tumors have mixed ductal and lobular patterns (74). Eventually, cells invade the blood and lymphatic systems to form distant metastases.

4.3.1 Some Genetic Alterations in Breast cancer

5-10% of breast cancers are hereditary (or familial). Two breast cancer susceptibility genes have been identified. BRCA1 and BRCA2 (breast cancer and ovarian cancer 1) are tumor suppressor genes involved in DNA repair, recombination and transcription (353). However, in sporadic cancers, mutations and LOH of other tumor suppressor genes such as *Rb*, *p53, p16* or *E-cadherin* are more common (40, 95, 110, 315). Alternatively, oncogenic amplification of *c-erbB2*, *c-myc* and *CCND1* (cyclin D1) was observed in many breast tumors (76, 105, 120, 306). AIB1, a steroid-receptor co-activator is another amplified oncogene (8).

Hormones like estrogen and progesterone play an important role in breast cancer. Normal breast epithelium and one third of breast cancers are hormone-responsive (64). ER- and PR-negative breast tumors contain more malignant and dedifferentiated cells. They are less responsive to treatment and have a worse prognostic (93).

Overall, malignant progression of breast cancer involves various steps, such as increased cellular proliferation, restrained apoptosis and genomic instability. Metastasis occurs when malignant cells leave the primary cancer to form other tumors at distant specific sites. In order to do so, metastatic cells must undergo epithelial-mesenchymal transition (EMT). These dedifferentiated cells have lost cell-cell attachment and cell-substratum adhesion, and have acquired cell locomotion, proteolysis and the ability to survive and proliferate at distant sites (51).

cDNA Microarray technology is a novel tool that determines the gene expression profile of cancers. These arrays will be useful in finding genes associated with certain cellular phenotypes and predicting clinical outcome of breast cancers as well as identifying novel markers and therapeutic targets (208, 346).

4.3.2 Importance of Novel Prognostic and Predictive Markers

Fifty to sixty percent of the 225,000 North-American women diagnosed with breast cancer in 2002 (310, 311) had node-negative tumors meaning that there are no detectable metastases in the adjacent lymph node (stage 1). However, 30% of women with this type of tumor are predicted to die from recurrence of the disease (149). It is therefore imperative to develop new prognostic and predictive markers to better define the group of these early-stage node-negative women who would benefit from adjuvant systemic therapy.

A prognostic marker will predict patient outcome irrespective of the treatment given (ex. TNM stages), whereas a predictive marker will indicate the response to a specific therapy (64). The most studied predictive factor is the status of the estrogen receptor (ER); it is well documented that ER^+ tumors will respond better to hormonal therapy with tamoxifen compared to ER^- tumors (251). However, ER status may be considered a weak prognostic factor (98). Progesterone receptor (PgR) status is also routinely analyzed as a predictive marker. Usually, the other marker analyzed is the c-erbB-2/HER-2 oncogene, which is amplified in approximately 30% of breast cancers (306). HER-2 is considered an independent weak prognostic marker for node-positive tumors, while results vary for node-negative tumors (149, 237). HER-2 status is known to be predictive of the response to specific therapies. However, the assays and types of assays used to determine HER-2

levels still have to be standardized (98, 149). Recently a study revealed the prognostic value of c-myc over HER-2 in node-negative breast carcinomas (291, 292).

Attempts to find relevant biomarkers are underway, but most have not been included in clinical analyses (60, 149, 294). Therefore in order to decrease the chance of recurrence in node-negative patients, it is imperative to develop standardized protocols and interpretate criteria that will help in improving diagnosis and the management of those women at risk and those not.

5- CDP/Cux Transcription Factor

CDP/Cux/Cut is an evolutionary conserved family of transcription factors that contains a Cut homeodomain with one or more Cut Repeats (CR) that function as DNA binding domains. Cut was initially identified in *Drosophila melanogaster* and is preserved amongst metazoans. The CCAAT-displacement activity was originally found in the sea urchin, as a factor that displaced a positive regulator of sperm histone H2B transcription (18). Cut proteins have since been isolated from several mammalian species, including human (CCAAT-displacement protein or CDP), dog (Cut-like homeobox or Clox), mouse (Cut homeobox Cux-1) and rat (CDP-2) (6, 245, 345, 381). Recently, a second Cux gene, Cux-2 was found in mouse and chicken (272, 327).

5.1 Homeodomain Transcription Factors

Transcription factors are proteins that interact with DNA via specific binding domains, and affect the process of transcription. Some DNA binding motifs have been identified, and examples include helix-turn-helix, zinc-fingers or basic leucine zippers.

Homeodomain containing proteins play an important role in developmental programs (226). The homeodomain is a 60 amino-acid region that recognizes DNA via a helix-turn-helix motif (252). It makes contact with a TAAT (ATTA) core in both the minor and major groove, via specific residues located at the amino-terminus of the homeodomain and via the 3rd helix (or "recognition helix"). The 9th residue of helix III is responsible for some sequence specificity between different homeodomain containing proteins by interacting with nucleotides upstream of the TAAT sequence (330). Cooperative dimerization between homeodomain containing factors was found to further increase their DNA binding specificity (372). Alternatively, sequence specificity can be improved by the presence of a second DNA binding domain in some homeobox proteins such as Pou, Lim, Paired and Cut (115).

5.2 *Drosophila melanogaster* Cut

The name Cut came from a mutation responsible for a 'cut wing' phenotype in *Drosophila melanogaster*. Several viable (*kinked femur* and *cut wing*) as well as lethal (lethal I, II and III) mutations have been identified in the *cut* locus. Cut expression in several tissues is controlled by a series of distal enhancers located at various distances from the gene (154, 155). The viable as well as lethal I and III mutations were caused by the insertion of a gypsy transposable element (insulator) between the different enhancers and the core promoter, thus preventing Cut expression in various tissues. Short deletions or point mutations within the coding sequence of *cut* were responsible for lethal II mutations, resulting in embryonic lethality. In tissues where it is usually expressed, *cut* was detected in precursor cells as well as in certain progeny cells, supporting a role in determining cell-type specificity (29, 30, 211). Sustained expression of *cut* in the adult fly would suggest

that not only is *cut* important for the acquisition of the developed phenotype, but also for its maintenance (30, 213). As an example, in the peripheral nervous system, lethal *cut* mutations caused cells to initiate the wrong developmental program such that external sensory organs became internal (chordotonal) sensory organs (33). In contrast, ectopic expression of *cut* in embryos triggered precursor cells to differentiate into external sensory organs (31, 213). The study of the role of *cut* in limb (wing, femur) development also uncovered a genetic interaction between *cut*, Notch and Wingless. At the dorsal-ventral border of the wing, *Cut* is a downstream effector of Notch activation, where it would maintain Wingless expression and suppress that of the Notch ligands, Delta and Serrate (244).

5.3 Mammalian CDP/Cux

The human CUTL1 gene spans at least 340 kb and contains 33 exons. Various splicing events and polyadenylation at two different sites can occur. The full length CDP/Cux protein contains sequences encoded by exon 2 to exon 24. Different isoforms resulting from differential splicing in exons 15 and 16 have been detected, but the functional significance of the existence remains unknown (383). An alternatively spliced transcript codes for a protein called CASP (CDP/Cux alternatively spliced protein), that contains the coiled-coil region. A unique carboxy-terminus (from exons 25 to 33), was also identified (202). CASP was recently found to be a golgi membrane protein, but its function remains unclear (121).

5.3.1 DNA Binding and Modulation of CDP/Cux Activity

DNA Binding of CDP/Cux

The fact that CDP/Cux contains four DNA binding domains; three Cut Repeats (CR) and one Cut Homeodomain (HD), is suggestive of a variety of DNA binding sequences, conformational modes and target promoters (5, 11, 131). The way CDP/Cux binds to DNA is still not fully understood. Recent studies showed that individual Cut Repeats cannot bind to DNA independently but will cooperate with another CR or the HD (234). Each different combination of DNA binding domain prefers different consensus sequences and has different binding kinetics. CR1CR2 binds transiently (rapid off and on rates) to DNA but with high affinity to two C(A/G)AT motifs arranged as direct or inverted repeats. This binding activity is responsible for the CCAAT-displacement activity of the full length CDP/Cux p200 protein originally described (234). In the context of p200, DNA binding by CR3HD is inactive. Stable DNA binding of CDP/Cux involves one or more CR with the HD, such as CR3HD or CR2CR3HD. Their preferred binding sequences are similar to ATCGAT. This stable DNA binding activity increases as cells progress from G1 to S phase of the cell cycle (see figure 3).

Modulation of CDP/Cux Activity

CDP/Cux activity is finely regulated at the post-translational level. Two such modifications were shown to inhibit its DNA binding and therefore transcriptional activity. The first is phosphorylation of the CRs and the HD, which can occur at various amino acids, by various kinases (70, 72, 287). The second was shown to be acetylation of the HD by P/CAF, although this observation remains controversial as it could not be repeated in Dr. Nepveu's laboratory (200). Instead, acetylation of CDP/Cux by P/CAF was shown to prevent its phosphorylation by cyclin A/cdk1 (Muzzin O. and Santaguida S., manuscript in

preparation). An inhibitory domain was also identified at the amino terminus of CDP/Cux (amino acids 1 to 101). This region was found to be sufficient to inhibit CDP/Cux DNA binding as well as its ability to stimulate transcription (Reynal L., Wang Y. and Truscott M., manuscript in preparation). It is currently unknown whether proteolytic processing is responsible for the removal of the inhibitory region or whether a post-translational modification, such as phosphorylation inhibits its action. As previously mentioned, for stable DNA binding to occur, the CCAAT displacement activity has to be turned off. This occurs via proteolytic cleavage of the full-length protein, resulting in an amino-terminally truncated isoform of 110 kDa that contains the DNA binding domains CR2CR3HD, and therefore possesses different DNA binding properties and transcriptional activities (235).

5.3.2 Transcriptional Properties of CDP/Cux

CDP/Cux was initially described as a transcriptional repressor that down-modulates genes in proliferating precursors cells that are then turned on in terminally differentiated cells (244). CDP/Cux can repress transcription via at least two mechanisms. The CCAAT displacement activity of CR1CR2 can compete with transcriptional activators for binding site occupancy. Two active repression domains were identified at the carboxy-terminus of CDP/Cux (216). These domains may directly recruit HDAC1 (histone deacetylase 1), a repressor of gene transcription (201). CDP/Cux was also reported to repress transcription of certain genes by binding to AT-rich sequences within matrix-attachment regions (MARs) (17, 54, 209, 210, 244, 319, 359).

Interestingly, CDP/Cux is involved in transcriptional activation of certain gene promoters in proliferating cells. CDP/Cux is part of the HiNF-D promoter complex that induces several *histone* gene promoters at the G1/S transition of the cell cycle (13, 14, 347-

349, 373). The DNA binding activity of HiNF-D is regulated in a cell cycle regulated manner in normal cells, where it increases in S phase. A recent report indicated that in late S-phase, CDP/Cux also represses the *histone H4* gene promoter using Rb as a co-repressor (126). In contrast, transformed cells display constitutive DNA binding by HiNF-D (144). CDP/Cux cooperates with the rat ITF2 transcription factor to activate the *tyrosine hydroxylase* gene promoter, but cannot activate it on its own (381). Recently, CDP/Cux was shown to activate the *DNA polymerase alpha*, *CAD*, and *cyclin A* gene promoters at the G1/S transition (334).

These results strongly suggest that, depending on the promoter context, and of the presence of specific binding partners, CDP/Cux could alternatively function as a transcriptional repressor or activator.

5.3.3 CDP/Cux Activity During Cell Cycle

As mentioned previously, many targets of CDP/Cux are genes that are repressed in proliferating cells and expressed as cells become terminally differentiated, when CDP/Cux activity ceases. For example, in the developing kidney, CDP/Cux is highly expressed in both mesenchymal and epithelial cells. At later stages of nephrogenesis, during cell cycle exit and terminal differentiation, CDP/Cux is down-regulated (351). In cell culture, CDP/Cux DNA binding oscillates during the cell cycle, peaking at the G1/S transition when it is dephosphorylated by the cdc25A phosphatase (71). This activity correlates temporally with histone H4 and DNA polymerase alpha expression (144, 334). At the same time, CDP/Cux down-modulates the CKI $p21^{waf1/cip1}$ gene promoter (71). CDP/Cux was also shown to decrease p27 expression, another CKI both *in vitro* and *in vivo* (191). Stable DNA binding at the G1/S transition results from proteolytic processing of the full-

length protein between CR1 and CR2 to generate the 110 kDa isoform. In the G2 phase of the cell cycle, CDP/Cux activity is then inhibited by phosphorylation of the Cut homeodomain by cyclin A/cdk1 (287). The current model for CDP/Cux regulation during the cell cycle is summarized in fig 3.

5.3.4 Knockout and Transgenic Mice

Four *cux-1* knockouts were generated. The first one, *cux-1*$^{\Delta CR1}$, in which the first CR was removed, displayed a very mild phenotype characterized by curly vibrissae, wavy hair and some pup loss due to the mother's impaired lactation (336). Two *cux-1* knockouts, *cux-1*$^{\Delta HD}$ and *CDP/Cux* $\Delta C^{-/-}$, in which the whole carboxy terminal region, including the homeodomain were deleted, were shown to have high perinatal lethality, male infertility, reduced growth and defect in hair follicle development (214, 304). In addition, a deficit in T and B cells as well as a surplus in myeloid cells was observed (304). At the molecular level, the DNA binding activity of HiNF-D was lost in mouse embryo fibroblasts derived from these mice or adults tissues (214). Another mouse in which gene targeting replaced the CR3 and HD revealed similar phenotypes, including delayed differentiation of the lung epithelia and altered hair follicle morphogenesis (102).

In contrast, transgenic mice over-expressing the cux-1 cDNA, under the control of the CMV enhancer/promoter, displayed multi-organ hyperplasia and organomegaly (191). An increased number of proliferative cells was observed in the adult transgenic kidney, which correlated with a decrease in the levels of the Cdk inhibitor p27. Eventually, these mice develop glomerulosclerosis and interstitial fibrosis, with an increase in type IV collagen in the mesangium (39).

These animals indicate an important role for CDP/Cux in tissue homeostasis in several organs.

5.3.5 CDP/Cux and Cancer

The CUTL1 gene was mapped to human chromosome 7, band q22, a chromosomal region frequently deleted or rearranged in several cancers, including acute myeloid leukemias, myelodysplastic syndrome and uterine leomyomas (194, 253, 290). LOH of CUTL1 was identified in 15% of human uterine leiomyomas and 18% of breast cancer analyzed (384, 385). In the latter, LOH of CUTL1 was associated with an increase in tumor size. Interestingly, 5 out the 12 tumors analyzed had lost the 5' end of the gene and retained the 3'end. Using an animal model, it was observed that transgenic mice overexpressing the polyomavirus Large T (LT) antigen under the control of the MMTV promoter developed mammary tumors and uterine leiomyomas (361). In support of the possibility that LT and CDP/Cux interact, LT and CDP/Cux could be co-immunoprecipitated. This interaction is a mechanism used by viral proteins to inactivate tumor suppressor genes, suggesting a potential involvement of CDP/Cux in some transformation events. From these studies, it was initially thought that CDP/Cux was a tumor suppressor gene located in 7q22. However, DNA sequencing of cDNA from 6 leiomyomas and 5 breast tumors did not reveal any point mutations or any gross rearrangements. Moreover, human uterine leiomyomas expressed more of the proteolytically derived p110 isoform than the normal adjacent myometrium (236). In addition, as mentioned earlier for transformed cells, HiNF-D DNA binding is no longer cell cycle regulated (144). In multiple myeloma, CUTL1 was found to be upregulated compared to normal plasma cells (83).

These results indicate that CDP/Cux might have a role in cancer. However, the nature of such role is currently not fully understood.

REFERENCES

1. **Achkar, C., Q. M. Gong, A. Frankfater, and A. S. Bajkowski.** 1990. Differences in targeting and secretion of cathepsins B and L by BALB/3T3 fibroblasts and Moloney murine sarcoma virus-transformed BALB/3T3 fibroblasts. J Biol Chem **265:**13650-4.

2. **Adams, J. M., S. Gerondakis, E. Webb, L. M. Corcoran, and S. Cory.** 1983. Cellular myc oncogene is altered by chromosome translocation to an immunoglobulin locus in murine plasmacytomas and is rearranged similarly in human Burkitt lymphomas. Proc Natl Acad Sci U S A **80:**1982-6.

3. **Alitalo, K., K. Saksela, R. Winqvist, R. Alitalo, J. Keski-Oja, M. Laiho, M. Ilvonen, S. Knuutila, and A. de la Chapelle.** 1985. Acute myelogenous leukaemia with c-myc amplification and double minute chromosomes. Lancet **2:**1035-9.

4. **Alitalo, K., and M. Schwab.** 1986. Oncogene amplification in tumor cells. Adv Cancer Res **47:**235-81.

5. **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 & Development **8:**245-57.

6. **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.

7. **Antequera, F., and A. Bird.** 1993. CpG islands. Exs **64:**169-85.

8. **Anzick, S. L., J. Kononen, R. L. Walker, D. O. Azorsa, M. M. Tanner, X. Y. Guan, G. Sauter, O. P. Kallioniemi, J. M. Trent, and P. S. Meltzer.** 1997. AIB1,

a steroid receptor coactivator amplified in breast and ovarian cancer. Science **277:**965-8.

9. **Arora, S., and S. S. Chauhan.** 2002. Identification and characterization of a novel human cathepsin L splice variant. Gene **293:**123-31.

10. **Asano, K., J. Clayton, A. Shalev, and A. G. Hinnebusch.** 2000. A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. Genes Dev **14:**2534-46.

11. **Aufiero, B., E. J. Neufeld, and S. H. Orkin.** 1994. Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein. Proc. Natl. Acad. Sci. USA **91:**7757-7761.

12. **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.

13. **Aziz, F., A. J. van Wijnen, J. L. Stein, and G. S. Stein.** 1998. HiNF-D (CDP-cut/CDC2/cyclin A/pRB-complex) influences the timing of IRF- 2-dependent cell cycle activation of human histone H4 gene transcription at the G1/S phase transition. J Cell Physiol **177:**453-64.

14. **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.

15. **Bakhshi, R., A. Goel, P. Seth, P. Chhikara, and S. S. Chauhan.** 2001. Cloning and characterization of human cathepsin L promoter. Gene **275:**93-101.

16. **Ballestar, E., and M. Esteller.** 2002. The impact of chromatin in human cancer: linking DNA methylation to gene silencing. Carcinogenesis **23:**1103-9.

17. **Banan, M., I. C. Rojas, W. H. Lee, H. L. King, J. V. Harriss, R. Kobayashi, C. F. Webb, and P. D. Gottlieb.** 1997. Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. J Biol Chem **272:**18440-52.

18. **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.

19. **Bargmann, C. I., M. C. Hung, and R. A. Weinberg.** 1986. Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. Cell **45:**649-57.

20. **Barrett, A. J., and H. Kirschke.** 1981. Cathepsin B, Cathepsin H, and cathepsin L. Methods Enzymol **80:**535-61.

21. **Barrett, A. J., and N. D. Rawlings.** 2001. Evolutionary lines of cysteine peptidases. Biol Chem **382:**727-33.

22. **Benavides, F., M. F. Starost, M. Flores, I. B. Gimenez-Conti, J. L. Guenet, and C. J. Conti.** 2002. Impaired hair follicle morphogenesis and cycling with abnormal epidermal differentiation in nackt mice, a cathepsin L-deficient mutation. Am J Pathol **161:**693-703.

23. **Beretta, L., A. C. Gingras, Y. V. Svitkin, M. N. Hall, and N. Sonenberg.** 1996. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. Embo J **15:**658-64.

24. **Berti, P. J., and A. C. Storer.** 1995. Alignment/phylogeny of the papain superfamily of cysteine proteases. J Mol Biol **246:**273-83.

25. **Bevec, T., V. Stoka, G. Pungercic, I. Dolenc, and V. Turk.** 1996. Major histocompatibility complex class II-associated p41 invariant chain fragment is a strong inhibitor of lysosomal cathepsin L. J Exp Med **183:**1331-8.

26. **Bird, A.** 2002. DNA methylation patterns and epigenetic memory. Genes Dev **16:**6-21.

27. **Bishop, J. M.** 1991. Molecular themes in oncogenesis. Cell **64:**235-48.

28. **Bivona, T. G., and M. R. Philips.** 2003. Ras pathway signaling on endomembranes. Curr Opin Cell Biol **15:**136-42.

29. **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 & Development **4:**1322-31.

30. **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.

31. **Blochlinger, K., L. Y. Jan, and Y. N. Jan.** 1991. Transformation of sensory organ identity by ectopic expression of Cut in Drosophila. Genes & Development **5:**1124-35.

32. **Bloom, J., and M. Pagano.** 2003. Deregulated degradation of the cdk inhibitor p27 and malignant transformation. Semin Cancer Biol **13:**41-7.

33. **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.

34. **Boehm, M., T. Yoshimoto, M. F. Crook, S. Nallamshetty, A. True, G. J. Nabel, and E. G. Nabel.** 2002. A growth factor-dependent nuclear kinase phosphorylates p27(Kip1) and regulates cell cycle progression. Embo J **21:**3390-401.

35. **Boettner, B., and L. Van Aelst.** 2002. The RASputin effect. Genes Dev **16:**2033-8.

36. **Bos, J. L.** 1989. ras oncogenes in human cancer: a review. Cancer Res **49:**4682-9.

37. **Bourne, H. R., D. A. Sanders, and F. McCormick.** 1990. The GTPase superfamily: a conserved switch for diverse cell functions. Nature **348:**125-32.

38. **Boyle, W. J., T. Smeal, L. H. Defize, P. Angel, J. R. Woodgett, M. Karin, and T. Hunter.** 1991. Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. Cell **64:**573-84.

39. **Brantley, J. G., M. Sharma, N. I. Alcalay, and G. B. Heuvel.** 2003. Cux-1 transgenic mice develop glomerulosclerosis and interstitial fibrosis. Kidney Int **63:**1240-8.

40. **Brenner, A. J., and C. M. Aldaz.** 1995. Chromosome 9p allelic loss and p16/CDKN2 in breast cancer and evidence of p16 inactivation in immortal breast epithelial cells. Cancer Res **55:**2892-5.

41. **Brinker, A., E. Weber, D. Stoll, J. Voigt, A. Muller, N. Sewald, G. Jung, K. H. Wiesmuller, and P. Bohley.** 2000. Highly potent inhibitors of human cathepsin L identified by screening combinatorial pentapeptide amide collections. Eur J Biochem **267:**5085-92.

42. **Brison, O.** 1983. [Gene amplification in eukaryotic cells]. Biochimie **65:**III-XI.

43. **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.

44. **Brown, M. S., J. Ye, R. B. Rawson, and J. L. Goldstein.** 2000. Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. Cell **100:**391-8.

45. **Buchkovich, K., L. A. Duffy, and E. Harlow.** 1989. The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. Cell **58:**1097-105.

46. **Calhoun, V. C., A. Stathopoulos, and M. Levine.** 2002. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. Proc Natl Acad Sci U S A **99:**9243-7.

47. **Calkhoven, C. F., and G. Ab.** 1996. Multiple steps in the regulation of transcription-factor level and activity. Biochem J **317 (Pt 2):**329-42.

48. **Calkhoven, C. F., C. Muller, and A. Leutz.** 2000. Translational control of C/EBPalpha and C/EBPbeta isoform expression. Genes Dev **14:**1920-32.

49. **Castiglioni, T., M. J. Merino, B. Elsner, T. T. Lah, B. F. Sloane, and M. R. Emmert-Buck.** 1994. Immunohistochemical analysis of cathepsins D, B, and L in human breast cancer. Hum Pathol **25:**857-62.

50. **Chambers, A. F., R. Colella, D. T. Denhardt, and S. M. Wilson.** 1992. Increased expression of cathepsins L and B and decreased activity of their inhibitors in metastatic, ras-transformed NIH 3T3 cells. Mol Carcinog **5:**238-45.

51. **Chambers, A. F., A. C. Groom, and I. C. MacDonald.** 2002. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer **2:**563-72.

52. **Chan, L. C., K. K. Karhi, S. I. Rayter, N. Heisterkamp, S. Eridani, R. Powles, S. D. Lawler, J. Groffen, J. G. Foulkes, M. F. Greaves, and et al.** 1987. A novel abl protein expressed in Philadelphia chromosome positive acute lymphoblastic leukaemia. Nature **325:**635-7.

53. **Chapman, H. A., R. J. Riese, and G. P. Shi.** 1997. Emerging roles for cysteine proteases in human biology. Annu Rev Physiol **59:**63-88.

54. **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.

55. **Chauhan, S. S., L. J. Goldstein, and M. M. Gottesman.** 1991. Expression of cathepsin L in human tumors. Cancer Res **51:**1478-81.

56. **Chauhan, S. S., N. C. Popescu, D. Ray, R. Fleischmann, M. M. Gottesman, and B. R. Troen.** 1993. Cloning, genomic organization, and chromosomal localization of human cathepsin L. J Biol Chem **268:**1039-45.

57. **Chen, C. H., D. P. von Kessler, W. Park, B. Wang, Y. Ma, and P. A. Beachy.** 1999. Nuclear trafficking of Cubitus interruptus in the transcriptional regulation of Hedgehog target gene expression. Cell **98:**305-16.

58. **Chen, X., J. Shen, and R. Prywes.** 2002. The luminal domain of ATF6 senses endoplasmic reticulum (ER) stress and causes translocation of ATF6 from the ER to the Golgi. J Biol Chem **277:**13045-52.

59. **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.

60. **Cheung, K. L., C. R. Graves, and J. F. Robertson.** 2000. Tumour marker measurements in the diagnosis and monitoring of breast cancer, vol. 26.

61. **Chiu, V. K., T. Bivona, A. Hach, J. B. Sajous, J. Silletti, H. Wiener, R. L. Johnson, 2nd, A. D. Cox, and M. R. Philips.** 2002. Ras signalling on the endoplasmic reticulum and the Golgi. Nat Cell Biol **4:**343-50.

62. **Chowdhury, S. F., J. Sivaraman, J. Wang, G. Devanathan, P. Lachance, H. Qi, R. Menard, J. Lefebvre, Y. Konishi, M. Cygler, T. Sulea, and E. O. Purisima.** 2002. Design of noncovalent inhibitors of human cathepsin L. From the 96-residue proregion to optimized tripeptides. J Med Chem **45:**5321-9.

63. **Ciechanover, A.** 1998. The ubiquitin-proteasome pathway: on protein death and cell life. Embo J **17:**7151-60.

64. **Clark, G. M.** 2001. Interpreting and integrating risk factors for patients with primary breast cancer. J Natl Cancer Inst Monogr:17-21.

65. **Collins, S., and M. Groudine.** 1982. Amplification of endogenous myc-related DNA sequences in a human myeloid leukaemia cell line. Nature **298:**679-81.

66. **Conaway, R. C., C. S. Brower, and J. W. Conaway.** 2002. Emerging roles of ubiquitin in transcription regulation. Science **296:**1254-8.

67. **Connor, M. K., R. Kotchetkov, S. Cariou, A. Resch, R. Lupetti, R. G. Beniston, F. Melchior, L. Hengst, and J. M. Slingerland.** 2003. CRM1/Ran-mediated nuclear export of p27(Kip1) involves a nuclear export signal and links p27 export and proteolysis. Mol Biol Cell **14:**201-13.

68. **Coquelle, A., E. Pipiras, F. Toledo, G. Buttin, and M. Debatisse.** 1997. Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell **89:**215-25.

69. **Coquelle, A., F. Toledo, S. Stern, A. Bieth, and M. Debatisse.** 1998. A new role for hypoxia in tumor progression: induction of fragile site triggering genomic rearrangements and formation of complex DMs and HSRs. Mol Cell **2:**259-65.

70. **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:**24862-24868.

71. **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 Journal **17:**4680-4694.

72. **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:**2561-2566.

73. **Cory, S.** 1986. Activation of cellular oncogenes in hemopoietic cells by chromosome translocation. Adv Cancer Res **47:**189-234.

74. **Cotran, R. S., Kumar, V., Robbins, S.L.** 1989. Robbins Pathologic Basis of Disease, 4th Edition ed. W.B. Saunders Company.

75. **Coulombe, R., P. Grochulski, J. Sivaraman, R. Menard, J. S. Mort, and M. Cygler.** 1996. Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. Embo J **15:**5492-503.

76. **Courjal, F., M. Cuny, J. Simony-Lafontaine, G. Louason, P. Speiser, R. Zeillinger, C. Rodriguez, and C. Theillet.** 1997. Mapping of DNA amplifications at 15 chromosomal localizations in 1875 breast tumors: definition of phenotypic groups. Cancer Res **57:**4360-7.

77. **Cygler, M., and J. S. Mort.** 1997. Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. Biochimie **79:**645-52.

78. **Dalla-Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce.** 1982. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci U S A **79:**7824-7.

79. **Dancey, J. E.** 2002. Clinical development of mammalian target of rapamycin inhibitors. Hematol Oncol Clin North Am **16:**1101-14.

80. **Danpure, C. J.** 1995. How can the product of a single gene be localized to more than one intracellular compartment? Trends In Cell Biology **5:**230-238.

81. **Datta, S. R., A. Brunet, and M. E. Greenberg.** 1999. Cellular survival: a play in three Akts. Genes Dev **13:**2905-27.

82. **de Klein, A., A. G. van Kessel, G. Grosveld, C. R. Bartram, A. Hagemeijer, D. Bootsma, N. K. Spurr, N. Heisterkamp, J. Groffen, and J. R. Stephenson.** 1982. A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. Nature **300:**765-7.

83. **De Vos, J., T. Thykjaer, K. Tarte, M. Ensslen, P. Raynaud, G. Requirand, F. Pellet, V. Pantesco, T. Reme, M. Jourdan, J. F. Rossi, T. Orntoft, and B. Klein.** 2002. Comparison of gene expression profiling between malignant and normal plasma cells with oligonucleotide arrays. Oncogene **21:**6848-57.

84. **de-The, G., A. Geser, N. E. Day, P. M. Tukei, E. H. Williams, D. P. Beri, P. G. Smith, A. G. Dean, G. W. Bronkamm, P. Feorino, and W. Henle.** 1978. Epidemiological evidence for causal relationship between Epstein-Barr virus and Burkitt's lymphoma from Ugandan prospective study. Nature **274:**756-61.

85. **DeCaprio, J. A., J. W. Ludlow, D. Lynch, Y. Furukawa, J. Griffin, H. Piwnica-Worms, C. M. Huang, and D. M. Livingston.** 1989. The product of the retinoblastoma susceptibility gene has properties of a cell cycle regulatory element. Cell **58:**1085-95.

86. **Delaria, K., L. Fiorentino, L. Wallace, P. Tamburini, E. Brownell, and D. Muller.** 1994. Inhibition of cathepsin L-like cysteine proteases by cytotoxic T-lymphocyte antigen-2 beta. J Biol Chem **269:**25172-7.

87. **Denhardt, D. T., A. H. Greenberg, S. E. Egan, R. T. Hamilton, and J. A. Wright.** 1987. Cysteine proteinase cathepsin L expression correlates closely with the metastatic potential of H-ras-transformed murine fibroblasts. Oncogene **2:**55-9.

88. **Dever, T. E.** 2002. Gene-specific regulation by general translation factors. Cell **108:**545-56.

89. **Dhillon, N., and R. T. Kamakaka.** 2002. Breaking through to the other side: silencers and barriers. Curr Opin Genet Dev **12:**188-92.

90. **Di Fiore, P. P.** 2003. Signal transduction: life on Mars, cellularly speaking. Nature **424:**624-5.

91. **Di Fiore, P. P., J. H. Pierce, M. H. Kraus, O. Segatto, C. R. King, and S. A. Aaronson.** 1987. erbB-2 is a potent oncogene when overexpressed in NIH/3T3 cells. Science **237:**178-82.

92. **DiCiommo, D., B. L. Gallie, and R. Bremner.** 2000. Retinoblastoma: the disease, gene and protein provide critical leads to understand cancer. Semin Cancer Biol **10:**255-69.

93. **Dickson, R. B., and G. M. Stancel.** 2000. Estrogen receptor-mediated processes in normal and cancer cells. J Natl Cancer Inst Monogr:135-45.

94. **Dohchin, A., J. I. Suzuki, H. Seki, M. Masutani, H. Shiroto, and Y. Kawakami.** 2000. Immunostained cathepsins B and L correlate with depth of invasion and different metastatic pathways in early stage gastric carcinoma. Cancer **89:**482-7.

95. **Done, S. J., N. C. Arneson, H. Ozcelik, M. Redston, and I. L. Andrulis.** 1998. p53 mutations in mammary ductal carcinoma in situ but not in epithelial hyperplasias. Cancer Res **58:**785-9.

96. **Dong, J. M., E. M. Prence, and G. G. Sahagian.** 1989. Mechanism for selective secretion of a lysosomal protease by transformed mouse fibroblasts. J Biol Chem **264:**7377-83.

97. **Downward, J.** 2003. Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer **3:**11-22.

98. **Duffy, M. J.** 2001. Biochemical markers in breast cancer: which ones are clinically useful? Clin Biochem **34:**347-52.

99. **Duffy, M. J.** 1996. Proteases as prognostic markers in cancer. Clinical Cancer Research **2:**613-8.

100. **Dufour, E., and B. Ribadeau-Dumas.** 1988. Proteolytic specificity of chicken cathepsin L on bovine beta-casein. Biosci Rep **8:**185-91.

101. **Dyson, N., K. Buchkovich, P. Whyte, and E. Harlow.** 1989. Cellular proteins that are targetted by DNA tumor viruses for transformation. Princess Takamatsu Symp **20:**191-8.

102. **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.

103. **Epling-Burnette, P. K., R. Garcia, F. Bai, S. Ismail, T. P. Loughran, Jr., J. Y. Djeu, R. Jove, and S. Wei.** 2002. Carboxy-terminal truncated STAT5 is induced by interleukin-2 and GM-CSF in human neutrophils. Cell Immunol **217:**1-11.

104. **Erickson, A. H.** 1989. Biosynthesis of lysosomal endopeptidases. J Cell Biochem **40:**31-41.

105. **Escot, C., C. Theillet, R. Lidereau, F. Spyratos, M. H. Champeme, J. Gest, and R. Callahan.** 1986. Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas. Proc Natl Acad Sci U S A **83:**4834-8.

106. **Esteller, M.** 2003. Cancer epigenetics: DNA methylation and chromatin alterations in human cancer. Adv Exp Med Biol **532:**39-49.

107. **Foekens, J. A., J. Kos, H. A. Peters, M. Krasovec, M. P. Look, N. Cimerman, M. E. Meijer-van Gelder, S. C. Henzen-Logmans, W. L. van Putten, and J. G. Klijn.** 1998. Prognostic significance of cathepsins B and L in primary human breast cancer. J Clin Oncol **16:**1013-21.

108. **Foulkes, N. S., and P. Sassone-Corsi.** 1992. More is better: activators and repressors from the same gene. Cell **68:**411-4.

109. **Friend, S. H., R. Bernards, S. Rogelj, R. A. Weinberg, J. M. Rapaport, D. M. Albert, and T. P. Dryja.** 1986. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature **323:**643-6.

110. **Fung, Y. K., and A. T'Ang.** 1992. The role of the retinoblastoma gene in breast cancer development. Cancer Treat Res **61:**59-68.

111. **Gaba, A., Z. Wang, T. Krishnamoorthy, A. G. Hinnebusch, and M. S. Sachs.** 2001. Physical evidence for distinct mechanisms of translational control by upstream open reading frames. Embo J **20:**6453-63.

112. **Gal, S., and M. M. Gottesman.** 1988. Isolation and sequence of a cDNA for human pro-(cathepsin L). Biochem J **253:**303-6.

113. **Gal, S., and M. M. Gottesman.** 1986. The major excreted protein (MEP) of transformed mouse cells and cathepsin L have similar protease specificity. Biochem Biophys Res Commun **139:**156-62.

114. **Ganem, D., and H. E. Varmus.** 1987. The molecular biology of the hepatitis B viruses. Annu Rev Biochem **56:**651-93.

115. **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.

116. **Gelb, B. D., G. P. Shi, H. A. Chapman, and R. J. Desnick.** 1996. Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. Science **273:**1236-8.

117. **Geyer, P. K., and I. Clark.** 2002. Protecting against promiscuity: the regulatory role of insulators. Cell Mol Life Sci **59:**2112-27.

118. **Giandomenico, V., M. Simonsson, E. Gronroos, and J. Ericsson.** 2003. Coactivator-dependent acetylation stabilizes members of the SREBP family of transcription factors. Mol Cell Biol **23:**2587-99.

119. **Gille, H., and J. Downward.** 1999. Multiple ras effector pathways contribute to G(1) cell cycle progression. J Biol Chem **274:**22033-40.

120. **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.

121. **Gillingham, A. K., and S. Munro.** 2003. Long coiled-coil proteins and membrane traffic. Biochim Biophys Acta **1641:**71-85.

122. **Gingras, A. C., S. P. Gygi, B. Raught, R. D. Polakiewicz, R. T. Abraham, M. F. Hoekstra, R. Aebersold, and N. Sonenberg.** 1999. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev **13:**1422-37.

123. **Gingras, A. C., B. Raught, and N. Sonenberg.** 2001. Regulation of translation initiation by FRAP/mTOR. Genes Dev **15:**807-26.

124. **Gottesman, M. M.** 1978. Transformation-dependent secretion of a low molecular weight protein by murine fibroblasts. Proc Natl Acad Sci U S A **75:**2767-71.

125. **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.

126. **Gupta, S., M. X. Luong, S. A. Bleuming, A. Miele, M. Luong, D. Young, E. S. Knudsen, A. J. Van Wijnen, J. L. Stein, and G. S. Stein.** 2003. Tumor suppressor pRB functions as a co-repressor of the CCAAT displacement protein (CDP/cut) to regulate cell cycle controlled histone H4 transcription. J Cell Physiol **196:**541-56.

127. **Hamad, N. M., J. H. Elconin, A. E. Karnoub, W. Bai, J. N. Rich, R. T. Abraham, C. J. Der, and C. M. Counter.** 2002. Distinct requirements for Ras oncogenesis in human versus mouse cells. Genes Dev **16:**2045-57.

128. **Hanahan, D., and R. A. Weinberg.** 2000. The hallmarks of cancer. Cell **100:**57-70.

129. **Harada, H., T. Fujita, M. Miyamoto, Y. Kimura, M. Maruyama, A. Furia, T. Miyata, and T. Taniguchi.** 1989. Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to the same regulatory elements of IFN and IFN-inducible genes. Cell **58:**729-39.

130. **Harada, H., K. Willison, J. Sakakibara, M. Miyamoto, T. Fujita, and T. Taniguchi.** 1990. Absence of the type I IFN system in EC cells: transcriptional activator (IRF-1) and repressor (IRF-2) genes are developmentally regulated. Cell **63:**303-12.

131. **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.

132. **Harbeck, N., U. Alt, U. Berger, R. Kates, A. Kruger, C. Thomssen, F. Janicke, H. Graeff, and M. Schmitt.** 2000. Long-term follow-up confirms prognostic impact of PAI-1 and cathepsin D and L in primary breast cancer. Int J Biol Markers **15:**79-83.

133. **Harbeck, N., U. Alt, U. Berger, A. Kruger, C. Thomssen, F. Janicke, H. Hofler, R. E. Kates, and M. Schmitt.** 2001. Prognostic impact of proteolytic factors (urokinase-type plasminogen activator, plasminogen activator inhibitor 1, and cathepsins B, D, and L) in primary breast cancer reflects effects of adjuvant systemic therapy. Clin Cancer Res **7:**2757-64.

134. **Harris, H., O. J. Miller, G. Klein, P. Worst, and T. Tachibana.** 1969. Suppression of malignancy by cell fusion. Nature **223:**363-8.

135. **Harvey, J. J.** 1964. An unidentified virus which causes the rapid production of tumors in mice. Nature **204:**1104-1105.

136. **Heidtmann, H. H., U. Salge, K. Havemann, H. Kirschke, and B. Wiederanders.** 1993. Secretion of a latent, acid activatable cathepsin L precursor by human non-small cell lung cancer cell lines. Oncol Res **5:**441-51.

137. **Heinrich, J., M. Bosse, H. Eickhoff, W. Nietfeld, R. Reinhardt, H. Lehrach, and K. Moelling.** 2000. Induction of putative tumor-suppressing genes in Rat-1 fibroblasts by oncogenic Raf-1 as evidenced by robot-assisted complex hybridization. J Mol Med **78:**380-8.

138. **Hellman, A., E. Zlotorynski, S. W. Scherer, J. Cheung, J. B. Vincent, D. I. Smith, L. Trakhtenbrot, and B. Kerem.** 2002. A role for common fragile site induction in amplification of human oncogenes. Cancer Cell **1:**89-97.

139. **Hengst, L., V. Dulic, J. M. Slingerland, E. Lees, and S. I. Reed.** 1994. A cell cycle-regulated inhibitor of cyclin-dependent kinases. Proc Natl Acad Sci U S A **91:**5291-5.

140. **Hengst, L., and S. I. Reed.** 1996. Translational control of p27Kip1 accumulation during the cell cycle. Science **271:**1861-4.

141. **Hershko, A., and A. Ciechanover.** 1998. The ubiquitin system. Annu Rev Biochem **67:**425-79.

142. **Hingorani, S. R., and D. A. Tuveson.** 2003. Ras redux: rethinking how and where Ras acts. Curr Opin Genet Dev **13:**6-13.

143. **Holley, R. W.** 1975. Control of growth of mammalian cells in cell culture. Nature **258:**487-90.

144. **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, J. L. Stein, and G. S. Stein.** 1990. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science **247:**1454-7.

145. **Honey, K., T. Nakagawa, C. Peters, and A. Rudensky.** 2002. Cathepsin L regulates CD4+ T cell selection independently of its effect on invariant chain: a role in the generation of positively selecting peptide ligands. J Exp Med **195:**1349-58.

146. **Hoppe, T., M. Rape, and S. Jentsch.** 2001. Membrane-bound transcription factors: regulated release by RIP or RUP. Curr Opin Cell Biol **13:**344-8.

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

148. **Irving, J. A., S. S. Shushanov, R. N. Pike, E. Y. Popova, D. Bromme, T. H. Coetzer, S. P. Bottomley, I. A. Boulynko, S. A. Grigoryev, and J. C. Whisstock.** 2002. Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation. J Biol Chem **277:**13192-201.

149. **Isaacs, C., V. Stearns, and D. F. Hayes.** 2001. New prognostic factors for breast cancer recurrence. Semin Oncol **28:**53-67.

150. **Ishida, N., T. Hara, T. Kamura, M. Yoshida, K. Nakayama, and K. I. Nakayama.** 2002. Phosphorylation of p27Kip1 on serine 10 is required for its binding to CRM1 and nuclear export. J Biol Chem **277:**14355-8.

151. **Ishidoh, K., and E. Kominami.** 1995. Procathepsin L degrades extracellular matrix proteins in the presence of glycosaminoglycans in vitro. Biochem Biophys Res Commun **217:**624-31.

152. **Ishidoh, K., and E. Kominami.** 2002. Processing and activation of lysosomal proteinases. Biol Chem **383:**1827-31.

153. **Ishidoh, K., T. Towatari, S. Imajoh, H. Kawasaki, E. Kominami, N. Katunuma, and K. Suzuki.** 1987. Molecular cloning and sequencing of cDNA for rat cathepsin L. FEBS Lett **223:**69-73.

154. **Jack, J., and Y. DeLotto.** 1995. Structure and regulation of a complex locus: the cut gene of Drosophila. Genetics **139:**1689-700.

155. **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.

156. **James, R., and R. A. Bradshaw.** 1984. Polypeptide growth factors. Annu Rev Biochem **53:**259-92.

157. **Jean, D., N. Guillaume, and R. Frade.** 2002. Characterization of human cathepsin L promoter and identification of binding sites for NF-Y, Sp1 and Sp3 that are essential for its activity. Biochem J **361:**173-84.

158. **Jenuwein, T., and C. D. Allis.** 2001. Translating the histone code. Science **293:**1074-80.

159. **Jerala, R., E. Zerovnik, J. Kidric, and V. Turk.** 1998. pH-induced conformational transitions of the propeptide of human cathepsin L. A role for a molten globule state in zymogen activation. Journal of Biological Chemistry **273:**11498-504.

160. **Jia, J., K. Amanai, G. Wang, J. Tang, B. Wang, and J. Jiang.** 2002. Shaggy/GSK3 antagonizes Hedgehog signalling by regulating Cubitus interruptus. Nature **416:**548-52.

161. **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.

162. **Jones, P. A., and S. B. Baylin.** 2002. The fundamental role of epigenetic events in cancer. Nat Rev Genet **3:**415-28.

163. **Joseph, L., S. Lapid, and V. Sukhatme.** 1987. The major ras induced protein in NIH3T3 cells is cathepsin L. Nucleic Acids Res **15:**3186.

164. **Joseph, L. J., L. C. Chang, D. Stamenkovich, and V. P. Sukhatme.** 1988. Complete nucleotide and deduced amino acid sequences of human and murine preprocathepsin L. An abundant transcript induced by transformation of fibroblasts. J Clin Invest **81:**1621-9.

165. **Julien, E., and W. Herr.** 2003. Proteolytic processing is necessary to separate and ensure proper cell growth and cytokinesis functions of HCF-1. Embo J **22:**2360-9.

166. **Kane, S. E.** 1993. Mouse procathepsin L lacking a functional glycosylation site is properly folded, stable, and secreted by NIH 3T3 cells. J Biol Chem **268:**11456-62.

167. **Katunuma, N., H. Tsuge, M. Nukatsuka, and M. Fukushima.** 2002. Structure-based development of cathepsin L inhibitors and therapeutic applications for prevention of cancer metastasis and cancer-induced osteoporosis. Adv Enzyme Regul **42:**159-72.

168. **Khazaie, K., T. J. Dull, T. Graf, J. Schlessinger, A. Ullrich, H. Beug, and B. Vennstrom.** 1988. Truncation of the human EGF receptor leads to differential

transforming potentials in primary avian fibroblasts and erythroblasts. Embo J **7:**3061-71.

169. **Kim, K., J. Cai, S. Shuja, T. Kuo, and M. J. Murnane.** 1998. Presence of activated ras correlates with increased cysteine proteinase activities in human colorectal carcinomas. Int J Cancer **79:**324-33.

170. **Kim, S. Y., A. Herbst, K. A. Tworkowski, S. E. Salghetti, and W. P. Tansey.** 2003. Skp2 regulates Myc protein stability and activity. Mol Cell **11:**1177-88.

171. **Kirschke, H., J. Langner, B. Wiederanders, S. Ansorge, and P. Bohley.** 1977. Cathepsin L. A new proteinase from rat-liver lysosomes. Eur J Biochem **74:**293-301.

172. **Kirsten, W. J. a. M., L.A.** 1967. Morphologic responses to a murine erythroblastosis virus. Journal of National Cancer Institute **39:**311-335.

173. **Knudson, A. G., Jr.** 1971. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A **68:**820-3.

174. **Koepp, D. M., J. W. Harper, and S. J. Elledge.** 1999. How the cyclin became a cyclin: regulated proteolysis in the cell cycle. Cell **97:**431-4.

175. **Koromilas, A. E., A. Lazaris-Karatzas, and N. Sonenberg.** 1992. mRNAs containing extensive secondary structure in their 5' non-coding region translate efficiently in cells overexpressing initiation factor eIF-4E. Embo J **11:**4153-8.

176. **Kozak, M.** 1987. At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. J Mol Biol **196:**947-50.

177. **Kozak, M.** 2002. Emerging links between initiation of translation and human diseases. Mamm Genome **13:**401-10.

178. **Kozak, M.** 1999. Initiation of translation in prokaryotes and eukaryotes. Gene **234:**187-208.

179. **Kozak, M.** 2001. New ways of initiating translation in eukaryotes? Mol Cell Biol **21:**1899-907.

180. **Kozak, M.** 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell **44:**283-92.

181. **Kozak, M.** 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene **299:**1-34.

182. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem **266:**19867-70.

183. **Krasilnikov, M. A.** 2000. Phosphatidylinositol-3 kinase dependent pathways: the role in control of cell growth, survival, and malignant transformation. Biochemistry (Mosc) **65:**59-67.

184. **Kristie, T. M., J. L. Pomerantz, T. C. Twomey, S. A. Parent, and P. A. Sharp.** 1995. The cellular C1 factor of the herpes simplex virus enhancer complex is a family of polypeptides. J Biol Chem **270:**4387-94.

185. **Kung, H. J., C. Boerkoel, and T. H. Carter.** 1991. Retroviral mutagenesis of cellular oncogenes: a review with insights into the mechanisms of insertional activation. Curr Top Microbiol Immunol **171:**1-25.

186. **Lah, T. T., E. Kalman, D. Najjar, E. Gorodetsky, P. Brennan, R. Somers, and I. Daskal.** 2000. Cells producing cathepsins D, B, and L in human breast carcinoma and their association with prognosis. Hum Pathol **31:**149-60.

187. **Land, H., L. F. Parada, and R. A. Weinberg.** 1983. Cellular oncogenes and multistep carcinogenesis. Science **222:**771-8.

188. **Land, H., L. F. Parada, and R. A. Weinberg.** 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature **304:**596-602.

189. **Lane, D. P., and L. V. Crawford.** 1979. T antigen is bound to a host protein in SV40-transformed cells. Nature **278:**261-3.

190. **Lazzarino, D., and C. A. Gabel.** 1990. Protein determinants impair recognition of procathepsin L phosphorylated oligosaccharides by the cation-independent mannose 6-phosphate receptor. J Biol Chem **265:**11864-71.

191. **Ledford, A. W., J. G. Brantley, G. Kemeny, T. L. Foreman, S. E. Quaggin, P. Igarashi, S. M. Oberhaus, M. Rodova, J. P. Calvet, and G. B. Vanden Heuvel.** 2002. Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia. Dev Biol **245:**157-71.

192. **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.

193. **Lemaire, R., R. M. Flipo, D. Monte, T. Dupressoir, B. Duquesnoy, J. Y. Cesbron, A. Janin, A. Capron, and R. Lafyatis.** 1994. Synovial fibroblast-like cell transfection with the SV40 large T antigen induces a transformed phenotype and permits transient tumor formation in immunodeficient mice. J Rheumatol **21:**1409-19.

194. **Lemieux, N., X. X. Zhang, D. Dufort, and A. Nepveu.** 1994. Assignment of the human homologue of the Drosophila Cut homeobox gene (CUTL1) to band 7q22 by fluorescence in situ hybridization. Genomics **24:**191-3.

195. **Lenarcic, B., and T. Bevec.** 1998. Thyropins--new structurally related proteinase inhibitors. Biol Chem **379:**105-11.

196. **Leonardi, A., B. Turk, and V. Turk.** 1996. Inhibition of bovine cathepsins L and S by stefins and cystatins. Biol Chem Hoppe Seyler **377:**319-21.

197. **Leto, G., F. M. Tumminello, G. Pizzolanti, G. Montalto, M. Soresi, and N. Gebbia.** 1997. Lysosomal cathepsins B and L and Stefin A blood levels in patients with hepatocellular carcinoma and/or liver cirrhosis: potential clinical implications. Oncology **54:**79-83.

198. **Levicar, N., J. Kos, A. Blejec, R. Golouh, I. Vrhovec, S. Frkovic-Grazio, and T. T. Lah.** 2002. Comparison of potential biological markers cathepsin B, cathepsin L, stefin A and stefin B with urokinase and plasminogen activator inhibitor-1 and clinicopathological data of breast carcinoma patients. Cancer Detect Prev **26:**42-9.

199. **Levine, M., and R. Tjian.** 2003. Transcription regulation and animal diversity. Nature **424:**147-51.

200. **Li, S., B. Aufiero, R. L. Schiltz, and M. J. Walsh.** 2000. Regulation of the homeodomain CCAAT displacement/cut protein function by histone acetyltransferases p300/CREB-binding protein (CBP)- associated factor and CBP. Proc Natl Acad Sci U S A **97:**7166-71.

201. **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. Journal of Biological Chemistry **274:**7803-15.

202. **Lievens, P. M., C. Tufarelli, J. J. Donady, A. Stagg, and E. J. Neufeld.** 1997. CASP, a novel, highly conserved alternative-splicing product of the CDP/cut/cux gene, lacks cut-repeat and homeo DNA-binding domains, and interacts with full-length CDP in vitro. Gene **197:**73-81.

203. **Lin, A., J. Frost, T. Deng, T. Smeal, N. al-Alawi, U. Kikkawa, T. Hunter, D. Brenner, and M. Karin.** 1992. Casein kinase II is a negative regulator of c-Jun DNA binding and AP-1 activity. Cell **70:**777-89.

204. **Lin, C. R., W. S. Chen, W. Kruiger, L. S. Stolarsky, W. Weber, R. M. Evans, I. M. Verma, G. N. Gill, and M. G. Rosenfeld.** 1984. Expression cloning of human EGF receptor complementary DNA: gene amplification and three related messenger RNA products in A431 cells. Science **224:**843-8.

205. **Linnevers, C., S. P. Smeekens, and D. Bromme.** 1997. Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes. FEBS Lett **405:**253-9.

206. **Littlewood-Evans, A. J., G. Bilbe, W. B. Bowler, D. Farley, B. Wlodarski, T. Kokubo, T. Inaoka, J. Sloane, D. B. Evans, and J. A. Gallagher.** 1997. The osteoclast-associated protease cathepsin K is expressed in human breast carcinoma. Cancer Res **57:**5386-90.

207. **Liu, C. W., M. J. Corboy, G. N. DeMartino, and P. J. Thomas.** 2003. Endoproteolytic activity of the proteasome. Science **299:**408-11.

208. **Liu, E. T.** 2003. Classification of cancers by expression profiling. Curr Opin Genet Dev **13:**97-103.

209. **Liu, J., A. Barnett, E. J. Neufeld, and J. P. Dudley.** 1999. Homeoproteins CDP and SATB1 interact: potential for tissue-specific regulation. Mol Cell Biol **19:**4918-26.

210. **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.

211. **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.

212. **Luciano, R. L., and A. C. Wilson.** 2002. An activation domain in the C-terminal subunit of HCF-1 is important for transactivation by VP16 and LZIP. Proc Natl Acad Sci U S A **99:**13403-8.

213. **Ludlow, C., R. Choy, and K. Blochlinger.** 1996. Functional analysis of Drosophila and mammalian Cut proteins In flies. Developmental Biology **178:**149-159.

214. **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.

215. **Magee, T., and C. Marshall.** 1999. New insights into the interaction of Ras with the plasma membrane. Cell **98:**9-12.

216. **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. Molecular & Cellular Biology **16:**5346-57.

217. **Malek, N. P., H. Sundberg, S. McGrew, K. Nakayama, T. R. Kyriakides, J. M. Roberts, and T. R. Kyriakidis.** 2001. A mouse knock-in model exposes sequential proteolytic pathways that regulate p27Kip1 in G1 and S phase. Nature **413:**323-7.

218. **Marcotrigiano, J., A. C. Gingras, N. Sonenberg, and S. K. Burley.** 1997. X-ray studies of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Nucleic Acids Symp Ser:8-11.

219. **Marcu, K. B., S. A. Bossone, and A. J. Patel.** 1992. myc function and regulation. Annu Rev Biochem **61:**809-60.

220. **Marcu, K. B., L. J. Harris, L. W. Stanton, J. Erikson, R. Watt, and C. M. Croce.** 1983. Transcriptionally active c-myc oncogene is contained within NIARD, a DNA sequence associated with chromosome translocations in B-cell neoplasia. Proc Natl Acad Sci U S A **80:**519-23.

221. **Mason, R. W., S. Gal, and M. M. Gottesman.** 1987. The identification of the major excreted protein (MEP) from a transformed mouse fibroblast cell line as a catalytically active precursor form of cathepsin L. Biochem J **248:**449-54.

222. **Mason, R. W., G. D. Green, and A. J. Barrett.** 1985. Human liver cathepsin L. Biochem J **226:**233-41.

223. **Mason, R. W., and S. D. Massey.** 1992. Surface activation of pro-cathepsin L. Biochem Biophys Res Commun **189:**1659-66.

224. **Massoglia, S., A. Gray, T. J. Dull, S. Munemitsu, H. J. Kun, J. Schlessinger, and A. Ullrich.** 1990. Epidermal growth factor receptor cytoplasmic domain mutations trigger ligand-independent transformation. Mol Cell Biol **10:**3048-55.

225. **Maubach, G., K. Schilling, W. Rommerskirch, I. Wenz, J. E. Schultz, E. Weber, and B. Wiederanders.** 1997. The inhibition of cathepsin S by its propeptide--specificity and mechanism of action. Eur J Biochem **250:**745-50.

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

227. **McGrath, M. E.** 1999. The lysosomal cysteine proteases. Annu Rev Biophys Biomol Struct **28:**181-204.

228. **McIntyre, G. F., and A. H. Erickson.** 1993. The lysosomal proenzyme receptor that binds procathepsin L to microsomal membranes at pH 5 is a 43-kDa integral membrane protein. Proc Natl Acad Sci U S A **90:**10588-92.

229. **McIntyre, G. F., G. D. Godbold, and A. H. Erickson.** 1994. The pH-dependent membrane association of procathepsin L is mediated by a 9-residue sequence within the propeptide. J Biol Chem **269:**567-72.

230. **Menard, R., E. Carmona, S. Takebe, E. Dufour, C. Plouffe, P. Mason, and J. S. Mort.** 1998. Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro. J Biol Chem **273:**4478-84.

231. **Mendez, R., and J. D. Richter.** 2001. Translational control by CPEB: a means to the end. Nat Rev Mol Cell Biol **2:**521-9.

232. **Methot, N., and K. Basler.** 2001. An absolute requirement for Cubitus interruptus in Hedgehog signaling. Development **128:**733-42.

233. **Molinari, E., M. Gilman, and S. Natesan.** 1999. Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency in vivo. Embo J **18:**6439-47.

234. **Moon, N. S., G. Berube, and A. Nepveu.** 2000. CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain. Journal of Biological Chemistry **275:**31325-31334.

235. **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. Molecular & Cellular Biology **21:**6332-6345.

236. **Moon, N. S., W. Rong Zeng, P. Premdas, M. Santaguida, G. Berube, and A. Nepveu.** 2002. Expression of N-terminally truncated isoforms of CDP/CUX is increased in human uterine leiomyomas. Int J Cancer **100:**429-32.

237. **Mori, I., Q. Yang, and K. Kakudo.** 2002. Predictive and prognostic markers for invasive breast cancer. Pathol Int **52:**186-94.

238. **Nakagawa, H., R. B. Chadwick, P. Peltomaki, C. Plass, Y. Nakamura, and A. de La Chapelle.** 2001. Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of H19-associated CTCF-binding sites in colorectal cancer. Proc Natl Acad Sci U S A **98:**591-6.

239. **Nakagawa, H., G. J. Nuovo, E. E. Zervos, E. W. Martin, Jr., R. Salovaara, L. A. Aaltonen, and A. de la Chapelle.** 2001. Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. Cancer Res **61:**6991-5.

240. **Nakagawa, T., W. Roth, P. Wong, A. Nelson, A. Farr, J. Deussing, J. A. Villadangos, H. Ploegh, C. Peters, and A. Y. Rudensky.** 1998. Cathepsin L: critical role in Ii degradation and CD4 T cell selection in the thymus. Science **280:**450-3.

241. **Nakagawa, T. Y., W. H. Brissette, P. D. Lira, R. J. Griffiths, N. Petrushova, J. Stock, J. D. McNeish, S. E. Eastman, E. D. Howard, S. R. Clarke, E. F. Rosloniec, E. A. Elliott, and A. Y. Rudensky.** 1999. Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. Immunity **10:**207-17.

242. **Nakashima, S., K. Morinaka, S. Koyama, M. Ikeda, M. Kishida, K. Okawa, A. Iwamatsu, S. Kishida, and A. Kikuchi.** 1999. Small G protein Ral and its downstream molecules regulate endocytosis of EGF and insulin receptors. Embo J **18:**3629-42.

243. **Neel, B. G., W. S. Hayward, H. L. Robinson, J. Fang, and S. M. Astrin.** 1981. Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion. Cell **23:**323-34.

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

245. **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.

246. **Nishida, Y., K. Kohno, T. Kawamata, K. Morimitsu, M. Kuwano, and I. Miyakawa.** 1995. Increased cathepsin L levels in serum in some patients with ovarian cancer: comparison with CA125 and CA72-4. Gynecol Oncol **56:**357-61.

247. **Nishimura, Y., T. Kawabata, K. Furuno, and K. Kato.** 1989. Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. Arch Biochem Biophys **271:**400-6.

248. **Nomura, T., and Y. Fujisawa.** 1997. Processing properties of recombinant human procathepsin L. Biochem Biophys Res Commun **230:**143-6.

249. **Oda, A., H. Wakao, and H. Fujita.** 2002. Calpain is a signal transducer and activator of transcription (STAT) 3 and STAT5 protease. Blood **99:**1850-2.

250. **Oki, M., and R. T. Kamakaka.** 2002. Blockers and barriers to transcription: competing activities? Curr Opin Cell Biol **14:**299-304.

251. **Osborne, C. K., M. G. Yochmowitz, W. A. Knight, 3rd, and W. L. McGuire.** 1980. The value of estrogen and progesterone receptors in the treatment of breast cancer. Cancer **46:**2884-8.

252. **Otting, G., Y. Q. Qian, M. Billeter, M. Muller, M. Affolter, W. J. Gehring, and K. Wuthrich.** 1990. Protein--DNA contacts in the structure of a homeodomain--DNA complex determined by nuclear magnetic resonance spectroscopy in solution. Embo J **9:**3085-92.

253. **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.

254. **Palombella, V. J., and T. Maniatis.** 1992. Inducible processing of interferon regulatory factor-2. Mol Cell Biol **12:**3325-36.

255. **Pardee, A. B.** 1989. G1 events and regulation of cell proliferation. Science **246:**603-8.

256. **Park, I. C., S. Y. Lee, D. G. Jeon, J. S. Lee, C. S. Hwang, B. G. Hwang, S. H. Lee, W. S. Hong, and S. I. Hong.** 1996. Enhanced expression of cathepsin L in metastatic bone tumors. J Korean Med Sci **11:**144-8.

257. **Pause, A., N. Methot, Y. Svitkin, W. C. Merrick, and N. Sonenberg.** 1994. Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. Embo J **13:**1205-15.

258. **Pelletier, J., and N. Sonenberg.** 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature **334:**320-5.

259. **Pestova, T. V., and V. G. Kolupaeva.** 2002. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. Genes Dev **16:**2906-22.

260. **Pestova, T. V., I. B. Lomakin, J. H. Lee, S. K. Choi, T. E. Dever, and C. U. Hellen.** 2000. The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature **403:**332-5.

261. **Piazza, F., J. Valens, E. Lagasse, and C. Schindler.** 2000. Myeloid differentiation of FdCP1 cells is dependent on Stat5 processing. Blood **96:**1358-65.

262. **Piluso, D., P. Bilan, and J. P. Capone.** 2002. Host cell factor-1 interacts with and antagonizes transactivation by the cell cycle regulatory factor Miz-1. J Biol Chem **277:**46799-808.

263. **Plebani, M., L. Herszenyi, R. Cardin, G. Roveroni, P. Carraro, M. D. Paoli, M. Rugge, W. F. Grigioni, D. Nitti, R. Naccarato, and et al.** 1995. Cysteine and serine proteases in gastric cancer. Cancer **76:**367-75.

264. **Poiesz, B. J., F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, and R. C. Gallo.** 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. Proc Natl Acad Sci U S A **77:**7415-9.

265. **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.

266. **Polgar, L.** 1989. Transition state stabilization and enzymic rate acceleration. Acta Biochim Biophys Hung **24:**25-32.

267. **Preiss, T., and M. W. Hentze.** 1999. From factors to mechanisms: translation and translational control in eukaryotes. Curr Opin Genet Dev **9:**515-21.

268. **Preiss, T., M. Muckenthaler, and M. W. Hentze.** 1998. Poly(A)-tail-promoted translation in yeast: implications for translational control. Rna **4:**1321-31.

269. **Price, M. A., and D. Kalderon.** 1999. Proteolysis of cubitus interruptus in Drosophila requires phosphorylation by protein kinase A. Development **126:**4331-9.

270. **Price, M. A., and D. Kalderon.** 2002. Proteolysis of the Hedgehog signaling effector Cubitus interruptus requires phosphorylation by Glycogen Synthase Kinase 3 and Casein Kinase 1. Cell **108:**823-35.

271. **Prior, I. A., A. Harding, J. Yan, J. Sluimer, R. G. Parton, and J. F. Hancock.** 2001. GTP-dependent segregation of H-ras from lipid rafts is required for biological activity. Nat Cell Biol **3:**368-75.

272. **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:**22624-22634.

273. **Rape, M., T. Hoppe, I. Gorr, M. Kalocay, H. Richly, and S. Jentsch.** 2001. Mobilization of processed, membrane-tethered SPT23 transcription factor by CDC48(UFD1/NPL4), a ubiquitin-selective chaperone. Cell **107:**667-77.

274. **Rascle, A., J. A. Johnston, and B. Amati.** 2003. Deacetylase activity is required for recruitment of the basal transcription machinery and transactivation by STAT5. Mol Cell Biol **23:**4162-73.

275. **Ray, A., P. Ray, N. Guthrie, A. Shakya, D. Kumar, and B. K. Ray.** 2003. Protein kinase A signaling pathway regulates transcriptional activity of SAF-1 by unmasking its DNA-binding domains. J Biol Chem **278:**22586-95.

276. **Reinheckel, T., J. Deussing, W. Roth, and C. Peters.** 2001. Towards specific functions of lysosomal cysteine peptidases: phenotypes of mice deficient for cathepsin B or cathepsin L. Biological Chemistry **382:**735-41.

277. **Rescheleit, D. K., W. J. Rommerskirch, and B. Wiederanders.** 1996. Sequence analysis and distribution of two new human cathepsin L splice variants. FEBS Lett **394:**345-8.

278. **Riese, R. J., R. N. Mitchell, J. A. Villadangos, G. P. Shi, J. T. Palmer, E. R. Karp, G. T. De Sanctis, H. L. Ploegh, and H. A. Chapman.** 1998. Cathepsin S activity regulates antigen presentation and immunity. J Clin Invest **101:**2351-63.

279. **Rodrigues, G. A., and M. Park.** 1993. Dimerization mediated through a leucine zipper activates the oncogenic potential of the met receptor tyrosine kinase. Mol Cell Biol **13:**6711-22.

280. **Rodriguez-Viciana, P., P. H. Warne, R. Dhand, B. Vanhaesebroeck, I. Gout, M. J. Fry, M. D. Waterfield, and J. Downward.** 1994. Phosphatidylinositol-3-OH kinase as a direct target of Ras. Nature **370:**527-32.

281. **Roth, W., J. Deussing, V. A. Botchkarev, M. Pauly-Evers, P. Saftig, A. Hafner, P. Schmidt, W. Schmahl, J. Scherer, I. Anton-Lamprecht, K. Von Figura, R. Paus, and C. Peters.** 2000. Cathepsin L deficiency as molecular defect of furless: hyperproliferation of keratinocytes and pertubation of hair follicle cycling. Faseb J **14:**2075-86.

282. **Rothman, J. E.** 1987. Protein sorting by selective retention in the endoplasmic reticulum and Golgi stack. Cell **50:**521-2.

283. **Rozman, J., J. Stojan, R. Kuhelj, V. Turk, and B. Turk.** 1999. Autocatalytic processing of recombinant human procathepsin B is a bimolecular process. FEBS Lett **459:**358-62.

284. **Ruley, H. E.** 1983. Adenovirus early region 1A enables viral and cellular transforming genes to transform primary cells in culture. Nature **304:**602-6.

285. **Salghetti, S. E., A. A. Caudy, J. G. Chenoweth, and W. P. Tansey.** 2001. Regulation of transcriptional activation domain function by ubiquitin. Science **293:**1651-3.

286. **Salminen, A., and M. M. Gottesman.** 1990. Inhibitor studies indicate that active cathepsin L is probably essential to its own processing in cultured fibroblasts. Biochem J **272:**39-44.

287. **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.

288. **Santamaria, I., G. Velasco, M. Cazorla, A. Fueyo, E. Campo, and C. Lopez-Otin.** 1998. Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas. Cancer Res **58:**1624-30.

289. **Sarnow, P., Y. S. Ho, J. Williams, and A. J. Levine.** 1982. Adenovirus E1b-58kd tumor antigen and SV40 large tumor antigen are physically associated with the same 54 kd cellular protein in transformed cells. Cell **28:**387-94.

290. **Scherer, S. W., E. J. Neufeld, P. M. Lievens, S. H. Orkin, J. Kim, and L. C. Tsui.** 1993. Regional localization of the CCAAT displacement protein gene (CUTL1) to 7q22 by analysis of somatic cell hybrids. Genomics **15:**695-696.

291. **Schlotter, C. M., U. Vogt, U. Bosse, B. Mersch, and K. Wassmann.** 2003. C-myc, not HER-2/neu, can predict recurrence and mortality of patients with node-negative breast cancer. Breast Cancer Res **5:**R30-6.

292. **Schmitt, F. C., and J. S. Reis-Filho.** 2003. c-myc, not her-2/neu, can predict the prognosis of breast cancer patients: how novel, how accurate, and how significant? Breast Cancer Res **5:**188-91.

293. **Schneider, R., V. I. Agol, R. Andino, F. Bayard, D. R. Cavener, S. A. Chappell, J. J. Chen, J. L. Darlix, A. Dasgupta, O. Donze, R. Duncan, O. Elroy-Stein, P. J. Farabaugh, W. Filipowicz, M. Gale, Jr., L. Gehrke, E. Goldman, Y. Groner, J. B. Harford, M. Hatzglou, B. He, C. U. Hellen, M. W. Hentze, J. Hershey, P. Hershey, T. Hohn, M. Holcik, C. P. Hunter, K. Igarashi, R. Jackson, R. Jagus, L. S. Jefferson, B. Joshi, R. Kaempfer, M. Katze, R. J. Kaufman, M. Kiledjian, S. R. Kimball, A. Kimchi, K. Kirkegaard, A. E. Koromilas, R. M. Krug, V. Kruys, B. J. Lamphear, S. Lemon, R. E. Lloyd, L. E. Maquat, E. Martinez-Salas, M. B. Mathews, V. P. Mauro, S. Miyamoto, I. Mohr, D. R. Morris, E. G. Moss, N. Nakashima, A. Palmenberg, N. T. Parkin, T. Pe'ery, J. Pelletier, S. Peltz, T. V. Pestova, E. V. Pilipenko, A. C. Prats, V. Racaniello, G. S. Read, R. E. Rhoads, J. D. Richter, R. Rivera-Pomar, T. Rouault, A. Sachs, P. Sarnow, G. C. Scheper, L. Schiff, D. R. Schoenberg, B. L. Semler, A. Siddiqui, T. Skern, N. Sonenberg, S. M. Tahara, A. A. Thomas, J. J. Toulme, J. Wilusz, E. Wimmer, G. Witherell, and M. Wormington.** 2001. New ways of initiating translation in eukaryotes. Mol Cell Biol **21:**8238-46.

294. **Schnitt, S. J.** 2001. Traditional and newer pathologic factors. J Natl Cancer Inst Monogr:22-6.

295. **Seidah, N. G., and A. Prat.** 2002. Precursor convertases in the secretory pathway, cytosol and extracellular milieu. Essays Biochem **38:**79-94.

296. **Selivanova, G.** 2001. Mutant p53: the loaded gun. Curr Opin Investig Drugs **2:**1136-41.

297. **Shatkin, A. J.** 1976. Capping of eucaryotic mRNAs. Cell **9:**645-53.

298. **Sheahan, K., S. Shuja, and M. J. Murnane.** 1989. Cysteine protease activities and tumor development in human colorectal carcinoma. Cancer Res **49:**3809-14.

299. **Shen, J., X. Chen, L. Hendershot, and R. Prywes.** 2002. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. Dev Cell **3:**99-111.

300. **Shi, G. P., J. S. Munger, J. P. Meara, D. H. Rich, and H. A. Chapman.** 1992. Molecular cloning and expression of human alveolar macrophage cathepsin S, an elastinolytic cysteine protease. J Biol Chem **267:**7258-62.

301. **Shi, G. P., A. C. Webb, K. E. Foster, J. H. Knoll, C. A. Lemere, J. S. Munger, and H. A. Chapman.** 1994. Human cathepsin S: chromosomal localization, gene structure, and tissue distribution. J Biol Chem **269:**11530-6.

302. **Shtivelman, E., B. Lifshitz, R. P. Gale, and E. Canaani.** 1985. Fused transcript of abl and bcr genes in chronic myelogenous leukaemia. Nature **315:**550-4.

303. **Shuja, S., and M. J. Murnane.** 1996. Marked increases in cathepsin B and L activities distinguish papillary carcinoma of the thyroid from normal thyroid or thyroid with non-neoplastic disease. Int J Cancer **66:**420-6.

304. **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.

305. **Sivaparvathi, M., M. Yamamoto, G. L. Nicolson, Z. L. Gokaslan, G. N. Fuller, L. A. Liotta, R. Sawaya, and J. S. Rao.** 1996. Expression and immunohistochemical localization of cathepsin L during the progression of human gliomas. Clin Exp Metastasis **14:**27-34.

306. **Slamon, D. J., G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich, and W. L. McGuire.** 1987. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science **235:**177-82.

307. **Smale, S. T., and J. T. Kadonaga.** 2003. The RNA Polymerase II Core Promoter. Annu Rev Biochem.

308. **Smith, S. M., and M. M. Gottesman.** 1989. Activity and deletion analysis of recombinant human cathepsin L expressed in Escherichia coli. J Biol Chem **264:**20487-95.

309. **Smith, S. M., S. E. Kane, S. Gal, R. W. Mason, and M. M. Gottesman.** 1989. Glycosylation of procathepsin L does not account for species molecular-mass differences and is not required for proteolytic activity. Biochem J **262:**931-8.

310. **Society, A. C.** 2002. Cancer facts and Figures 2002.

311. **Society, C. C.** 2002. Canadian Cancer Statistics 2002.

312. **Sol-Church, K., J. Frenck, and R. W. Mason.** 2000. Cathepsin Q, a novel lysosomal cysteine protease highly expressed in placenta. Biochem Biophys Res Commun **267:**791-5.

313. **Sol-Church, K., J. Frenck, and R. W. Mason.** 2000. Mouse cathepsin M, a placenta-specific lysosomal cysteine protease related to cathepsins L and P. Biochim Biophys Acta **1491:**289-94.

314. **Sol-Church, K., J. Frenck, D. Troeber, and R. W. Mason.** 1999. Cathepsin P, a novel protease in mouse placenta. Biochem J **343 Pt 2:**307-9.

315. **Sommers, C. L.** 1996. The role of cadherin-mediated adhesion in breast cancer. J Mammary Gland Biol Neoplasia **1:**219-29.

316. **Sporn, M. B., and A. B. Roberts.** 1985. Autocrine growth factors and cancer. Nature **313:**745-7.

317. **Spotts, G. D., S. V. Patel, Q. Xiao, and S. R. Hann.** 1997. Identification of downstream-initiated c-Myc proteins which are dominant-negative inhibitors of transactivation by full-length c-Myc proteins. Mol Cell Biol **17:**1459-68.

318. **Stark, G. R., M. Debatisse, E. Giulotto, and G. M. Wahl.** 1989. Recent progress in understanding mechanisms of mammalian DNA amplification. Cell **57:**901-8.

319. **Stunkel, W., Z. Huang, S. H. Tan, M. J. O'Connor, and H. U. Bernard.** 2000. Nuclear matrix attachment regions of human papillomavirus type 16 repress or activate the E6 promoter, depending on the physical state of the viral DNA. Journal of Virology **74:**2489-501.

320. **Sukhova, G. K., Y. Zhang, J. H. Pan, Y. Wada, T. Yamamoto, M. Naito, T. Kodama, S. Tsimikas, J. L. Witztum, M. L. Lu, Y. Sakara, M. T. Chin, P. Libby, and G. P. Shi.** 2003. Deficiency of cathepsin S reduces atherosclerosis in LDL receptor-deficient mice. J Clin Invest **111:**897-906.

321. **Suminami, Y., S. Nawata, and H. Kato.** 1998. Biological role of SCC antigen. Tumour Biol **19:**488-93.

322. **Taniguchi, S., Y. Nishimura, T. Takahashi, T. Baba, and K. Kato.** 1990. Augmented excretion of procathepsin L of a fos-transferred highly metastatic rat cell line. Biochem Biophys Res Commun **168:**520-6.

323. **Tansey, W. P.** 2001. Transcriptional activation: risky business. Genes Dev **15:**1045-50.

324. **Tao, K., N. A. Stearns, J. Dong, Q. L. Wu, and G. G. Sahagian.** 1994. The proregion of cathepsin L is required for proper folding, stability, and ER exit. Archives of Biochemistry & Biophysics **311:**19-27.

325. **Tarun, S. Z., Jr., and A. B. Sachs.** 1996. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. Embo J **15:**7168-77.

326. **Tarun, S. Z., Jr., S. E. Wells, J. A. Deardorff, and A. B. Sachs.** 1997. Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. Proc Natl Acad Sci U S A **94:**9046-51.

327. **Tavares, A. T., T. Tsukui, and J. C. Izpisua Belmonte.** 2000. Evidence that members of the Cut/Cux/CDP family may be involved in AER positioning and polarizing activity during chick limb development. Development **127:**5133-5144.

328. **Thomssen, C., P. Oppelt, F. Janicke, K. Ulm, N. Harbeck, H. Hofler, W. Kuhn, H. Graeff, and M. Schmitt.** 1998. Identification of low-risk node-negative breast cancer patients by tumor biological factors PAI-1 and cathepsin L. Anticancer Res **18:**2173-80.

329. **Thomssen, C., M. Schmitt, L. Goretzki, P. Oppelt, L. Pache, P. Dettmar, F. Janicke, and H. Graeff.** 1995. Prognostic value of the cysteine proteases cathepsins B and cathepsin L in human breast cancer. Clin Cancer Res **1:**741-6.

330. **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.

331. **Troen, B. R., D. Ascherman, D. Atlas, and M. M. Gottesman.** 1988. Cloning and expression of the gene for the major excreted protein of transformed mouse fibroblasts. A secreted lysosomal protease regulated by transformation. J Biol Chem **263:**254-61.

332. **Troen, B. R., S. S. Chauhan, D. Ray, and M. M. Gottesman.** 1991. Downstream sequences mediate induction of the mouse cathepsin L promoter by phorbol esters. Cell Growth Differ **2:**23-31.

333. **Troen, B. R., S. Gal, and M. M. Gottesman.** 1987. Sequence and expression of the cDNA for MEP (major excreted protein), a transformation-regulated secreted cathepsin. Biochem J **246:**731-5.

334. **Truscott, M., L. Raynal, P. Premdas, B. Goulet, L. Leduy, G. Berube, and A. Nepveu.** 2003. CDP/Cux stimulates transcription from the DNA polymerase alpha gene promoter. Mol Cell Biol **23:**3013-28.

335. **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.

336. **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.

337. **Turk, B., I. Dolenc, B. Lenarcic, I. Krizaj, V. Turk, J. G. Bieth, and I. Bjork.** 1999. Acidic pH as a physiological regulator of human cathepsin L activity. European Journal of Biochemistry **259:**926-32.

338. **Turk, B., I. Dolenc, V. Turk, and J. G. Bieth.** 1993. Kinetics of the pH-induced inactivation of human cathepsin L. Biochemistry **32:**375-80.

339. **Turk, B., D. Turk, and V. Turk.** 2000. Lysosomal cysteine proteases: more than scavengers. Biochim Biophys Acta **1477:**98-111.

340. **Turk, B., V. Turk, and D. Turk.** 1997. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biological Chemistry **378:**141-50.

341. **Turk, D., G. Guncar, M. Podobnik, and B. Turk.** 1998. Revised definition of substrate binding sites of papain-like cysteine proteases. Biol Chem **379:**137-47.

342. **Turk, V., and W. Bode.** 1991. The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett **285:**213-9.

343. **Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, and et al.** 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. Nature **309:**418-25.

344. **Vagner, S., B. Galy, and S. Pyronnet.** 2001. Irresistible IRES. Attracting the translation machinery to internal ribosome entry sites. EMBO Rep **2:**893-8.

345. **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.

346. **van 't Veer, L. J., H. Dai, M. J. van de Vijver, Y. D. He, A. A. Hart, R. Bernards, and S. H. Friend.** 2003. Expression profiling predicts outcome in breast cancer. Breast Cancer Res **5:**57-8.

347. **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.

348. **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. Journal of Cellular Biochemistry **66:**512-23.

349. **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. Proceedings of the National Academy of Sciences of the United States of America **93:**11516-21.

350. **Vandel, L., and T. Kouzarides.** 1999. Residues phosphorylated by TFIIH are required for E2F-1 degradation during S-phase. Embo J **18:**4280-91.

351. **Vanden Heuvel, G. B., 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 Int **50:**453-61.

352. **Velicescu, M., D. J. Weisenberger, F. A. Gonzales, Y. C. Tsai, C. T. Nguyen, and P. A. Jones.** 2002. Cell division is required for de novo methylation of CpG islands in bladder cancer cells. Cancer Res **62:**2378-84.

353. **Venkitaraman, A. R.** 2002. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell **108:**171-82.

354. **Villadangos, J. A., R. A. Bryant, J. Deussing, C. Driessen, A. M. Lennon-Dumenil, R. J. Riese, W. Roth, P. Saftig, G. P. Shi, H. A. Chapman, C. Peters, and H. L. Ploegh.** 1999. Proteases involved in MHC class II antigen presentation. Immunol Rev **172:**109-20.

355. **Vogel, J. L., and T. M. Kristie.** 2000. The novel coactivator C1 (HCF) coordinates multiprotein enhancer formation and mediates transcription activation by GABP. Embo J **19:**683-90.

356. **Voges, D., P. Zwickl, and W. Baumeister.** 1999. The 26S proteasome: a molecular machine designed for controlled proteolysis. Annu Rev Biochem **68:**1015-68.

357. **von der Lehr, N., S. Johansson, S. Wu, F. Bahram, A. Castell, C. Cetinkaya, P. Hydbring, I. Weidung, K. Nakayama, K. I. Nakayama, O. Soderberg, T. K. Kerppola, and L. G. Larsson.** 2003. The F-box protein Skp2 participates in c-Myc proteosomal degradation and acts as a cofactor for c-Myc-regulated transcription. Mol Cell **11:**1189-200.

358. **Votta, B. J., M. A. Levy, A. Badger, J. Bradbeer, R. A. Dodds, I. E. James, S. Thompson, M. J. Bossard, T. Carr, J. R. Connor, T. A. Tomaszek, L. Szewczuk, F. H. Drake, D. F. Veber, and M. Gowen.** 1997. Peptide aldehyde inhibitors of cathepsin K inhibit bone resorption both in vitro and in vivo. J Bone Miner Res **12:**1396-406.

359. **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.

360. **Watanabe, M., T. Higashi, M. Hashimoto, I. Tomoda, S. Tominaga, N. Hashimoto, S. Morimoto, Y. Yamauchi, H. Nakatsukasa, M. Kobayashi, and et**

al. 1987. Elevation of tissue cathepsin B and L activities in gastric cancer. Hepatogastroenterology **34:**120-2.

361. **Webster, M., N. Martin-Soudant, A. Nepveu, R. D. Cardiff, and W. Muller.** 1998. The induction of uterine leiomyomas and mammary tumors in transgenic mice expressing the Polyomavirus (PyV) large T (LT) antigen is associated with ability of PyV LT antigen to form specific complexes with Retinoblastoma and CUTL1 family members.Oncogene, in the press.

362. **Weihofen, A., and B. Martoglio.** 2003. Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides. Trends Cell Biol **13:**71-8.

363. **Weinberg, R. A.** 1990. The retinoblastoma gene and cell growth control. Trends Biochem Sci **15:**199-202.

364. **Wells, A., and J. M. Bishop.** 1988. Genetic determinants of neoplastic transformation by the retroviral oncogene v-erbB. Proc Natl Acad Sci U S A **85:**7597-601.

365. **Welss, T., J. Sun, J. A. Irving, R. Blum, A. I. Smith, J. C. Whisstock, R. N. Pike, A. von Mikecz, T. Ruzicka, P. I. Bird, and H. F. Abts.** 2003. Hurpin is a selective inhibitor of lysosomal cathepsin L and protects keratinocytes from ultraviolet-induced apoptosis. Biochemistry **42:**7381-9.

366. **Werness, B. A., A. J. Levine, and P. M. Howley.** 1990. Association of human papillomavirus types 16 and 18 E6 proteins with p53. Science **248:**76-9.

367. **Whiteside, S. T., P. King, and S. Goodbourn.** 1994. A truncated form of the IRF-2 transcription factor has the properties of a postinduction repressor of interferon-beta gene expression. J Biol Chem **269:**27059-65.

368. **Whiteside, S. T., K. V. Visvanathan, and S. Goodbourn.** 1992. Identification of novel factors that bind to the PRD I region of the human beta-interferon promoter. Nucleic Acids Res **20:**1531-8.

369. **Whyte, P., H. E. Ruley, and E. Harlow.** 1988. Two regions of the adenovirus early region 1A proteins are required for transformation. J Virol **62:**257-65.

370. **Wilkinson, M. G., and J. B. Millar.** 2000. Control of the eukaryotic cell cycle by MAP kinase signaling pathways. Faseb J **14:**2147-57.

371. **Wilson, A. C., M. G. Peterson, and W. Herr.** 1995. The HCF repeat is an unusual proteolytic cleavage signal. Genes Dev **9:**2445-58.

372. **Wilson, D. S., and C. Desplan.** 1995. Homeodomain proteins. Cooperating to be different. Curr Biol **5:**32-4.

373. **Wu, F., and A. S. Lee.** 2002. CDP and AP-2 mediated repression mechanism of the replication-dependent hamster histone H3.2 promoter. J Cell Biochem **84:**699-707.

374. **Wysocka, J., and W. Herr.** 2003. The herpes simplex virus VP16-induced complex: the makings of a regulatory switch. Trends Biochem Sci **28:**294-304.

375. **Xing, R., A. K. Addington, and R. W. Mason.** 1998. Quantification of cathepsins B and L in cells. Biochem J **332 (Pt 2):**499-505.

376. **Xu, Y. H., N. Richert, S. Ito, G. T. Merlino, and I. Pastan.** 1984. Characterization of epidermal growth factor receptor gene expression in malignant and normal human cell lines. Proc Natl Acad Sci U S A **81:**7308-12.

377. **Yang, S. H., P. Shore, N. Willingham, J. H. Lakey, and A. D. Sharrocks.** 1999. The mechanism of phosphorylation-inducible activation of the ETS-domain transcription factor Elk-1. Embo J **18:**5666-74.

378. **Ye, J., U. P. Dave, N. V. Grishin, J. L. Goldstein, and M. S. Brown.** 2000. Asparagine-proline sequence within membrane-spanning segment of SREBP triggers intramembrane cleavage by site-2 protease. Proc Natl Acad Sci U S A **97:**5123-8.

379. **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.

380. **Yeyeodu, S., K. Ahn, V. Madden, R. Chapman, L. Song, and A. H. Erickson.** 2000. Procathepsin L self-association as a mechanism for selective secretion. Traffic **1:**724-37.

381. **Yoon, S. O., and D. M. Chikaraishi.** 1994. Isolation of two E-box binding factors that interact with the rat tyrosine hydroxylase enhancer. Journal of Biological Chemistry **269:**18453-62.

382. **Yudt, M. R., and J. A. Cidlowski.** 2001. Molecular identification and characterization of a and b forms of the glucocorticoid receptor. Mol Endocrinol **15:**1093-103.

383. **Zeng, R. W., E. Soucie, N. Sung Moon, N. Martin-Soudant, G. Bérubé, L. Leduy, and A. Nepveu.** 2000. Exon/intron structure and alternative transcripts of the CUTL1 gene. Gene **241:**75-85.

384. **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:**2355-2365.

385. **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:**2015-2021.

386. **Zhong, H., M. J. May, E. Jimi, and S. Ghosh.** 2002. The phosphorylation status of nuclear NF-kappa B determines its association with CBP/p300 or HDAC-1. Mol Cell **9:**625-36.

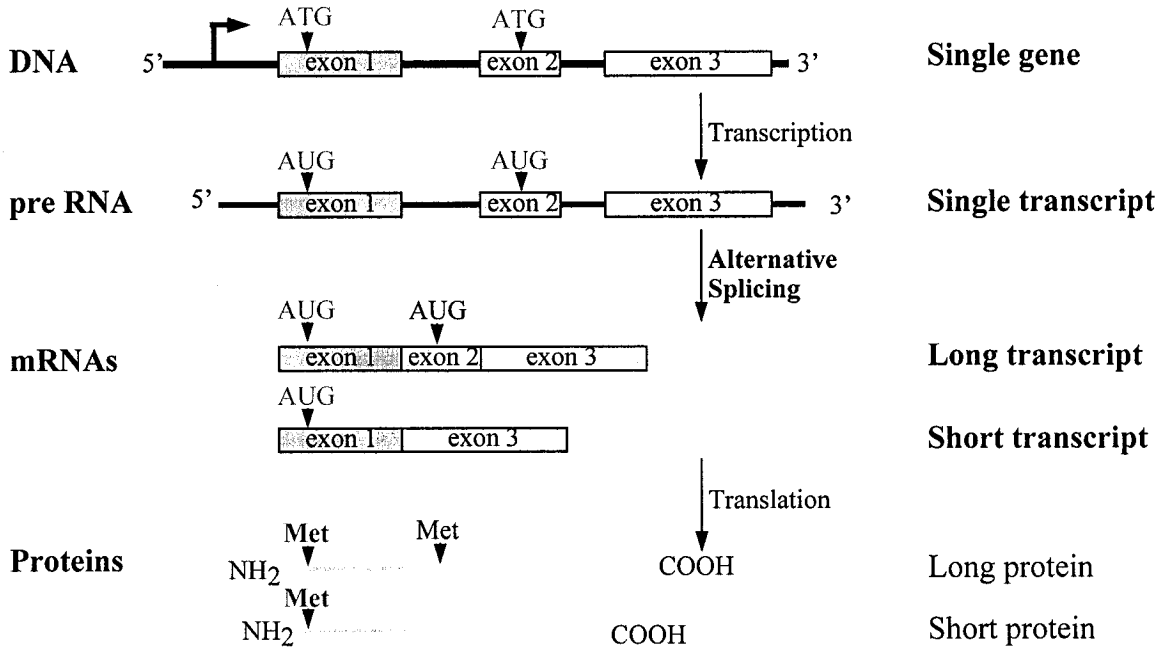
387. **Zhu, J., F. Shibasaki, R. Price, J. C. Guillemot, T. Yano, V. Dotsch, G. Wagner, P. Ferrara, and F. McKeon.** 1998. Intramolecular masking of nuclear import signal on NF-AT4 by casein kinase I and MEKK1. Cell **93:**851-61.

388. **zur Hausen, H.** 1989. Papillomaviruses in anogenital cancer as a model to understand the role of viruses in human cancers. Cancer Res **49:**4677-81.

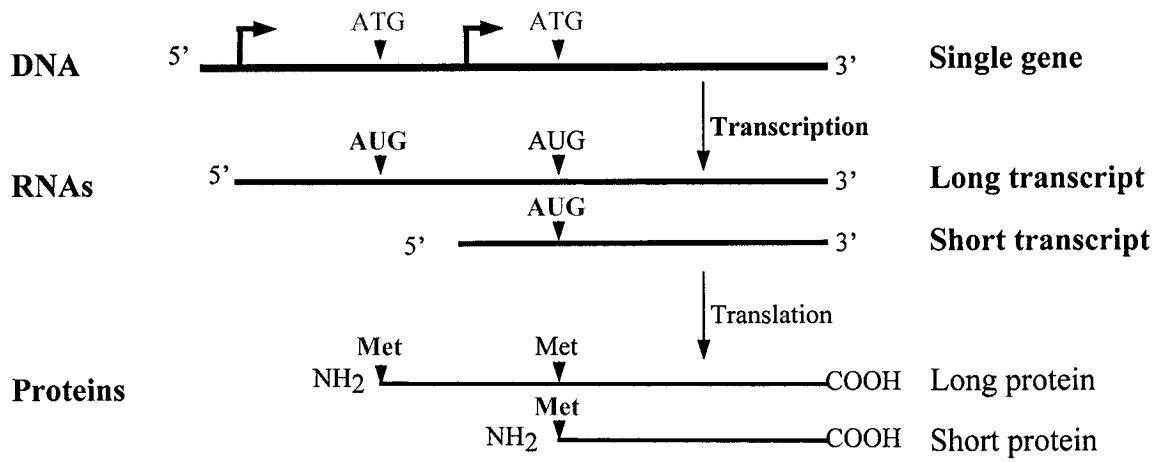
Figure 1: Alternative mechanisms to generate various protein isoforms from a single gene.

(A) Alternative splicing of the hnRNA (or pre-RNA) allows for the generation of mRNAs that are free of introns and contain various exons combinations. This can give rise to different transcripts that can be devoid of internal AUGs (as exemplified in this figure). In some cases, the expelled exon(s) could have encoded a localization signal, a DNA binding domain or a protein-protein interaction domain. (B) Multiple transcription initiation sites also yields transcripts of different sizes. This mechanism can be used to generate proteins with different properties, as mentioned above. It can also be employed to produce RNAs with less structured 5'UTR, or without upstream ORFs or initiating AUG (as shown in this figure). (C) Multiple translation initiation sites allows for the synthesis of proteins of various lengths with diverse properties from a single transcript. In this case, the transcript contains two (or more) in-frame translation initiation sites. The use of the alternative AUG can occur via leaky scanning, ribosome shunting or internal ribosomal entry site. It must be noted that these mechanisms presented are not mutually exclusive.

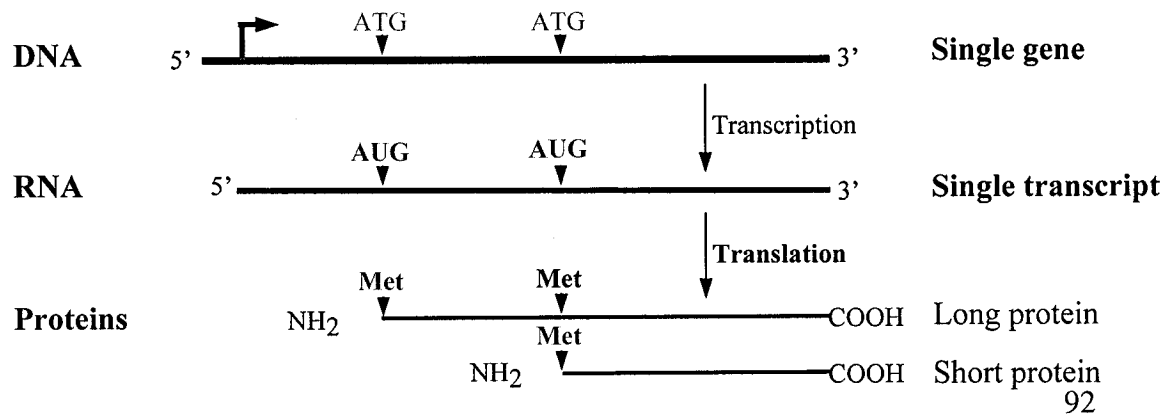
A- Alternative Splicing



B-Multiple Transcription Initiation Sites



C- Multiple Translation Initiation Sites

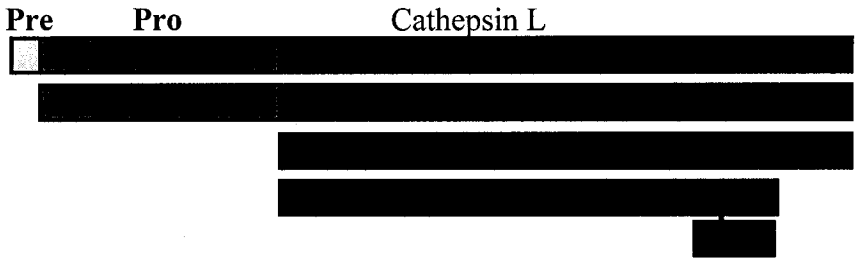


(Adapted from C.J. Danpure, 1995)

Figure 2: Schematic representation of cathepsin L and its intracellular route.

(A) Schematic of the pre-pro-enzyme cathepsin L. The pre region corresponds to the signal peptide that allows the translating protein to enter the endoplasmic reticulum (ER). The pro domain is a chaperone important for the folding of the protease and for inhibiting its activity in unwanted compartments. In a cell-type specific manner, carboxy-terminal cleavage of the mature cathepsin L also occurs. (B) During its translation, cathepsin L is directed within the ER by the signal peptide. Once inside, cathepsin L is glycosylated. In the golgi vesicules, the oligosaccharides are trimmed and obtain higher mannose residues. Once phosphorylated, they form the M6P recognition marker that binds to the MPR. The proenzyme is then targeted to a prelysosomes compartment, where the receptor dissociated from the enzyme and is recycled back to the golgi. In some cases, such as in transformed cells, cathepsin L can be directly secreted.

A


Pre
Pro
Cathepsin L

B

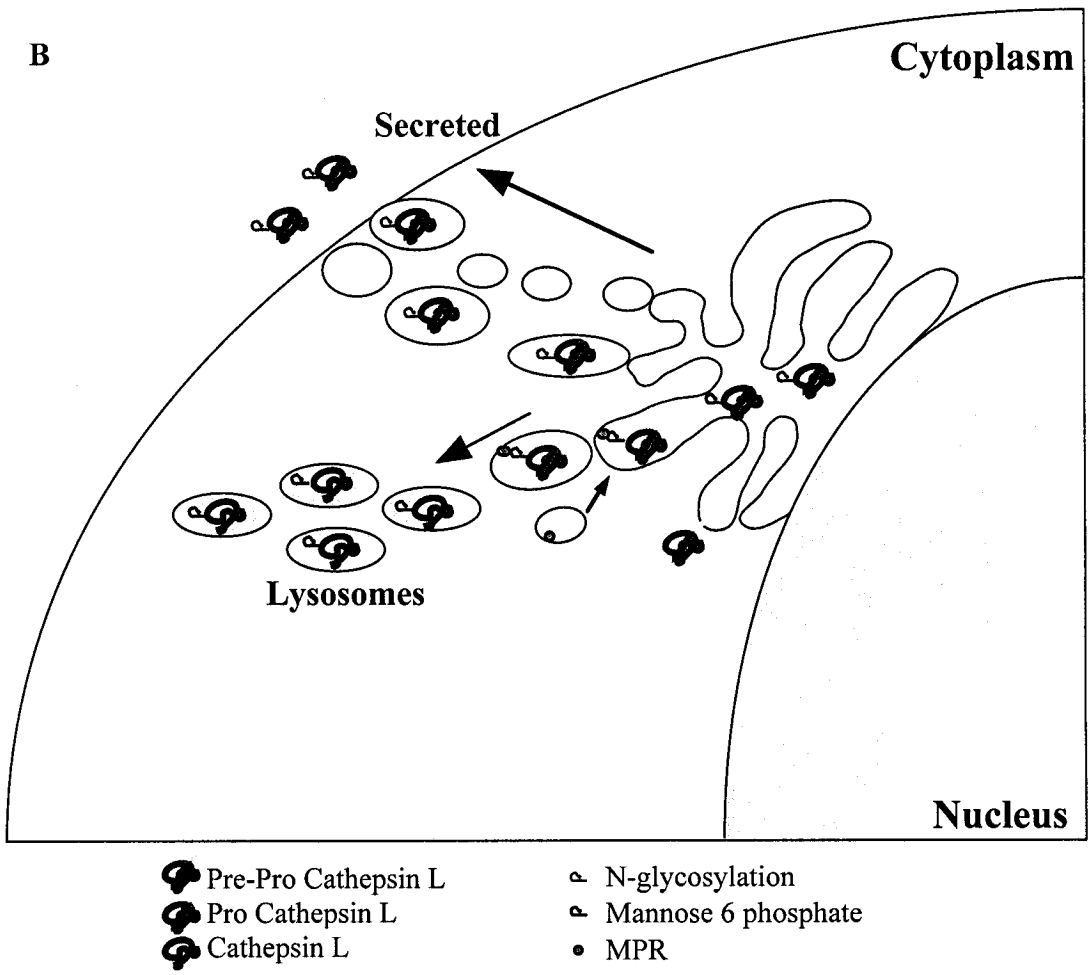
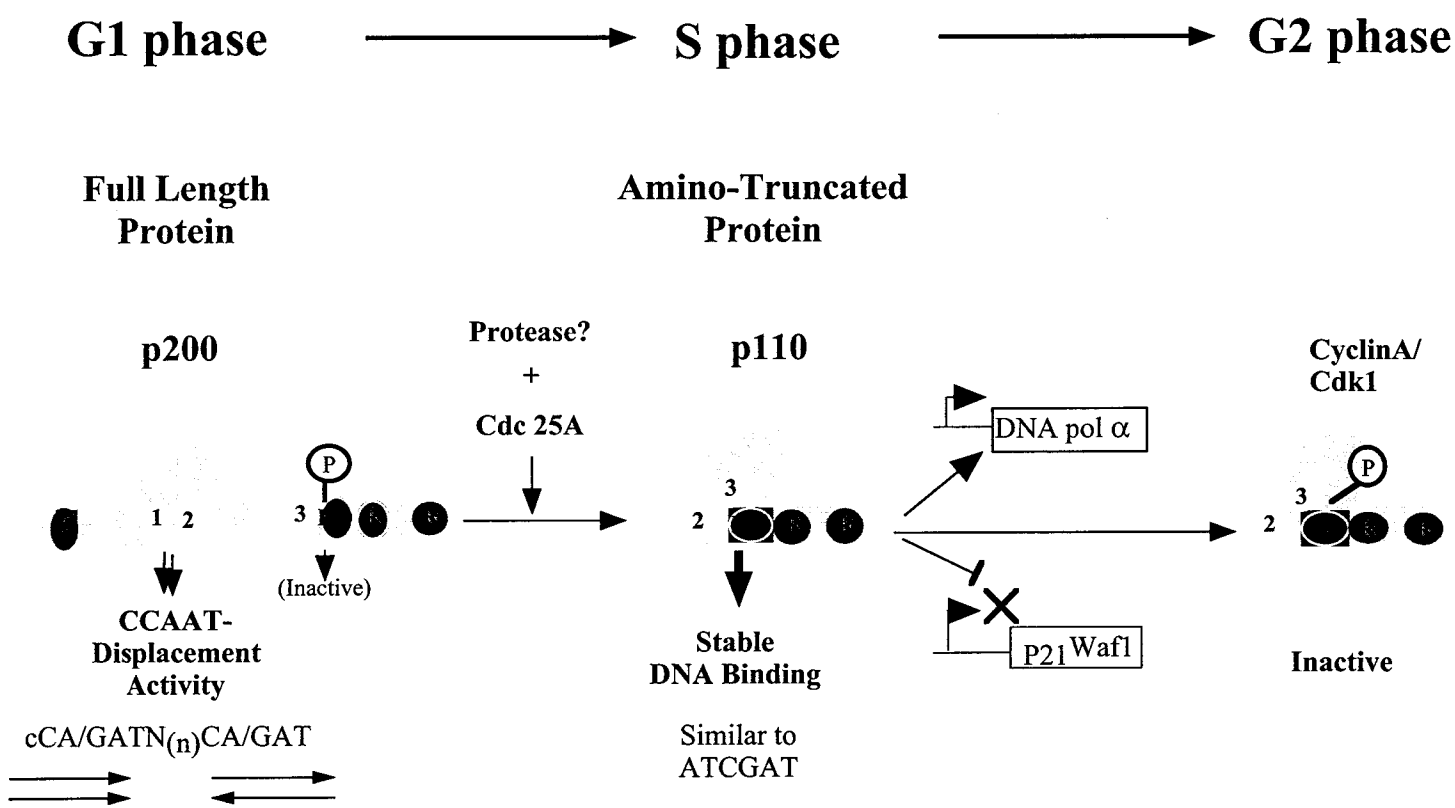

Cytoplasm
Secreted
Lysosomes
Nucleus
Pre-Pro Cathepsin L
Pro Cathepsin L
Cathepsin L
N-glycosylation
Mannose 6 phosphate
MPR

Figure 3: Alternative modes of DNA binding and regulation of CDP/Cux during the cell cycle.

The CCAAT displacement activity of the full length CDP/Cux (p200) involves the DNA binding domains CR1CR2, whereas stable DNA binding requires CR3HD. During the G1/S transition, the stable DNA binding activity is increased following two events. CDP/Cux is dephosphorylated by the Cdc25A phosphatase and is proteolytically processed to generate an amino-truncated isoform of 110kDa (p110). The p110 isoform can activate the *DNA polymerase alpha* gene promoter and repress that of the $P21^{Waf1/Cip1}$. At the G2 phase of the cell cycle, phosphorylation by cyclin A/Cdk1 of CDP/Cux inhibits its DNA binding activity.


G1 phase
S phase
G2 phase
Full Length Protein
Amino-Truncated Protein
p200
Protease?
+
Cdc 25A
p110
CyclinA/ Cdk1
DNA pol α
P
1 2
3
(Inactive)
CCAAT-Displacement Activity
cCA/GATN(n)CA/GAT
3
2
Stable DNA Binding
Similar to ATCGAT
P21Waf1
Inactive

Chapter II A Cathepsin L Isoform that Is Devoid of a Signal Peptide Localizes to the Nucleus in S phase and Processes the CDP/Cux Transcription Factor

Preface

This project originally focused on CDP/Cux expression and activity in cancer cells. My colleague, Dr. Nam Sung Moon, demonstrated that the full-length CDP/Cux transcription factor, p200, was proteolytically processed between CR1 and CR2 into an amino-truncated p110 isoform at the G1/S transition of the cell cycle. This p110 isoform, in contrast to p200, could bind stably to DNA and activate the *DNA polymerase alpha* gene promoter. In electrophoretic mobility shift assays (EMSA) p110 generates a faster migrating complex than p200. My initial experiments found that cancer cells displayed a more intense p110 retarded complex. Dr. Nam-Sung Moon also showed that proteolytic processing of CDP/Cux is tightly regulated in normal cells, whereas in my experiments, cancer cells display constitutive processing of p110. To better understand this phenomenon, subsequent investigations shifted focus to identify the protease that is responsible for the processing of CDP/Cux and to understand the importance of its altered regulation in cancer cells. While these studies were pursued concurrently, I chose to present the identification of the protease in the first chapter of results and in the next chapter, the characterization of CDP/Cux in cancer cells. These results will show that cathepsin L processing of CDP/Cux is activated in many cancer cells. In the last chapter of results I describe another mechanism that enables certain cancer cells to express more amino-terminally truncated CDP/Cux

isoforms. This mechanism takes place at the transcriptional level and involves the initiation of transcription at a different start site located within the gene itself. Although this work was published prior to the other two chapters, it will be presented last because the significance of these findings are clearer with an understanding of the overall regulation of CDP/Cux. When considered in a more general context, it thus becomes obvious that cancer cells will use alternate mechanisms to constitutively express short CDP/Cux isoforms.

In this chapter, cathepsin L is identified as the protease responsible for the proteolytic cleavage of CDP/Cux. As cathepsin L was known to be a lysosomal enzyme, its presence in the nucleus had to be ascertained. The mechanism for its translocation to the nucleus was also investigated. In addition, I provide functional evidence that p110 plays a key role in cell cycle progression by accelerating entry into S phase.

Abstract

The sub-class of cysteine proteases termed lysosomal cathepsins has long thought to be primarily involved in end-stage protein breakdown within lysosomal compartments. Furthermore, few specific protein substrates for these proteases have been identified. We show here that the 'lysosomal' cysteine protease, cathepsin L, functions in the regulation of cell cycle progression through proteolytic processing of the transcription factor CDP/Cux within the nucleus. Cathepsin L processes recombinant CDP/Cux to yield a functional protein fragment that can form a stable complex with DNA and modulate transcriptional activity. Forced expression of cathepsin L increased CDP/Cux processing and gene ablation of cathepsin L reduced CDP/Cux processing *in situ*. Furthermore, catalytically active cathepsin L is localized to the nucleus during the G1-S transition as detected by immunofluorescence imaging and labeling using activity based probes. Trafficking of cathepsin L to the nucleus is accomplished through a novel mechanism involving the synthesis of a truncated protease that is devoid of its signal peptide. Interestingly, forced expression of the processed CDP/Cux isoform is able to accelerate entry into S phase, and circumvent the G1-block that ensues following treatment of cells with the cysteine protease inhibitor, E64-d. Overall these results uncover a novel role for cysteine proteases in the control of cell cycle progression.

Introduction

Mammalian cysteine proteases of the papain family are targeted to the lumen of the endoplasmic reticulum via a signal peptide within their pro-domain and subsequently are glycosylated, phosphorylated and trafficked to the lysosomes (Chapman et al., 1997a). These proteases have long been thought to function exclusively in the terminal degradation of proteins in the lysosomes, but a number of recent observations have suggested that distinct cathepsins may play specific, non-redundant, physiological functions (reviewed in (Reinheckel et al., 2001; Turk et al., 2000). For example, specific phenotypes sometimes involving a particular tissue or cell-type have been observed following gene inactivation of different cathepsin genes. Cathepsin B has recently been shown to mediate protection from autolysis at the cell surface of cytotoxic T-cells and natural killer cells, and cathepsin L has been found to catalyze pro-hormone processing in secretory vesicles of neuroendocrine cells (Balaji et al., 2002; Yasothornsrikul et al., 2003). Taken together, these observations support the notion that cysteine proteases may also be involved in the regulatory processing of distinct protein substrates. The intriguing possibility that a cathepsin could be present in the nucleus was suggested from the findings that a nuclear serpin, MENT (Myeloid and Erythroid Nuclear Termination stage-specific protein), can efficiently inhibit a nuclear papain-like cysteine protease and cause a block in cellular proliferation (Irving et al., 2002). However, there is at present no direct evidence for the presence of a cathepsin in the nucleus.

The CDP/Cux/Cut transcription factors are a family of homeodomain proteins that are conserved among metazoans (reviewed in (Nepveu, 2001)). Genetic analysis in *Drosophila melanogaster* indicated that *cut* mediates phenotypic effects in a large number of tissues

(Bodmer et al., 1987; Braun, 1940; Jack et al., 1991; Jack, 1985; Liu et al., 1991). Inactivation of *cux-1* by gene targeting in the mouse has revealed several phenotypes including growth retardation, delayed differentiation of lung epithelia, altered hair follicle morphogenesis, male infertility and a deficit in T and B cells (Ellis et al., 2001; Luong et al., 2002; Sinclair et al., 2001). In contrast to the small size of *cux-1* knock-out mice, transgenic mice expressing *cux-1* displayed multi-organ hyperplasia and organomegaly (Ledford et al., 2002). Thus, from genetic studies both in *Drosophila* and the mouse, it is clear that the CDP/Cux/Cut gene plays an important role in tissue homeostasis in several organs.

A role for CDP/Cux in cell cycle progression, in particular at the G1/S transition, was inferred from a number of evidence. In particular, CDP/Cux binding to DNA increases at the G1/S transition (Coqueret et al., 1998; Holthuis et al., 1990; van Wijnen et al., 1991; van Wijnen et al., 1996; van Wijnen et al., 1989). In S phase, CDP/Cux was shown to repress the $p21^{WAF1/CIP1}$ gene and to stimulate expression of reporter plasmids carrying promoters from the *DNA pol α*, *carbamoyl-phosphate synthase/aspartate carbamoyltransferase/dihydroorotase* (*CAD*), and *cyclin A* genes (Coqueret et al., 1998; Moon et al., 2001; Truscott et al., 2003). Increase in DNA binding at the G1/S transition involves two regulatory events: dephosphorylation of the Cut homeodomain by the Cdc25A phosphatase, and a specific proteolytic cleavage (Coqueret et al., 1998; Moon et al., 2001). As cells progress into S phase, a fraction of CDP/Cux p200 molecules is proteolytically processed into N-terminally truncated proteins of 110 kDa. While CDP/Cux p200 only transiently binds to DNA and carries the CCAAT-displacement activity, p110 makes stable interaction with DNA (Moon et al., 2000; Moon et al., 2001). Importantly, p110 but not p200 can stimulate expression of the endogenous of the *DNA pol α* gene (Truscott et al., 2003). Thus, cell cycle-dependent processing of CDP/Cux serves to generate a p110 isoform with distinct DNA

binding and transcriptional properties. However, the specific mechanisms and enzymes involved in proteolytic processing of CDP/Cux within the nucleus have remained largely unknown.

In the present study, we have used a number of strategies to identify the protease responsible for the S phase-specific proteolytic processing of CDP/Cux. Small molecule activity based probes (ABPs) were used to track the activation of papain family cysteine proteases during cell cycle progression. One cysteine protease, cathepsin L, localized to the nucleus at the G1/S transition and was capable of processing CDP/Cux *in vitro* and *in vivo* to produce the physiologically relevant p110 isoform. Mechanistic studies suggest that translation initiation at downstream AUG sites allows the synthesis of N-terminally truncated cathepsin L isoforms that are devoid of a signal peptide and are targeted to the nucleus as cells progress into the cell cycle. These findings uncover a hitherto unsuspected function for the 'lysosomal' cysteine protease cathepsin L and suggest a novel mechanism for trafficking of this protease to the nucleus.

Results

Protease inhibitors directed against cathepsins inhibit CDP/Cux processing. As a first step to identify the protease responsible for CDP/Cut processing, we tested the effect of various protease inhibitors on the expression and activity of the CDP/Cux processed isoform, p110. NIH3T3 cells were transfected with a vector, MCH, expressing the full length CDP/Cux protein with Myc and HA epitope tags at its amino- and carboxy-termini, respectively. Transfected cells were treated with various protease inhibitors for four hours prior to harvesting. Nuclear extracts were tested by Western blot analysis with an antibody that recognizes the carboxy-terminal epitope, and in electrophoretic mobility shift assay with a probe specific for the CDP/Cux p110, as previously described (Moon et al., 2001). Based on steady-state levels of CDP/Cux p110 (Fig. 1A) and the reduction in the intensity of the retarded complex (Fig. 1B), processing was inhibited by MG132 and E-64-D, and to a minor extent by ALLN, cathepsin inhibitor I and chloroquine. In contrast, lactacystin, PD150606 and EGTA had no effect, suggesting that neither the proteasome nor calpain were involved. Thus, this pattern of inhibition suggested that a cathepsin-like protease is involved in the processing of CDP/Cux.

Forced expression of CDP/Cux p110 overcomes a G1-block imposed by a cysteine protease inhibitor. Treatment of cells with a cysteine protease inhibitor was previously reported to prevent their progression into S phase (Mellgren, 1997). The molecular basis for this G1 block is still unclear. Since processing of the CDP/Cux transcription factor was shown to take place at the G1/S transition, we tested whether forced expression of the processed CDP/Cux isoform would be able to circumvent the block imposed by a cysteine protease inhibitor (Moon et al., 2001). Populations of NIH3T3 stably carrying an empty vector

or a vector expressing the 878-1505 CDP/Cux protein were generated by retroviral infection. Cells were synchronized into G0 by serum starvation for 3 days, and then stimulated to re-enter into the cell cycle by adding fresh medium with 10% serum. At the indicated times, cells were harvested and their cell cycle distribution was analyzed by FACS analysis following staining of the DNA with propidium iodide. Cells expressing CDP/Cux 878-1505 started to enter in S phase at 14 hour following serum addition, whereas the control population of cells entered in S phase at 16 hours (Fig. 1C and D). Thus, expression of CDP/Cux 878-1505 shortens the G1 phase coming out of quiescence. In the presence of the cysteine protease inhibitor E-64d, control cells did not progress into S phase at 16 hours, and at 18 hours only 8% of these cells had progressed into S phase. These results confirm that a cysteine protease inhibitor can cause a delay in G1 (Fig 1E). Interestingly, cells expressing CDP/Cux 878-1505 progressed into S phase at 14 hours, and at 16 and 18 hours a large fraction of these cells had progressed into S phase (Fig. 1F). These results indicate that CDP/Cux 878-1505 can overcome the G1 block that is imposed by inhibition of cysteine proteases.

Cathepsin L can proteolytically process CDP/Cux *in vitro*. Proteolytic processing was previously mapped to the linker region in between the CR1 and CR2 domains of CDP/Cux, and a recombinant protein containing these domains fused to a nuclear localization signal was shown to be processed following transfection into NIH3T3 cells (Moon et al., 2001). To verify whether a cathepsin cysteine protease could cleave CDP/Cux, purified CR1+CR2 recombinant protein was incubated *in vitro* with a panel of cathepsins. Cleaved peptides were rapidly generated following incubation with recombinant cathepsin L, K, F and V (Fig. 2A). Since CDP/Cux is expressed in all tissues and is processed in a similar manner in cells of various origins (Moon et al., 2001), it is likely that the protease responsible for its

processing is also ubiquitously expressed. Among the ubiquitously expressed cathepsins that were able to cleave CR1+CR2 *in vitro*, cathepsin L was the most efficient. We therefore verified whether CR1+CR2 was cleaved similarly following transfection into NIH3T3 or incubation *in vitro* with cathepsin L. The same vector, pTri-Ex-2, served for the expression of CR1+CR2 in bacteria and in mammalian cells. As seen in Fig. 2B, the CR1+CR2 peptides generated by cathepsin L *in vitro* displayed similar mobility on SDS-PAGE as the primary peptides produced following transfection into Ras-FR3T3 cells (compare lane 1 with lanes 4-6). N-terminal sequencing of the cleaved peptides revealed that processing occurred at three positions *in vitro*, following residues Q^{643}, S^{747} and S^{755} (Fig. 2D). To assess whether cathepsin L could cleave CDP/Cux at neutral pH, reactions were carried at various pH using a larger recombinant CDP/Cux substrate (Fig. 2C). Although cleavage was more efficient at pH5.5, substantial cleavage was still observed at pH 7.0 and 7.5.

The steady-state level of the CDP/Cux processed isoform correlates with cathepsin L expression. Co-expression of cathepsin L with Myc-Cux-HA led to an increase in the steady-state level of CDP/Cux p110 both in NIH3T3 (Fig. 3A, compare lane 1 with lanes 2 and 3) and CV-1 cells (Fig. 3A, compare lane 4 with lanes 5 and 6). Addition of E-64d inhibited, at least partially, the production of p110 in cells (Fig. 3A, compare lane 7 with 6). Importantly, co-transfection with a catalytically inactive mutant of cathepsin L, Cat L^{C25S}, did not lead to an increase in CDP/Cux p110 (Fig. 3A, compare lane 11 with 9 and 10). Proteolytic processing of Myc-Cux-HA was then assessed in mouse embryo fibroblasts (MEF) derived from the cathepsin L knockout (Roth et al., 2000). Transfections were carried in parallel in three types of MEF cells: Cat $L^{-/-}$, Cat $B^{-/-}$., and wild type. Proteolytic processing of Myc-Cux-HA was much less efficient in Cat $L^{-/-}$ MEF (Fig. 3B, compare

lanes 2, 3 and 4 with lanes 5 and 6). Yet, a faint band corresponding to p110 was still observed in Cat $L^{-/-}$ MEF (Fig. 3B, lanes 2-4). Altogether, results from gain- and loss-of-function analysis are consistent with the notion that cathepsin L is responsible for the proteolytic processing of CDP/Cux. These results do not, however, exclude the possibility that other proteases can also proteolytically cleave CDP/Cux, as demonstrated from the presence of low amount of CDP/Cux processed isoforms in Cat $L^{-/-}$ MEF cells (Fig. 3B, lanes 2-4).

The cathepsin L heavy chain is present at higher level in the nuclear fraction of NIH3T3 cells that are enriched in S phase. Cathepsin L has been characterized as a protease that localizes to the lysosomes (Erickson, 1989; McGrath, 1999). However, some early data from immunofluorescence studies with anti-cathepsin L antibodies revealed a weak signal within or at the surface of the nucleus (see Fig. 2 in (Portnoy et al., 1986)). We therefore investigated whether a fraction of cathepsin L could be found in the nucleus. As a first approach, we used small molecule activity based probes (ABPs) to radioactively label active papain-family cathepsins directly within cells (Bogyo et al., 2000). Since expression of CDP/Cux p110 in NIH3T3 cells was previously shown to increase in S phase, the experiment was performed using populations of NIH3T3 cells synchronized at different points of the cell cycle (Moon et al., 2001). To help determine which of the many ABP-associated signals in NIH3T3 cells may correspond to cathepsin L, the experiment was performed in parallel with Cat $L^{-/-}$ MEF cells. NIH3T3 cells and Cat $L^{-/-}$ MEF were serum-starved and then re-stimulated to enter the cell cycle. At the indicated time, cells were labeled for one hour with the ABP, ^{125}I-JPM-OET, after which cytoplasmic and nuclear fractions were prepared and analyzed by SDS-PAGE and autoradiography (Fig. 4). As ex-

pected, a number of bands were observed in the cytoplasmic fraction of both NIH3T3 and the Cat $L^{-/-}$ MEF cells (Fig. 4A). Interestingly, in the nuclear fractions, two bands were present in the samples from NIH3T3 cells but were absent in the samples from Cat $L^{-/-}$ MEF cells (Fig. 4B). To verify whether these bands indeed represented cathepsin L, samples from the 14h time points were analyzed by immunoprecipitation with an anti mouse cathepsin L antibody prior to analysis by SDS-PAGE. Two bands corresponding to the single and heavy chains of cathepsin L were present in the nuclear fraction of NIH3T3 cells, but were not observed in the nuclear fraction of Cat $L^{-/-}$ MEF cells (Fig. 4C). Thus, active cathepsin L can be detected in the nuclear fraction of NIH3T3 cells, and the activity of cathepsin L in the nuclear fraction increased with the progression of cells into S-phase.

Indirect immunofluorescence demonstrates that the sub-cellular localization of cathepsin L varies during the cell cycle. In order to analyze the sub-cellular localization of cathepsin L, indirect immunofluorescence was performed on NIH3T3 cells using antibodies raised against murine cathepsin L (Chapman et al., 1997b). Essentially two staining patterns were observed in continuously growing NIH3T3 cells (Fig. 5B). Approximately 80% of the cells exhibited a strong cytoplasmic signal consistent with lysosomal localization of cathepsin L (Erickson, 1989; McGrath, 1999). In 20% of the cells, a signal was also observed over the nuclear region. The observation of different staining patterns in the same population of cells suggested that the sub-cellular localization of cathepsin L may vary throughout the cell cycle. To verify this hypothesis, NIH3T3 cells were synchronized by serum starvation-stimulation. Throughout the cell cycle, cells were analyzed by indirect immunofluorescence and cell cycle distribution was monitored by FACS analysis after staining with propidium iodide. The most frequent staining pattern for each time point is illus-

trated in panel A, while the proportion of each staining pattern is provided in panel B after examination of at least 200 cells at each time point (Fig. 5A and B). We observed an increase in the proportion of cells with a nuclear staining pattern at 12h, 13h, and 14h following serum addition when a large fraction of cells progressed through the S and G2 phases of the cell cycle. Synchronization using an alternate method, the double thymidine block procedure, confirmed that a nuclear cathepsin L staining was increased in populations of cells enriched in S phase (data not shown; see also Fig. 5D). To verify whether the signal over the nuclear area originated from the surface of or from within the nucleus, serial images at various depths were obtained by confocal microscopy (Fig. 5C). The results unequivocally show localization of the signal within the nucleus. To ensure that the distinct nuclear foci were due to authentic cathepsin L protein and not to an unexpected antibody cross-reaction, we examined the localization of cathepsin L with two other polyclonal cathepsin L antibodies directed against distinct regions of the cathepsin L protein (see Material and Methods). All anti-cathepsin L antibodies revealed an S phase-specific nuclear staining pattern (data not shown).

To unambiguously verify whether cathepsin L can localize to the nucleus, we engineered a vector, Cat L-HA, expressing cathepsin L with the influenza virus hematoglutinin (HA) epitope tag inserted at position 216 of the cathepsin L protein sequence (accession number XM_122577). The crystal structure of cathepsin L predicted that an epitope tag inserted at this position would reside within an external loop of the mature cathepsin L protein (Coulombe et al., 1996). We have confirmed that cathepsin L with this tag targets to endosomes and is secreted similarly to wt protein (A. E., unpublished observations). Moreover, the Cat L-HA protein expressed in the yeast *Pichia Pastoris* was found to be enzymatically active (see below and Fig. 7C). NIH3T3 cells were transfected with the Cat L-HA

expression vector, synchronized by double thymidine block or serum starvation/stimulation, and were analyzed by indirect immunofluorescence using an anti-HA antibody. NIH3T3 cells transfected with the Cat L-HA vector showed some cells that were visible in phase contrast that did not display any signal with the anti-HA antibody (Fig. 5D). We surmise that cells with an HA signal correspond to transfected cells, whereas cells with no HA signal represent the untransfected cells present in the population. In agreement with the results obtained with the endogenous cathepsin L, a proportion of the cells expressing Cat L-HA displayed a signal in the nucleus, and the proportion of cells with a nuclear signal increased following enrichment in S phase with double thymidine block or serum stimulation (Fig. 5D). That the signal over the nuclear area was truly intra-nuclear was demonstrated from scanning confocal microscopy (Fig. 5E). Thus, the results with 3 different polyclonal antibodies against cathepsin L, and with the Cat L-HA recombinant protein revealed that cathepsin L can localize to the nucleus, and that this process is regulated in a cell cycle dependent manner.

Translation can initiate at internal start sites within the cathepsin L mRNA. We investigated a number of mechanisms that could account for the presence of cathepsin L in the nucleus. Interfering with lysosome targeting either pharmacologically or by introducing a point mutation in the cathepsin L coding sequence to prevent glycosylation increased cathepsin L secretion but did not affect nuclear targeting, in agreement with earlier reports (Chapman et al., 1997b). We then considered the possibility that an alternatively spliced or initiated cathepsin L mRNA may encode a protein devoid of the signal peptide, however, RNase mapping and reverse-transcriptase polymerase chain reaction (RT-PCR) analyses with various oligonucleotide primers failed to reveal the existence of novel cathepsin L mRNA (data not shown). Moreover, expression of cathepsin L mRNA did not vary during

the cell cycle (data not shown). These results were in agreement with our earlier findings showing a nuclear signal upon transfection of a full-length cDNA clone expressing Cat L-HA (Fig. 5D and E). Interestingly, using the cathepsin L cDNA in an *in vitro* transcription/translation system, we detected two bands of faster mobility that could correspond to cathepsin L proteins generated as the result of translation initiation at sites downstream from the first AUG start codon (Fig 6A, lane 2). To verify this hypothesis, we performed *in vitro* mutagenesis to remove putative translation initiation sites from the cathepsin L cDNA. Interestingly, the 5' portion of the murine cathepsin L cDNA contains 7 AUG codons, corresponding to methionine 1, 56, 58, 75, 77, 83 and 111. This represents a higher than usual concentration of methionine codons. Mutation of the first methionine codon from AUG to UUC (M1) prevented expression of the slowest species, confirming that this band correspond to the full-length cathepsin L (Fig 6A, lane 3). Yet, the two faster migrating species were still produced by this mutant, indicating that translation indeed could start at downstream positions. Replacement of the methionine codons 56 and 58, 75 and 77 with UUC greatly reduced expression of these shorter species, however, a weak band could still be observed (Fig 6A, lane 4). Additional replacements at methionine codons 83 and 111 caused the disappearance of this band, but new faster species appeared (Fig. 6A, lanes 5-6). However, in the context of a cDNA in which the methionine codon 1 was left intact, replacement of methionine codons 56, 58, 75, 77 and 83 was sufficient to eliminate expression of shorter Cathepsin L species (Fig. 6A, lanes 7-9). These results demonstrated that translation could start *in vitro* at several downstream AUG initiation sites within the cathepsin L mRNA, thereby allowing the synthesis of cathepsin L with distinct amino-termini.

Following transfection of the Cat L-HA and M1-HA plasmids into NIH3T3, we observed cathepsin L species of similar electrophoretic mobility as that seen in the *in vitro* system (Fig 6B, compare lanes 1-2 with lanes 4-5). These results demonstrated that truncated cathepsin L proteins are also present in cells upon forced expression of the cathepsin L mRNA. We then verified whether short cathepsin L proteins are expressed from the endogenous gene. Following *in vivo* labeling with ^{35}S-methionine and immunoprecipitation with Cathepsin L antibodies, we detected a faint band, which co-migrated with a recombinant short Cathepsin L species (Fig 6B, compare lane 11 with lanes 8, 9 and 10). Altogether these results suggest that translation *in vivo* can start at downstream initiation sites within the endogenous Cathepsin L mRNA.

N-terminally truncated cathepsin L species are excluded from a membrane fraction *in vitro* and exhibit a distinct subcellular localization *in vivo*. Cathepsin L proteins that start at amino acids 56, 58, 75 or 77 would not include a signal peptide, and consequently should be excluded from the endoplasmic reticulum. To verify this hypothesis, *in vitro* translation was performed in the presence of a membrane fraction. Proteins were then fractionated to distinguish proteins that are internalized during translation from those proteins that are not inserted into a membrane. As expected, the full-length cathepsin L was found mostly if not exclusively in the membrane fraction (Fig 6C, lanes 2 and 3). In contrast, the shorter cathepsin L produced by the mutant M1 were not associated with the membrane fraction and were found instead in the supernatant (Fig 6C, lanes 5-6). To compare the subcellular localization of the full-length and short cathepsin L isoforms, NIH3T3 cells were transfected with either wild type or mutated Cat L-HA vectors and indirect immunofluorescence was performed using anti-HA antibodies. Only a fraction of cells on any given slide generated a signal, indicating that the signal obtained with the anti-HA antibody originated

from transfected cells (data not shown). As seen previously, cells transfected with the wild type Cat L vector displayed two different staining patterns with 20% of cells showing a distinct nuclear signal (Fig. 6D). Remarkably, a strong nuclear signal was never observed in cells transfected with the mutant M56-111-HA that expresses only the full-length cathepsin L isoform (Fig. 6D). In contrast, staining within the nucleus was observed in all cells transfected with the M1-Cat L-HA vector (Fig 6D). To confirm that nuclear staining was due to the expression of N-terminally truncated cathepsin L species, we engineered a vector in which the 5' end of the Cat L-HA cDNA was truncated up to nucleotide 247, thereby deleting the first AUG as well as the first 47 codons. In the resulting clone, the cathepsin L open reading frame was predicted to start at the AUG site corresponding to codon 56 of the wild type cathepsin L mRNA. This vector was therefore called AUG-56. 100% of the cells transfected with this vector displayed a strong nuclear staining pattern, thereby confirming that N-terminally truncated cathepsin L species can localize to the nucleus (Fig 6D).

N-terminally truncated cathepsin L species stimulate the processing and transcriptional activity of CDP/Cux. The distinct sub-cellular localization of the cathepsin L mutants suggested that only the N-terminally truncated cathepsin L proteins could co-localize with CDP/Cux and modulate its activity. Indeed, co-expression of either wild type or M1 Cat L-HA in NIH3T3 cells stimulated processing of the Myc-Cux-HA protein (Fig. 7A, compare lanes 3-4 and 5-6 with lane 2). In contrast, increased processing was not observed with the M56-111-HA mutant that does not localize to the nucleus (Fig. 7A, compare lane 9 with lanes 7 and 8). Overall, stimulation of CDP/Cux processing correlated with the expression of shorter cathepsin L-HA isoforms as seen in Fig. 7B. To confirm that the HA-tagged cathepsin L was enzymatically active, a full-length cathepsin L-HA protein was ex-

pressed in the yeast *Pichia Pastoris*, as previously described (Menard et al., 1998). A cleaved peptide of the expected molecular mass was generated upon incubation of the purified CR1+CR2 recombinant protein with medium from a Cat L-HA transformed *Pichia* clone (Fig. 7C, compare lane 3 with lanes 1 and 2). Processing of CR1+CR2 was inhibited in the presence of the cathepsin inhibitor E-64d (Fig. 7C, lane 4), and no processing was observed when the medium was taken from the original, untransformed, yeast (Fig. 7C, lane 5). Altogether, these results indicate that the HA-tagged cathepsin L enzyme is catalytically active, and that N-terminally truncated cathepsin L-HA molecules can localize to the nucleus and cleave CDP/Cux.

Processing of CDP/Cux p200 was previously shown to generate an isoform, p110, that displays distinct DNA binding properties (Moon et al., 2001; Truscott et al., 2003). We therefore tested whether co-expression of cathepsin L with CDP/Cux p200 could change its biochemical activities. When wild type cathepsin L was co-expressed with CDP/Cux p200, we observed an increase in the steady state level of CDP/Cux p110 (Fig. 7D, lane 3), and in the corresponding retarded complex as seen in electrophoretic mobility shift assay (Fig. 7D, lane 7). Similarly, co-expression of an N-terminally truncated cathepsin L species, Cat L^{AUG56}, also led to an increase in CDP/Cux p110 expression and DNA binding activity (Fig. 7D, lanes 12 and 17). In contrast, the M56-111 cathepsin L mutant that expresses only the full-length cathepsin L isoform was unable to stimulate CDP/Cux p110 expression and DNA binding activity (Fig. 7D, lanes 13 and 18). In summary, our results demonstrate that translation initiation at downstream start sites within the cathepsin L mRNA allows the synthesis of Cathepsin L isoforms that are devoid of a signal peptide (Fig. 6A and B), do not transit through the endoplasmic reticulum (Fig. 6C), can localize to the nucleus (Fig.

6D), cleave the CDP/Cux transcription factor (Fig. 7A) and modify its DNA binding properties (Fig. 7D).

Discussion

Cell cycle research in the last decade has revealed a crucial role for the proteasome in the regulated transition from one phase of the cell cycle to another. Expression of many cell cycle regulatory proteins is controlled in large part by their degradation (DeSalle and Pagano, 2001; Hershko, 1997; Peters, 2002). In addition to this broad role of proteolysis in establishing novel patterns of protein expression, our results indicate that limited proteolysis or proteolytic processing can serve in a more subtle way to modulate the biochemical properties of specific proteins in the nucleus. In the case of the CDP/Cux transcription factor, proteolytic processing serves to generate an isoform that is capable of making a stable interaction with DNA and engage in distinct regulatory functions (Moon 01, Truscott 03). In particular, CDP/Cux p110, but not p200, was found to stimulate expression of reporters carrying the promoters from the *DNA pol α*, *CAD* and *cyclin A* genes (Moon et al., 2001; Truscott et al., 2003). In addition, we showed in the present study that CDP/Cux p110 was able to accelerate the G1 to S transition following exit from quiescence (Fig. 1). This property of CDP/Cux p110 is not limited to experimental conditions of serum starvation/re-stimulation, but was also observed following a nocodazole-block in G2/M and in continuously proliferating cells fractionated by cell elutriation (Sansregret et al., submitted). Thus, proteolytic processing of CDP/Cux is likely to represent an important regulatory step at the onset of S phase.

The presence of active cathepsin L in the nucleus raises a number of questions regarding its processing and activity outside of the ER and lysosomal compartments. As a signal peptide is present at the N-terminus of cathepsin L, the protein should be targeted to the endoplasmic reticulum during its synthesis. However, our results revealed the existence of N-

terminally truncated cathepsin L isoforms that are devoid of the signal peptide. The presence of faster migrating species that are generated at low efficiency from the cathepsin L mRNA in the *in vitro* transcription/translation system as well as in cells suggested that smaller cathepsin L isoforms may be expressed (Fig. 6). Mutational analysis indicated that these smaller cathepsin L isoforms were the products of translation initiation at downstream AUG sites by a mechanism involving either leaky scanning or internal ribosome entry (Fig. 6). Importantly, only N-terminally truncated cathepsin L isoforms were able to translocate to the nucleus and stimulate CDP/Cux processing (Fig. 6). In contrast, the replacement of 7 internal AUG codons with alanine codons lead to the expression of a full-length cathepsin L isoform that did not generate a signal in the nucleus and could not stimulate CDP/Cux processing (Fig. 6). Therefore, we conclude that translation initiation at downstream AUG sites within the cathepsin L mRNA is the first requirement in the chain of events that lead to the presence of active cathepsin L in the nucleus.

A major obstacle to the correct folding and processing of N-terminally truncated cathepsin L outside of the endoplasmic reticulum and lysosomes presumably would be the neutral pH and different redox-potential of the cytoplasmic and nuclear milieu. While most studies on this subject have focused on activation of cathepsin L inside of the lysosome, proper maturation of this protease elsewhere is not impossible in light of the demonstrated autocatalytic processing of cathepsin L *in vitro* (Menard et al., 1998). In this study, autocatalytic processing required that the pH be reduced to 5.3, a situation that is unlikely to be found in the cytoplasm or the nucleus. However, we speculate that the truncated prodomain of shorter cathepsin L species does not bind as tightly to the active groove, thereby allowing autocatalytic processing at higher pH. In support for this notion, a truncated form of cathepsin B that lacks the signal peptide and part of the propeptide (52 residues) was

shown to fold properly and become active (Mehtani et al., 1998). As the cytosolic chaperone machinery is recruited while the nascent protein is still on the ribosome, we envisage the additional possibility that cytosolic chaperones somehow serve as substitutes for the proregion (Frydman, 2001). Furthermore, recent evidence suggests that a redox potential might be available in the cytosol, as members of the protein disulfide isomerase (PDI) family were found in non-ER locations, including the cytosol and the nucleus (reviewed in (Turano et al., 2002)). In relation to this, a number of studies have also raised the intriguing possibility that cell cycle progression from the G1 to the S phase may be associated with overall cellular (or nuclear) metabolic state/redox status (Menon et al., 2003; Zheng et al., 2003).

We envisage that cathepsin L isoforms that are devoid of a signal peptide are released into the cytoplasm and eventually make their way to the nucleus by a mechanism that remains to be defined. In this regard, examination of the cathepsin L sequence revealed the presence of a lysine-rich sequence that resembles the canonical nuclear localization signal. However, extensive mutation of this sequence did not prevent nuclear targeting in cells that were synchronized in S phase (data not shown). This finding leads us to speculate that cathepsin L molecules associate with a partner protein in order to travel to the nucleus.

Can cathepsin L be a functional enzyme at neutral pH? Since optimal enzymatic activity by cathepsins is achieved at a pH 5.5 (Turk et al., 1997), we can reasonably assume that cathepsins might only be weakly active in the nuclear milieu (Turk et al., 1999; Turk et al., 1993). Yet, using partially purified cathepsin L from two independant sources, rat liver and the yeast Pichia pastoris, we have shown that cathepsin L is able to process a CDP/Cux substrate at neutral pH, although a reduction in activity was observed at pH 7.5 relative to pH 5.5 (Fig. 2C and data not shown). We emphasize that limited cathepsin L activity in the

nucleus is entirely compatible with a role in the proteolytic processing of specific nuclear proteins. In contrast, optimal activity of cathepsins in the acidic environment of the lysosome is necessary for the terminal degradation of proteins. Thus, the sub-optimal pH that prevails in the nucleus should not be taken as an obstacle but rather as an important element that enables cathepsin L to play a role in the limited proteolysis of nuclear proteins. Future studies should investigate the possibility that other cysteine proteases might be present in the nucleus. In agreement with this notion, we observed a small amount of the CDP/Cux p110 isoform in cat L-/- MEF cells (Fig. 3B). It is possible that redundant protease activities co-exist in the nucleus in a somewhat constitutive manner; alternatively, gene inactivation of cathepsin L may lead to the compensatory induction of other proteases like cathepsin F for example. We note that the coding sequences of cathepsin F and V contain, like cathepsin L, several AUG codons that are located downstream of the first AUG codons.

We consider the possibility that some of the phenotypes already described for the various cathepsins mutant mice might involve a biochemical activity distinct from that which has been characterized in the lysosomes. For example, we note that the CDP/Cux knockout, the cathepsin L knockout and the furless mutant display strikingly similar phenotypes. Indeed, each of these mutant mice exhibits smaller size at birth, a high rate of mortality early in life and defective hair follicle cycle. In the case of CDP/Cux, this phenotype was not entirely surprising in light of the documented role of Drosophila *cut* in the development of bristles (Blochlinger et al., 1991; Bodmer et al., 1987). In contrast, it was not immediately apparent why inactivation of cathepsin L would lead to a tissue-specific defect in the hair follicle (Benavides et al., 2002; Roth et al., 2000). It was suggested that cathepsin L may be involved in the activation of extracellular degrading metalloproteinases, or in the prote-

olytic processing of growth factors or their receptors, in a way that would modulate the balance between proliferation and differentiation in the skin (Roth et al., 2000). Although we cannot exclude that the molecular and cellular basis for the altered hair follicle cycling may be entirely different for the cat L-/- and cux-/- mutant mice, we deem worthy of investigation the notion that the role of cathepsin L in the control of proliferation and differentiation involves the proteolytic processing of some nuclear proteins.

Acknowledgements

We are grateful to Dr. Mirek Cygler for advice on the structure of cathepsin L, To Dr. Robert Ménard for advice on *Pichia pastoris* and to Dr. Eric Chevet for the canine pancreatic microsomal membranes and technical advice. A.N. is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec. B. G. is the recipient of a fellowship from the Research Institute of the McGill University Health Center. This research was supported by grants #MT-11590 from the Canadian Institute of Health Research of Canada and #014288 from the National Cancer Institute of Canada to A.N. and by The Sandler Family Program in Basic Sciences to M.B.

Experimental Procedures

Cell culture. NIH3T3, CV-1, MEF and MEF Cat B -/- cells were cultured in DMEM medium supplemented with 10%FBS. MEF Cat L -/- cells were cultured in RPMI medium supplemented with 10%FBS. To obtain cells in G0 using the serum starvation/stimulation method, NIH3T3 cells were maintained for 2.5 days in DMEM (0h). The medium was then changed for DMEM plus 10% FBS and cells were harvested at different time points. Synchronization in G1/S was also performed using the double thymidine procedure (55). FACS analysis was performed as previously described (32). Transfections were done using ExGene500 (MBI Fermentas) or GeneJuice (Novagen) according to the manufacturer's instructions. When specified, various cell-permeable protease inhibitors were added to cells 4 hours prior lysis (Most inhibitors were purchased from Calbiochem. ALLN, 20 μM; MG132, 20 μM; Lactacystin, 10 μM; E-64D (or EST), 20 μM (or 200 μM); PD501616 20 μM; EGTA, 1mM; Cathepsin Inhibitor I, 20 μM; and Chloroquine, 100 μM). For population of NIH3T3 cells over-expressing CDP/Cux 878-1505, 50 000 cells were first plate in a 6 well plate. pREV or pREV-878-1505 expressing viruses were added to the cells and the plates were spun at 300g for 1 h at 32°C. 48 h later, cells were trypsinized and replated at lower densities, and the next day, infected cells were selected with hygromycin (0.3 mg/ml) for 10 days.

Preparation of cell extracts. Nuclear extracts were prepared according to the procedure of Lee et al. (56) 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). Nuclei were then resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM MgCl2, 420 mM $NaCl_2$, 0.2 mM EDTA, 1 mM DTT) and incubated at 4 °C for 30 min. After 15

min. of centrifugation, the supernatant were collected. Buffers A and C were supplemented with protease inhibitor mix tablet (Roche), MG132 and DTT. For total extracts, cells were lysed in Buffer AE (0.6% SDS, 50 mM Tris pH 8, 100 mM NaCl, 2 mM EDTA and protease inhibitors (Roche)).

Immunoprecipitation. Polyclonal antibodies for Cathepsin L (Chapman et al., 1997b) were incubated with cell extracts in AE buffer supplemented with 4% triton and 8 mM iodoacetamide for 90 min. Protein G Agarose beads (Gibco BRL) were added for an extra hour. IP's were washed with total lysis buffer 3 times. Immunoprecipitated products were resolved on a 12% polyacrylamide gel and analyzed by Western blot using an HA antibody or autoradiography..

Cathepsin activity labeling. Cathepsin activity was measured using a small molecule ABP as previously described (Bogyo et al., 2000). Briefly, NIH3T3 and MEF Cat $L^{-/-}$cells were synchronized by serum-starvation and restimulation. One hour before each time point, cells were labeled with 10^6 counts/ml of a ^{125}Iodine-labeled E64 analog, ^{125}I-JPM-OET. Extracts and immunoprecipitations were done as mentioned using cathepsin L antibodies.

Western blot analysis. Protein extracts (10 μg for transfected NIH3T3, MEF and MEF cat $B^{-/-}$ cells or 50, 100 and 150 μg for transfected MEF cat L-/- and 25 μg for non transfected cells) were separated by electrophoresis on 6% (CDP/Cux), 8% (CR1+2) or 12% (Cathepsin L) polyacrylamide gels. Western blot analysis with aHA, aMyc , α861 were performed as previously described (Moon) (Coqueret et al., 1998).

Expression and purification of CR1CR2 fusion proteins. The bacterial expression vectors pTriEX (invitrogen) expressing CR1+2-NLS-HA-Myc or amino acids 612-1336-HA

were introduced into the BL21(DE3) of *E.coli* and induced with IPTG. The fusion proteins were purified Ni-NTA agarose (Qiagen using procedures provided by the suppliers).

CDP/Cux in vitro cleavage assay and cleavage site determination. 200 ng of bacterially purified CR1CR2 were incubated with 200 pM of recombinant cathepsin L in reaction buffer I (50mM Acetate buffer, 5mM DTT, pH 5.5). Similar assay was performed in the presence of the recombinant cathepsins B, H, K, F, V and S. In one experiment, in vitro cleavage of CDP/Cux was tested under different pH conditions. In this experiment, the assay was performed in 50 mM MES (at pH range of 5.5-6.5) or HEPES (at pH 7.0 and 7.5) supplemented with 5 mM DTT. In order to determine the cleavage site, recombinant CDP/Cut cleavage products were blotted on to a PVDF membrane and subjected to N-terminal sequencing by Edman degradation.

Electrophoretic mobility shift assay (EMSA). 1 μg of nuclear extracts from transfected NIH3T3 cells were incubated in a final volume of 30 μl for 5 min. at room temperature in binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 1 mM DTT, 5 mM EDTA, 5% glycerol, 25 mM NaCl) with 60 ng of poly dIdC and 30 μg of bovine serum albumin (BSA) as non specific competitors. 10 pg (5×10^{-16} mol) of radiolabeled oligonucleotide (TCGAGACGGT**ATCGAT**AAGCTTCTTTTC) was added and further incubated for 15 min. at room temperature. Samples were loaded on a 4% polyacrylamide gel (29:1) and separated by electrophoresis at 8V/cm in 0.5X Tris-Borate EDTA. Gels were dried and visualized by autoradiography.

Plasmid Construction. Plasmid maps and sequences will be provided upon request. Wild-type mouse cathepsin L was inserted into pCDNA3 vector using BamHI sites. Two PCR-amplification were performed to obtain the pcDNA3-Cat L-Ha tagged construct. The first

primer set included BamH1 and Nhe I sites (nt 14 to 34 + Bam HI: ACTGGGATC-CGAGTTTGACTTC and nt 716 to 693 + Nhe I: ACTGGCTAGCCTCGGCTCTGTATTTACAAGATCC), respectively, and the second primer set included the HA tag-Nhe I and Not I site respectively (nt 717 to 736 +HA+ Nhe I: ACTGGCTAGCTATCCATACGACGTACCTGACTACGCATTCGCTGTGGCTAAT-GACAC and nt 1175 to 1151 + Not I: ACTGGCGGCCGCTCAATGATTCAAGTAC-CATGGTTTG). The PCR products were digested and ligated to a Bam HI/Not I digested pCDNA3 vector. Cathepsin L mutants were made by site directed mutagenesis using oligonucleotides containing the following substitutions. The seven AUG codons (Met 1, 56, 58, 75, 77, 82, 111) were mutated to UUC (Phe) and the UGC codon (Cys 25) was mutated to UCC (Ser). Oligonucleotide sequences are available upon request. AUG 56 was constructed by PCR using a primer overlapping AUG 56 and using Cat L-HA as a template with the same 3' nt1175 to 1151 + Not I primer. All PCR were sequenced.

***Pichia pastori* Expression system.** Cat L-HA was introduced into the pPIC9 vector and recombinant yeast were generated and selected using the manufacturer's instructions (Invitrogen). Recombinant and wild type *P. Pastori* were induced for 48 hours with 0.5% MeOH, spun for 2 min, and the supernantant (media) was kept at –80°C. For *in Vitro* assays, 1 µl of media was incubated with the recombinant CR1+2 protein for 10 min at 37°C at pH 7.0. Protein loading buffer was added to stop the reaction. Samples were boiled and loaded on an 8% SDS-polyacrylamide gel and analyzed by Western blot using a Myc antibody.

Immunofluorescence. NIH3T3 cells were plated on a cover slip and transfected or not with 5 µg of wild type or mutant pCDNA3-CatL-HA. Cells were fixed with 2% paraformaldehyde and quenched in 50 mM of NH_4Cl. Cells were solubilized (95% PBS + 5% FBS

+ 0.5% Triton X-100) and incubated with αCatL (1:20000, (Chapman et al., 1997b); 1:200, Santa Cruz D20; 1:200, Santa Cruz M19 or αHA (1:20000) for 1 hr at RT. Secondary antibody (anti- mouse alexa 488 1:1000 or anti-rabbit alexa 594) was added for 30 min. Cells were visualized using a RETIGA 1300 digital camera (QIMAGING) and a Zeiss AxioVert 135 microscope with a 63x objective. Images were analyzed using Northern Eclipse version 6.0 (Empix Imaging).

***In Vitro* transcription-translation.** *In Vitro* transcription-translation was performed with ^{35}S-labelled methionine-cysteine labeling mix (Amersham) using rabbit reticulocyte lysate - TNT Kit from Promega in the absence or presence of canine pancreatic stripped microsomal membranes. Membranes were isolated as described in (Kim et al., 2002) with minor modifications. Microsomes were isolated by centrifugation in a TL100.1 rotor (50 000 rpm for 5 min) through 100 µl of a 0.5M sucrose cushion (1X PBS, 100 mM KAc, 50 mM HEPES pH 7.4, 5 mM $MgAc_2$). Protein in the supernatant were precipitated with 2% TCA and washed once with EtOH before loading. Membranes were washed with 2 mM EDTA and recentrifuged as described above. Pellets were dissolved in protein loading buffer and analyzed by SDS-PAGE (12%). All gels were processed by fluorography using Amplify (Amersham), dried and exposed.

References

Balaji, K. N., Schaschke, N., Machleidt, W., Catalfamo, M., et al. (2002). Surface cathepsin B protects cytotoxic lymphocytes from self-destruction after degranulation. J Exp Med *196*, 493-503.

Benavides, F., Starost, M. F., Flores, M., Gimenez-Conti, I. B., et al. (2002). Impaired hair follicle morphogenesis and cycling with abnormal epidermal differentiation in nackt mice, a cathepsin L-deficient mutation. Am J Pathol *161*, 693-703.

Blochlinger, K., Jan, L. Y., and Jan, Y. N. (1991). Transformation of sensory organ identity by ectopic expression of Cut in Drosophila. Genes Dev *5*, 1124-1135.

Bodmer, R., Barbel, S., Shepherd, S., Jack, J. W., et al. (1987). Transformation of sensory organs by mutations of the cut locus of D. melanogaster. Cell *51*, 293-307.

Bogyo, M., Verhelst, S., Bellingard-Dubouchaud, V., Toba, S., et al. (2000). Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. Chem Biol *7*, 27-38.

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

Brigid Hogan, R. B., Frank Costantini, Elizabeth Lacy (1994). Manipulating the Mouse Embryo, A Laboratory Manual, Second Edition edn, Cold Spring Harbour Laboratory Press).

Chapman, H. A., Riese, R. J., and Shi, G. P. (1997a). Emerging roles for cysteine proteases in human biology. Annu Rev Physiol *59*, 63-88.

Chapman, R. L., Kane, S. E., and Erickson, A. H. (1997b). Abnormal glycosylation of procathepsin L due to N-terminal point mutations correlates with failure to sort to lysosomes. J Biol Chem *272*, 8808-8816.

Coqueret, O., Berube, G., and Nepveu, A. (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*, 4680-4694.

Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., et al. (1996). Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. Embo J *15*, 5492-5503.

DeSalle, L. M., and Pagano, M. (2001). Regulation of the G1 to S transition by the ubiquitin pathway. FEBS Lett *490*, 179-189.

Ellis, T., Gambardella, L., Horcher, M., Tschanz, S., et al. (2001). The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev *15*, 2307-2319.

Erickson, A. H. (1989). Biosynthesis of lysosomal endopeptidases. J Cell Biochem *40*, 31-41.

Goulet, B., Watson, P., Poirier, M., Leduy, L., et al. (2002). Characterization of a tissue-specific CDP/Cux isoform, p75, activated in breast tumor cells. Cancer Res *62*, 6625-6633.

Henderson, J. E. (1997). Nuclear targeting of secretory proteins. Mol Cell Endocrinol *129*, 1-5.

Hershko, A. (1997). Roles of ubiquitin-mediated proteolysis in cell cycle control. Curr Opin Cell Biol *9*, 788-799.

Holthuis, J., Owen, T. A., van Wijnen, A. J., Wright, K. L., et al. (1990). Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science *247*, 1454-1457.

Irving, J. A., Shushanov, S. S., Pike, R. N., Popova, E. Y., et al. (2002). Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation. J Biol Chem *277*, 13192-13201.

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

Jack, J. W. (1985). Molecular organization of the cut locus of Drosophila melanogaster. Cell *42*, 869-876.

Kim, S. J., Mitra, D., Salerno, J. R., and Hegde, R. S. (2002). Signal sequences control gating of the protein translocation channel in a substrate-specific manner. Dev Cell *2*, 207-217.

Ledford, A. W., Brantley, J. G., Kemeny, G., Foreman, T. L., et al. (2002). Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia. Dev Biol *245*, 157-171.

Lee, K. A. W., Bindereif, A., and Green, M. R. (1988). A small-scale procedure for preparation of nuclear extracts that support efficient transcription and Pre-mRNA splicing. Gene Anal Techn *5*, 22-31.

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

Luong, M. X., van der Meijden, C. M., Xing, D., Hesselton, R., et al. (2002). Genetic ablation of the CDP/Cux protein C terminus results in hair cycle defects and reduced male fertility. Mol Cell Biol *22*, 1424-1437.

McGrath, M. E. (1999). The lysosomal cysteine proteases. Annu Rev Biophys Biomol Struct *28*, 181-204.

Mehtani, S., Gong, Q., Panella, J., Subbiah, S., et al. (1998). In vivo expression of an alternatively spliced human tumor message that encodes a truncated form of cathepsin B. Subcellular distribution of the truncated enzyme in COS cells. J Biol Chem *273*, 13236-13244.

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*, 555-558.

Menard, R., Carmona, E., Takebe, S., Dufour, E., et al. (1998). Autocatalytic processing of recombinant human procathepsin L. Contribution of both intermolecular and unimolecular events in the processing of procathepsin L in vitro. J Biol Chem *273*, 4478-4484.

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

Moon, N. S., Premdas, P., Truscott, M., Leduy, L., et al. (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.

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

Peters, J. M. (2002). The anaphase-promoting complex: proteolysis in mitosis and beyond. Mol Cell *9*, 931-943.

Portnoy, D. A., Erickson, A. H., Kochan, J., Ravetch, J. V., et al. (1986). Cloning and characterization of a mouse cysteine proteinase. J Biol Chem *261*, 14697-14703.

Reinheckel, T., Deussing, J., Roth, W., and Peters, C. (2001). Towards specific functions of lysosomal cysteine peptidases: phenotypes of mice deficient for cathepsin B or cathepsin L. Biological Chemistry *382*, 735-741.

Roth, W., Deussing, J., Botchkarev, V. A., Pauly-Evers, M., et al. (2000). Cathepsin L deficiency as molecular defect of furless: hyperproliferation of keratinocytes and pertubation of hair follicle cycling. Faseb J *14*, 2075-2086.

Sinclair, A. M., Lee, J. A., Goldstein, A., Xing, D., et al. (2001). Lymphoid apoptosis and myeloid hyperplasia in CCAAT displacement protein mutant mice. Blood *98*, 3658-3667.

Stein, G. S., and Borun, T. W. (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.

Truscott, M., Raynal, L., Premdas, P., Goulet, B., et al. (2003). CDP/Cux stimulates transcription from the DNA polymerase alpha gene promoter. Mol Cell Biol *23*, 3013-3028.

Turk, B., Dolenc, I., Lenarcic, B., Krizaj, I., et al. (1999). Acidic pH as a physiological regulator of human cathepsin L activity. Eur J Biochem *259*, 926-932.

Turk, B., Dolenc, I., Turk, V., and Bieth, J. G. (1993). Kinetics of the pH-induced inactivation of human cathepsin L. Biochemistry *32*, 375-380.

Turk, B., Turk, D., and Turk, V. (2000). Lysosomal cysteine proteases: more than scavengers. Biochim Biophys Acta *1477*, 98-111.

Turk, B., Turk, V., and Turk, D. (1997). Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biological Chemistry *378*, 141-150.

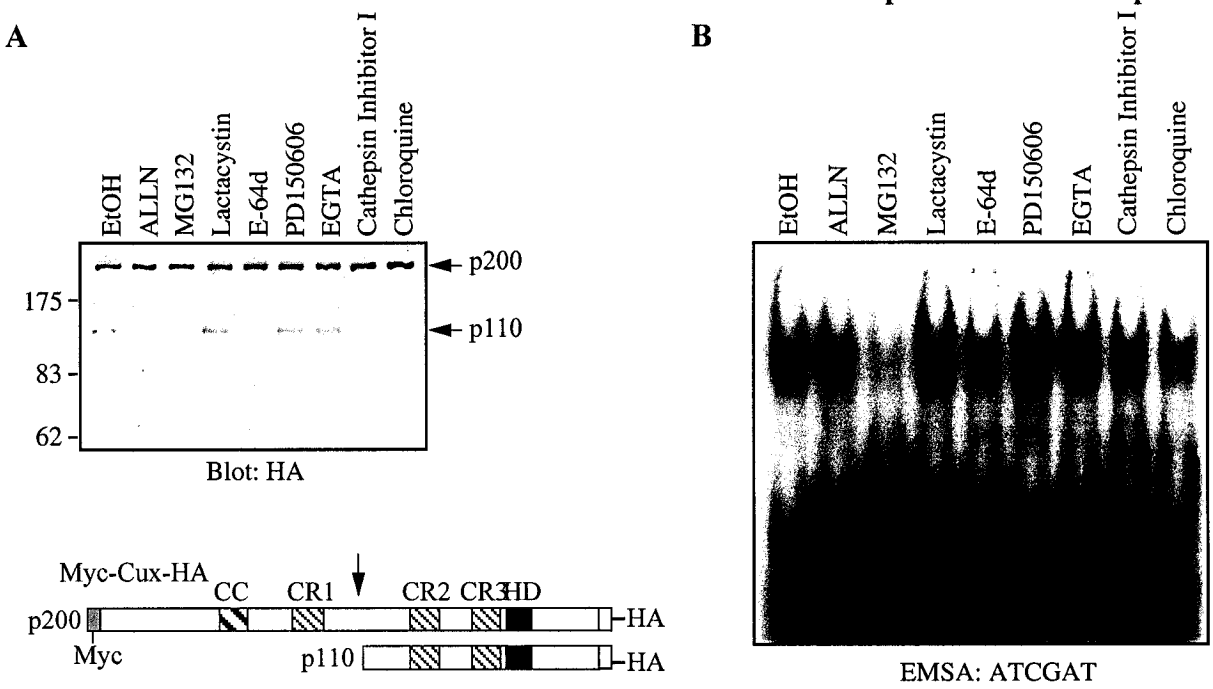
van Wijnen, A. J., Choi, T. K., Owen, T. A., Wright, K. L., et al. (1991). Involvement of the cell cycle-regulated nuclear factor HiNF-D in cell growth control of a human H4 histone gene during hepatic development in transgenic mice. Proc Natl Acad Sci U S A *88*, 2573-2577.

van Wijnen, A. J., van Gurp, M. F., de Ridder, M. C., Tufarelli, C., et al. (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*, 11516-11521.

van Wijnen, A. J., Wright, K. L., Lian, J. B., Stein, J. L., et al. (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*, 15034-15042.

Yasothornsrikul, S., Greenbaum, D., Medzihradszky, K. F., Toneff, T., et al. (2003). Cathepsin L in secretory vesicles functions as a prohormone-processing enzyme for production of the enkephalin peptide neurotransmitter. Proc Natl Acad Sci U S A.

Figure 1: CDP/Cux processing by a papain family cysteine protease is required for its function in cell cycle regulation. (A and B) Unsynchronized NIH3T3 cells were transfected with a vector expressing Myc-Cut-HA. After 18 hours, cells were treated with various cell-permeable protease inhibitors, and after an additional 4 hours, nuclear extracts were prepared and analyzed in Western blots with anti-HA antibodies (A), and in EMSA with the ATCGAT probe (B). Indicated is a schematic representation of the full length, HA and Myc tagged Cux protein. The evolutionarily conserved domains of the CDP/Cux protein are indicated (CC, coiled-coil; CR1, CR2 and CR3: Cut repeats 1, 2 and 3; HD, homeodomain) (C and D) Stable populations of NIH3T3 cells carrying a p110 expression vector or the empty vector were synchronized by serum starvation/re-stimulation. In two series of plates (E and F), the cysteine protease inhibitor E-64d was added to a final concentration of 200 μM at the same time as the fresh medium with 10% serum. At the indicated times, cells were harvested and cell cycle distribution was monitored by fluorescence-activated cell sorting (FACS) analysis after staining of the DNA with propidium iodide.


A
EtOH
ALLN
MG132
Lactacystin
E-64d
PD150606
EGTA
Cathepsin Inhibitor I
Chloroquine
175
83
62
p200
p110
Blot: HA
Myc-Cux-HA
CC
CR1
CR2
CR3
HD
p200
Myc
HA
p110
HA
B
EtOH
ALLN
MG132
Lactacystin
E-64d
PD150606
EGTA
Cathepsin Inhibitor I
Chloroquine
EMSA: ATCGAT

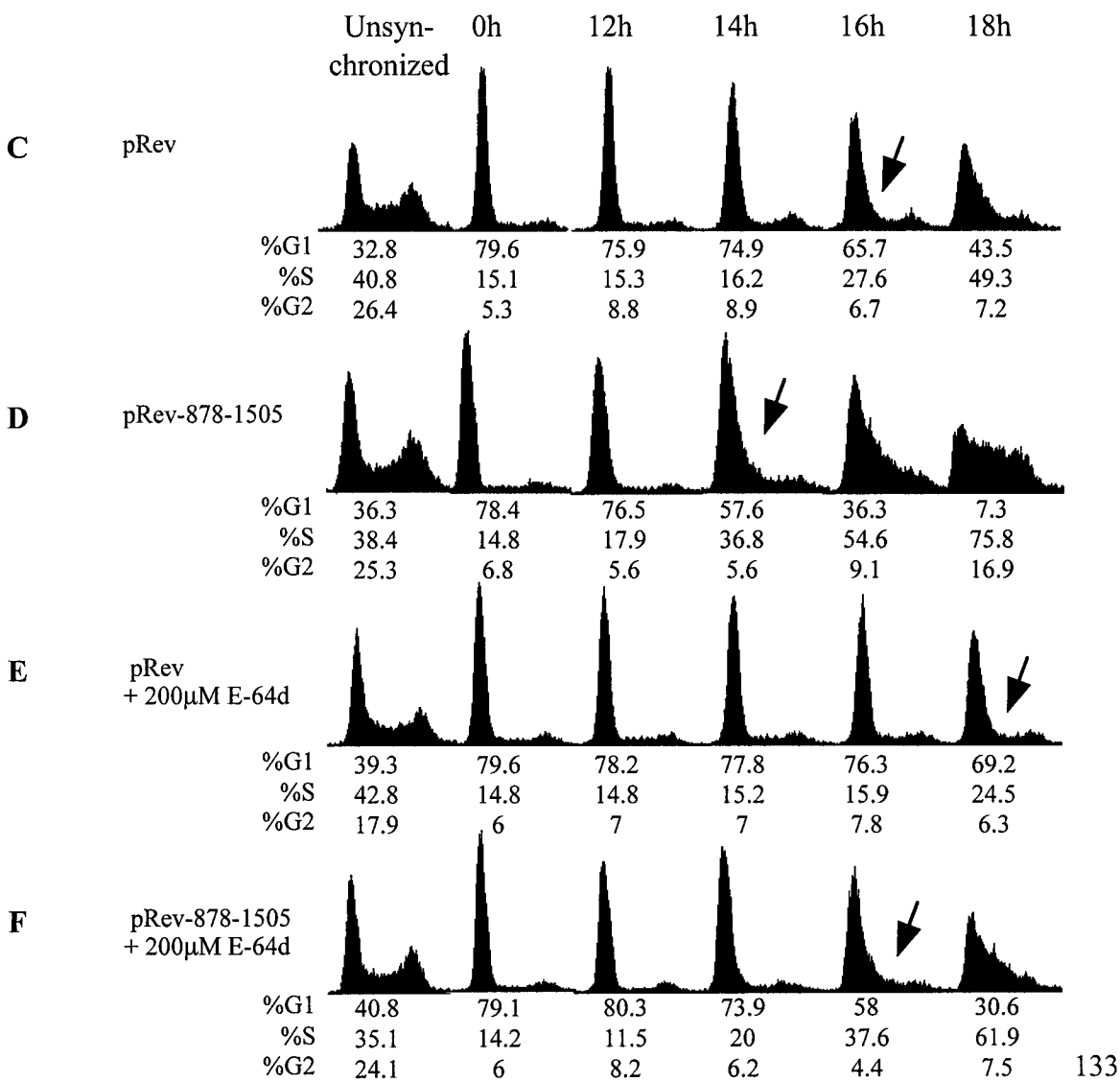
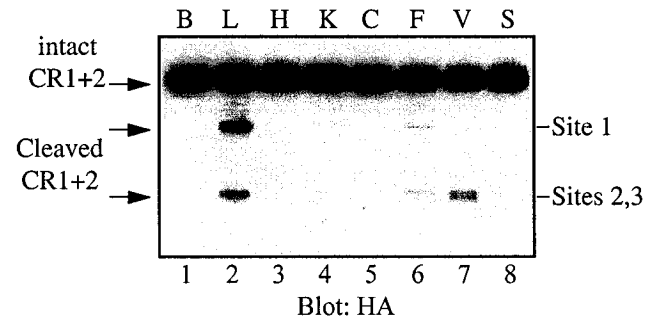
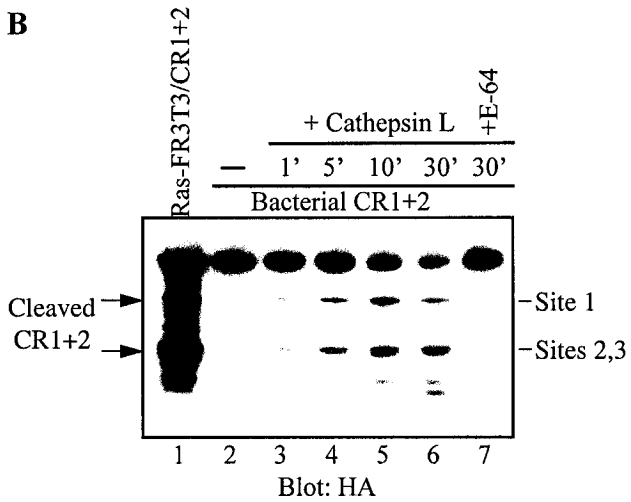

Unsyn-chronized
0h
12h
14h
16h
18h
C
pRev
%G1 32.8 79.6 75.9 74.9 65.7 43.5
%S 40.8 15.1 15.3 16.2 27.6 49.3
%G2 26.4 5.3 8.8 8.9 6.7 7.2
D
pRev-878-1505
%G1 36.3 78.4 76.5 57.6 36.3 7.3
%S 38.4 14.8 17.9 36.8 54.6 75.8
%G2 25.3 6.8 5.6 5.6 9.1 16.9
E
pRev
+ 200µM E-64d
%G1 39.3 79.6 78.2 77.8 76.3 69.2
%S 42.8 14.8 14.8 15.2 15.9 24.5
%G2 17.9 6 7 7 7.8 6.3
F
pRev-878-1505
+ 200µM E-64d
%G1 40.8 79.1 80.3 73.9 58 30.6
%S 35.1 14.2 11.5 20 37.6 61.9
%G2 24.1 6 8.2 6.2 4.4 7.5

Figure 2. CDP/Cux is proteolytically processed by cathepsin L *in vitro* to produce the endogenous p110 fragment. (A) Purified, bacterially expressed HA-CR1+CR2 protein (200ng) was incubated with various recombinant cathepsins (0.2ng) for 10' at RT. Reactions were quenched by the addition of sample buffer and samples were analyzed by Western blot using anti-HA tag antibodies. The fragments resulting from cleavage at sites 1, 2 and 3 are indicated at right. (B) Purified CR1+CR2-HA protein (200ng) was incubated with recombinant cathepsin L (0.2ng) for 1, 5' 10' and 30' at RT (lanes 2-6). As a control, 100 μM E-64 was added to the enzyme simultaneously with the substrate and incubated for 30' at RT (lane 7). A nuclear extract from CR1+CR2-HA Ras-3T3 transfectants probed with the anti-HA antibody is shown to compare the band pattern for the recombinant cathepsins to that observed for CDP-Cut processing *in vitro* (lane 1). (C) Recombinant CDP/Cux 612-1336-HA was purified from bacteria and incubated with cathepsin L at various pH for the indicated times. Samples were analyzed by western blot using anti-HA tag antibodies. (D) Schematic representation of the His, HA-tagged internal CR1+CR2 construct. The processed sites observed *in vitro* are indicated with arrows. Sequences of the cleavage sites identified by N-terminal sequencing of the protein fragments are indicated.

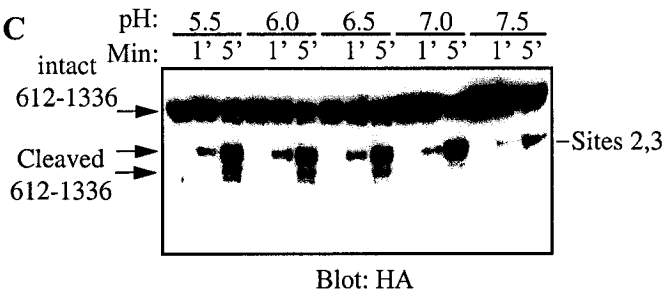
A



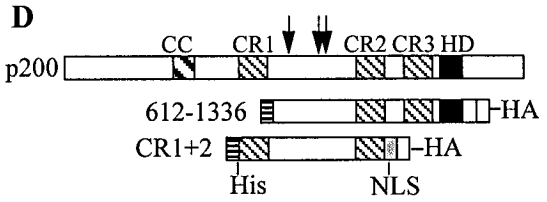
B



C



D



Cleavage site 1 - NRLFQ^{643} EVPK

Cleavage site 2,3 - PKLLS^{747} TSPMPTVS^{755} SYPP

Figure 3: CDP/Cux is proteolytically processed by cathepsin L *in vivo* to produce the endogenous p110 fragment in NIH 3T3 and CV-1 cells, but not in cathepsin L deficient cells.

(A) NIH3T3 (left panel) and CV1 (right panel) cells were transfected with the Myc-Cux-HA vector expressing the CDP/Cux protein with Myc and HA epitope tags at its amino- and carboxy-terminus, respectively, with or without increasing amounts of a wild-type cathepsin L vector (lanes 2, 3, 5, 6, 7 and 10) or with one where the active cysteine 25 was mutated for a serine (lane 9). In lane 7, cells were incubated in the presence of 20 μM of E-64d for 2 hours prior to harvesting. Nuclear extracts were resolved on a 6% denaturing polyacrylamide gel and analyzed by immunoblotting with anti-HA antibodies. (B) Wild type MEF, MEF cat L-/- and MEF cat B-/- cells were transfected with the Myc-Cux-HA vector (see diagram in Fig. 1). Nuclear extracts were analyzed by immunoblotting using an HA antibody.

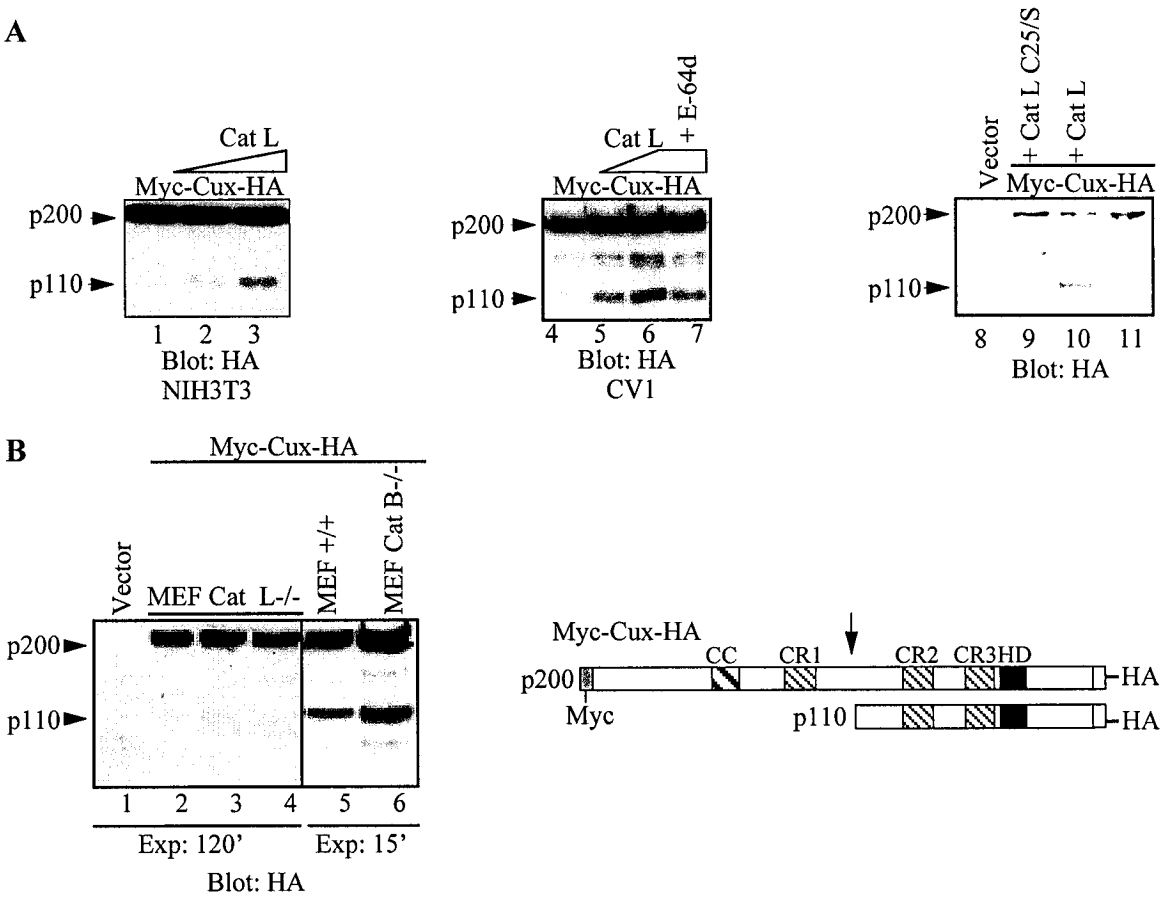
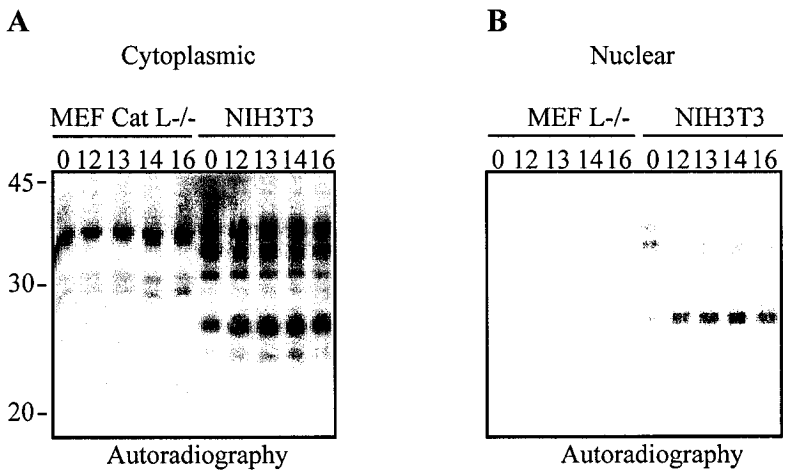

A
Cat L
Myc-Cux-HA
p200
p110
1 2 3
Blot: HA
NIH3T3
Cat L + E-64d
Myc-Cux-HA
p200
p110
4 5 6 7
Blot: HA
CV1
Vector
+ Cat L C25/S
+ Cat L
Myc-Cux-HA
p200
p110
8 9 10 11
Blot: HA
B
Myc-Cux-HA
Vector
MEF Cat L-/-
MEF +/+
MEF Cat B-/-
p200
p110
1 2 3 4 5 6
Exp: 120'
Exp: 15'
Blot: HA
Myc-Cux-HA
CC
CR1
CR2
CR3HD
p200
HA
Myc
p110
HA

Figure 4: Cathepsin L Activity in the nucleus increases in a cell cycle dependant manner.

(A) and (B) MEF cat L-/- and NIH3T3 cells were incubated for 72hr in the absence of serum (time point O). Serum-containing medium was added and cells were isolated after 12, 13, 14, and 16 hour following serum addition. At each time point cells were incubated with 10^6 counts/ml ^{125}I-JPM-OET for 1hr at 37C. Cells were subsequently lysed in hypotonic buffer and cytosolic and nuclear fractions were prepared as described in the materials and methods. Samples (50μg/lane) were analyzed on a 14% SDS-PAGE. (C) ^{125}I-JPM-OET labeled nuclear and cytosolic extracts from the 14hr time point were subjected to immunoprecipitation using an anti mouse cathepsin L antibody (IP CL lanes) and control normal mouse serum (IP 0 lanes). Full-length Cathepsin L single chain form (CLSC) and the truncated heavy chain form (CLHC) are labeled with arrows at right.


A
Cytoplasmic
MEF Cat L-/-
NIH3T3
0 12 13 14 16 0 12 13 14 16
45-
30-
20-
Autoradiography
B
Nuclear
MEF L-/-
NIH3T3
0 12 13 14 16 0 12 13 14 16
Autoradiography

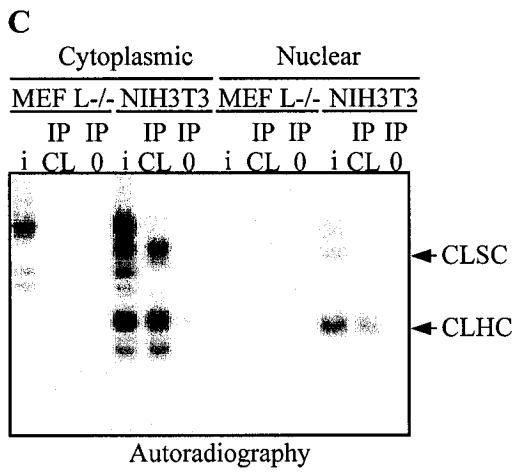
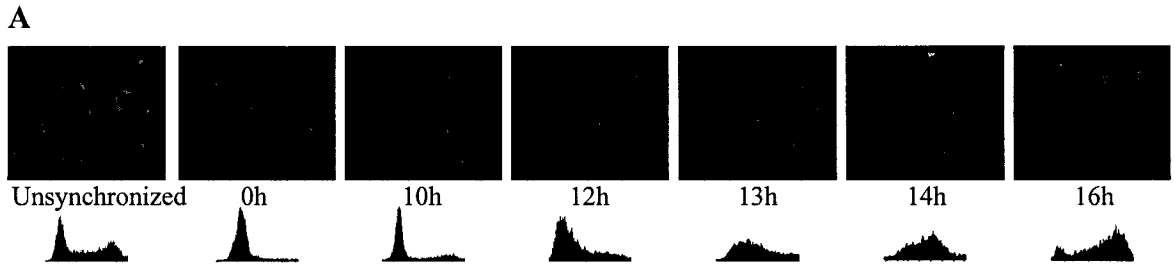

C
Cytoplasmic
Nuclear
MEF L-/-
NIH3T3
MEF L-/-
NIH3T3
IP i
IP CL
IP 0
IP i
IP CL
IP 0
IP i
IP CL
IP 0
IP i
IP CL
IP 0
CLSC
CLHC
Autoradiography

Figure 5: Evidence from indirect immunofluorescence that cathepsin L translocates to the nucleus in NIH 3T3 cells in a cell cycle dependent manner.

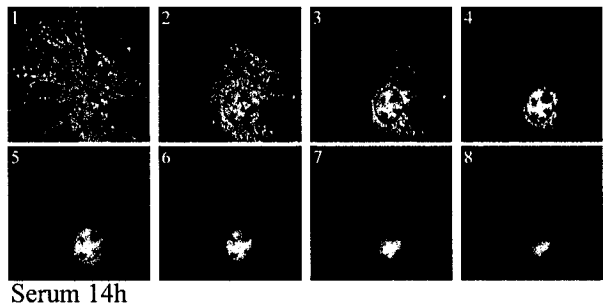
(A) NIH3T3 cells were plated on a glass cover-slip, and synchronized by serum starvation/stimulation, as described in Materials and Methods. Indirect immunofluorescence was performed using a cathepsin L antibody and the cell cycle distribution was monitored by fluorescence-activated cell sorting (FACS) analysis after staining of the DNA with propidium iodide. (B) Two different patterns of staining were observed. The relative percentage of cells exhibiting each staining pattern was calculated upon examination of a minimum of 200 cells. (C) Confocal microscopy images of 14h serum stimulated NIH3T3 cells. Cathepsin L was detected by indirect immunofluorescence assay and analyzed by eight-step Z-position sectional scanning of the cell. (D) NIH3T3 cells were plated on a glass cover-slip, transfected with a vector expressing cathepsin L containing an internal HA epitope (Cat L-HA). Cells were either maintained as an unsynchronized cell population or were synchronized using either of two methods: double thymidine block followed by a 2 hour-release or serum stimulation for 13 hours following a 3-day starvation period. (E) Confocal microscopy images, as in (C).



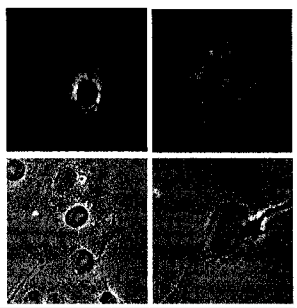
B

Unsyn	79%	21%
0h	80%	20%
10h	50%	50%
12h	25%	75%
13h	10%	90%
14h	10%	90%
16h	88%	12%

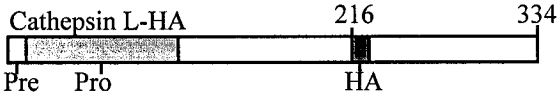
C



D



Unsyn.	80%	20%
13h serum	60%	40%
2h post-thymidine	60%	40%



E

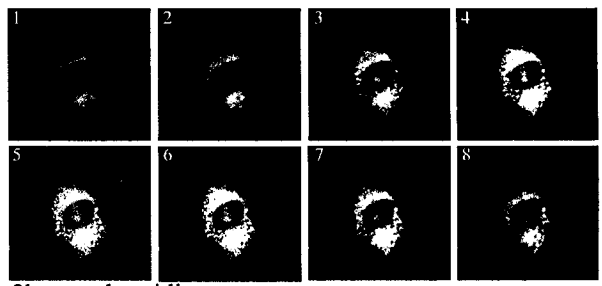


Figure 6: Translation initiation at an internal start site allows the expression of a cathepsin L isoform devoid of the signal peptide

(A) Wild type or mutated cathepsin L-HA vectors were expressed *in vitro* using the TNT-Promega transcription-translation system in the presence of ^{35}S-labelled methionine and cysteine. The mutation numbers refer to the methionine codons that were mutated (AUG to UUC): M1-HA, the first AUG codon was mutated; M56,58,75,77,83,111-HA: the corresponding methionine codons were mutated. Cat L-HA proteins expressed *in vitro* were resolved on a 12% SDS-PAGE and revealed by autoradiography. (B) Wild type or mutated cat L-HA vectors were expressed *in vitro* using the TNT-Promega transcription-translation system, and *in vivo* in NIH3T3 cells. 48 hours after transfection, cat L species were detected by autoradiography or immunoblotting with an HA antibody, as indicated. Samples from NIH3T3 cells were first immunoprecipitated with cat L antibodies. (C) *In vitro* transcription-translation with ^{35}S-labelled methionine was performed using the cat L-HA or M1-HA vectors without (input) or with microsomal membranes. Following the reactions, membranes and supernatant were separated as described in material and methods and cat L isoforms were detected by autoradiography as in (A). (D) Wild type or mutated cat L-HA vectors were introduced into NIH3T3 and indirect immunofluorescence was performed using an HA antibody. Representative cells are shown. In the cases where more than one staining pattern was observed, the proportion of cells exhibiting each pattern is provided.

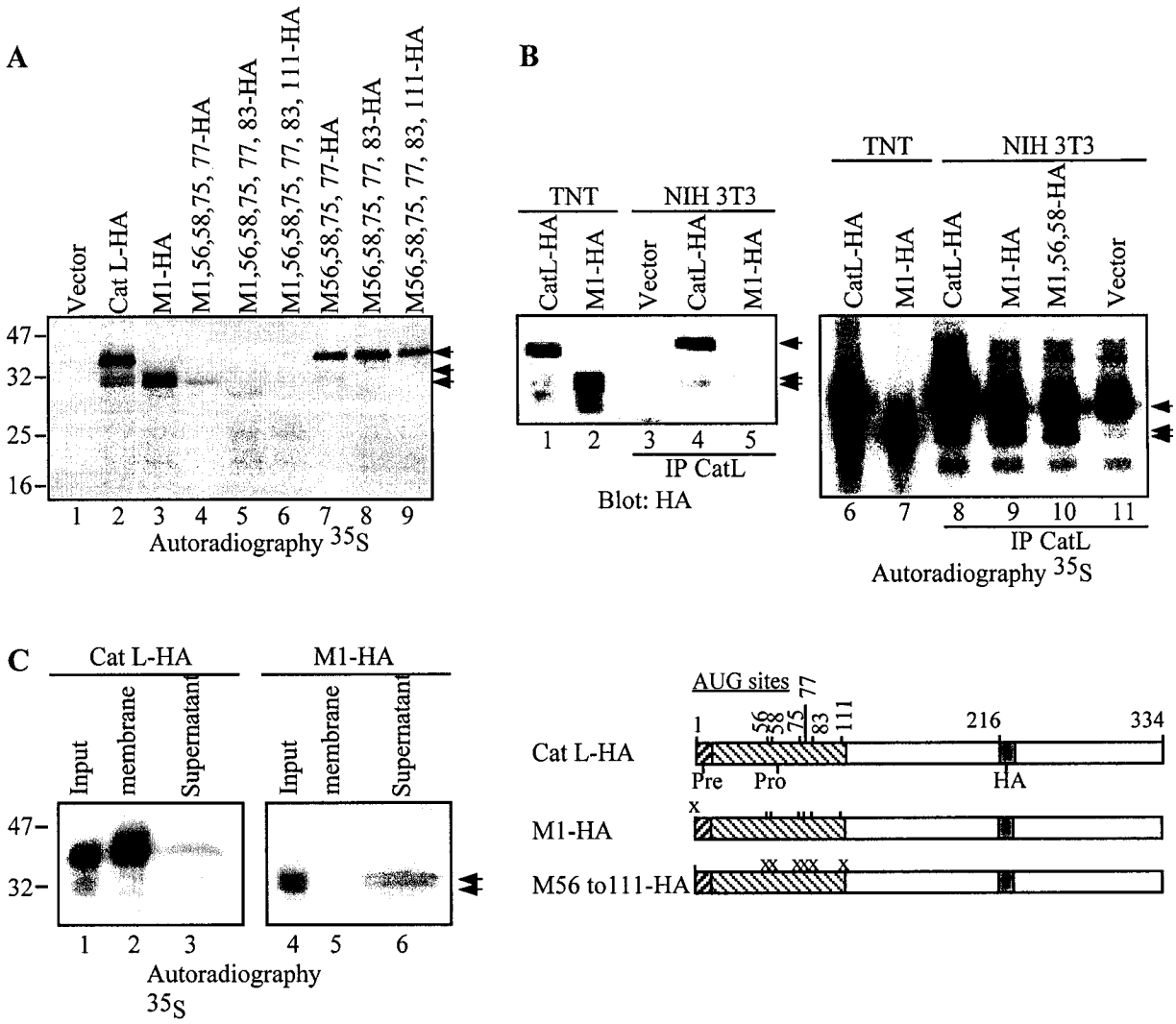
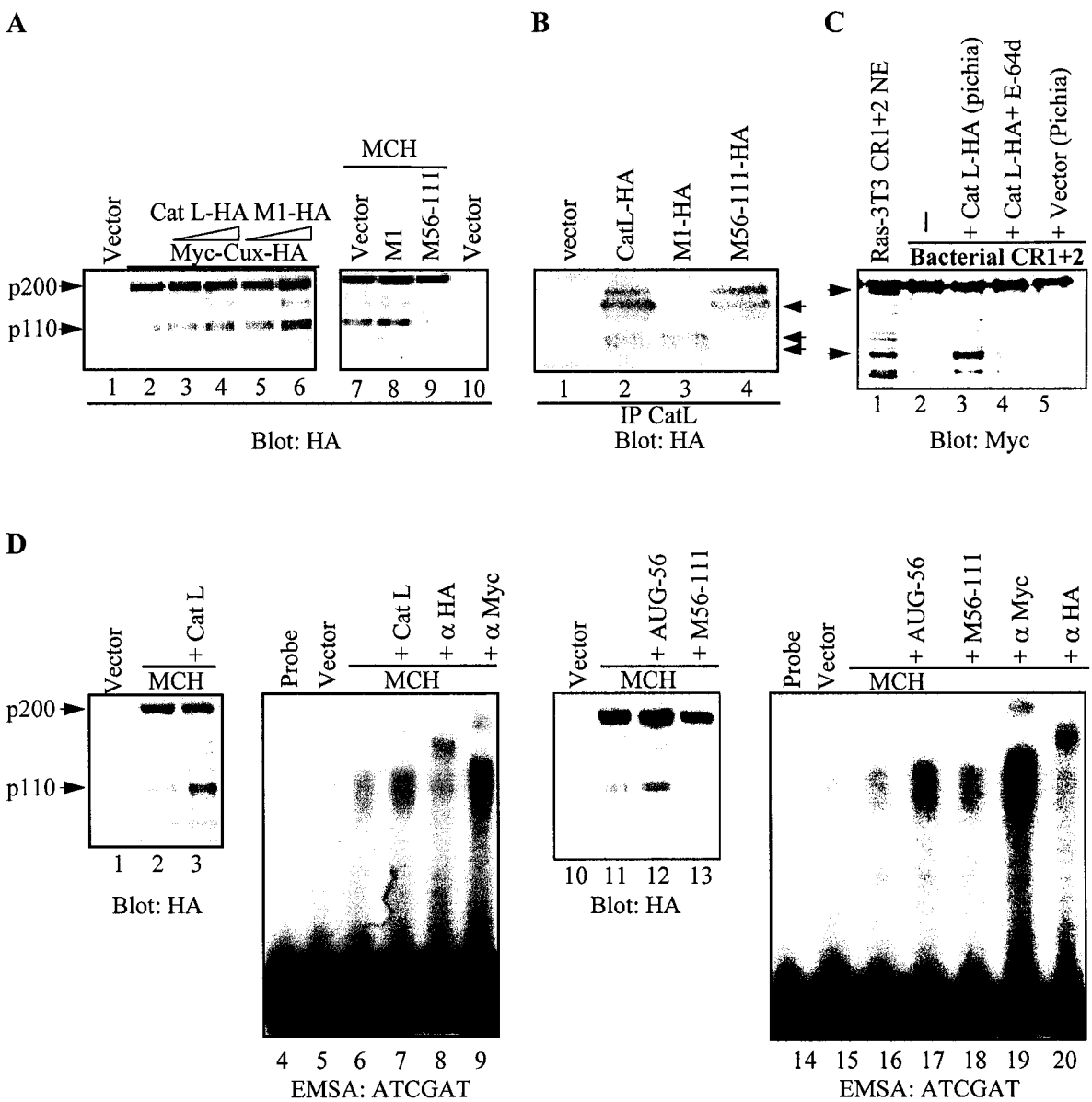

A
Vector
Cat L-HA
M1-HA
M1,56,58,75, 77-HA
M1,56,58,75, 77, 83-HA
M1,56,58,75, 77, 83, 111-HA
M56,58,75, 77-HA
M56,58,75, 77, 83-HA
M56,58,75, 77, 83, 111-HA
47-
32-
25-
16-
1 2 3 4 5 6 7 8 9
Autoradiography 35S
B
TNT
NIH 3T3
CatL-HA
M1-HA
Vector
CatL-HA
M1-HA
1 2 3 4 5
IP CatL
Blot: HA
TNT
NIH 3T3
CatL-HA
M1-HA
CatL-HA
M1-HA
M1,56,58-HA
Vector
6 7 8 9 10 11
IP CatL
Autoradiography 35S
C
Cat L-HA
M1-HA
Input
membrane
Supernatant
Input
membrane
Supernatant
47-
32-
1 2 3 4 5 6
Autoradiography
35S
AUG sites
1
56
58
75
77
83
111
216
334
Cat L-HA
Pre
Pro
HA
M1-HA
M56 to111-HA
D
Cat L-HA
M56, 58, 75, 77, 83,111 -HA
M1-HA
AUG 56-HA
80%
20%
100%
100%
100%

Figure 7: Short cathepsin L isoforms can process CDP/Cux and increase its DNA binding activity.

(A) NIH3T3 cells were transfected with the Myc-Cux-HA vector with or without increasing amounts of wild type or mutated cat L-HA vectors. For detection of the Myc-Cux-HA protein, nuclear extracts were resolved on a 6% denaturing polyacrylamide gel and analyzed by immunoblotting with anti-HA antibodies. (B) In order to detect cat L-HA proteins, total extracts were first submitted to immunoprecipitation with cat L antibodies prior to immunoblotting with HA antibodies. This procedure was used to eliminate a non-specific band that co-migrates with the short cat L species. (C) 250 ng of bacterially purified CR1+2-HA protein was incubated with 1µl of media from either wild type or cat L-HA expressing *Pichia pastori* in the presence or absence of E-64d for 10 min at 37°C. Reactions were quenched by the addition of sample buffer and samples were analyzed by western blot using anti-HA antibodies. A nuclear extract from CR1+CR2-HA Ras-3T3 transfectants probed with the anti-HA antibody is shown to compare the band pattern for the recombinant cathepsins to that observed for CDP-Cut processing *in vitro* (lane 1). (D) NIH3T3 cells were transfected with Myc-Cux-HA vector (MCH) with or without increasing amounts of wild type or mutated cat L vectors. Nuclear extracts were submitted to Western Blot analysis as described in (A) (Lanes 1,2,3 and 10 to 13). EMSA (Lanes 4 to 9 and 14 to 20) were performed as described in Figure 1.


A
B
C
MCH
Vector
Cat L-HA M1-HA
Myc-Cux-HA
Vector
M1
M56-111
Vector
p200
p110
1 2 3 4 5 6 7 8 9 10
Blot: HA
vector
CatL-HA
M1-HA
M56-111-HA
1 2 3 4
IP CatL
Blot: HA
Ras-3T3 CR1+2 NE
–
+ Cat L-HA (pichia)
+ Cat L-HA+ E-64d
+ Vector (Pichia)
Bacterial CR1+2
1 2 3 4 5
Blot: Myc
D
Vector
+ Cat L
MCH
p200
p110
1 2 3
Blot: HA
Probe
Vector
+ Cat L
+ α HA
+ α Myc
MCH
4 5 6 7 8 9
EMSA: ATCGAT
Vector
+ AUG-56
+ M56-111
MCH
10 11 12 13
Blot: HA
Probe
Vector
+ AUG-56
+ M56-111
+ α Myc
+ α HA
MCH
14 15 16 17 18 19 20
EMSA: ATCGAT

Chapter III Elevated Cathepsin L Expression In Cancer Cells Correlates with Increased Proteolytic Processing of the CDP/Cux Transcription Factor

Preface

As more evidence for the involvement of CDP/Cux in cell cycle progression becomes available, we can begin to understand the potential role for CDP/Cux in cancer as previous observations had already suggested. It was shown that p110 steady state levels were increased in the majority of uterine leiomyomas analyzed compared to their normal counterparts. In addition, CDP/Cux is the DNA binding counterpart of the HiNF-D complex. This complex, which regulates histone gene expression in S phase, is constitutively active in transformed cells. In chapter II, the protease responsible for the cleavage of the full length CDP/Cux into the p110 isoform was identified as cathepsin L. This protease has long been known to be over-expressed and secreted in many cancers. Moreover, cathepsin L expression was shown to correlate with the metastatic potential of certain tumors.

In this chapter, the increased processing of CDP/Cux in transformed cells is first described, followed by the corroboration of this increased cleavage with cathepsin L activity. I will argue that the role of cathepsin L in cancer is not limited to its extra-cellular functions but also involves some activities within the cells, particularly in the nucleus.

Abstract

The CDP/Cux transcription factor was previously shown to acquire distinct DNA binding and transcriptional properties following a proteolytic processing event that takes place at the G1/S transition of the cell cycle. We have recently obtained evidence implicating a nuclear isoform of cathepsin L in the proteolytic processing of CDP/Cux. In the present study, we show that CDP/Cux DNA binding activity is increased in several transformed cells as compared to their normal counterparts and that this correlates with the increased processing of CDP/Cux. As previously reported, cathepsin L expression was found to be elevated in ras-transformed 3T3 cells as compared to the parental 3T3 cells. Importantly, *in vivo* labeling experiments using activity based probes demonstrated the increased activity of cathepsin L in the nuclear fraction of ras-3T3 cells, and ectopic expression of cathepsin L in 3T3 cells brought their capacity to process CDP/Cux close to that of ras-3T3 cells. These results raise the interesting notion that the role of cathepsin L in cancer may not be limited to its extracellular activities but may also involve intracellular functions specifically in the nucleus. More generally, there is emerging evidence to suggest that proteolytic processing may represent an important epigenetic mechanism in cancer. As site-proteolysis has the potential to generate truncated protein isoforms with novel biochemical properties, we propose that aberrant expression or activity of intracellular proteases may lead to the activation of some proto-oncogenes and inactivation of some tumor suppressors.

Introduction

CDP/Cux (CCAAT-displacement protein/cut homeobox) belongs to a family of transcription factors present in all metazoans and involved in the control of proliferation and differentiation (reviewed in (1)). From genetic studies both in Drosophila and the mouse, the CDP/Cux/Cut gene is believed to play various roles in the development and homeostasis of several tissues (2-15). Transcription initiation at three possible start sites, alternative splicing and proteolytic processing contribute to generate multiple CDP/Cux isoforms (p200, p110, p75 and CASP) that contain various combinations of four DNA binding domains and, therefore, display distinct DNA binding and transcriptional activities (16-19). A role for CDP/Cux in cell cycle progression has been inferred from a number of reports. The histone nuclear factor D (HiNF-D), which was later found to include CDP/Cux as its DNA binding partner, was shown to be up-regulated in S phase in normal cells (20-24). CDP/Cux was also found to repress the promoter of the $p21^{WAF1/CIP1}$ gene and to stimulate the expression of the DNA polymerase α, CAD and cyclin A genes (25, 26). The increase in the stable DNA binding activity of CDP/Cux at the G1/S transition was shown to result from at least two post-translational modifications: dephosphorylation of the Cut homeodomain by the Cdc25A phosphatase, and proteolytic cleavage of CDP/Cux p200 between CR1 and CR2 to generate the amino-terminally truncated CDP/Cux p110 (17, 25). During the G2 phase of the cell cycle, cyclin A/Cdk1 was reported to phosphorylate two serine residues in the region of the Cut homeodomain and thus inhibit DNA binding by CDP/Cux p110 (27). Altogether these studies indicate that CDP/Cux fulfills an important function in cell cycle progression, specifically in S phase, and that it is important to down-modulate CDP/Cux once the genome has been replicated.

The gene coding for CDP/Cux in human, CUTL1, was mapped to 7q22, a chromosomal region that is frequently rearranged in several types of human tumors (28-35). Loss-of-heterozygosity studies originally implicated CUTL1 as a putative tumor suppressor gene since it was located within the smallest commonly deleted region in uterine leiomyomas and breast tumors (36, 37). This notion was reinforced by the detection of specific complexes of CDP/Cux and Polyomavirus large T (PyV LT) antigen in both leiomyomas and mammary tumors of female transgenic mice expressing the PyV LT transgene (38). However, no mutation was found within the CUTL1 gene and cDNAs in ovarian cancers and uterine leiomyomas, and expression analysis in uterine leiomyomas revealed that CDP/Cux was expressed in all tumor samples (39, 40). Moreover, N-terminally truncated CDP/Cux isoforms were often expressed at higher level in the leiomyoma than in adjacent normal myometrium (39). Thus, not only did these findings indicate that CUTL1 is not the tumor suppressor on 7q22, but the frequent increase in smaller CDP/Cux isoforms suggested that molecular events associated with the truncation of CDP/Cux proteins may be selected in uterine leiomyomas.

We have recently presented evidence implicating a nuclear isoform of cathepsin L in the proteolytic processing of CDP/Cux (Goulet et al., unpublished observations; see attached manuscript). Cathepsin L is a ubiquitously expressed cysteine protease that localizes primarily to the lysosomes and can be secreted (41-43). The growing protein is first targeted into the endoplasmic reticulum (ER) via a signal peptide (44). However, mutagenesis and expression analyses *in vitro* and *in vivo* indicated that translation initiation within the cathepsin L mRNA can take place with low efficiency at two internal start sites, thereby generating cathepsin L species devoid of the N-terminal signal peptide. Immuno-fluorescence imaging and labeling using activity based probes revealed that N-terminally

truncated cathepsin L could traffic to the nucleus during the S and G2 phases of the cell cycle. CDP/Cux peptides of identical electrophoretic mobility were generated *in vivo* and with purified cathepsin L *in vitro*. Moreover, ectopic expression of cathepsin L increased CDP/Cux processing, whereas less processing was observed in mouse embryo fibroblasts derived from the cathepsin L knock-out ($MEF^{cat\ L-/-}$). These results uncovered a novel role for cathepsin L in the processing of a nuclear protein.

Expression of cysteine cathepsins was reported to be elevated in many types of cancers (reviewed in (45-47)). In particular, tumor promoters like phorbol ester and certain oncogenes like ras, v-src, SV40 Large T, and raf were shown to induce cathepsin L expression and secretion (48-51). In fact, cathepsin L was independently isolated as the major excreted protein (MEP) of transformed cells (42). Moreover, the metastatic potential of transformed cell lines were shown to correlate with cathepsin L expression levels (52). Abundant literature has extended this observation to human tumors originating from various tissues. Notably, human colorectal carcinomas containing K-ras mutations demonstrated greater increases in cathepsin L activity (53-59). Overall, it is generally admitted that the role of cathepsin L in cancer would be related to its activities outside the cells once it has been secreted.

In the present study, we have used tissue culture systems to investigate CDP/Cux expression and activity in transformed cells. We found that proteolytic processing of CDP/Cux is activated in many transformed cell lines and that this results at least in part from the increased cathepsin L expression and activity in the nucleus of these cells. These results raise the interesting notion that the implication of cathepsin L in cancer might not be limited to its extracellular activities but also involve intracellular functions specifically in the nucleus.

Material and Methods

Cell Culture. 3T3, ras-3T3 i.e. transformed with a constitutively active Ha-ras-V12 (Ref Chang J. virol), 293, HeLa, PANC1 were cultured in DMEM medium supplemented with 10%FBS. All breast tumor cell lines and MDCK were a generous gift from Dr. Morag Park. MCF7, MDA 231, T47D and Hs578T cells were cultured in DMEM medium supplemented with 5% FBS. SkBr3 and MDCK cells were cultured in DMEM medium supplemented with 10%FBS. MDA 436 cells were cultured in Leibovitz medium supplemented with 15% FBS and 5 mM insuline. HMEC cells were purchased from clonetics and cultured using the manufacturer's medium and instructions. LoVo C5, a human colon carcinoma cell line a generous gift from Dr. Nabil Seidah, were grown in MEM medium supplemented with 10% FBS. All transfections were done using Exgene500 (Fermentas) or Gene Juice (Novagen) according to manufacturer's instructions. When indicated, 50 μM of E-64d was added to the cells for 4 hours before harvesting. To obtain cells in G0 using the serum starvation/stimulation method, cells were maintained for 3 days in DMEM (SS). The medium was then changed for DMEM plus 10% FBS with or without 200 μM E-64d (Calbiochem) and cells were harvested at different times after serum stimulation.

Preparation of cell extract. 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). Nuclei were then resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM MgCl2, 420 mM $NaCl_2$, 0.2 mM EDTA, 1 mM DTT) and incubated at 4 °C for 30 min. After 15 min. of centrifugation, the supernatant were collected. For total cell extracts, cells were lysed in RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% DOC,

0.1% SDS), incubated on ice for 15 min and passed 5 times through a 22G1 syringe. All buffers were supplemented with protease inhibitor mix tablets (Roche) and DTT**Electrophoretic mobility shift assay (EMSA).** 5 µg of nuclear extracts were incubated in a final volume of 30 µl for 5 minutes at room temperature in binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 1 mM DTT, 5 mM EDTA, 5% glycerol, 25 mM NaCl) with 60 ng of poly dIdC and 30 µg of bovine serum albumin (BSA) as non specific competitors. 10 pg of radiolabelled oligonucleotide was added and further incubated for 15 min. at room temperature. Samples were loaded on a 4% polyacrylamide gels (29:1) and separated by electrophoresis at 8V/cm in 1X Tris-Glycine. Gels were dried and visualized by autoradiography. For EMSA using the OCT oligonucleotide, the nuclear extracts were incubated 5 min at room temperature in the binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 1 mM DTT, 5 mM EDTA, 5% glycerol, 25 mM NaCl) with 4 µg of poly dIdC and 30 µg of BSA as non specific competitors. For antibody supershift experiments, nuclear extracts were incubated 30 min at room temperature with 0.5µg of antibody prior to the incubation with the radiolabeled oligonucleotide. Competition experiments were performed by adding cold oligonucleotide (10 ng, 50 ng, 100 ng and 200 ng) before adding the radiolabeled probe.

Oligonucleotides. ATCGAT: TCGAGACGGTATCGATAAGCTTCTTTTC, NF-Y ATCG:, TCGAGAAAAGAACAACCAATCACCGATC, OCT: GATCCGAGCTTCACCTTATTTGCATAAGCGATTGA

Pulse-chase labeling. 3T3 and ras-3T3 cells were transfected with a vector encoding a CDP/Cut species beginning upstream of CR2 (pXM/CDP 878). Two days after transfection, medium was removed from cells and replaced with medium containing

0.1mCi/ml of ^{35}S-labelled methionine. Cells were labeled for 4 hours, then washed with PBS and incubated with cold medium. At different point times, cells were harvested and lysed in RIPA buffer (50 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1 mM DTT and proteinase inhibitors tablets (Roche). Immunoprecipitation was performed using HA antibody (BabCo) and protein G agarose (Amersham Pharmacia Biotech). The immunoprecipitated product and a portion of the lysate were resolved by electrophoresis on an 8% SDS polyacrylamide gel followed by fixing and washing with Amplify fluorographic enhancer (Amersham Pharmacia Biotech). The gel was dried and exposed to film.

Counterflow centrifugal elutriation. Cells were grown until 80% confluency, and were resuspended in 10 ml of cold PBS + 1% FBS. Cells (around 2.10^8 cells) were submitted to counterflow centrifugal elutriation (Beckman Coulter Avanti TM J-20 centrifuge using a JE-5.0 rotor) at constant speed (2460 rpm). NIH3T3, ras-NIH3T3 and ras-FR3T3 cells were equilibrated at a flow rate of 16 ml/min and eluted at a flow rate from 17 ml/min to 40 ml/min. Hela cells were equilibrated at a flow rate of 20 ml/min and eluted with a flow rate from 22 ml/min to 50 ml/min. To monitor cell cycle distribution, an aliquot of each elutriation fraction was fixed, stained with propidium iodide and submitted to fluorescence activated cell sorting (FACS) analysis.

Purification of His-CR1+2. pTriex2 vector encoding histidine-CR1+2-NLS-HA proteins were 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 manufacturer's instructions.

***In Vitro* processing assay.** 250 ng of purified his-CR1+2 (from bacteria or mammalian cells) was incubated with 25 μg of whole cell extracts from 3T3, ras-3T3 or 3T3 cells

transfected with Cathepsin L expressing vector, in buffer S (50 mM MES, pH 6.8, 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) for 20 min at 37°C. Protein loading buffer was added, the sample were boiled and loaded on an 8% SDS polyacrylamide gel, followed by Western blot analysis using HA antibody.

Immunoprecipitation. Polyclonal antibodies for Cathepsin L (39) were incubated with cell extracts in AE buffer supplemented with 4% triton and 8 mM iodoacetamide for 1h30. Protein G Agarose beads (Gibco BRL) were added for an extra hour. IP's were washed with total lysis buffer 3 times. Immunoprecipitated products were resolved on a 12% polyacrylamide gel and analyzed autoradiography.

Immunofluorescence. 3T3 and ras-3T3 cells were plated on a coverslip. 2 days after, cells were fixed with 3.7% paraformaldehyde for 10 min. After 2 washes with 1X PBS, cells were quenched for 10 min in 50 mM of NH_4Cl, solubilized for 10 min (95% PBS + 5% FBS + 0.5% Triton X-100) and incubated with αCatL (1:20000) for 1 hr at RT. After extensive washing, the secondary antibody (anti-rabbit alexa 488) was incubated for 30 min at RT in the dark. Cells were visualized using a RETIGA 1300 digital camera (QIMAGING) and a Zeiss AxioVert 135 microscope with a 63x objective. Images were analysed using Northern Eclipse version 6.0 (Empix Imaging).

Plasmid description: pXM/CDP (a.a 1 to 1505), pXM/CDP Δ 733-766 was and pXM/CDP 878-1505 constructs were described previously. also described previously (NS paper). To construct pXM/CDP 938-1505, pXM/CDP was digested with AflIII-NotI and pXM/Myc-HA (Ref MS paper) was digested with MluI and NotI. The two were ligated together. pXM/CDP Δ 918-938 was constructed using a pXM/Myc-HA vector digested

with XhoI-NotI ligated to the first insert from a XhoI-PvuI digest of pXM/CDP and a second insert from a AflIII-NotI digest of pXM/CDP. pTriEx 2.1/His CR1+2-NLS-HA was described previously (catL paper).

Cathepsin activity. Cathepsin activity was measured as previously described (60). Briefly, 3T3, ras-3T3 and MEF cat $L^{-/-}$ cells were labeled with 10^6CPM/ml of ^{125}I-JPM-OET for 1hr at 37^{O}C. Cells were washed with PBS and subjected to hypotonic lysis. Cytosolic and nuclear fractions from each time point were analyzed by SDS-PAGE. Gels were fixed, dried and analyzed using a Typhoon phosphor Imager. In some cases labeled extracts were subjected to immunoprecipitation using cathepsin L antibodies.

Results

CDP/Cux expression and DNA binding activity are increased in ras-transformed 3T3 cells. DNA binding by the CDP/Cux transcription factor was compared in ras-transformed and normal 3T3 cells. The CDP/Cux transcription factor was previously shown to display two DNA binding activities involving the full-length isoform, p200, and the processed isoform, p110 (17, 61). Electrophoretic mobility shift assays (EMSA) were performed using as a DNA probe the 29FP oligonucleotide, which contains binding sites for the two isoforms of CDP/Cux. As a control, EMSA were performed in parallel with oligonucleotides containing a binding site for the Oct transcription factor. Oct DNA binding activity was similar in the two cell lines (Fig. 1A, lanes 3 and 4). In contrast, DNA binding by the two main CDP/Cux isoforms was increased in ras-3T3 cells as compared to the parental 3T3 cell line (Fig. 1A, lanes 1 and 2). The two retarded complexes seen in lane 2 correspond to the p110 (1) and the p200 (2) isoforms, as demonstrated from supershift analysis with antibodies raised against distinct portions of CDP/Cux and competition with cold oligonucleotides. An antibody raised against the N-terminal region that is unique to the p200 isoform was able to supershift the retarded complex #2 to generate a slower migrating complex at the top of the gel, #5 (Fig. 1A, lane 6). However, the N-terminal antibody did not affect the migration of the p110 complex, #1 (Fig. 1A, lane 6). In contrast, antibodies raised against C-terminally located epitopes, α861 and α1300, were able to supershift both retarded complexes: complex #1 was supershifted into complex #3, and complex #2 was supershifted into complex #4 (Fig. 1A, compare lane 5 with lanes 7 and 8). An unrelated antibody raised against the glutathione-S-transferase (αGST) did not affect either retarded complex (Fig. 1A, lane 9). In agreement with the results obtained

with antibodies, the two retarded complexes were efficiently competed away with the same, cold, oligonucleotides (Fig. 1 A, lanes 10-13). The NFY-ATCG oligonucleotides, which contain an optimal binding site for p200 but a low affinity binding site for p110, competed away the #2 retarded complex at low concentration and the #1 retarded complex only at higher concentration (Fig. 1 A, lanes 14-17).

ras-3T3 cells express relatively higher level of short CDP/Cux proteins. From Western blot analysis, the ratio of full-length and truncated CDP/Cux isoforms was strikingly different in ras3T3 cells as compared to the parental 3T3 cells (Fig. 1B). The p110 isoform was more abundant in ras-3T3 cells, and an additional shorter species was also present in these cells (Fig. 1B). These results suggest either that processing of CDP/Cux is more efficient in ras-3T3 cells or that the processed CDP/Cux proteins are more stable in these cells. Notwithstanding its molecular basis, increased p110 expression in ras-3T3 cells is likely to be responsible for the elevated p110 DNA binding activity. In contrast, the steady-state level of the p200 isoform in the two cell lines did not correlate well with the p200 DNA binding activity (Fig. 1, compare A and B). We speculate that post-translational mechanisms may modulate p200 activity.

Processing of CDP/Cux is increased in ras-transformed 3T3 cells. The increase in steady-state level of the CDP/Cux processed isoform in ras-3T3 cells may alternatively result from accelerated synthesis or prolonged half-life of p110. To test these hypotheses, we independently assessed p110 steady-state level upon introduction of an exogenous p200 species, and we measured the half-life of a recombinant protein corresponding to p110. First, 3T3 and ras-3T3 cells were transfected with a vector, CDP, expressing the full length CDP/Cux protein with Myc and the influenza virus hematoglutinin (HA) epitope tags at its amino- and carboxy-termini, respectively. Nuclear extracts were then analyzed by

immunoblotting with the anti-HA antibody that recognizes the C-terminus. As seen in Fig. 2A, a higher level of the processed p110 isoform was observed in ras-3T3 cells. Secondly, a vector expressing a recombinant protein that corresponds to p110, CDP878, was introduced into each cell line, and the half-life of the protein was measured by pulse-chase analysis as described in Material and Methods. As seen in Fig. 2B, the half-life of CDP878 was not significantly different in the two cell lines. Altogether, results from pulse-chase labeling, transfection studies and endogenous CDP/Cux expression indicated that processing was increased in ras-3T3 cells.

Processing of CDP/Cux is increased in many tumor cell lines. To verify whether CDP/Cux processing was often activated in transformed cells, the CDP expression vector was introduced into a panel of tumor cell lines and processing was assessed by immunoblotting with the anti-HA antibody. Short CDP/Cux protein species were generated at higher levels in several breast tumor cell lines including SkBr3, Hs578T and MDA231 (Fig. 3A, left panel). In contrast, CDP/Cux processing was relatively weaker in human mammary epithelial cells (HMEC) and the TD47 and MCF7 cell lines, which are derived from breast tumors but have retained the capacity to differentiate *in vitro* and are thus considered less transformed than other breast tumor cell lines. CDP/Cux processing was also elevated in PANC1, 293 and HeLa cell lines but was reduced in MDCK, LoVoC5 and FR3T3 cells. Overall, elevated CDP/Cux processing was observed in many, but not all, tumor cell lines (Fig. 3A, right panel).

A shorter CDP/Cux processed species is generated in epithelial cells. In most cell lines of epithelial origin, we observed in addition to p110 another processed isoform with an apparent molecular mass of 95 kDa (Fig. 3A). Generation of the p95 isoform did not require the production of p110, as it was produced from a deletion mutant, Δ733-766, in

which the p110 processing site had been removed (Fig. 3B, lanes 1 and 2). The p95 isoform migrated closely to a recombinant protein whose amino-terminus corresponded to amino acid 938 of CDP (Fig. 3B, lanes 3 and 4). In accordance with this finding, the p95 isoform was not generated upon expression of a deletion mutant, Δ918-938, in which the corresponding amino acids had been deleted (Fig. 3B, lanes 5 and 6). The p95 isoform was generated upon co-expression of CDP with cathepsin L (Fig. 3C, lanes 2-4), and its expression was reduced following a 4-hour treatment with the cysteine protease inhibitor, E-64d (Fig. 3C, lanes 1 and 5). Altogether, these results indicate that that p95 is generated by a cathepsin L-dependent processing event that takes place in the region between amino acids 918 and 938. Intriguingly, however, cathepsin L is unable to generate the p95 isoform when CDP and cathepsin L are co-expressed in fibroblastic cells like NIH3T3 and FR3T3 (Goulet et al., unpublished observations; and data not shown). We speculate that the cleavage event responsible for the production of p95 is regulated by a cell-type specific post-translational modification.

Expression of CDP/Cux p110 is not regulated in a cell cycle-dependent manner in transformed cells. Processing of CDP/Cux was previously shown to be regulated during cell cycle progression in NIH3T3 cells (17). Processing was either very weak or non existent in G1, but became more intense as cells progressed into S phase. To verify whether expression of CDP/Cux p110 remained cell cycle-dependent in transformed cells, fractions of cells in different phases of the cell cycle were obtained by counterflow centrifugal elutriation. CDP/Cux expression was analyzed by immunoblotting and the DNA content of cells in each fraction was measured by fluorescence activated cell sorting (FACS). In NIH3T3 cells, little or no p110 was detected in the first two fractions which contain the smallest cells in G1 (Fig. 4, NIH panel). Expression of p110 became apparent

in fraction 3 and increased steadily from fraction up to fractions 6 and 7. Since p200 expression also increased in the same fractions, it appears that p110 expression is dependent, at least to some extent, on a threshold level of p200. However, upon forced expression of p200 the proteolytic processing event itself was found also to be cell cycle-regulated (17). In contrast, in the three transformed cell lines tested here, HeLa, ras-NIH3T3 and ras-FR3T3, expression of both p200 and p110 was relatively constant throughout the cell cycle (Fig. 4). We conclude that expression of CDP/Cux p110 is constitutive in transformed cells.

Ras-3T3 cells display elevated nuclear cathepsin L expression and activity. We have recently obtained evidence implicating a nuclear isoform of cathepsin L in the processing of CDP/Cux (Goulet et al., unpublished observations). Interestingly, cathepsin L is over-expressed in many transformed cells in tissue culture as well as in human cancers (reviewed in the introduction). We therefore considered the possibility that the increased processing of CDP/Cux in transformed cells might be related to the elevated expression of cathepsin L in these cells. Indeed, from Western blot analysis, cathepsin L was expressed at higher level in ras-3T3 cells than in the parental 3T3 cells (Fig. 5 A). To measure cathepsin L activity in specific cellular compartments, we used a recently described method to radioactively label cathepsins directly within cells (60). 3T3, ras-3T3 cells and cat $L^{-/-}$ mouse embryo fibroblasts ($MEF^{Cat\ L-/-}$) were labeled for one hour with a ^{125}Iodine-labeled E64 analog, ^{125}I-JPM-OET. Cytoplasmic and nuclear fractions were then prepared and analyzed by SDS-PAGE and autoradiography (Fig. 5B). A band of the expected molecular mass for the heavy chain of cathepsin L was observed in the cytoplasmic and nuclear fraction of both 3T3 and ras-3T3 cells, but was absent in the $MEF^{Cat\ L-/-}$ sample. This band was clearly more intense in the nuclear fraction of ras-3T3 cells than in the corresponding

fraction of 3T3 cells (Fig. 5B, lanes 1 and 2). To confirm that this band could represent cathepsin L, nuclear fractions were submitted to immunoprecipitation with a cathepsin L antibody prior to SDS-PAGE. Again, this band was present at higher intensity in the nuclear fraction of ras-3T3 cells than in the 3T3 sample, and importantly, was absent in the nuclear fraction of MEF cat $L^{-/-}$ cells (Fig. 5C). Thus, higher cathepsin L activity can be detected in the nuclear fraction of ras-3T3 cells than in 3T3 cells. As a third approach to analyze the expression of cathepsin L, indirect immunofluorescence was performed on 3T3 and ras-3T3 cells using antibodies raised against murine cathepsin L (41). Although this method can be considered semi-quantitative at best, we clearly observed a stronger signal in ras-3T3 cells with the anti-cathepsin L antibody (Fig. 5D). In summary, results from immunoblotting, cell labeling with a radiolabeled inhibitor, immunoprecipitation and indirect immunofluorescence confirm that cathepsin L is expressed at higher level in ras-3T3 cells. Importantly, a higher concentration of active cathepsin L appears to be present in ras-3T3 cells.

Forced expression of cathepsin L in 3T3 cells brings their capacity to process CDP/Cux close to that of ras-3T3 cells. If indeed the increased capacity of ras-3T3 cells to process CDP/Cux is due to their higher expression of cathepsin L, we would predict that over-expression of cathepsin L in normal 3T3 cells should bring their capacity to process CDP/Cux up to the level observed in ras-3T3 cells. To test this hypothesis, a small recombinant CDP/Cux protein, CR1+CR2, containing the site of processing fused to a nuclear localization signal was expressed into 3T3 and ras-3T3 cells in the presence or absence of cathepsin L expression vector. Following transient transfection, the CR1+CR2 protein was more efficiently cleaved in ras-3T3 than in 3T3 cells (Fig. 6A, lanes 1 and 2). Treatment of ras-3T3 cells with the cysteine protease inhibitor E-64d reduced cleavage of

CR1+CR2 (Fig. 6A, lanes 4 and 5). Co-expression of cathepsin L in 3T3 cells increased the cleavage of CR1+CR2, however, not quite to the level observed in ras-3T3 cells. As a second test, we employed an *in vitro* processing assay using a bacterially expressed CR1+CR2 protein and whole cell extracts as a source of protease. As a control to ensure that processing was similar *in vitro* and *in vivo*, the same vector, pTri-Ex-2, served for the expression of CR1+CR2 in bacteria and in mammalian cells. Indeed, peptides of similar mobility were generated *in vitro* and *in vivo* (Fig. 6B, compare lane 5 with lanes 2-4). Processing of CR1+CR2 was more efficient using extracts from ras-3T3 cells than from 3T3 cells (Fig. 6B, lanes 1 to 3). Interestingly, upon ectopic expression of cathepsin L in 3T3 cells, whole cell extracts from these cells were able to process CR1+CR2 as well if not better than the extracts from ras-3T3 cells (Fig. 6B, lanes 2 and 4). These results strongly suggest that the higher ability of transformed cells to process CDP/Cux is related to their increased expression of cathepsin L. However, the discrepancy between the results obtained in the *in vivo* and *in vitro* assays (compare Fig. 6A and 5B) suggests that cathepsin L expression level is not the only factor at play.

Discussion

The CDP/Cux p110 isoform is generated by a proteolytic processing event that takes place at the G1/S transition in normal cells (17). Whereas the full-length p200 isoform can only transiently bind to DNA and carry the CCAAT-displacement activity, CDP/Cux p110 is able to make a stable interaction with DNA and is involved in distinct regulatory transactions (17, 26, 61). In particular, CDP/Cux p110 is able to stimulate expression of the DNA pol α gene and of other genes whose expression is up-regulated in S phase (26). These findings suggest that CDP/Cux p110 might play a role in cell cycle progression in promoting the G1/S transition. Indeed, ectopic expression of CDP/Cux p110 was found to shorten passage through the G1 phase in NIH3T3 cells (Sansregret, manuscript in preparation). Interestingly, an increase in CDP/Cux p110 expression was previously reported in uterine leiomyomas, which represent the most frequent benign tumors in humans (39). These results raised the possibility that regulation of CDP/Cux expression might be altered in cancer cells. In the present study, we have use tissue culture systems to investigate CDP/Cux expression and activity in transformed cells. The DNA binding activity of CDP/Cux p110 was found to be elevated in several transformed cell line (Fig. 1A and 2C). When observed, the elevation in DNA binding activity correlated with an increase in the steady-state level of CDP/Cux p110 (Fig. 1B and 2D). Ectopic expression together with pulse-chase labeling experiments *in vivo* demonstrated that proteolytic processing of CDP/Cux is more efficient in ras-transformed 3T3 cells than in their normal counterparts (Fig. 2A and 2B). Moreover, whereas little or no processing was detected during the G1 phase in NIH3T3, constitutive processing was observed in three transformed cell lines (Fig. 4).

Altogether our results indicated that proteolytic processing of CDP/Cux is activated in many cancer cells. In parallel, we confirmed, as was previously reported, that cathepsin L expression and activity was increased in ras-transformed cells (48-51)(Fig. 5). Moreover, while previous studies reported the overall increase in cathepsin L expression as well as its secretion from cancer cells, our results revealed in addition that cathepsin L expression and activity were increased in the nucleus of ras-transformed cells (Fig. 5). The causal link between increased cathepsin L expression and CDP/Cux processing in cancer cells received support from the observation that ectopic expression of cathepsin L stimulated CDP/Cux processing (Fig. 6). However, in this experiment we noted that the stimulation in CDP/Cux processing was more pronounced when processing was assayed in an *in vitro* system than *in vivo* (compare results in Fig. 6 A and B). We conclude that the level of cathepsin L expression, albeit important, is not the only factor controlling the extent of CDP/Cux processing in cells. We consider likely that the regulated transport of cathepsin L to the nucleus, a mechanism that has yet to be investigated, plays an important role in this process. Alternatively, it is possible that post-translational modifications of CDP/Cux may modulate its processing.

The link between higher cathepsin L expression and the metastatic properties of cancer cells is well established. However, recent studies using cell permeable protease inhibitors have raised the possibility that intracellular functions of cathepsin L might also contribute to the tumorigenic phenotype. For example, inhibition of insulin-like growth factor-1 processing was reported to adversely affect receptor-mediated signaling (62). In another study using a multistage mouse model of cancer, the angiogenic properties of pancreatic islet tumours were differently affected by cell permeable and cell-impermeable

cysteine protease inhibitors, raising the possibility that intracellular activities of cathepsins might influence the angiogenic process (M. Bogyo and D. Hanahan, personal communication).

The recent discovery that cathepsin L isoforms devoid of a signal peptide can localize to the nucleus, and the identification of a nuclear substrate raises the interesting notion that the nuclear activities of cathepsin L might also promote the acquisition of a cancerous phenotype. We envisage that whereas extra-cellular cathepsin L can only contribute to the metastasis process, nuclear cathepsin L would have the ability to affect the transcriptional program of cells by changing the properties of transcriptional factors. In this manner, higher nuclear cathepsin L activity could potentially affect various stages of tumorigenesis including hyperplasia, angiogenesis, cell movement and invasion. The findings that CDP/Cux processing is often elevated in uterine leiomyomas, which are benign tumors of smooth muscle cell origin, suggests that cathepsin L might participate in the early steps of tumorigenesis (39).

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 (63-66). Proteolytic processing of transcription factors was found to redirect the localization of these proteins or to generate specific isoforms with distinct biochemical properties (66-71). The potential role of intra-cellular proteases in cancer is emerging from a number of studies (Fig. 7). Activation of v-Src was shown to increase the protein synthesis of calpain II, which in turn, is responsible for the increased cell migration and proliferation (72). Cyclin E, an important regulator of cell cycle progression was shown to be proteolytically processed into hyperactive truncated

isoforms in cancer cells (73, 74). Calpain was recently found to be the enzyme responsible for this cleavage (75). Oncogenic Ras was shown to increase levels and activity of the intracellular form of Notch-1 via up-regulation of Presenilin 1. From the studies cited above, we envisage that cancer is not only acquired via genetic changes but that epigenetic mechanisms also contribute to cancer promotion. As site-proteolysis has the potential to generate truncated protein isoforms with novel biochemical properties, we propose that aberrant expression or activity of intracellular proteases may lead to the activation of some proto-oncogenes and inactivation of some tumor suppressors.

Acknowledgements

A.N. is the recipient of a scholarship from the Fonds de la Recherche en Santé du Québec. B. G. is the recipient of a fellowship from the Research Institute of the McGill University Health Center. This research was supported by grant #011303 from the Canadian Breast Cancer Research Initiative to A.N.

References

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

2. Jack, J. W. Molecular organization of the cut locus of Drosophila melanogaster. Cell, *42:* 869-876, 1985.

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

4. Jack, J. and DeLotto, Y. Effect of wing scalloping mutations on cut expression and sense organ differentiation in the Drosophila wing margin. Genetics, *131:* 353-363, 1992.

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

6. Liu, S. and Jack, J. Regulatory interactions and role in cell type specification of the Malpighian tubules by the cut, Kruppel, and caudal genes of Drosophila. Developmental Biology, *150:* 133-143, 1992.

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

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

9. Braun, W. The effect of puncture on the developing wing of several mutants of *Drosophila melanogaster.* J. Exp. Zool., *84:* 325-350, 1942.

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

11. Ellis, T., Gambardella, L., Horcher, M., Tschanz, S., Capol, J., Bertram, P., Jochum, W., Barrandon, Y., and Busslinger, M. The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev, *15:* 2307-2319., 2001.

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

13. Luong, M. X., van der Meijden, C. M., Xing, D., Hesselton, R., Monuki, E. S., Jones, S. N., Lian, J. B., Stein, J. L., Stein, G. S., Neufeld, E. J., and van Wijnen, A. J. Genetic ablation of the CDP/Cux protein C terminus results in hair cycle defects and reduced male fertility. Mol Cell Biol, *22:* 1424-1437, 2002.

14. Tufarelli, C., Fujiwara, Y., Zappulla, D. C., and Neufeld, E. J. 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, 1998.

15. Ledford, A. W., Brantley, J. G., Kemeny, G., Foreman, T. L., Quaggin, S. E., Igarashi, P., Oberhaus, S. M., Rodova, M., Calvet, J. P., and Vanden Heuvel, G. B. Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia. Dev Biol, *245:* 157-171, 2002.

16. Zeng, R. W., Soucie, E., Sung Moon, N., Martin-Soudant, N., Bérubé, G., Leduy, L., and Nepveu, A. Exon/intron structure and alternative transcripts of the CUTL1 gene. Gene, *241:* 75-85, 2000.

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

18. Lievens, P. M., Tufarelli, C., Donady, J. J., Stagg, A., and Neufeld, E. J. CASP, a novel, highly conserved alternative-splicing product of the CDP/cut/cux gene, lacks cut-repeat and homeo DNA-binding domains, and interacts with full-length CDP in vitro. Gene, *197:* 73-81, 1997.

19. Goulet, B., Watson, P., Poirier, M., Leduy, L., Berube, G., Meterissian, S., Jolicoeur, P., and Nepveu, A. Characterization of a tissue-specific CDP/Cux isoform, p75, activated in breast tumor cells. Cancer Research, *62:* 6625-6633, 2002.

20. van Wijnen, A. J., Wright, K. L., Lian, J. B., Stein, J. L., and Stein, G. S. 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:* 15034-15042., 1989.

21. Holthuis, J., Owen, T. A., van Wijnen, A. J., Wright, K. L., Ramsey-Ewing, A., Kennedy, M. B., Carter, R., Cosenza, S. C., Soprano, K. J., Lian, J. B., Stein, J. L., and Stein, G. S. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science, *247:* 1454-1457., 1990.

22. van Wijnen, A. J., Choi, T. K., Owen, T. A., Wright, K. L., Lian, J. B., Jaenisch, R., Stein, J. L., and Stein, G. S. 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:* 2573-2577., 1991.

23. Wright, K. L., Dell'Orco, R. T., van Wijnen, A. J., Stein, J. L., and Stein, G. S. Multiple mechanisms regulate the proliferation-specific histone gene transcription factor HiNF-D in normal human diploid fibroblasts. Biochemistry, *31:* 2812-2818., 1992.

24. van Wijnen, A. J., van Gurp, M. F., de Ridder, M. C., Tufarelli, C., Last, T. J., Birnbaum, M., Vaughan, P. S., Giordano, A., Krek, W., Neufeld, E. J., Stein, J. L., and Stein, G. S. 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, *93:* 11516-11521, 1996.

25. Coqueret, O., Berube, G., and 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, *17:* 4680-4694, 1998.

26. Truscott, M., Raynal, L., Premdas, P., Goulet, B., Leduy, L., Berube, G., and Nepveu, A. CDP/Cux stimulates transcription from the DNA polymerase alpha gene promoter. Molecular and Cellular Biology, *23:* 3013-3028, 2003.

27. Santaguida, M., Ding, Q., Berube, G., Truscott, M., Whyte, P., and Nepveu, A. Phosphorylation of the CCAAT displacement protein (CDP)/Cux transcription

factor by cyclin A-Cdk1 modulates its DNA binding activity in G(2). Journal of Biological Chemistry, *276:* 45780-45790, 2001.

28. Lemieux, N., Zhang, X. X., Dufort, D., and Nepveu, A. Assignment of the human homologue of the Drosophila Cut homeobox gene (CUTL1) to band 7q22 by fluorescence in situ hybridization. Genomics, *24:* 191-193, 1994.

29. Scherer, S. W., Neufeld, E. J., Lievens, P. M., Orkin, S. H., Kim, J., and Tsui, L. C. Regional localization of the CCAAT displacement protein gene (CUTL1) to 7q22 by analysis of somatic cell hybrids. Genomics, *15:* 695-696, 1993.

30. Berkerkarauzum, S., Luleci, G., Ozbilim, G., Erdogan, A., Kuzucu, A., and Demircan, A. Cytogenetic Findings In Thirty Lung Carcinoma Patients. Cancer Genetics & Cytogenetics, *100:* 114-123, 1998.

31. Fischer, K., Brown, J., Scherer, S. W., Schramm, P., Stewart, J., Fugazza, G., Pascheberg, U., Peter, W., Tsui, L. C., Lichter, P., and Dohner, H. Delineation of genomic regions in chromosome band 7q22 commonly deleted in myeloid leukemias. Recent Results in Cancer Research, *144:* 46-52, 1998.

32. Ishwad, C. S., Ferrell, R. E., Hanley, K., Davare, J., Meloni, A. M., Sandberg, A. A., and Surti, U. Two discrete regions of deletion at 7q in uterine leiomyomas. Genes, Chromosomes & Cancer, *19:* 156-160, 1997.

33. Lundgren, R., Mandahl, N., Heim, S., Limon, J., Henrikson, H., and Mitelman, F. Cytogenetic analysis of 57 primary prostatic adenocarcinomas. Genes, Chromosomes & Cancer, *4:* 16-24, 1992.

34. Mertens, F., Johansson, B., Hoglund, M., and Mitelman, F. Chromosomal imbalance maps of malignant solid tumors: a cytogenetic survey of 3185 neoplasms. Cancer Research, *57:* 2765-2780, 1997.

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

36. Zeng, W. R., Scherer, S. W., Koutsilieris, M., Huizenga, J. J., Filteau, F., Tsui, L. C., and Nepveu, A. Loss Of Heterozygosity and Reduced Expression Of the Cutl1 Gene In Uterine Leiomyomas. Oncogene, *14:* 2355-2365, 1997.

37. Zeng, W. R., Watson, P., Lin, J., Jothy, S., Lidereau, R., Park, M., and Nepveu, A. 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:* 2015-2021, 1999.

38. Webster, M. A., Martinsoudant, N., Nepveu, A., Cardiff, R. D., and Muller, W. J. 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:* 1963-1972, 1998.

39. Moon, N. S., Rong Zeng, W., Premdas, P., Santaguida, M., Berube, G., and Nepveu, A. Expression of N-terminally truncated isoforms of CDP/CUX is increased in human uterine leiomyomas. International Journal of Cancer, *100:* 429-432, 2002.

40. Neville, P. J., Thomas, N., and Campbell, I. G. Loss of heterozygosity at 7q22 and mutation analysis of the CDP gene in human epithelial ovarian tumors. Int J Cancer, *91:* 345-349, 2001.

41. Portnoy, D. A., Erickson, A. H., Kochan, J., Ravetch, J. V., and Unkeless, J. C. Cloning and characterization of a mouse cysteine proteinase. J Biol Chem, *261:* 14697-14703., 1986.

42. Troen, B. R., Ascherman, D., Atlas, D., and Gottesman, M. M. Cloning and expression of the gene for the major excreted protein of transformed mouse fibroblasts. A secreted lysosomal protease regulated by transformation. J Biol Chem, *263:* 254-261., 1988.

43. Turk, B., Turk, V., and Turk, D. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. Biological Chemistry, *378:* 141-150, 1997.

44. Erickson, A. H. Biosynthesis of lysosomal endopeptidases. J Cell Biochem, *40:* 31-41., 1989.

45. Kane, S. E. and Gottesman, M. M. The role of cathepsin L in malignant transformation. Semin Cancer Biol, *1:* 127-136, 1990.

46. Krepela, E. Cysteine proteinases in tumor cell growth and apoptosis. Neoplasma, *48:* 332-349, 2001.

47. Lah, T. T. and Kos, J. Cysteine proteinases in cancer progression and their clinical relevance for prognosis. Biol Chem, *379:* 125-130, 1998.

48. Troen, B. R., Chauhan, S. S., Ray, D., and Gottesman, M. M. Downstream sequences mediate induction of the mouse cathepsin L promoter by phorbol esters. Cell Growth Differ, *2:* 23-31., 1991.

49. Lemaire, R., Flipo, R. M., Monte, D., Dupressoir, T., Duquesnoy, B., Cesbron, J. Y., Janin, A., Capron, A., and Lafyatis, R. Synovial fibroblast-like cell transfection with the SV40 large T antigen induces a transformed phenotype and permits

transient tumor formation in immunodeficient mice. J Rheumatol, *21:* 1409-1419., 1994.

50. Heinrich, J., Bosse, M., Eickhoff, H., Nietfeld, W., Reinhardt, R., Lehrach, H., and Moelling, K. Induction of putative tumor-suppressing genes in Rat-1 fibroblasts by oncogenic Raf-1 as evidenced by robot-assisted complex hybridization. J Mol Med, *78:* 380-388, 2000.

51. Joseph, L., Lapid, S., and Sukhatme, V. The major ras induced protein in NIH3T3 cells is cathepsin L. Nucleic Acids Res, *15:* 3186., 1987.

52. Denhardt, D. T., Greenberg, A. H., Egan, S. E., Hamilton, R. T., and Wright, J. A. Cysteine proteinase cathepsin L expression correlates closely with the metastatic potential of H-ras-transformed murine fibroblasts. Oncogene, *2:* 55-59, 1987.

53. Herszenyi, L., Plebani, M., Carraro, P., De Paoli, M., Roveroni, G., Cardin, R., Tulassay, Z., Naccarato, R., and Farinati, F. The role of cysteine and serine proteases in colorectal carcinoma. Cancer, *86:* 1135-1142., 1999.

54. Herszenyi, L., Plebani, M., Carraro, P., De Paoli, M., Roveroni, G., Rugge, M., Cardin, R., Naccarato, R., and Farinati, F. [Role and behavior of cathepsin B and cathepsin L in gastric cancer]. Orv Hetil, *136:* 1315-1318., 1995.

55. Strojnik, T., Zidanik, B., Kos, J., and Lah, T. T. Cathepsins B and L are markers for clinically invasive types of meningiomas. Neurosurgery, *48:* 598-605., 2001.

56. Dohchin, A., Suzuki, J. I., Seki, H., Masutani, M., Shiroto, H., and Kawakami, Y. Immunostained cathepsins B and L correlate with depth of invasion and different metastatic pathways in early stage gastric carcinoma. Cancer, *89:* 482-487., 2000.

57. Park, I. C., Lee, S. Y., Jeon, D. G., Lee, J. S., Hwang, C. S., Hwang, B. G., Lee, S. H., Hong, W. S., and Hong, S. I. Enhanced expression of cathepsin L in metastatic bone tumors. J Korean Med Sci, *11:* 144-148., 1996.

58. Nishida, Y., Kohno, K., Kawamata, T., Morimitsu, K., Kuwano, M., and Miyakawa, I. Increased cathepsin L levels in serum in some patients with ovarian cancer: comparison with CA125 and CA72-4. Gynecol Oncol, *56:* 357-361., 1995.

59. Chauhan, S. S., Goldstein, L. J., and Gottesman, M. M. Expression of cathepsin L in human tumors. Cancer Res, *51:* 1478-1481., 1991.

60. Bogyo, M., Verhelst, S., Bellingard-Dubouchaud, V., Toba, S., and Greenbaum, D. Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. Chem Biol, *7:* 27-38, 2000.

61. Moon, N. S., Berube, G., and Nepveu, A. CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain. Journal of Biological Chemistry, *275:* 31325-31334, 2000.

62. Navab, R., Chevet, E., Authier, F., Di Guglielmo, G. M., Bergeron, J. J., and Brodt, P. Inhibition of endosomal insulin-like growth factor-I processing by cysteine proteinase inhibitors blocks receptor-mediated functions. J Biol Chem, *276:* 13644-13649, 2001.

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

64. Gervais, F. G., Xu, D., Robertson, G. S., Vaillancourt, J. P., Zhu, Y., Huang, J., LeBlanc, A., Smith, D., Rigby, M., Shearman, M. S., Clarke, E. E., Zheng, H., Van Der Ploeg, L. H., Ruffolo, S. C., Thornberry, N. A., Xanthoudakis, S., Zamboni, R.

J., Roy, S., and Nicholson, D. W. Involvement of caspases in proteolytic cleavage of Alzheimer's amyloid- beta precursor protein and amyloidogenic A beta peptide formation. Cell, *97:* 395-406, 1999.

65. Chan, Y. M. and Jan, Y. N. Roles for proteolysis and trafficking in notch maturation and signal transduction. Cell, *94:* 423-426, 1998.

66. Aza-Blanc, P., Ramirez-Weber, F. A., Laget, M. P., Schwartz, C., and Kornberg, T. B. Proteolysis that is inhibited by hedgehog targets Cubitus interruptus protein to the nucleus and converts it to a repressor. Cell, *89:* 1043-1053, 1997.

67. Welm, A. L., Timchenko, N. A., and Darlington, G. J. C/EBPalpha regulates generation of C/EBPbeta isoforms through activation of specific proteolytic cleavage. Molecular & Cellular Biology, *19:* 1695-1704, 1999.

68. Vogel, J. L. and Kristie, T. M. Autocatalytic proteolysis of the transcription factor-coactivator C1 (HCF): a potential role for proteolytic regulation of coactivator function. Proceedings of the National Academy of Sciences of the United States of America, *97:* 9425-9430, 2000.

69. Haze, K., Yoshida, H., Yanagi, H., Yura, T., and Mori, K. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. Molecular & Cellular Biology, *10:* 3787-3799, 1999.

70. Brown, M. S. and Goldstein, J. L. The SREBP pathway: regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. Cell, *89:* 331-340, 1997.

71. Fan, C. M. and Maniatis, T. Generation of p50 subunit of NF-kappa B by processing of p105 through an ATP-dependent pathway. Nature, *354:* 395-398, 1991.

72. Carragher, N. O., Westhoff, M. A., Riley, D., Potter, D. A., Dutt, P., Elce, J. S., Greer, P. A., and Frame, M. C. v-Src-induced modulation of the calpain-calpastatin proteolytic system regulates transformation. Mol Cell Biol, *22:* 257-269, 2002.

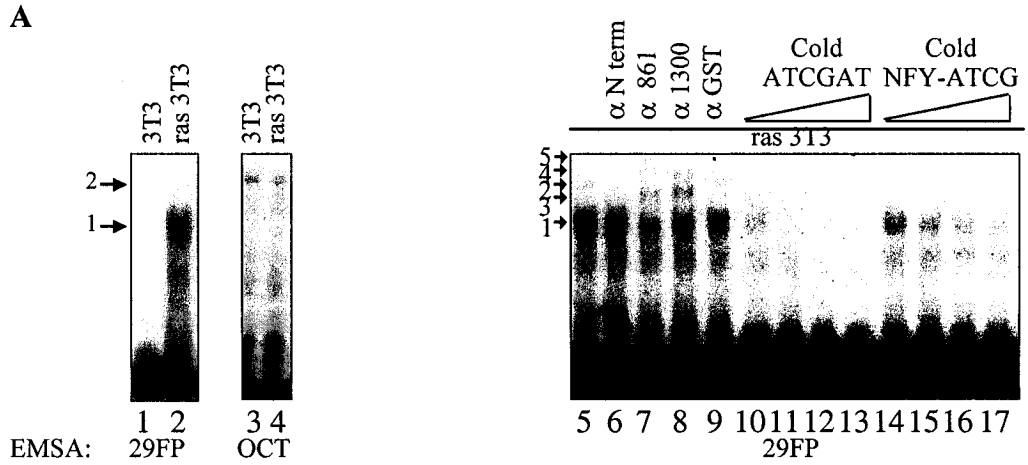
73. Porter, D. C., Zhang, N., Danes, C., McGahren, M. J., Harwell, R. M., Faruki, S., and Keyomarsi, K. Tumor-specific proteolytic processing of cyclin E generates hyperactive lower-molecular-weight forms. Mol Cell Biol, *21:* 6254-6269, 2001.

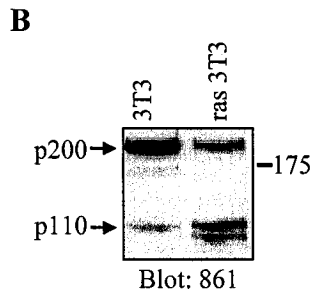
74. Harwell, R. M., Porter, D. C., Danes, C., and Keyomarsi, K. Processing of cyclin E differs between normal and tumor breast cells. Cancer Res, *60:* 481-489, 2000.

75. Wang, X. D., Rosales, J. L., Magliocco, A., Gnanakumar, R., and Lee, K. Y. Cyclin E in breast tumors is cleaved into its low molecular weight forms by calpain. Oncogene, *22:* 769-774, 2003.

Figure 1: Increased CDP/Cux DNA binding activity in ras-transformed 3T3 cells.

(A) Nuclear extracts (5mg) from 3T3 and ras-transformed 3T3 cells were incubated 15 min at room temperature with radiolabeled oligonucleotides containing the indicated binding sites: ATCGAT or OCT. Retarded complexes were resolved on a 4% non-denaturing polyacrylamide Tris-glycine gel and revealed by autoradiography. For supershift experiments, nuclear extracts of ras-3T3 (5 μg) were incubated 30 min at room temperature with the indicated antibodies prior to the incubation with oligonucleotides. Competition experiments were performed by adding increasing amounts (10 ng, 50 ng, 100 ng and 200 ng) of cold oligonucleotides containing the indicated sequence (see Material & Methods). The CDP/Cux retarded complexes are indicated by the numbered arrows: 1, p110; 2, p200; 3, supershift of p110 with α861 or α1300; 4, supershift of p200 with α861 or α1300; 5, supershift of p200 with αN-term. (B) Nuclear extracts were prepared from 3T3 and ras-3T3, and analyzed immunoblotting with anti-CDP/Cux α861 antibody. (C) Schematic representation of the full-length (p200) and proteolytically processed (p110) CDP/Cux isoforms. The regions recognized by each antibody are indicated.


A
3T3
ras 3T3
3T3
ras 3T3
2
1
1 2
3 4
EMSA:
29FP
OCT
α N term
α 861
α 1300
α GST
Cold
ATCGAT
Cold
NFY-ATCG
ras 313
5
4
3
2
3
1
5 6 7 8 9 10 11 12 13 14 15 16 17
29FP


B
3T3
ras 3T3
p200
p110
175
Blot: 861

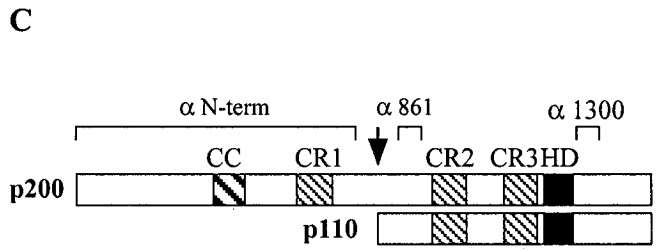
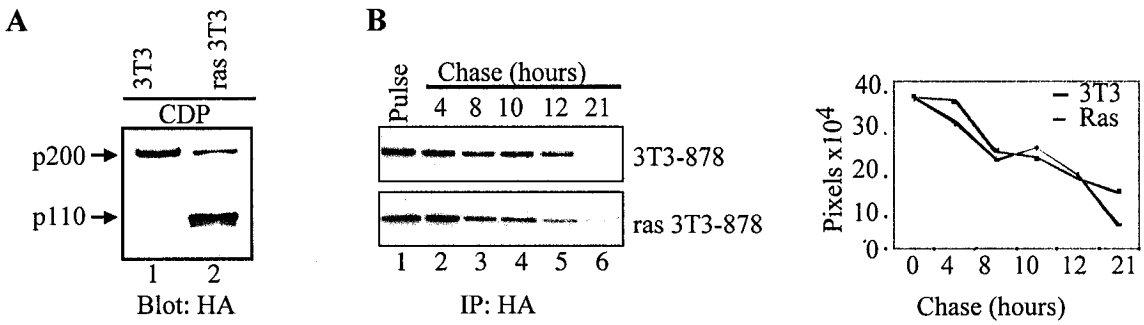

C
α N-term
α 861
α 1300
CC
CR1
CR2
CR3HD
p200
p110

Figure 2: Increased CDP/Cux processing in tumor cell lines.

(A) 3T3 and ras-3T3 cells were transfected with a vector expressing a CDP/Cux protein containing Myc and HA epitope tags at its amino- and carboxy termini, respectively (CDP: Myc-Cux-HA). Nuclear extracts were prepared and analyzed by immunoblotting with an anti-HA antibody. (B) A vector coding for an HA-tagged recombinant CDP/Cux p110 protein, 878, was introduced into 3T3 and ras-3T3 cells. 24 hours later, cells were labeled for 4 hours with ^{35}S-methionine. The medium was replaced for complete medium with cold methionine (chase). At the indicated time, total extracts were prepared, immunoprecipitated with anti-HA antibodies. Proteins were resolved on a SDS-polyacrylamide gel and detected by autoradiography. (C) Schematic representation of the recombinant proteins used in this figure and the proteins generated by proteolytic processing.


A
3T3
ras 3T3
CDP
p200
p110
1 2
Blot: HA
B
Pulse
Chase (hours)
4 8 10 12 21
3T3-878
ras 3T3-878
1 2 3 4 5 6
IP: HA
Pixels x10^4
40.
30.
20.
10.
0
3T3
Ras
0 4 8 10 12 21
Chase (hours)

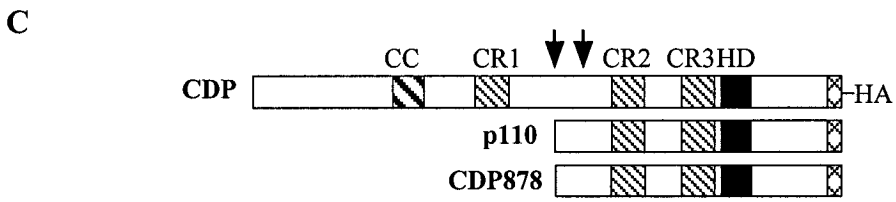

C
CC
CR1
CR2
CR3HD
CDP
HA
p110
CDP878

Figure 3: Processing of CDP/Cux is increased in many tumor cell lines.

(A) The indicated cell lines cells were transfected with the Myc-Cux-HA vector (CDP) and nuclear extracts were analyzed by immunoblotting with an anti-HA antibody. (B) Hs578T breast tumor cells were transfected with the indicated CDP/Cux expression vectors. Nuclear extracts were prepared and analyzed in Western blots with anti-HA. (C) CV-1 cells were transfected with a vector expressing CDP, and were treated or not for four hours with the cysteine protease inhibitor E-64d. Nuclear extracts were analyzed as mentioned above. (D) Schematic representation of the recombinant proteins used in this figure and the proteins generated by proteolytic processing.

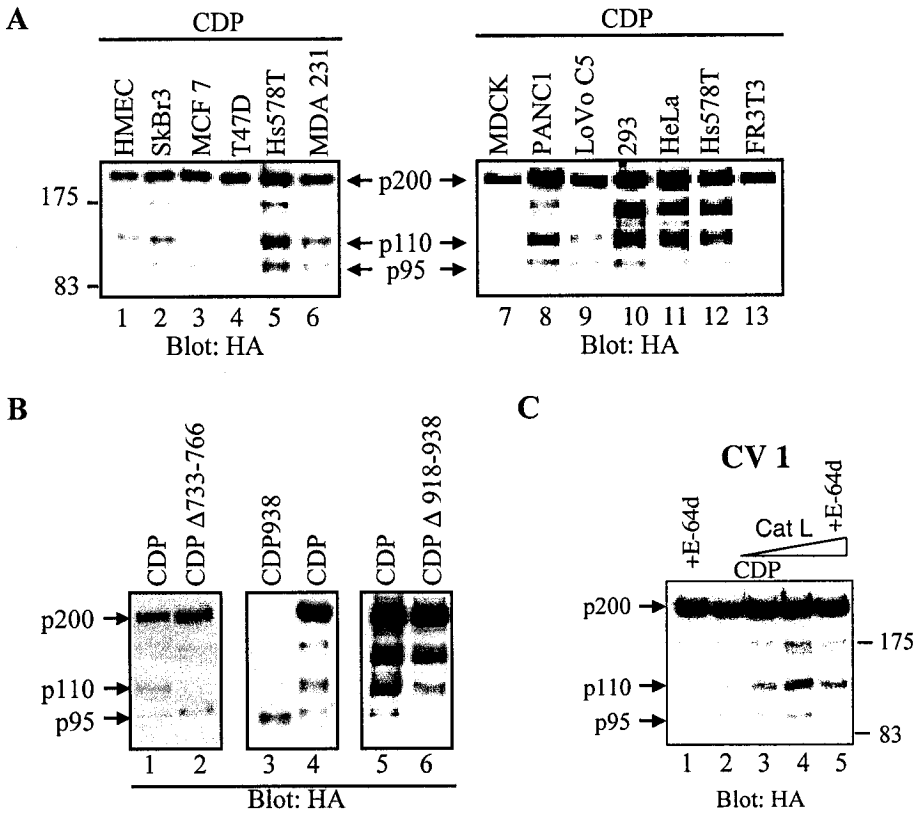

A
CDP
CDP
HMEC
SkBr3
MCF 7
T47D
Hs578T
MDA 231
MDCK
PANC1
LoVo C5
293
HeLa
Hs578T
FR3T3
175
83
p200
p110
p95
1 2 3 4 5 6
7 8 9 10 11 12 13
Blot: HA
Blot: HA
B
CDP
CDP Δ733-766
CDP938
CDP
CDP
CDP Δ 918-938
p200
p110
p95
1 2 3 4 5 6
Blot: HA
C
CV 1
+E-64d
Cat L
+E-64d
CDP
p200
p110
p95
175
83
1 2 3 4 5
Blot: HA
D
CC
CR1
CR2
CR3HD
CDP
HA
p110
p95
CDP Δ733-766
CDP938
CDP Δ918-938

Figure 4: CDP/Cux processing into p110 is not regulated in a cell cycle-dependent manner in transformed cells.

Approximately $2x10^8$ continuously cycling cells were subjected to counterflow centrifugal elutriation to obtain fractions of cells in different phases of the cell cycle. Cells from each fraction were analyzed for their DNA content by fluorescence activated cell sorting (FACS) and nuclear extracts were analyzed by immunoblotting with the anti-CDP/Cux α861 antibody.

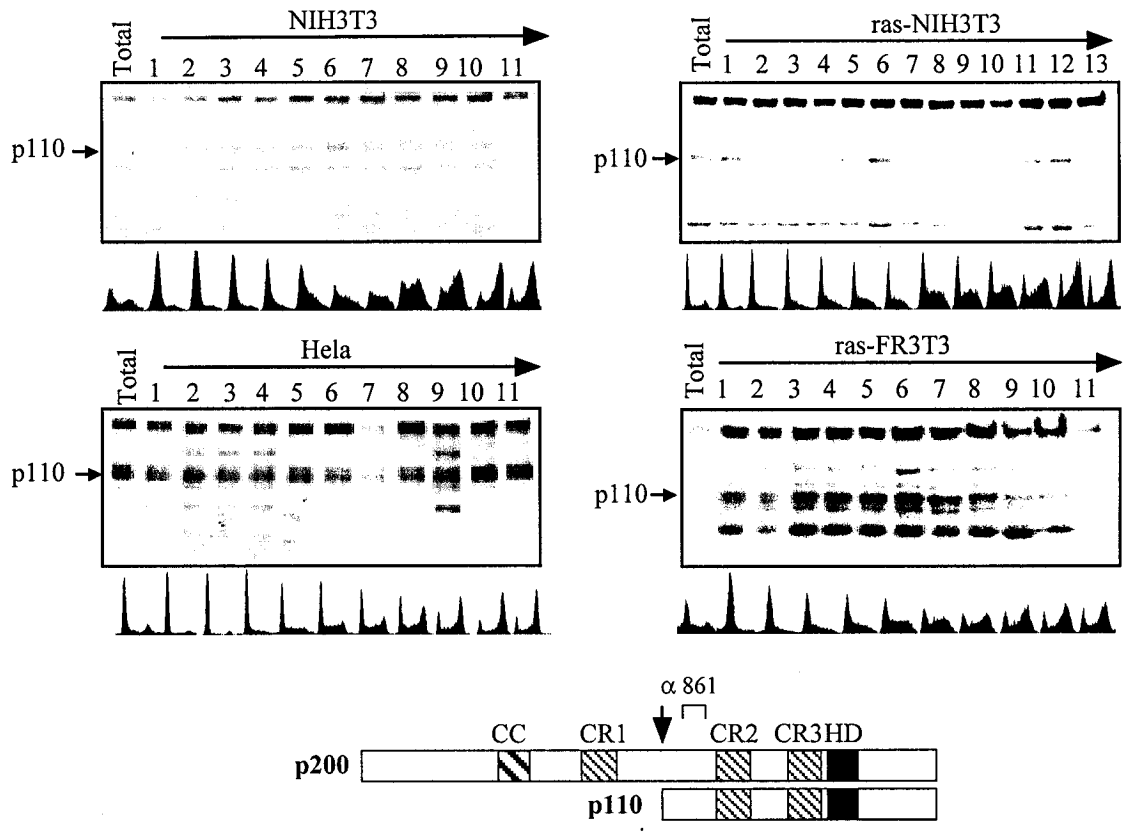

NIH3T3
Total
1 2 3 4 5 6 7 8 9 10 11
p110
ras-NIH3T3
Total
1 2 3 4 5 6 7 8 9 10 11 12 13
p110
Hela
Total
1 2 3 4 5 6 7 8 9 10 11
p110
ras-FR3T3
Total
1 2 3 4 5 6 7 8 9 10 11
p110
α 861
CC
CR1
CR2
CR3HD
p200
p110

Figure 5: Ras-transformed cells exhibit increased cathepsin L expression and activity. (A) Total extracts from 3T3 and ras-3T3 cells (25 μg) were resolved on an SDS-polyacrylamide gel and analyzed by immunoblotting with a cathepsin L antibody. As a loading control, the same membrane was immunoblotted with an actin antibody (below). (B) 3T3, ras-3T3 and MEF cat $L^{-/-}$ cells were incubated with 10^6 counts/ml ^{125}I-JPM-OET for 1 hour at 37°C. Cells were subsequently lysed in hypotonic buffer and cytoplasmic and nuclear fractions were prepared as described in the materials and methods. Proteins (50 μg/lane) were resolved on a 12% SDS-polyacrylamide and revealed by autoradiography. (C) ^{125}I-JPM-OET labeled nuclear extracts were subjected to immunoprecipitation using an anti mouse cathepsin L antibody. A Schematic representation of cathepsin L isoforms is shown below. (D) 3T3 and ras-3T3 cells were plated on a glass cover-slip and indirect immunofluorescence was performed using the anti-mouse cathepsin L antibody (top panel). Dapi staining of cell nuclei (bottom panel).

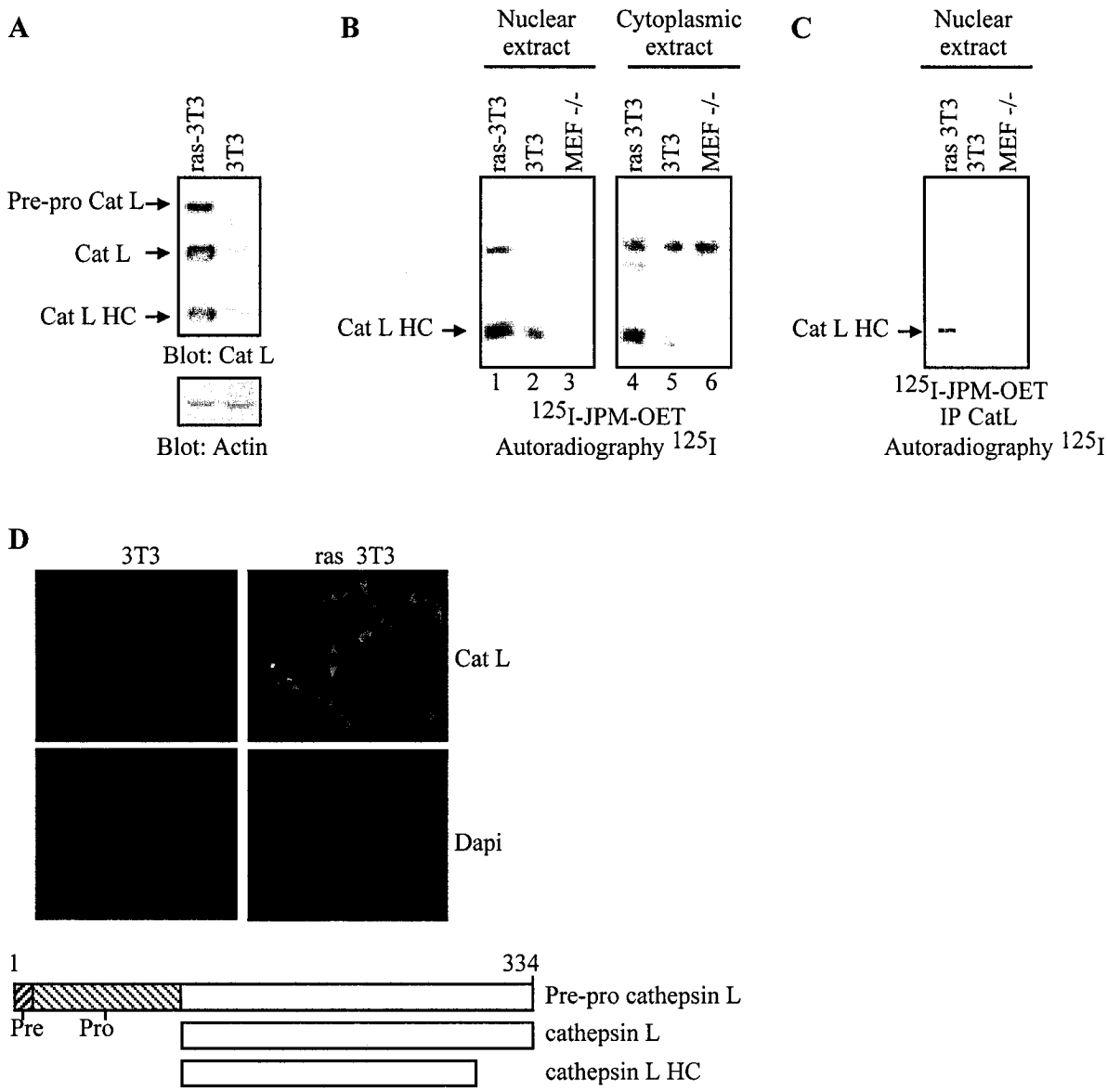
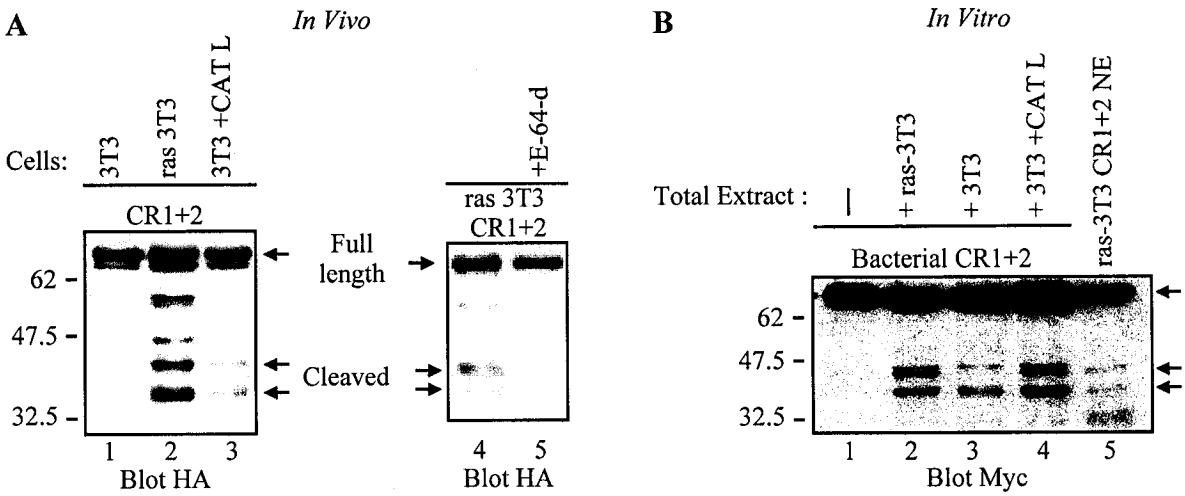

A
ras-3T3
3T3
Pre-pro Cat L
Cat L
Cat L HC
Blot: Cat L
Blot: Actin
B
Nuclear extract
Cytoplasmic extract
ras-3T3
3T3
MEF -/-
ras 3T3
3T3
MEF -/-
Cat L HC
1 2 3 4 5 6
125I-JPM-OET
Autoradiography 125I
C
Nuclear extract
ras 3T3
3T3
MEF -/-
Cat L HC
125I-JPM-OET
IP CatL
Autoradiography 125I
D
3T3
ras 3T3
Cat L
Dapi
1
334
Pre
Pro
Pre-pro cathepsin L
cathepsin L
cathepsin L HC

Figure 6: CDP/Cut is more efficiently processed *in vitro* using extracts from ras-3T3 cells as a source of protease.

The same vector, pTri-Ex-2, served to expression the CR1+CR2 region of CDP/Cux in bacteria and in mammalian cells (see diagram in C). (A) 3T3 and ras-3T3 were transfected with the CR1+2 expression vector. In lane 3, cells also received a vector expressing cathepsin L. In lane 5, cells were treated for 4 hours with 50 μM of the E-64d cysteine protease inhibitor prior to harvesting. Nuclear extracts were prepared and analyzed by immunoblotting with an anti-HA antibody. (B) 250 ng of bacterially purified His tagged CR1+2 protein were incubated 20 min at 37°C with 25 μg of total extracts from 3T3, ras-3T3 and 3T3 cells transfected with a cathepsin L expression vector (lanes 2, 3 and 4). Untreated CR1+CR2 is shown in lane 1. As a control for processing, lane 5 received nuclear extracts from ras-3T3 cells transfected with the CR1+CR2 vector. Proteins were then resolved on a SDS-polyacrylamide gel and immumobloted with αHA monoclonal antibody. (C) Schematic representation of CDP/Cux and the His- and HA-tagged internal CR1+CR2 construct.


A
In Vivo
Cells:
3T3
ras 3T3
3T3 +CAT L
+E-64-d
CR1+2
ras 3T3
CR1+2
62 -
47.5 -
32.5 -
Full length
Cleaved
1 2 3
Blot HA
4 5
Blot HA
B
In Vitro
Total Extract :
-
+ ras-3T3
+ 3T3
+ 3T3 +CAT L
ras-3T3 CR1+2 NE
Bacterial CR1+2
62 -
47.5 -
32.5 -
1 2 3 4 5
Blot Myc

C

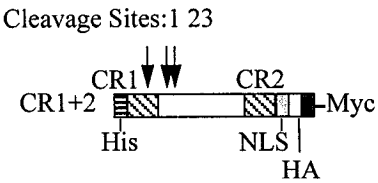
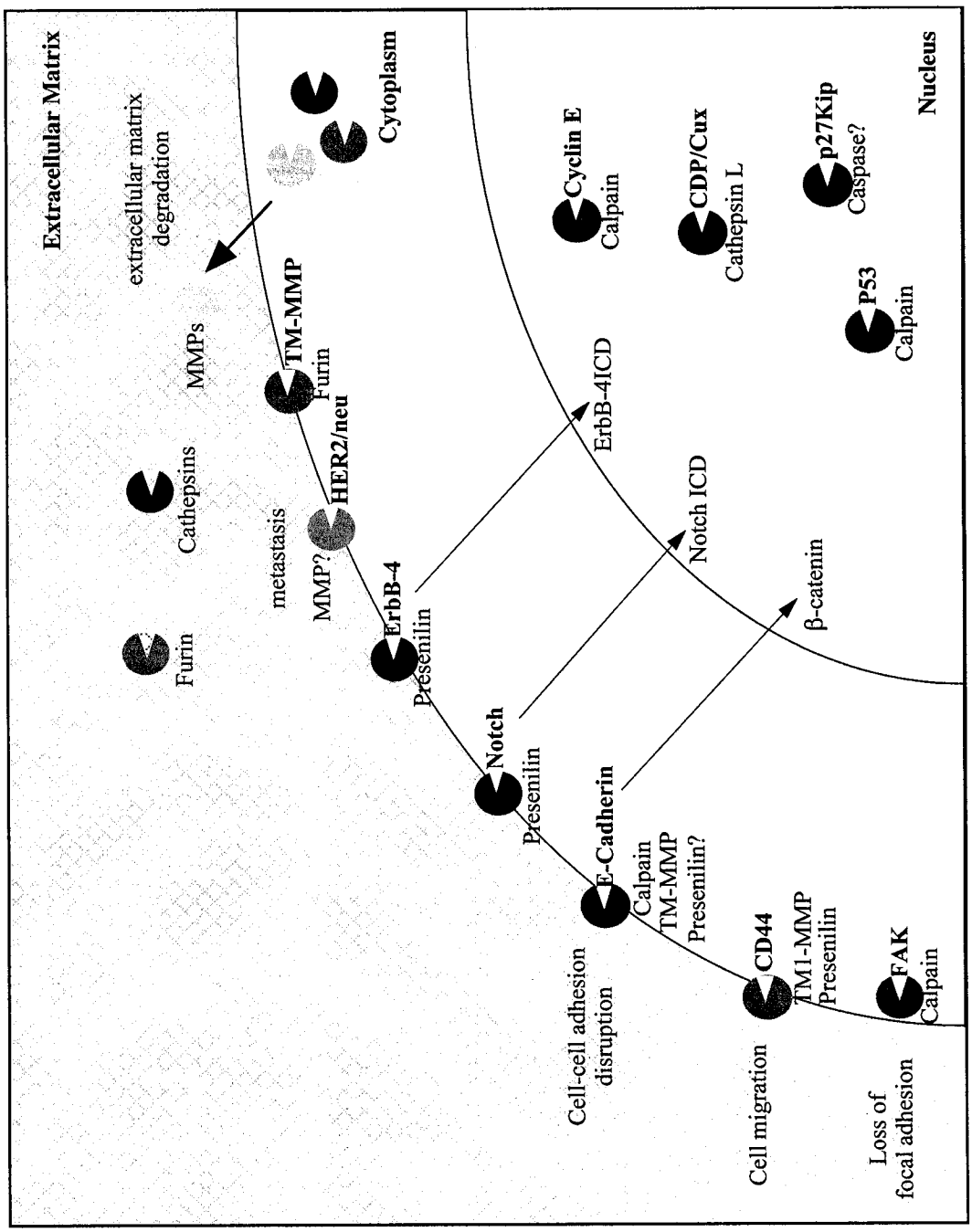

Cleavage Sites:1 23
CR1
CR2
CR1+2
Myc
His
NLS
HA

Figure 7: Intracellular proteases in cancer

The figure illustrates the intracellular proteases that have been reported to play a role in cancer.


Extracellular Matrix
extracellular matrix degradation
MMPs
Cathepsins
Furin
metastasis
MMP?
TM-MMP
Furin
HER2/neu
ErbB-4
Presenilin
Cytoplasm
Notch
Presenilin
E-Cadherin
Calpain
TM-MMP
Presenilin?
CD44
TM1-MMP
Presenilin
FAK
Calpain
Cell-cell adhesion disruption
Cell migration
Loss of focal adhesion
ErbB-4ICD
Notch ICD
β-catenin
Cyclin E
Calpain
CDP/Cux
Cathepsin L
P53
Calpain
p27Kip
Caspase?
Nucleus

Chapter IV Characterization of a Tissue-Specific CDP/Cux Isoform, p75, Activated in Breast Tumor Cells

Preface

While studying the proteolytic processing of CDP/Cux, RNase protection assays were performed using different probes to exclude the possibility that the truncated isoform results from the generation of a novel mRNA. A protected fragment corresponding to the size of exon 21 was detected using a probe containing exons 19 to 21, but no novel mRNA encoding for p110 was found. Further experiments led to the identification of a novel mRNA that is initiated within intron 20 of CUTL1. Importantly, the transcription of this mRNA is normally regulated in a tissue- and cell-specific manner but aberrantly transcribed in breast cancer cells.

CUTL1 is a large gene that codes for a protein containing four DNA binding domains, three Cut repeats and a homeodomain. Proteolytic processing generates a protein with two Cut repeats and the homeodomain. Other genes of the Cut homeodomain family encodes for proteins containing one, two or three Cut repeats, which could confer different DNA binding properties and biological outputs. This manuscript described the identification of the intron20-mRNA, the biochemical properties of p75, a novel short CDP/cux isoform that contains only one Cut repeat and the homeodomain and some its potential biological activities.

Abstract

Two isoforms of the CDP/Cux transcription factor have been characterized so far. The full-length protein, p200, contains four DNA binding domains, transiently binds to DNA and carries the CCAAT-displacement activity. The p110 isoform is generated by proteolytic processing at the G1/S transition and is capable of stable interaction with DNA. Here we demonstrate the existence of a shorter CDP/Cux isoform, p75, which contains only two DNA binding domains, Cut repeat 3 and the Cut homeodomain, and binds more stably to DNA. CDP/Cux p75 was able to repress a reporter carrying the promoter for the cyclin dependent kinase inhibitor p21 gene, and to activate a DNA polymerase α gene reporter. Expression of CDP/Cux p75 involved a novel mechanism: transcription initiation within intron 20. The intron 20-initiated mRNA (I20-mRNA) was expressed at higher level in the thymus and in CD4+/CD8+ and CD4+ T cells. I20-mRNA was expressed only weakly or not at all in normal human mammary epithelial cells (HMECs) and normal breast tissues, but was detected in many breast tumor cells lines and breast tumors. In invasive tumors a significant association was established between higher I20-mRNA expression and a diffuse infiltrative growth pattern (n=41, p=0.0137). In agreement with these findings, T47D breast cancer cells stably expressing p75 could not form tubule structures in collagen but rather developed as solid undifferentiated aggregates of cells. Altogether these results suggest that aberrant expression of the CDP/Cux p75 isoform in mammary epithelial cells may be associated with the process of tumorigenesis in breast cancer.

Introduction

CDP/Cux (CCAAT-displacement protein/cut homeobox) belongs to a family of transcription factors present in all metazoans and involved in the control of proliferation and differentiation (reviewed in (1)). In *Drosophila melanogaster*, a large number of phenotypes were found to be caused by insertion of transposable insulator sequences that interfered with the function of tissue-specific enhancers (2-6). The affected tissues included the wings ("cut wing"), legs, external sense organs, Malpighian tubules, tracheal system and some structures in the central nervous systems (2, 7-14). In higher vertebrates, there are two CDP/Cux genes called CDP-1 and CDP-2 in human, and Cux-1 and Cux-2 in mouse and chicken (15-17). While Cux-2 is expressed primarily in nervous tissues, Cux-1 is present in most tissues (15, 18, 19). The cux-1 gene has been the subject of three gene-targeting studies. The cux-1 knockout mice displayed phenotypes in various organs including curly whiskers, growth retardation, delayed differentiation of lung epithelia, altered hair follicle morphogenesis, male infertility, and a deficit in T and B cells (19-22). In contrast to the small size of the cux-1 knock-out mice, transgenic mice expressing Cux-1 under the control of the CMV enhancer/promoter displayed multi-organ hyperplasia and organomegaly (23). Thus, from genetic studies both in Drosophila and the mouse, it is clear that the CDP/Cux/Cut gene plays an important role in the development and homeostasis of several tissues.

In tissue culture, the expression and activity of CDP/Cux has been associated with cellular proliferation (24-26), the repression of genes that are turned on in terminally differentiated cells (27-32), and the regulation of matrix attachment regions (MARs) (33-37). CDP/Cux/Cut proteins may contain two to four DNA binding domains. All proteins con-

tain at least a Cut homeodomain (HD) and either one to three Cut repeats (CR1, CR2 and CR3). The *cut* superclass of homeobox genes has been divided into three classes: CUX, ONECUT and SATB (38). While the Drosophila Cut, human CDP and mouse Cux genes contain three Cut repeats, in each species there is also a ONECUT gene containing a single Cut repeat (15, 16, 39-41). SATB1 includes two Cut repeat-like domains and a divergent Cut-like homeodomain (42). The term CDP/Cux will be used in the remainder of this manuscript to designate mammalian proteins that contain three Cut repeats and a Cut homeodomain.

Individual Cut repeats cannot bind to DNA on their own but need to cooperate with a second Cut repeat or with the Cut homeodomain (43). Two CDP/Cux DNA binding activities have been reported in cells. CDP/Cux p200 binds transiently to DNA, like the CR1CR2 domains, and carries the CCAAT-displacement activity (15, 31, 43, 44). At the G1/S transition of the cell cycle, proteolytic cleavage of p200 generates CDP/Cux p110, which contains CR2CR3HD and exhibits distinct DNA binding specificity and kinetics (45). In particular, p110 is able to make a stable interaction with DNA.

In the present study, we describe a novel CDP/Cux isoform, p75, that is encoded by a mRNA initiated within intron 20 (I20-mRNA). Interestingly, this novel isoform displays DNA binding properties distinct from that of the previously characterized p200 and p110 CDP/Cux isoforms. While expression of the I20-mRNA is restricted to certain tissues or cells, we found that its expression was activated in breast tumor cell lines and in primary human breast tumors. These results, together with the finding that the p110 isoform is expressed at higher level in uterine leiomyomas, suggest that alternative mechanisms may be selected in cancer cells to favor expression of short CDP/Cux isoforms (46).

Material and Methods

RNA preparation: RNA was prepared using Trizol^{TM} purchased from GibcoBRL according to manufacturer's instructions and treated with RNase free DNase at 37°C for 30 min.

RNase mapping: The riboprobes were prepared as previously described (39). Briefly, 1 μg of template DNA, transcription buffer (200 mM PIPES, 2 M NaCl, 5 mM EDTA), 10 mM DTT, RNasin (40 units), 500 mM ATP, CTP, GTP, 12 mM UTP, 50 μCi α-^{32}P-UTP and T7 RNA polymerase (69 units) (Pharmacia) were combined and then incubated for 1 h at 37°C. After 1 h, 500 mM UTP was added and further incubated for 5 min. The riboprobe was then treated with RNase free DNase at 37°C for 15 min, extracted with chloropane and run through a Sephadex G50 spun column. Forty μg of total RNA was annealed to 8x10^5 cpm of labeled riboprobe at 54°C for 16 h in 80% formamide, 0.4 M NaCl, 0.4 M piperazine-N,N-bis (2-ethanesulfonic acid) (PIPES) (pH 6.4), 1 mM EDTA. RNA-RNA hybrids were digested with 30 U of RNase T2 (GibcoBRL) per ml at 37°C for 1 h. Hybrids were then precipitated with 20 μg of tRNA, 295 μl 4 M guanidine thiocyanate and 590 μl of isopropanol. Pellets were resuspended in 80% formamide, 1x TBE and 0.1% xylene cyanol + bromophenol blue, denatured and electrophoresed on 4% acrylamide-8M urea gel. Gels were dried and visualized by autoradiography.

Reverse transcriptase-PCR: Reverse transcriptase-polymerase chain reaction (RT-PCR) was performed on Human Multiple Tissue cDNA (MTC^{TM}) of normalized, first-strand cDNA preparations derived from different adult human tissues, purchased directly from Clonetech. MTC cDNA ranges in size from 0.1 to at least 6 Kbp and is virtually free of genomic DNA (Supplier information). cDNA from mouse tissues, thymocytes, breast tumor cell lines and breast tumor samples (from the Manitoba Breast Tumor Bank) were prepared

using Superscript^{TM}II RNaseH$^-$Reverse transcriptase (Gibco BRL) according to the manufacturer's instructions.

Primers used: Fi20 (human nt-40 to –18 within intron 20) GCTATTTTCAGGCACGGTTTCTC, B22 (human nt 3630-3609 and mouse nt 3345-3324): TCCACATTGTTGGGGTCGTTC, F19 (human nt 3021-3041): AGAAAGGCCGAGAACCCTTCA, Fi20m (mouse nt-111 to -88): CGACGGTCCCCTTCTGGAATGG and F18 (mouse nt 2411-2447): CAAGCGCTGAGTCCC

Primers were labeled in a final volume of 50 µl, containing 5µl of 10X kinase buffer (70 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$ and 5 mM DTT), 15 units of T4 polynucleotide kinase and 0.8 mCi γ ^{32}P-ATP and incubated at 37°C for 1 h. The labeled primer was then run through a Sephadex G25 spun column. PCR was performed in a final volume of 30 µl, containing 1 ng cDNA, 1.5 mM $MgCl_2$, 3 µl standard 10x PCR buffer (200 mM Tris-HCl pH 8.4, 500 mM KCl), 0.45 µM of each primer, 0.12 mM dNTPs, and 1 units of Taq polymerase (Gibco BRL). An initial step of 4 minutes at 95°C was followed by 25 cycles of 45 s of denaturation at 95°C, 50s of annealing at 61°C, and 60 s of extension at 72°C, followed by a final extension step of 7 minutes at 72°C. Pilot tests have been done to make sure the PCR reaction does not reach its plateau (not shown). PCR products were separated on a 4% non-denaturing polyacrylamide gel and visualized by autoradiography.

Cell culture: HeLa, HEL, 293 and NIH3T3 cells were cultured in DMEM medium supplemented with 10%FBS. All breast tumor cell lines were a generous gift from Dr. Morag Park. MCF7, MDA 231, MDA468, T47D, Hs578T, MDA 435s, BT 549 cells were cultured in DMEM medium supplemented with 5% FBS. SkBr3 cells were cultured in

DMEM medium supplemented with 10% FBS. MDA 436 cells were cultured in Leibovitz medium supplemented with 15% FBS and 10 mg/ml insulin. MCF 10A and MCF 12A cells were cultured in 50% DMEM-F12 medium supplemented with 5% HS, 10 mg/ml insulin (Gibco BRL), 0.5 mg/ml hydrocortisol (Sigma), 0.1 mg/ml cholera enterotoxin (Gibco BRL), and 20ng/ml EGF (Boehringer Mannheim). HMEC cells were purchased from Clonetics and cultured using the manufacturer's medium and instructions. Transfections were done using ExGene500 (MBI Fermentas) according to manufacturer's instructions.

T47D collagen assay: T47D cells were transfected with 10 μg of pMX or pMX-p75 along with 1 μg of pSV-NEO. Stable expressing lines were selected for 3 weeks with 400 μg/ml G418 (Gibco, BRL). Tubule forming assay was performed by adding $2x10^5$ cells/ml into 1.3 mg/ml of collagen in DMEM supplemented with 5% FBS as described in (40, 41). Cells were cultured for 10 days. Tubules were visualized using a RETIGA 1300 digital camera (QIMAGING, Burnaby, BC, Canada) and a Zeiss AxioVert 135 microscope with a 10x objective (Carl Zeiss Canada Ltd., Toronto, Canada). Cells in collagen were then fixed in 4% paraformaldehyde, embeded in paraffin and sectioned (8 μM). Sections were stained with hematoxylin and eosin. Images were acquired using PixCell II™ LCM system (Arcturus Engineering Inc. Mountain View, USA) using a 40x objective.

Preparation of nuclear cell extract: Nuclear extracts were prepared according to the procedure of Lee et al. (42) 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). Nuclei were then resuspended in Buffer C (20 mM Hepes, pH 7.9, 25 % glycerol 1.5mM MgCl2, 420 mM $NaCl_2$, 0.2 mM EDTA, 1 mM DTT) and incubated at 4 °C for 30

min. After 15 min. of centrifugation, the supernatant were collected. Buffers A and C were supplemented with protease inhibitor mix tablet (Roche) and DTT.

Preparation of mouse thymus extract: Thymus protein extracts were prepared by homogenization 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 (Roche)).

Expression and purification of CDP/Cut fusion proteins. The bacterial expression vectors pET-15b (Novagen) expressing CR2CR3HD and CR3HD were introduced into the BL21(DE3) of *E.coli* and induced with IPTG. The fusion proteins were purified by affinity chromatography using procedures provided by the suppliers.

Electromobility shift assay (EMSA): 1 μg of nuclear extracts from transfected mammalian cells were incubated in a final volume of 30 μl for 5 min. at room temperature in binding buffer (10 mM Tris-HCl, pH 7.5, 1 mM $MgCl_2$, 1 mM DTT, 5 mM EDTA, 5% glycerol, 25 mM NaCl) with 60 ng of poly dIdC and 30 μg of bovine serum albumin (BSA) as non specific competitors. 10 pg (5×10^{-16} mol) of radiolabelled oligonucleotide (TCGAGACGGT**ATCGAT**AAGCTTCTTTTC) was added and further incubated for 15 min. at room temperature. Samples were loaded on a 4% polyacrylamide gels (29:1) and separated by electrophoresis at 8V/cm in 0.5X Tris-Borate EDTA. Gels were dried and visualized by auto-radiography. DNA binding affinity was performed using a fixed amount of DNA and increasing concentration of proteins. For off-rate experiments, 1 μg of nuclear extracts were incubated with 60 pg 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.

Luciferase Assay: This assay was performed as previously described (37). Briefly, HS578T cells were plated at $1x10^5$ cells per 22.1 mm well. The next day, 0.5 μg of reporter DNA and increasing amount of effector DNA were transfected using ExGen500 (MBI, Fermentas). Cells were harvested 40 hours later: cells from 3 wells served to prepare separate cytoplasmic extracts for luciferase assay. Because the internal control plasmid is itself often repressed by CDP/Cux, as a control for transfection efficiency the purified b-galactosidase protein (Sigma) was included in the transfection mix. The luciferase activity was then normalized based on β-galactosidase activity.

Immunofluorescence: NIH3T3 cells were plated on a coverslip and transfected with 5 μg of pMX-Intron 20-HA. 2 days after, cells were fixed with 100% of methanol of 2 min. After 2 washes with 1X PBS, cells were quenched for 10 min in 50 mM of NH_4Cl, solubilized for 10 min (95% PBS + 5% FBS + 0.5% Triton X-100) and incubated with αHA (1:10000) for 1 hr at RT. After extensive washing, the secondary antibody (anti- mouse alexa 488 1:1000) was incubated for 30 min at RT in the dark. Cells were visualized using a RETIGA 1300 digital camera (QIMAGING, Burnaby, BC, Canada) and a Zeiss AxioVert 135 microscope with a 63x objective. Images were analysed using Northern Eclipse version 6.0 (Empix Imaging, Missisauga, Canada).

Plasmid Construction: For the expression of human Intron 20, PCR-amplification was performed to produce a fragment starting within intron 20 using a primer at the 5'end which includes XhoI and NotI sites ACTGCTCGAGCGGCCGCTTTTAGCAGAATGCCCTCATG) and a 3' end primer (nt

3862-3841: GTTTTTGGTGACGGGTATGGC). The product was digested with XhoI and BstXI (nt 3625 of CDP/Cux) and ligated together with a BstXI-NotI fragment that includes nt 3625 to 4551 of CDP/Cux to an XhoI-NotI digested pMX139 vector. To generate pMX139-Intron-20-HA, a fragment BstXI-XhoI from pMX-Intron-20 was ligated together with a BstXI-NotI fragment from pMX-Myc-HSCDP-HA into a pMX vector previously digested with NotI-XhoI. In order to generate a mouse intron 20 construct, an XhoI-BamHI (human nt 3936 and mouse nt 3651) fragment of pMX-human i20 was ligated to a BamHI-NotI fragment from mouse pMX-CDP/Cux.

Human Breast Cancer Specimen Analysis: A cohort of 41 invasive ductal carcinomas was selected from the Manitoba Breast Tumor Bank with two subgroups. All cases are processed uniformly to produce matched mirror image parafin and frozen tissue blocks. Tumor pathology and characteristics can therefore be assessed directly in high quality paraffin sections from tissue immediately adjacent to frozen tissue sections used for RNA extraction and RT-PCR analysis (43). The first group comprised invasive ductal carcinomas (n=21) showing large cohesive clusters of tumor cells forming nests or glandular arrangements, without a diffuse or infiltrating growth pattern. The second subgroup (n=21) comprised invasive tumors selected for a diffuse infiltrating growth pattern. These included 'Mixed Ductal & Lobular Carcinomas' (n=9) with a significant lobular component or a 'lobular' pattern of growth but with either focal glandular formation and/or ductal type cytological features and invasive lobular carcinomas (n=11) (44, 45).

Results

A CDP/Cux mRNA is initiated within intron 20 and is expressed in a tissue-specific manner. RNase mapping using a riboprobe containing exons 19, 20 and 21 generated a smaller protected fragment than anticipated with RNA samples from certain sources, notably HeLa cells and placenta (Fig. 1A). This result indicated the existence of an alternative CDP/Cux transcript that contains exon 21, but not exon 20. The 5' end of the novel transcript was cloned from placenta by the method of rapid amplification of cDNA ends (RACE) using as reverse primers two successive oligonucleotides from exon 21. DNA sequencing analysis showed that the sequence upstream of exon 21 originated from intron 20 and extended at least 500 nt upstream of the intron 20/exon21 junction. To exclude the possibility that genomic DNA that was still present in our RNA preparations, reverse-transcriptase polymerase chain reaction (RT-PCR) analysis was performed using as forward primers oligonucleotides from exon 19 or intron 20 and as a reverse primer, an oligonucleotide from exon 22. First-strand cDNA preparations derived from different adult human tissues (Clonetech) were used as a source of material (Fig. 1C). We obtained fragments of 609 bp with the exons 19 and 22 primers and of 474 with the intron 20 and exon 22 primers. The latter corresponds to the size predicted for a mRNA containing sequences from intron 20, exon 21 and exon 22. This result confirmed the existence of a CDP/Cux mRNA that initiates upstream of exon 21 and does not contain exon 20. RT-PCR analysis of mouse tissue RNAs confirmed that a similar transcript is expressed at higher levels in the placenta and thymus of the mouse (Fig. 2 A). The I20-mRNA was expressed in mature and immature T cells, but at higher level in mature $CD4^+$ than in mature $CD8^+$ T cells (Fig.

2B). Altogether, these findings indicate that the I20-mRNA is expressed in a tissue and cell-type specific manner.

RT-PCR products of the expected sizes were obtained with forward primers situated approximately 500, 1500, 2500 nt upstream of exon 21 (data not shown). No product was obtained with oligonucleotides positioned 3000 and 3500 nt upstream of exon 21. RNase mapping analysis was then performed with a riboprobe containing nucleotides –2270 to –2978 upstream of exon 21. A unique protected fragment of approximately 200 nt was observed, suggesting that transcription may start at a position approximately 2.5 kbp upstream of the intron 20/exon 21 junction (Fig. 1B). We cannot, however, exclude the possibility that this position represents an alternative splice acceptor site. Altogether, our results define a novel CDP/Cux mRNA that is expressed in a tissue-specific manner and is initiated upstream of exon 21. We will refer to this novel transcript as the intron 20-containing mRNA (I20-mRNA).

I20-mRNA encodes for a CDP/Cux protein of 75 kDa that localizes to the nucleus and binds to DNA. The I20-mRNA contains a long 5'untranslated sequence followed by an open reading frame starting at the beginning of exon 21. An AUG codon is present at a position corresponding to nt 3224 of the HSCDP cDNA sequence. The sequence at this position, CCGAUGG, does not conform to the Kozak consensus. Yet, a protein was expressed in an *in vitro* transcription/translation system, and replacement of AUG for UUC completely eliminated translation (Fig 2A). Transfection of NIH3T3 cells with a mouse I20-mRNA expression vector gave rise to a novel protein of 75 kDa that co-migrated with a protein present in mouse thymus (Fig 3B). This protein was detected with the C-terminal α1300 but not the α23 N-terminal CDP/Cux antibody (3A, lane 3). In electrophoretic mobility shift assays (EMSA), nuclear extracts from transfected NIH3T3 cells

generated a retarded complex which could be supershifted with the α1300 CDP/Cux but not with an unrelated antibody (Fig. 3C, lanes 2-4). When NIH3T3 cells were transfected with a vector expressing p75 with an influenza virus hematoglutinin (HA) tag at its carboxy-terminus, a specific signal was detected by indirect immunofluorescence in the nucleus of transfected NIH3T3 cells (Fig. 3D). Altogether these results demonstrate that I20-mRNA codes for a CDP/Cux protein of 75 kDa that localizes to the nucleus and binds to DNA.

p75 makes a more stable interaction with DNA than p110 and displays similar transcriptional properties in reporter assays. The CDP/Cux p110 isoform contains CR2, CR3 and HD (CR2CR3HD) while the p75 isoform contains CR3 and HD (CR3HD). We compared the DNA binding properties of the two isoforms first using bacterially expressed his-tagged fusion proteins and then nuclear extracts from transfected mammalian cells. The purified CR2CR3HD and CR3HD his-tagged proteins exhibited similar DNA binding affinities, with apparent dissociation constants of (K_D(app.)) of 0.7 and 1.1 10^{-9} M, respectively (Fig. 4A, top panels). In contrast, their DNA binding kinetics were different. CR3HD bound more stably to DNA than CR2CR3HD (Fig. 4A, bottom panels). In agreement with these findings, the off rate of p110 was 0.8 min. and that of p75 was 6.15 min. (Fig. 4B). These results indicate that CDP/Cux p75 makes a more stable interaction with DNA than the p110 isoform that is expressed in S phase. To investigate the transcriptional properties of p75 CDP/Cux, reporter assays were performed in parallel using either p110 or p75 as effectors. The results from several experiments indicated that both proteins, p110 and p75, were able similarly to repress the $p21^{WAF1/CIP1}$ reporter and stimulate expression from the DNA pol α reporter (Fig. 5). Whether stimulation of gene expression involves direct activation or repression of a down-modulator is currently under investigation. In

summary, p75 CDP/Cux localizes to the nucleus, binds to DNA, and is able to regulate transcription of target genes.

I20-mRNA and p75 are weakly or not expressed in HMEC cells, but are detected in many breast tumor cells lines. We analyzed expression of the CDP/Cux full length and I20-mRNAs in a panel of breast tumor cell lines and in human mammary epithelial cells (HMEC). In RNase protection and RT-PCR assays, a fragment corresponding to the I20-mRNA was detected in some but not all breast tumor cell lines (Fig. 6A and B). Similarly, Western blot analysis revealed that the 75 kDa protein was detected in some breast tumor cell lines but not in HMEC cells (Fig. 6C). We then compared expression of I20-mRNA in two pairs of cell lines. Expression of the I20-mRNA was higher in the tumorigenic Hs578T cell line than in its non tumorigenic counterpart, Hs578Bst (Fig. 6D, left panel) (53). A similar comparison in a pair of immortalized and notch-transformed mammary epithelial cell lines of mouse origin, HC11 and notch-HC11 (54), also showed that the mouse CDP/Cux I20-mRNA was expressed at a higher level in the transformed line (Fig. 6D, right panel). Altogether, these findings suggest that expression of the CDP/Cux I20-mRNA and p75 protein is activated in some breast cancer cells.

T47D-p75 stable cell lines do not form tubules in collagen. To evaluate whether p75 may be able to confer to mammary epithelial cells properties that are associated with cellular transformation, we generated T47D cell lines stably expressing p75. Although T47D cells are derived from a breast tumor, they have retained the capability to differentiate and form tubules in collagen (48, 49). For this reason, they represent a good cellular model in which to investigate the effect of putative oncogenes. Interestingly, we observed that the T47D clones expressing p75 could no longer form tubules in collagen (Fig.7A and 7B). Moreover, the colonies generated by the p75-T47D clones were not hollow cysts but

instead compact aggregates of cells which were devoid of a central lumen (Fig. 7C). These results indicate that upon forced expression of CDP/Cux p75, T47D cells loose their ability to form an organized epithelial sheet.

The I20-mRNA is expressed in some breast carcinomas but not in normal breast tissue. Using RT-PCR analysis, the I20-mRNA was not detected in RNA isolated from a reduction mammoplasty tissue sample from a woman without known breast pathology (Fig. 8A, right panel). This result is in accordance with the findings that the I20-mRNA was not expressed in normal mouse mammary glands (see Fig. 2A). However, a strong I20-mRNA signal was observed in three cases of breast cancer, tumors C8921D, A168A and C8961B (Fig. 8A, left panel). Two of these tumors, C8921D and C8961B are lobular carcinomas, whereas A168A was classified as a mixed ductal-lobular carcinoma but with a very diffuse growth pattern. All other tumors showing low or no I20-mRNA expression were classified as ductal carcinomas.

Higher I20-mRNA expression in invasive carcinoma is associated with a more diffuse growth pattern and the lack of tubular differentiation. The above results raised the possibility that I20-mRNA expression may be associated with a more diffuse growth pattern. To further examine this hypothesis, we analyzed CDP/Cux mRNA expression in an expanded panel of invasive carcinomas that were selected on the basis of their classification as either ductal, lobular or mixed lobular/ductal carcinomas. Higher I20-mRNA expression levels was significantly associated with invasive lobular and invasive mixed lobular/ductal carcinomas compared to invasive ductal carcinomas {mean (sd), lobular and mixed; n=20, 115200 (84770) vs ductal n=21, 45510 (43360), p=0.0137, Mann Whitney test} (55). These results suggest that the I20-mRNA is expressed at a higher level in a subset of breast tumors that exhibit a more diffuse growth pattern compared to tumors that exhibit the abil-

ity to form cohesive clusters and tubules. These results are in agreement with our tissue culture assays showing that mammary epithelial cells loose their ability to form tubules in collagen upon forced expression of the I20-mRNA.

Discussion

The full length 200 kDa CDP/Cux protein was previously found to be proteolytically processed in S phase to generate an amino-terminally truncated isoform of 110 kDa (45). Here, we showed that an alternate mechanism can serve to generate an amino-terminally truncated isoform of 75 kDa: transcription initiation within an intron. Thus, alternative mechanisms can serve to generate amino-terminally truncated CDP/Cux isoforms (see Fig. 9). We speculate that the existence of two modes of regulation enables the production of short Cut proteins in response to different signaling or developmental cues. Interestingly, while the ONECUT genes contain only one Cut repeat, it appears that the more complex Cux genes have the potential to encode proteins with either one (p75), two (p110) or three Cut repeats (p200). We favor the opinion that there is a functional advantage associated with the ability to express proteins with variable numbers of Cut repeats. Various combinations of Cut repeats and the Cut homeodomain were found to exhibit different DNA binding properties (43, 56-59). CDP/Cux p200 and p110 clearly displayed distinct DNA binding and transcriptional properties (43, 45). The differences between p110 and p75 appear to be more subtle. We have not detected differences in binding affinity or specificity, but DNA binding was found to be more stable in the case of p75. This property could make of p75 a more potent transcription factor than p110. The two isoforms were found to behave similarly in a reporter assay (see Fig. 5), but it is possible that subtle differences would not be revealed in transient assays where proteins are overexpressed. Another potential difference between p75 and p110 concerns the proteins with which they can interact, as the presence or absence of Cut repeat 2 may allow interaction with different

partners. Future studies should investigate the protein-protein interaction capabilities of the various CDP/Cux isoforms.

We have shown that forced expression of CDP/Cux p75 isoform in T47D cells inhibited the formation of tubules in a tissue culture assay and that elevated I20-mRNA expression in breast tumors was associated with a more diffuse infiltrative growth pattern and the lack of tubule structures (Fig. 7 and 8). These findings would warrant a more extensive study to verify whether CDP/Cux I20-mRNA or p75 protein expression could be used in the future as a molecular marker to augment grading of breast tumors. Interestingly, activated forms of the Notch4 receptors were previously shown to prevent branching morphogenesis in mammary epithelial cells (60, 61). That Notch and CDP/Cux p75 can confer similar phenotypes onto mammary epithelial cells raises the possibility that the receptor and the transcription factor function in the same pathway. This notion is in agreement with the reported interactions between the *notch* pathway and *cut* in *Drosophila melanogaster* (reviewed in (1)). Indeed, a number of genetic studies indicated that activation of the Notch pathways leads to the induction of Cut which would function as a downstream effector of this pathway (8, 62-66). It will be interesting to verify whether a similar connection exists in the mammary gland between Notch and CDP/Cux and, if so, which of the CDP/Cux isoform functions downstream of the Notch signaling pathway. Activation of the Notch pathway in human cancers has so far been documented only in T cell leukemia (67). However, both the murine Notch1 and Notch4/int3 have been found to represent common sites of provirus integration in mouse mammary tumors (54, 68). Moreover, recombinant truncated Notch proteins were shown to transform HC11 and MCF-10A mammary epithelial cells *in vitro* (54, 69). These findings raise the possibility that activation of the Notch pathway might also take place in some human breast cancers but would avoid detection by the cur-

rently available tools (70). In this respect, the confirmation of CDP/Cux as a downstream targets of Notch in higher vertebrates could provide a molecular marker for Notch activation in human cancers.

Acknowledgements

We acknowledge the expert technical assistance of Ms. Ginette Massé in the separation of T cell populations by FACS sorting and Ms. Jo-Ann Bader for her expertise in paraffin embeding, sectioning and hematoxylin/eosin staining. This research was funded by an operating grant to A.N. from the Canadian Breast Cancer Research Initiative. A.N. is the recipient of a scholarship from Le Fonds de la Recherche en Santé du Québec. B.G. was the recipient of a studentship from the Research Institute of the McGill University Health Center.

References

1. Nepveu, A. Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development. Gene, *270:* 1-15., 2001.
2. Jack, J., Dorsett, D., Delotto, Y., and Liu, S. Expression of the cut locus in the Drosophila wing margin is required for cell type specification and is regulated by a distant enhancer. Development, *113:* 735-747, 1991.
3. Jack, J. and DeLotto, Y. Structure and regulation of a complex locus: the cut gene of Drosophila. Genetics, *139:* 1689-1700, 1995.
4. Modolell, J., Bender, W., and Meselson, M. Drosophila melanogaster mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. Proc Natl Acad Sci U S A, *80:* 1678-1682, 1983.
5. Cai, H. N. and Levine, M. The gypsy insulator can function as a promoter-specific silencer in the Drosophila embryo. EMBO J., *16:* 1732-1741, 1997.
6. Dorsett, D. Distance-independent inactivation of an enhancer by the suppressor of Hairy-wing DNA-binding protein of Drosophila. Genetics, *134:* 1135-1144, 1993.
7. Jack, J. W. Molecular organization of the cut locus of Drosophila melanogaster. Cell, *42:* 869-876, 1985.
8. Jack, J. and DeLotto, Y. Effect of wing scalloping mutations on cut expression and sense organ differentiation in the Drosophila wing margin. Genetics, *131:* 353-363, 1992.
9. Liu, S., McLeod, E., and Jack, J. Four distinct regulatory regions of the cut locus and their effect on cell type specification in Drosophila. Genetics, *127:* 151-159, 1991.

10. Liu, S. and Jack, J. Regulatory interactions and role in cell type specification of the Malpighian tubules by the cut, Kruppel, and caudal genes of Drosophila. Dev. Biol., *150:* 133-143, 1992.

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

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

13. Braun, W. The effect of puncture on the developing wing of several mutants of *Drosophila melanogaster.* J. Exp. Zool., *84:* 325-350, 1942.

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

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

16. Valarche, I., Tissier-Seta, J. P., Hirsch, M. R., Martinez, S., Goridis, C., and Brunet, J. F. The mouse homeodomain protein Phox2 regulates Ncam promoter activity in concert with Cux/CDP and is a putative determinant of neurotransmitter phenotype. Development, *119:* 881-896, 1993.

17. Quaggin, S. E., Vandenheuvel, G. B., Golden, K., Bodmer, R., and Igarashi, P. Primary Structure, Neural-Specific Expression, and Chromosomal Localization Of Cux-2, a Second Murine Homeobox Gene Related to Drosophila Cut. J. Biol. Chem., *271:* 22624-22634, 1996.

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

19. Ellis, T., Gambardella, L., Horcher, M., Tschanz, S., Capol, J., Bertram, P., Jochum, W., Barrandon, Y., and Busslinger, M. The transcriptional repressor CDP (Cutl1) is essential for epithelial cell differentiation of the lung and the hair follicle. Genes Dev, *15:* 2307-2319., 2001.

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

21. Luong, M. X., van der Meijden, C. M., Xing, D., Hesselton, R., Monuki, E. S., Jones, S. N., Lian, J. B., Stein, J. L., Stein, G. S., Neufeld, E. J., and van Wijnen, A. J. Genetic ablation of the CDP/Cux protein C terminus results in hair cycle defects and reduced male fertility. Mol Cell Biol, *22:* 1424-1437, 2002.

22. Tufarelli, C., Fujiwara, Y., Zappulla, D. C., and Neufeld, E. J. 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, 1998.

23. Ledford, A. W., Brantley, J. G., Kemeny, G., Foreman, T. L., Quaggin, S. E., Igarashi, P., Oberhaus, S. M., Rodova, M., Calvet, J. P., and Vanden Heuvel, G. B. Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia. Dev Biol, *245:* 157-171, 2002.

24. Holthuis, J., Owen, T. A., van Wijnen, A. J., Wright, K. L., Ramsey-Ewing, A., Kennedy, M. B., Carter, R., Cosenza, S. C., Soprano, K. J., Lian, J. B., Stein, J. L., and

Stein, G. S. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science, *247:* 1454-1457., 1990.

25. van Wijnen, A. J., Choi, T. K., Owen, T. A., Wright, K. L., Lian, J. B., Jaenisch, R., Stein, J. L., and Stein, G. S. 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:* 2573-2577., 1991.

26. Coqueret, O., Berube, G., and Nepveu, A. 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:* 4680-4694, 1998.

27. Pattison, S., Skalnik, D. G., and Roman, A. 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:* 2013-2022, 1997.

28. Lawson, N. D., Khannagupta, A., and Berliner, N. Isolation and Characterization Of the Cdna For Mouse Neutrophil Collagenase - Demonstration Of Shared Negative Regulatory Pathways For Neutrophil Secondary Granule Protein Gene Expression. Blood, *91:* 2517-2524, 1998.

29. van Gurp, M. F., Pratap, J., Luong, M., Javed, A., Hoffmann, H., Giordano, A., Stein, J. L., Neufeld, E. J., Lian, J. B., Stein, G. S., and van Wijnen, A. J. The CCAAT displacement protein/cut homeodomain protein represses osteocalcin gene transcription and forms complexes with the retinoblastoma protein-related protein p107 and cyclin A. Cancer Res, *59:* 5980-5988, 1999.

30. O'Connor, M. J., Stunkel, W., Koh, C. H., Zimmermann, H., and Bernard, H. U. The differentiation-specific factor CDP/Cut represses transcription and replication of

human papillomaviruses through a conserved silencing element [In Process Citation]. J Virol, *74:* 401-410, 2000.

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

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

33. Liu, J., Bramblett, D., Zhu, Q., Lozano, M., Kobayashi, R., Ross, S. R., and Dudley, J. P. The matrix attachment region-binding protein SATB1 participates in negative regulation of tissue-specific gene expression. Mol Cell Biol, *17:* 5275-5287, 1997.

34. Banan, M., Rojas, I. C., Lee, W. H., King, H. L., Harriss, J. V., Kobayashi, R., Webb, C. F., and Gottlieb, P. D. Interaction of the nuclear matrix-associated region (MAR)-binding proteins, SATB1 and CDP/Cux, with a MAR element (L2a) in an upstream regulatory region of the mouse CD8a gene. J Biol Chem, *272:* 18440-18452, 1997.

35. Chattopadhyay, S., Whitehurst, C. E., and Chen, J. 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:* 29838-29846, 1998.

36. Wang, Z., Goldstein, A., Zong, R. T., Lin, D., Neufeld, E. J., Scheuermann, R. H., and Tucker, P. W. 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:* 284-295, 1999.

37. Stunkel, W., Huang, Z., Tan, S. H., O'Connor, M. J., and Bernard, H. U. Nuclear matrix attachment regions of human papillomavirus type 16 repress or activate the E6 promoter, depending on the physical state of the viral DNA. J. Virol., *74:* 2489-2501, 2000.

38. Burglin, T. R. and Cassata, G. Loss and gain of domains during evolution of cut superclass homeobox genes. Int J Dev Biol, *46:* 115-123, 2002.

39. Blochlinger, K., Bodmer, R., Jack, J., Jan, L. Y., and Jan, Y. N. Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in Drosophila. Nature, *333:* 629-635, 1988.

40. Lemaigre, F. P., Durviaux, S. M., Truong, O., Lannoy, V. J., Hsuan, J. J., and Rousseau, G. G. 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:* 9460-9464, 1996.

41. Lannoy, V. J., Burglin, T. R., Rousseau, G. G., and Lemaigre, F. P. 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:* 13552-13562, 1998.

42. Dickinson, L. A., Dickinson, C. D., and Kohwi-Shigematsu, T. An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. J Biol Chem, *272:* 11463-11470, 1997.

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

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

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

46. Moon, N. S., Zeng, W. R., Premdas, P., Santaguida, M., Bérubé, G., and Nepveu, A. Expression of N-terminally truncated Isoforms of CDP/Cux is increased in human uterine leiomyomas. Int. J. Cancer In press.

47. Zeng, R. W., Soucie, E., Sung Moon, N., Martin-Soudant, N., Bérubé, G., Leduy, L., and Nepveu, A. Exon/intron structure and alternative transcripts of the CUTL1 gene. Gene, *241:* 75-85, 2000.

48. Keely, P. J. Ras and Rho protein induction of motility and invasion in T47D breast adenocarcinoma cells. Methods Enzymol, *333:* 256-266, 2001.

49. Keely, P. J., Fong, A. M., Zutter, M. M., and Santoro, S. A. Alteration of collagen-dependent adhesion, motility, and morphogenesis by the expression of antisense alpha 2 integrin mRNA in mammary cells. J Cell Sci, *108 (Pt 2):* 595-607, 1995.

50. Hiller, T., Snell, L., and Watson, P. H. Microdissection RT-PCR analysis of gene expression in pathologically defined frozen tissue sections. Biotechniques, *21:* 38-40, 42, 44, 1996.

51. Pereira, H., Pinder, S. E., Sibbering, D. M., Galea, M. H., Elston, C. W., Blamey, R. W., Robertson, J. F., and Ellis, I. O. Pathological prognostic factors in breast cancer. IV: Should you be a typer or a grader? A comparative study of two histological prognostic features in operable breast carcinoma. Histopathology, *27:* 219-226, 1995.

52. Ellis, I. O., Galea, M., Broughton, N., Locker, A., Blamey, R. W., and Elston, C. W. Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. Histopathology, *20:* 479-489, 1992.

53. Hackett, A. J., Smith, H. S., Springer, E. L., Owens, R. B., Nelson-Rees, W. A., Riggs, J. L., and Gardner, M. B. Two syngeneic cell lines from human breast tissue: the aneuploid mammary epithelial (Hs578T) and the diploid myoepithelial (Hs578Bst) cell lines. J Natl Cancer Inst, *58:* 1795-1806, 1977.

54. Dievart, A., Beaulieu, N., and Jolicoeur, P. Involvement of Notch1 in the development of mouse mammary tumors. Oncogene, *18:* 5973-5981, 1999.

55. Mann, H. B. and Whitney, D. R. On a Test of Whether One of Two Random Variables Is Stochastically Larger Than the Other. Annals of Mathematical Statistics, *18:* 50-60, 1947.

56. Harada, R., Berube, G., Tamplin, O. J., Denis-Larose, C., and Nepveu, A. DNA-binding specificity of the cut repeats from the human cut-like protein. Mol. Cell. Biol., *15:* 129-140, 1995.

57. Harada, R., Dufort, D., Denis-Larose, C., and Nepveu, A. Conserved cut repeats in the human cut homeodomain protein function as DNA binding domains. J. Biol. Chem., *269:* 2062-2067, 1994.

58. Aufiero, B., Neufeld, E. J., and Orkin, S. H. Sequence-specific DNA binding of individual Cut repeats of the human CCAAT displacement/Cut homeodomain protein. Proc. Natl. Acad. Sci. USA, *91:* 7757-7761, 1994.

59. Andres, V., Chiara, M. D., and Mahdavi, V. A new bipartite DNA-binding domain: cooperative interaction between the cut repeat and homeo domain of the cut homeo proteins. Genes Dev., *8:* 245-257, 1994.

60. Uyttendaele, H., Soriano, J. V., Montesano, R., and Kitajewski, J. Notch4 and Wnt-1 proteins function to regulate branching morphogenesis of mammary epithelial cells in an opposing fashion. Dev Biol, *196:* 204-217, 1998.

61. Soriano, J. V., Uyttendaele, H., Kitajewski, J., and Montesano, R. Expression of an activated Notch4(int-3) oncoprotein disrupts morphogenesis and induces an invasive phenotype in mammary epithelial cells in vitro. Int J Cancer, *86:* 652-659, 2000.

62. Neumann, C. J. and Cohen, S. M. A hierarchy of cross-regulation involving Notch, wingless, vestigial and cut organizes the dorsal/ventral axis of the Drosophila wing. Development, *122:* 3477-3485, 1996.

63. Thomas, U., Jonsson, F., Speicher, S. A., and Knust, E. Phenotypic and molecular characterization of SerD, a dominant allele of the Drosophila gene Serrate. Genetics, *139:* 203-213, 1995.

64. Micchelli, C. A., Rulifson, E. J., and Blair, S. S. 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:* 1485-1495, 1997.

65. de Celis, J. F., Garcia-Bellido, A., and Bray, S. J. Activation and function of Notch at the dorsal-ventral boundary of the wing imaginal disc. Development, *122:* 359-369, 1996.

66. Doherty, D., Feger, G., Younger-Shepherd, S., Jan, L. Y., and Jan, Y. N. Delta is a ventral to dorsal signal complementary to Serrate, another Notch ligand, in Drosophila wing formation. Genes Dev., *10:* 421-434, 1996.

67. Ellisen, L. W., Bird, J., West, D. C., Soreng, A. L., Reynolds, T. C., Smith, S. D., and Sklar, J. TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms. Cell, *66:* 649-661, 1991.

68. Gallahan, D. and Callahan, R. The mouse mammary tumor associated gene INT3 is a unique member of the NOTCH gene family (NOTCH4). Oncogene, *14:* 1883-1890, 1997.

69. Imatani, A. and Callahan, R. Identification of a novel NOTCH-4/INT-3 RNA species encoding an activated gene product in certain human tumor cell lines. Oncogene, *19:* 223-231, 2000.

70. Callahan, R. and Raafat, A. Notch signaling in mammary gland tumorigenesis. J Mammary Gland Biol Neoplasia, *6:* 23-36, 2001.

Figure 1: Identification of a novel CDP/Cux transcript starting within intron 20

(A) RNase mapping analysis was performed using a riboprobe encompassing exons 19 to 21 of the CUTL1 gene and total RNA samples from 3 cell lines (HeLa, HEL and 293), placenta and tRNA as a control. The undigested riboprobe as well as the major protected fragments are indicated by arrows. Shown below is a schematic representation of the riboprobe and the two protected fragments. The black bars on top of the probe indicate the position of CR2 and CR3 coding sequences. **(B)** RNase mapping analysis was performed using a riboprobe derived from the CUTL1 intron 20 and total RNA samples from various cells lines, placenta and tRNA as a control. **(C)** CDP/Cux (upper panel) and intron 20-initiated mRNA (I20-mRNA) (middle panel) were analyzed by reverse-transcriptase polymerase chain amplification (RT-PCR) using Human Multiple Tissue cDNA (MTC^{TM}) of normalized, first-strand cDNA preparations derived from different adult human tissues (Clontech) and the indicated oligonucleotides as primers (see Material and Methods). GAPDH mRNA was used as a control for the quantity of mRNA/cDNA in each sample.

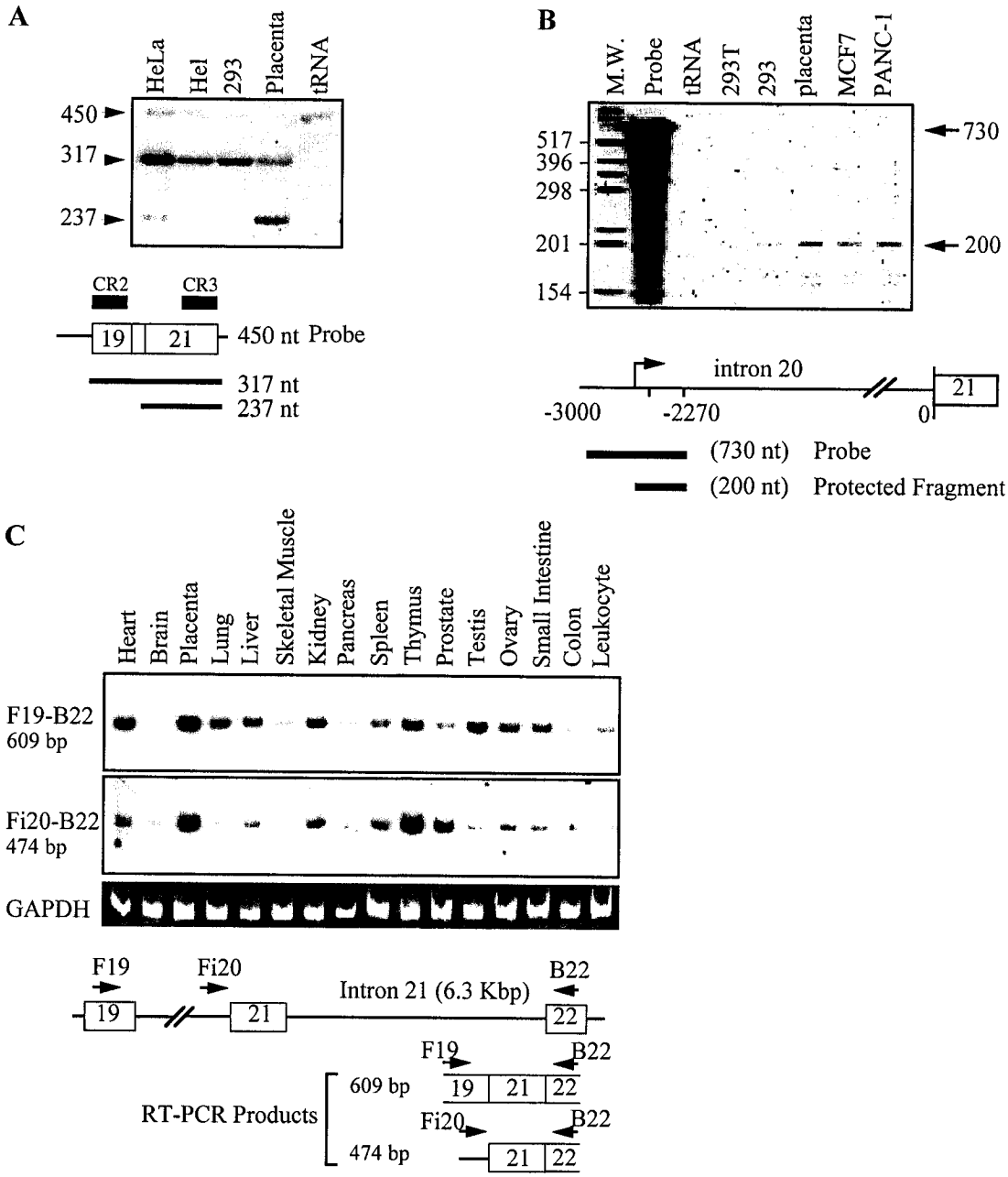

A
HeLa
Hel
293
Placenta
tRNA
450
317
237
CR2
CR3
19
21
450 nt Probe
317 nt
237 nt
B
M.W.
Probe
tRNA
293T
293
placenta
MCF7
PANC-1
517
396
298
201
154
730
200
intron 20
-3000
-2270
0
21
(730 nt) Probe
(200 nt) Protected Fragment
C
Heart
Brain
Placenta
Lung
Liver
Skeletal Muscle
Kidney
Pancreas
Spleen
Thymus
Prostate
Testis
Ovary
Small Intestine
Colon
Leukocyte
F19-B22
609 bp
Fi20-B22
474 bp
GAPDH
F19
Fi20
B22
Intron 21 (6.3 Kbp)
19
21
22
RT-PCR Products
609 bp
474 bp

Figure 2: Identification of I20-mRNA in mouse tissues and mouse thymocytes

(A) Intron 20-initiated mRNA (I20-mRNA) was analyzed by RT-PCR using first-strand cDNA derived from various mouse tissues and the indicated oligonucleotides. The amplified fragments are 900 bp for the full length CDP/Cux mRNA and 600 bp for the intron 20-initiated mRNA. GAPDH mRNA was used as a control for the quantity of mRNA/cDNA in each sample. **(B)** I20-mRNA was analyzed by RT-PCR as described above using first-strand cDNA preparations derived double positive (CD4+/CD8+) or single positive (CD4+ or CD8+) mouse thymocytes.

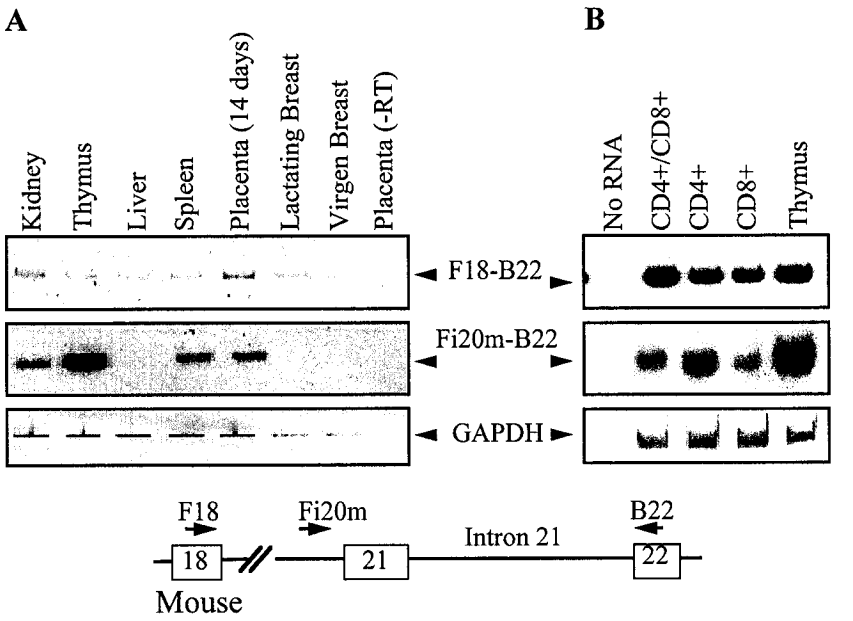

A
B
Kidney
Thymus
Liver
Spleen
Placenta (14 days)
Lactating Breast
Virgen Breast
Placenta (-RT)
No RNA
CD4+/CD8+
CD4+
CD8+
Thymus
F18-B22
Fi20m-B22
GAPDH
F18
Fi20m
B22
Intron 21
18
21
22
Mouse

Figure 3: The I20-mRNA codes for a nuclear protein of 75 kDa that binds to ATCGAT

(A) A cDNA fragment for the I20-mRNA was inserted into the pcDNA3.1 vector and the first ATG was replaced for TTC. The resulting plasmids were tested in an *in vitro* transcription/translation system (Promega) in the presence of ^{35}S-labeled methionine. **(B)** The pcDNA3.1-p75 vector was introduced into NIH3T3 cells. Nuclear extracts from transfected NIH3T3 cells (5 μg) and total extracts from mouse thymus (500 μg) were analyzed in Western blots with anti-CDP/Cux 23 and 1300 antibodies. **(C)** Nuclear extracts from transfected NIH3T3 cells were analyzed in EMSA with the ATCGAT probe. Note that the small amount of proteins used from transfected cells precludes the detection of endogenous CDP/Cux proteins in immunoblots and EMSA. **(D)** An HA-tagged was inserted at the C-terminus of p75 and the resulting vector was introduced intro NIH3T3 cells. Indirect immunofluorescence was performed using an HA antibody.

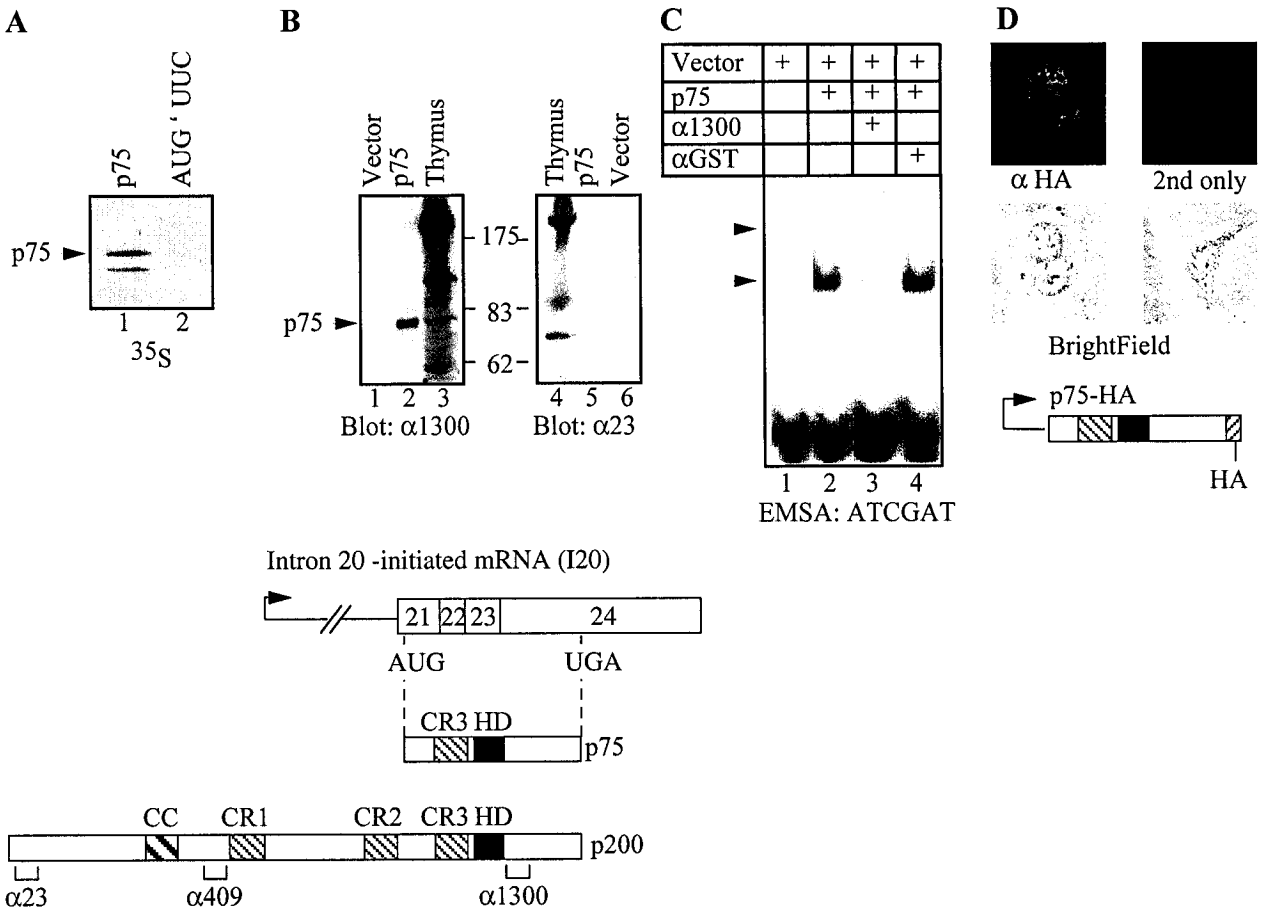
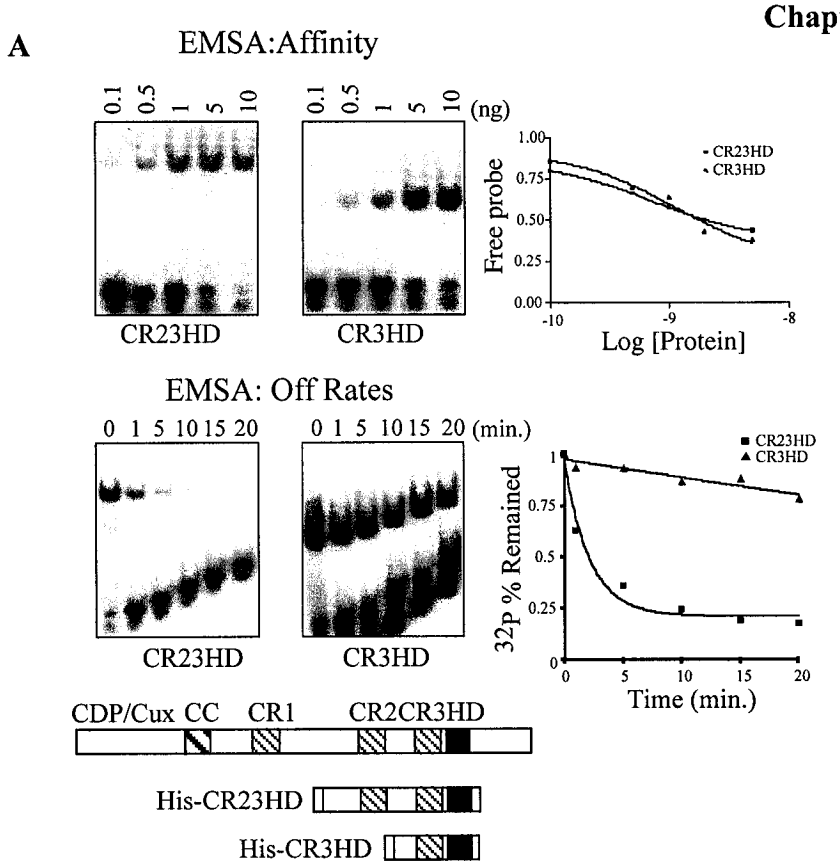

A
p75
AUG ' UUC
p75
1 2
35S
B
Vector
p75
Thymus
Thymus
p75
Vector
175
83
62
p75
1 2 3
Blot: α1300
4 5 6
Blot: α23
C
Vector + + + +
p75 + + +
α1300 +
αGST +
1 2 3 4
EMSA: ATCGAT
D
α HA
2nd only
BrightField
p75-HA
HA
Intron 20 -initiated mRNA (I20)
21 22 23 24
AUG
UGA
CR3 HD
p75
CC
CR1
CR2
CR3 HD
p200
α23
α409
α1300

Fig. 4: p75 binds more stably to DNA than the proteolytically processed p110 isoform

(A) The DNA binding affinity and stability were measured using purified His-tagged proteins containing two Cut repeats (CR2CR3HD), as in p110, or one Cut repeat (CR3HD), as in p75. (Top panels) Binding affinity: ten pM of radiolabeled oligonucleotides containing the ATCGAT sequence was incubated with increasing concentrations of CR2CR3HD or CR3HD fusion proteins. The incubations took place at room temperature until the equilibrium was reached (15 min.) and the samples were resolved by electrophoresis on a non-denaturing polyacrylamide gel. To obtain the apparent dissociation constant, $K_{D(app)}$, the percentage of free DNA (relative to the amount in the lane with no protein added) was plotted against the log of protein concentration. (Bottom panels) Off rates: 1 μg 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. **(B)** Hs578T cells were transfected with vectors expressing recombinant p110 and p75 CDP/Cux isoforms. Nuclear extracts were used to measure off rates as described above.


A
EMSA:Affinity
0.1 0.5 1 5 10 (ng)
CR23HD
CR3HD
Free probe
Log [Protein]
EMSA: Off Rates
0 1 5 10 15 20 (min.)
32p % Remained
Time (min.)
CDP/Cux CC CR1 CR2CR3HD
His-CR23HD
His-CR3HD

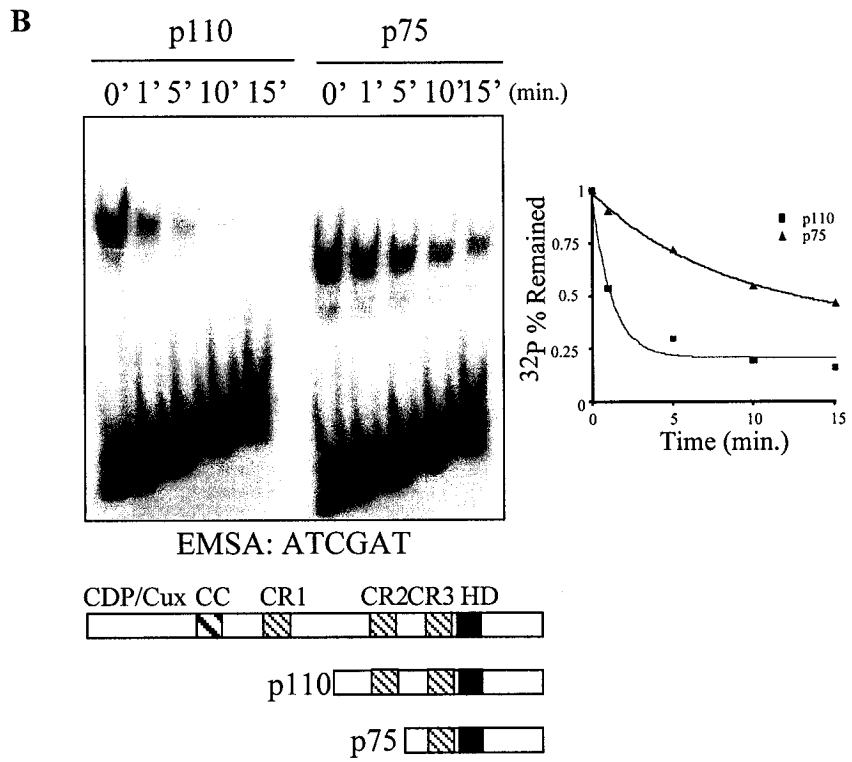

B
p110
p75
0' 1' 5' 10' 15' 0' 1' 5' 10'15' (min.)
32p % Remained
Time (min.)
EMSA: ATCGAT
CDP/Cux CC CR1 CR2CR3 HD
p110
p75

Fig. 5: p75 can repress the expression of a $P21^{waf1/cip1}$ reporter and stimulate the expression of a DNA polymerase α reporter

NIH3T3 cells were transfected with 500 ng of a $P21^{Cip1/Waf1}$/luciferase reporter construct **(A)** and Hs578T, with a DNA polymerase α/luciferase reporter **(B)**, together with either an empty vector or a vector encoding for p110 or p75 CDP/Cux. Two days later, 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 β-galactosidase activity from an internal control.

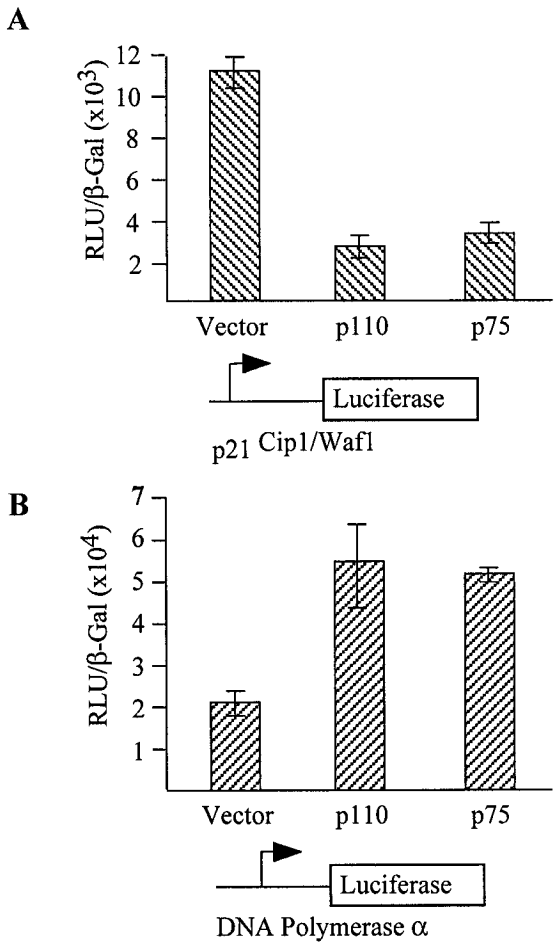

A
RLU/β-Gal (x10³)
12
10
8
6
4
2
Vector
p110
p75
Luciferase
p21 Cip1/Waf1
B
RLU/β-Gal (x10⁴)
7
6
5
4
3
2
1
Vector
p110
p75
Luciferase
DNA Polymerase α

Fig. 6: Intron 20-initiated mRNA is detected in breast tumor cell lines but weakly or not in normal human mammary epithelial cells (HMEC)

(A) RNase mapping analysis was performed using a riboprobe encompassing exons 19 to 21 of the CUTL1 gene (as described in Fig. 1A) and total RNA samples from breast tumor cell lines and a primary culture of human mammary epithelial cells (HMEC). The undigested riboprobe as well as the major protected fragments are indicated by arrows. A schematic representation of the riboprobe and the two protected fragments is shown in Fig. 1A. **(B)** CDP/Cux (upper panel) and intron 20-initiated mRNA (I20-mRNA) (middle panel) were analyzed by RT-PCR using total RNA from breast tumor cell lines and HMEC cells. The assay was performed as described in Fig. 1C. GAPDH mRNA was used as a control for the quantity of mRNA/cDNA in each sample. **(C)** 25 μg of nuclear extracts from breast tumor cell lines and HMEC were analyzed in Western blots using anti-CDP/Cux 1300 and 409 antibodies. As controls to identify the p75 protein, 5 μg of nuclear extracts were taken from pcDNA3.1-p75 transfected NIH3T3 (lane p75). Coomassie-blue staining of the membranes are shown to compare the amounts of proteins in each sample. The epitopes recognized by the 1300 and 409 antibodies are shown in Fig. 3. **(D)** RT-PCR analysis of full length and I20 CDP/Cux mRNA in a pair of immortalized and transformed cell lines of human origin, Hs578Bst and Hs578T (left panel) and of mouse origin, HC11 and notch-HC11 (right panel). The assay was performed as in Fig. 1C for the human cell lines and as in Fig. 2 for the mouse cell lines, except that the two forward primers were included together with the same backward primer.

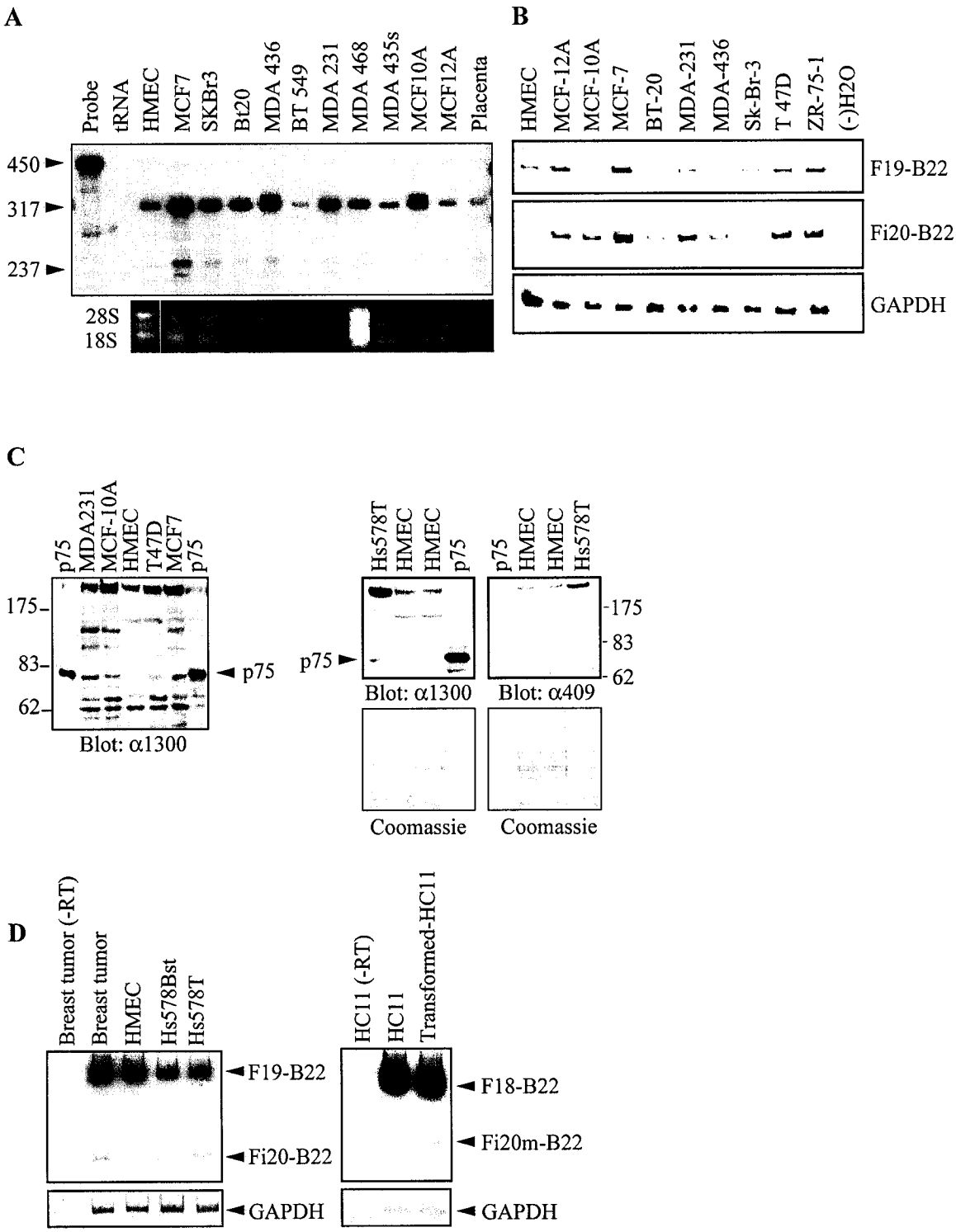
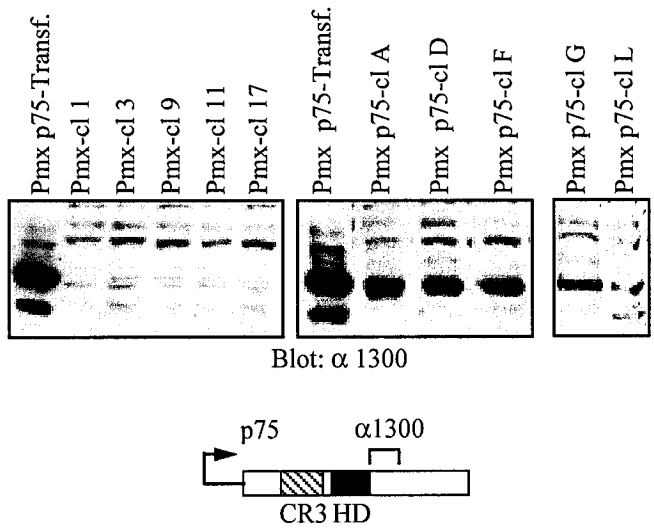

A
Probe
tRNA
HMEC
MCF7
SKBr3
Bt20
MDA 436
BT 549
MDA 231
MDA 468
MDA 435s
MCF10A
MCF12A
Placenta
450
317
237
28S
18S
B
HMEC
MCF-12A
MCF-10A
MCF-7
BT-20
MDA-231
MDA-436
Sk-Br-3
T 47D
ZR-75-1
(-)H2O
F19-B22
Fi20-B22
GAPDH
C
p75
MDA231
MCF-10A
HMEC
T47D
MCF7
p75
175
83
62
p75
Blot: α1300
Hs578T
HMEC
HMEC
p75
p75
HMEC
HMEC
Hs578T
175
83
62
Blot: α1300
Blot: α409
Coomassie
Coomassie
D
Breast tumor (-RT)
Breast tumor
HMEC
Hs578Bst
Hs578T
HC11 (-RT)
HC11
Transformed-HC11
F19-B22
Fi20-B22
GAPDH
F18-B22
Fi20m-B22
GAPDH

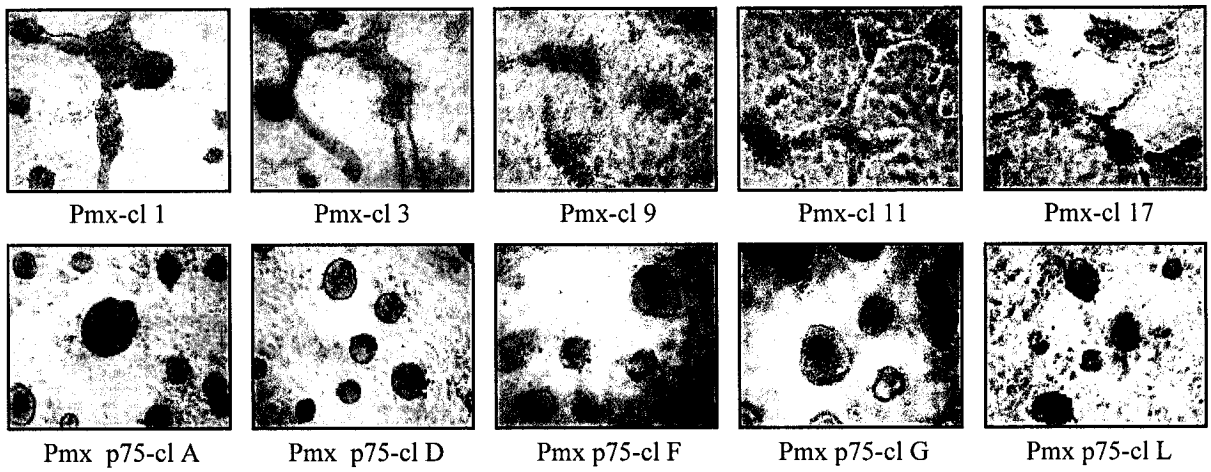
Fig. 7: T47D cell line stably expressing p75 cannot form tubule in collagen

(A) Cells from the T47D breast tumor cell line were transfected with pMX-p75 or the empty vector, pMX139, along with a neomycin expressing vector. Following a 3 week-selection in the presence of G-418, stably transfected cell clones were expanded and analyzed for p75 CDP/Cux expression by Western blotting with the CDP/Cux1300 antibodies. A diagram of the p75 expression vector is shown below. **(B)** Stably transfected T47D clones were analyzed for tubulogenesis in collagen assay. Tubules and cysts were photographed as described in material and methodes using a 10x objective. **(C)** Hematoxylin and eosin staining of T47D-pMX-p75 clone D cysts sectionned from paraffin were visualized as described in material and methodes using a 40x objective. The other clones showed the same morphology.

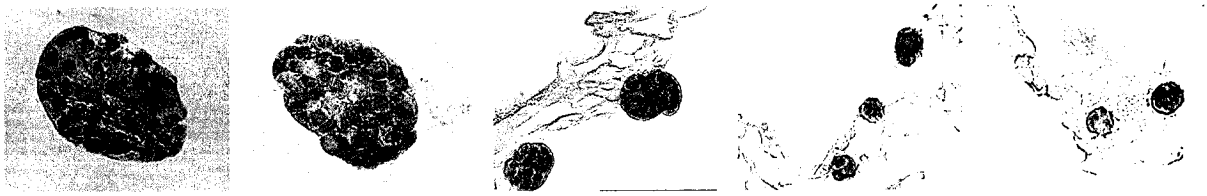
A



B



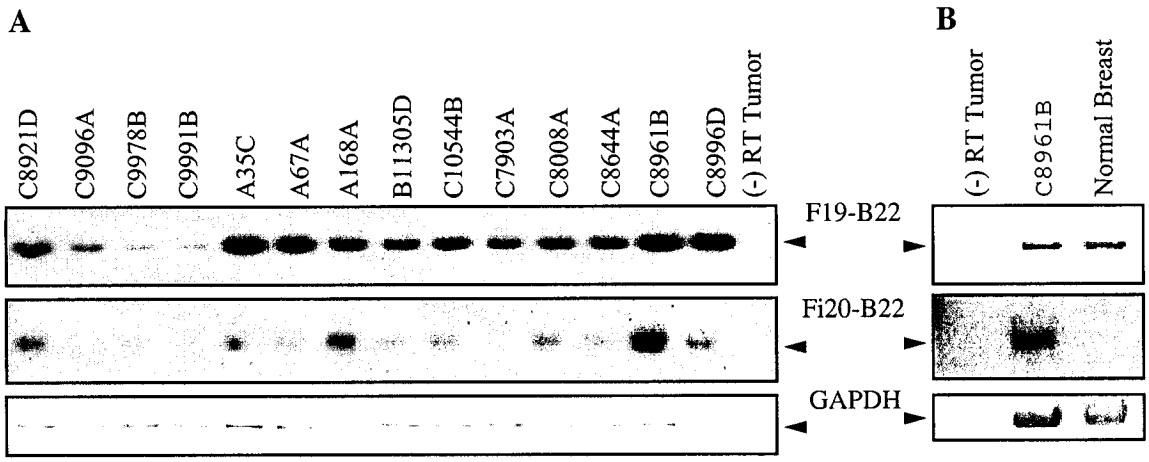
C

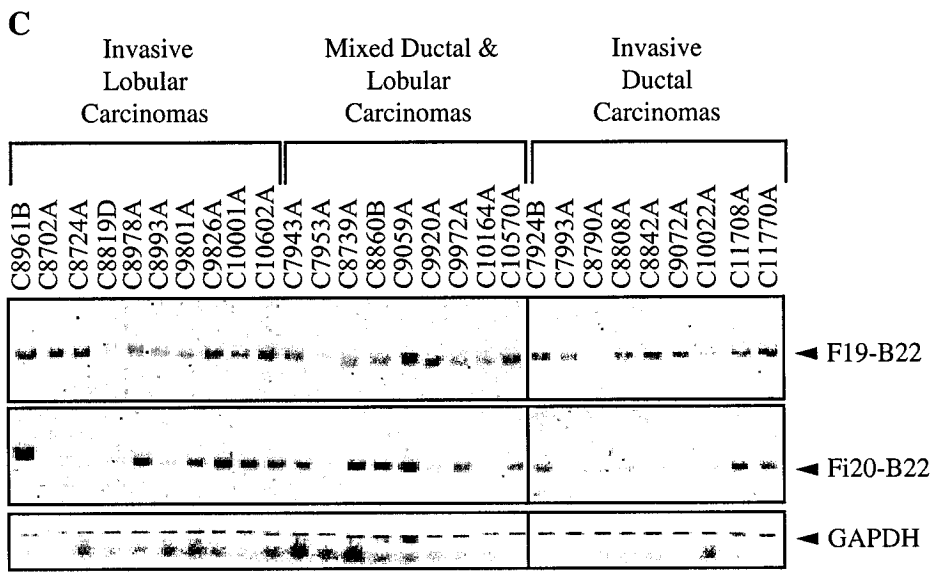


Pmx p75-cl D

Fig. 8: RT-PCR analysis of CDP/Cux mRNA expression in normal breast tissue and breast tumor.

(A) Total mRNA was isolated from a panel of 14 human breast tumors. RT-PCR was performed with either the F19 and B22 primers (upper panel) or the Fi20 and B22 primers (lower panel). Equal mRNA/cDNA amounts were verified using GAPDH mRNA amplification as a control. **(B)** Similar RT-PCR assays were performed using RNA isolated from a reduction mammoplasty tissue sample from three women without known breast pathology. **(C)** Total mRNA was isolated from a panel of 32 human breast tumors divided in three classes: invasive lobular carcinomas (ILC); invasive ductal and lobular mixed carcinomas (IDLC) and invasive ductal carcinomas (IDC). RT-PCR was performed as described in (A).


A
B
C8921D
C9096A
C9978B
C9991B
A35C
A67A
A168A
B11305D
C10544B
C7903A
C8008A
C8644A
C8961B
C8996D
(-) RT Tumor
(-) RT Tumor
C8961B
Normal Breast
F19-B22
Fi20-B22
GAPDH


C
Invasive Lobular Carcinomas
Mixed Ductal & Lobular Carcinomas
Invasive Ductal Carcinomas
C8961B
C8702A
C8724A
C8819D
C8978A
C8993A
C9801A
C9826A
C10001A
C10602A
C7943A
C7953A
C8739A
C8860B
C9059A
C9920A
C9972A
C10164A
C10570A
C7924B
C7993A
C8790A
C8808A
C8842A
C9072A
C10022A
C11708A
C11770A
F19-B22
Fi20-B22
GAPDH

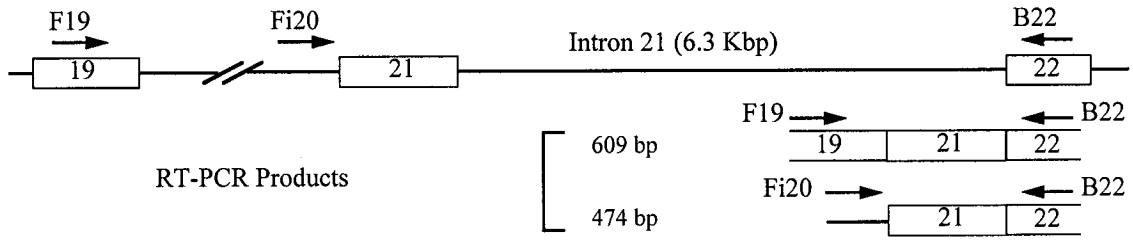
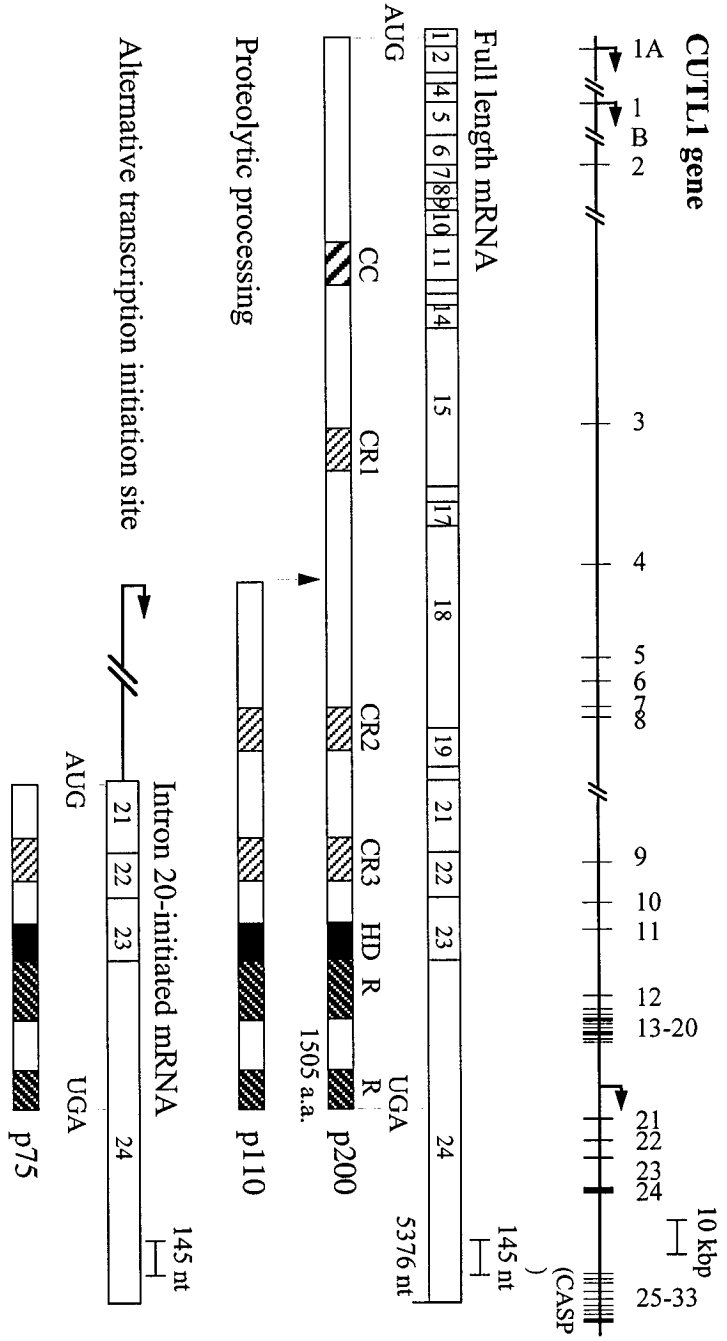

F19
Fi20
Intron 21 (6.3 Kbp)
B22
19
21
22
RT-PCR Products
609 bp
474 bp
F19
B22
19
21
22
Fi20
B22
21
22

Fig. 9: Schematic representation of the CUTL1 gene, some of its mRNAs and proteins

The exon/intron structure of the gene is shown at the top. Below are two CDP/Cux mRNAs , the full length mRNA and the intron 20-initiated mRNA (I20), as well as the proteins encoded by these mRNAs: p200, p110 and p75. Note that two mechanisms can be used to generate amino-terminally truncated CDP/Cux isoforms: proteolytic processing of p200 to generate p110, or transcription initiation at a start site within intron 20 to produce p75. We speculate that the existence of two modes of regulation enables the production of short Cut proteins in response to different signals or developmental cues.


CUTL1 gene
1A
1
B
2
3
4
5
6
7
8
9
10
11
12
13-20
21
22
23
24
25-33
(CASP
)
10 kbp
145 nt
Full length mRNA
1
2
4
5
6
7
8β
10
11
14
15
17
18
19
21
22
23
24
AUG
CC
CR1
CR2
CR3
HD
R
R
UGA
5376 nt
Proteolytic processing
1505 a.a.
p200
p110
Alternative transcription initiation site
Intron 20-initiated mRNA
145 nt
21
22
23
24
AUG
UGA
p75

Chapter V- General Discussion

In this thesis, I have identified cathepsin L as the protease responsible for the proteolytic cleavage of CDP/Cux at the G1/S transition of the cell cycle. I have also shown that in transformed cells, short CDP/Cux isoforms can be generated by different mechanisms, which are alternative transcription initiation and increased proteolytic processing. The fact that the enzyme that cleaves CDP/Cux is a lysosomal enzyme raised many questions and opened a whole new field in cell cycle and cancer research. In this last section, I will discuss the significance of my results as well as some of the issues that still remain to be addressed.

5.1 The Short CDP/Cux Proteins

Until now, we have detected at least three different short endogenous CDP/Cux proteins that do not possess the amino-terminal region present in the full-length protein. We found that two CDP-Cux proteins were generated by proteolytic processing of the full length CDP/Cux: p110 and p95. The other one, p75, was produced from a novel mRNA transcript initiated within intron 20 (i20) of CUTL1. Another proteolytic event, carboxy-terminal cleavage was also observed to occur in a proliferation or cell cycle regulated manner (79). These C-terminally-truncated products were also detected following transient transfection of p200 or p110 in some breast tumor cell lines, such as MCF 7 and Hs578T. A student in the lab is currently studying these isoforms and found that they might be responsible for the transcriptional activation of the DNA polymerase alpha gene promoter rather than the p110 (Truscott M. *et al*, manuscript in preparation). As this region contains

two active repression domains, their removal might be necessary to convert CDP/Cux from a repressor to an activator. Whether proteolytic cleavage within the carboxy-terminal domain of CDP/Cux is also aberrantly regulated in cancer cells will have to be investigated. These short CDP/Cux isoforms are likely to play an important role in the cell.

5.1.1 The p75 Protein

The p75 isoform differs from p110 in many ways. First, it is expressed in a tissue-specific manner such as in thymus, prostate and placenta, but not in the normal breast. However, the i20 transcript and p75 protein were expressed in many breast tumor cell lines and in a subset of breast cancers (which are mainly epithelial cells) (chapter IV). It is tempting to speculate that p75 could be involved in regulating cell type specificity in mammalian cells. For example, in the thymus, p75 expression correlated with certain stages of the development of thymocytes (Chapter IV). Three of the CDP/Cux knockout mice generated showed general pleiotropic developmental abnormalities (23, 70, 104). However, as CR3HD is deleted, we cannot distinguish between the activities of p110 and p75. Regarding the link between p75 expression and cancer, the effect of p75 could be at the level of cellular differentiation rather than proliferation. This link could be similar to the example of the Notch receptor. The Notch receptor is involved in the regulation of development and in tissue homeostasis, and its role in cancer has also been well established (72).

When epithelial cells undergo transformation, they often "dedifferentiate" and can acquire a mesenchymal phenotype (14, 32). T47D cells that stably over-express p75, are unable to differentiate into tubules (EMT) when grown in collagen (chapter IV). In addition, epithelial-mesenchymal transitions were observed in a number of MCF12A clones upon selection for stable over-expression of p75. This EMT was transient and lost upon

treatment with trypsin. The cells returned to their normal morphology but proliferated slightly faster than cells stably transfected with empty vector. It would be interesting to test whether these T47D-p75 cells have become more aggressive, i.e. if they can form tumors in nude mice and, if so, would these metastasize.

We have generated lines of transgenic mice expressing p75 in order to test the oncogenic or differentiation potential of p75. Three types of constructs were generated for these purposes. In one, the transgene was under the control of the MMTV promoter. Another had the same promoter flanked with hprt sequences for specific integration at the hprt locus (32). The third line was generated with p75 under the control of a CMV enhancer and chicken β-actin promoter. So far, all mice were born with a normal phenotype and the usual mendelian frequencies. Mice from the CMV line displayed organomegaly in a number of tissues (such as the heart and the thymus) at an early age. A more in-depth analysis of this will be performed. These phenotypes are in accordance with the ones observed in the *cux1* transgenic mice (59). Interestingly, one mouse from the hprt line showed a defect in kidney development; no nephrons were formed. This phenotype is consistent with the fact that p75 inhibits tubule formation. However, as of yet, we have not observed additional mice with this abnormality. These mice are currently being monitored for tumor development. Eventually, they should be crossed with other transgenic mice overexpressing known oncogenes (such as ras, erbB2/neu, c-myc) to measure the rates of tumor formation.

In light of my results (chapter IV), we are currently generating antibodies raised against the amino-terminus of p75. Once characterized, they will be used to perform immunohistochemistry on breast tumors collected by the MUHC Breast Tumor Bank. We will correlate the staining of p75 with some of the known pathological features of tumors.

The staining of p75 may emerge as a novel biological marker for certain breast cancers, and be used as a predictive or prognostic tool.

Another comparison that remains to be made is the *in vivo* DNA binding characteristics of p75 and p110. As p110 may be expressed in cells that harbor p75 (for example in the breast tumor cell lines), the question arises if these isoforms compete for the same binding sites. From studies using the Tandem Affinity Purification tool (Tap-tag), p75 and p110 seem to have different binding partners (Leduy L. and Nepveu A.) (15, 93). In accordance with this result and if indeed p75 is involved in cellular differentiation, we would expect it to regulate different genes than p110. Another possibility could be that, *in vivo*, p75 may act as a dominant negative on some genes regulated by p110. This could happen if some of the essential co-factors recruited by p110 do not interact with p75, and as the latter binds more stably to DNA, it would be preferentially bound, preventing the action of p110 and its associated factors. There are various examples in the literature of two proteins from the same gene having divergent effects. Some isoforms of the C/EBPβ (CCAAT/enhancer binding protein) transcription factor have opposite transcriptional effects. The 35 kDa LAP protein (liver-enriched transcriptional activating protein) is, as its name indicates, a transcriptional activator, whereas the 21 kDa LIP protein (liver-enriched transcriptional inhibitory protein) functions as a dominant negative inhibitor. Both proteins are translated from the same mRNA (18). Interestingly, site-specific proteolysis of the full length C/EBPβ was also shown to give rise to protein analogous to LIP (125). It is also the case for two isoforms of the p63 regulator, a member of the p53 family of proteins. While TAp63alpha and DeltaNp63alpha arise from the same gene, transcription of the latter is initiated within intron 3 of the gene and as a consequence DeltaNp63alpha does not express

amino-terminal sequences (21). Both proteins have opposing regulatory functions on many genes in cancer and development (128). A last example is c-Myc, where shorter proteins are translated via leaky scanning and can act as dominant negatives of the full length protein (106).

5.1.2 The p110 Isoform

A role for CDP/Cux during the cell cycle initially came from studies indicating that it was the DNA binding subunit of the promoter complex HiNF-D, which regulates the transcription of histone H4 gene at the G1/S transition (120). Interestingly, in transformed cells, the DNA binding of HiNF-D was constitutively elevated through the cell cycle (37). I speculate that the isoform of CDP/Cux that is part of the HiNF-D complex might be p110 at the G1/S transition and in transformed cells. In a preliminary experiment, I observed that p110 could bind to the site II of the histone H4 promoter (data not shown) (118, 119). This hypothesis could be confirmed by doing supershift experiments with CDP/Cux antibodies that specifically recognize the p110 species in electrophoretic mobility shift assays using the purified HiNF-D complex (120). A panel of CDP/Cux antibodies has been sent to the laboratories of Drs. Stein and Van Wijnen in order to test this hypothesis.

Stable NIH-3T3 cells over-expressing p110 confirmed the role for this CDP/Cux isoform in cell cycle (chapter II). Interestingly, the more these cells were maintained in culture, the higher the DNA content, (i.e. multiples of 2n larger than 4n) indicating a role for CDP/Cux in endoreduplication (Goulet B. and Sansregret L., unpublished observations).

The potential role of p110 in cell cycle progression does not appear to be essential for life. The knockout mice are viable. The MEF *cux* $1^{-/-}$ cells do not have any obvious

defect in their cell cycle profile (although their behavior has not been extensively studied) (70) (Sansregret L., unpublished observations). This could be explained by the fact that the MEF cells may not divide as rapidly as NIH3T3 cells. A relevant observation is that by immunohistochemistry, CDP/Cux expression was mainly observed in epithelial cells (of the mammary gland, the colon, the uterus and the lung) and not in the surrounding mesenchymal cells like fibroblasts (Poirier M., Master Thesis, 2002). Again, mesenchymal cells do not multiply as much as epithelial cells. Alternatively, a second cux gene (*cux-2*) has been identified (96, 112), and could complement *cux-1*. Although cux-2 expression was not upregulated in the MEF *cux* $1^{-/-}$ cells (Gingras H. and Nepveu A.), and is not as efficiently processed into a p110 isoform its different DNA binding characteristic may still allow it to complement *cux-1*.

Western blot analysis using the 861 antibody revealed an increased amount of p110 in pregnant and lactating mouse mammary gland, in which epithelial cells are proliferating, compared to tissues isolated from virgin mouse. As the gland underwent involution (apoptosis), the levels of p110 dropped (data not shown). Interestingly, some p110 mice displayed a "big nipple" phenotype after weaning, suggesting that in the continued presence of p110, cells do not undergo apoptosis properly (Furtenbacher H., Nepveu A.). In agreement with this observation, CDP/Cux knockout mice exhibit increased apoptosis of thymocytes (104). These results would suggest that the presence of p110 might also protect cells from undergoing apoptosis (and vice-versa). Other result also indicated the presence of p110 in specifically proliferating cells. In growing MDA 231 breast tumor cells (two days in culture), Western blot and DNA precipitation indicated increased expression of p110 and p95 proteins as well as increased DNA binding activity. Yet at high confluence (6 days in culture; most cells are in G0/G1), the expression and DNA binding activity of

both isoforms were lost. Interestingly, the intensity and activity of a protein migrating at 75 kDa did not vary (data not shown). This would suggest that p75, in contrast to p110, is not regulated in a cell cycle-dependent manner.

5.2 CDP/Cux in Cancer

No clear role for CDP/Cux in cancer has been identified as of yet. Initially, LOH studies in leiomyomas and breast cancer samples implicated CUTL1 as a tumor suppressor gene located on 7q22 (135, 136). Further analysis of CDP/Cux expression in leiomyomas, as well as cell cycle studies in NIH3T3, pointed more towards CDP/Cux as being an oncogene (80)..

This discrepancy can be explained in two manners. The simplest way would be that the tumor suppressor gene at 7q22 is not CUTL1 but another gene. Moreover, no mutations were found in CUTL1 in breast cancers, or in CDP/Cux cDNAs in uterine leiomyomas (80).

Alternatively, different isoforms of CDP/Cux could produce these opposing effect. The tumor suppressor could be p200, whereas p110 would be the oncogene. There are data that support this alternative hypothesis. First of all, p200 was found to compete with p110 for a binding site within the DNA polymerase α promoter (Moon N.S., unpublished results). Thus, the CCAAT displacement activity of p200 could prevent activation of the DNA polymerase α gene promoter. Secondly, it has been very difficult to generate stable cell lines over-expressing p200. The over-expression of p200 might impede cell cycle progression. This was finally achieved by retroviral infections, which allowed much lower expression of the protein. Thirdly, I made an interesting observation when soft agar assays

were performed using the wild type ras-3T3 cells and those stably expressing either p110 or p200. Ras cells expressing the full-length protein generated the same number of colonies (or slightly fewer) than wild type. Ras-3T3 with stably integrated p110 formed twice as many clones compared to wild type. Subsequent subcutaneous injection of these Ras 3T3/p110 cells into nude mice formed larger tumors that had more angiogenesis (data not shown). These results suggest that Ras and p110 are oncogenes that can cooperate. In the future, crosses of p110 transgenic mice with other oncognene transgenics may further confirm this hypothesis.

CDP/Cux would not be the only protein with tumor suppressor and oncogene properties. The transcription factor E2F1 was initially shown to promote cell proliferation by activating the transcription of several S phase genes. However, when knockout animals were generated, these animals were viable and developed neoplasias with age. E2F1 became the first protein to function both as an oncogene and a tumor suppressor gene (42, 91, 105, 130). This dual activity was rationalized with the fact that E2F1 is able to be an activator and a repressor of transcription, and by its role in both proliferation and apoptosis (41). Another protein with dual tumorigenic activity is Ras. Ras is the most frequently activated oncogene in cancer. However, people are finding tumors in which wild type Ras is inactivated. In neuroblastomas, H-Ras overexpression leads to tumor regression (51). In lung carcinogenesis, K-Ras2 displayed tumor suppressor activity and was frequently lost during lung tumor progression (134). The mechanisms responsible for these observations are currently unknown. One possibility comes from the observation that, depending on the cell type, Ras can induce either cellular differentiation (as in PC12 cells) or proliferation. A last example is the membrane receptor Notch. Aberrant signaling by Notch has been associated with tumorigenesis (40, 133). Recently, a role for Notch 1 as a tumor suppressor

gene in mouse skin cancer models was proposed (85). Interestingly, all the proteins mentioned above were shown to have dual activities, depending on the cell type and the cell environment.

A series of transgenic mice were generated to study the role of CDP/Cux *in vivo*. Three transgenic mouse lines were generated that over-expressed different CDP/Cux isoforms (p200, p110 and p75) under the control of the MMTV promoter flanked by hprt sequences (32). The advantage of using the hprt locus was that all transgenes integrated at the same location in the genome. These mice are currently being backcrossed to FVB mice to facilitate the study of mammary glands. Indeed, the majority of the research done on this topic was performed in the FVB strain of mouse. Nevertheless, as backcrosses are progressing, three of the original p110-expressing mice were found to be sick.

One of them was a year and a half old and had enlarged mammary and salivary gland lymph nodes. Histopathology analysis revealed a lymphoma. Although MMTV-driven expression is supposedly restricted to mammary tissues, some oncogenes under its control also induced lymphomas. This occurred with other transcription factors, A-Myb and c-myc and the tyrosine receptor ErbB2 (17, 58, 109). Moreover, there are specific and inducible transcription factors that regulate MMTV expression in B cells (3, 69). The enhanced expression of CDP/Cux could prevent apoptosis in lymphoid cells, and/or induce proliferation of these cells (104). ChIP assays using specific CDP/Cux antibodies are revealing novel genes bound by p110 (Harada R. Nepveu A. unpublished observation). One of them is Bcl6, an important regulator lymphocytes differentiation that, when upregulated, can induce lymphomas (107).

Another MMTV-p110 mouse at one year and two months had a liver that was about ten times larger than normal. Furthermore, a tumor was detected on one of the lobes. The spleen was also unusually large in size.

The third MMTV-p110 mouse at one year had a pancreas that was full of tumors. The phenotypes of these mice add to increasing amount of evidence for a role for p110 in cell proliferation and in tumorigenicity. A mouse expressing the full-length isoform was sacrificed at six months of age because of illness. Among the ailments were an ocular prolapse and ascites. It also displayed lymphadenopathy and organomegaly of the spleen and kidneys. The organomegaly observed in these mice is in agreement with the results obtained with the *cux1* transgenic mouse (59).

One interesting question would be to know if CDP/Cux knockout animals are less prone to tumors. Crosses with transgenic mice overexpressing oncogenes or treatment of these and control animals with tumor promoting agents would provide evidence against or in favor of this hypothesis. However, one type of cancer these mice may get is myeloid leukemias. The $cux1^{\Delta HD/\Delta HD}$ mice showed an increase in myeloid cells and furthermore, 7q22 was deleted in acute myeloid leukemias and myeloid dysplasia (27, 43, 66, 67, 104, 113). This discrepancy points to cell- and/or isoform-specific roles for CDP/Cux proteins. Could this protein be an oncogene for certain types of cells and a tumor suppressor gene for others?

Knowing the importance of p110 in the cell cycle, the regulation of the cleavage of p200 in normal cells should be studied. Projects are currently underway to find a role for cyclin D/Cdk4 in the inhibition of amino-terminal proteolytic processing (Santaguida M. *et al.* Manuscript in preparation). Interestingly, cyclin E/Cdk2 stimulated the cleavage of CDP/Cux (Leduy L., Santaguida M. and Nepveu A.). That cyclin D is upregulated in some

cancers does not preclude an increase in the processing of p200 in these cells (30). As a matter of fact, recent reports indicate that cyclin D could act as a sink for the Cdk inhibitors (CKI) p21 and p27, therefore releasing cyclin E/Cdk2 from inhibition (103). It is currently not clear whether the effect of cyclin E/Cdk2 on CDP/Cux processing is direct or if it is via the activation of Cdc25A or both. As a preliminary results, Cdc25A was shown to enhance the proteolysis of CDP/Cux (Hebert S. Master thesis, 2001). Results therefore suggest a tight regulation of CDP/Cux cleavage, at the level of CDP/Cux itself (for example by phosphorylation) and via the modulation of cathepsin L levels.

5.3 The Novel Role of Cathepsin L in Cell Cycle Progression

5.3.1 Which Cathepsin L Isoform Translocates to the Nucleus?

A short cathepsin L isoform translocates to the nucleus in a cell cycle-regulated manner. In order to do so, translation must be initiated at downstream AUG sites to generate these shorter isoforms lacking a signal peptide.

Under normal circumstances, this short sequence located at the amino-terminus allows the protein to be routed to the endoplasmic reticulum (ER). Once the newly formed protein enters this compartment, it is glycosylated and folded with the help of resident chaperones. A correctly folded protein continues its path towards the trans golgi network (TGN), and eventually is either secreted, or sent to other compartments, such as the lysosomes. Misfolded proteins are retained in the ER to either be refolded or degraded (22).

In the ER, truncated cathepsin L species cannot fold properly. In contrast to the zymogene, a cathepsin L isoform that starts at methionine 56 and is fused to a signal peptide cannot not be secreted in *Pichia pastori*. The detection of AUG56 in total extracts

indicated that the enzyme accumulated inside the yeast (data not shown). The portion of the prodomain missing in short cathepsin L isoforms contained some residues important for the folding and glycosylation of the protein (11, 101, 111). When abnormal folding occurs, proteins aggregate and are eventually degraded (111).

Since shorter cathepsin L isoforms are devoid of a signal peptide and do not enter the ER, how does proper folding occur in a non-redox environment? On the one hand, recent evidence indicates that a redox potential might be available in the cytosol. Members of the protein disulfide isomerase (PDI) family were found in non-ER locations, including the cytosol and the nucleus (117). It is also becoming clear that some cytosolic chaperones, such as Hsp90 or Hsp33 displayed redox activities (31, 82, 90). Interestingly, the redox potential of the mammalian cell changes during the cell cycle, where a more oxidative state may be necessary for the G1/S transition (78). This could help the proper folding of all the necessary factors that are newly synthesized. On the other hand, the folding machinery in the cytosol is somewhat different than in the ER. The cytosol contains chaperones that function while the nascent protein is still on the ribosome, and others that form barrel structures and provide the fully synthesized proteins with a favorable environment for folding (28). Therefore it may be possible that the cytosolic chaperones somehow serve as substitutes for the proregion. It must be noted that aggregation of unfolded or partially folded intermediates is sometimes observed in crowded environments of the cells. Following transient expression of cathepsin L AUG 56, some cells displayed the diffuse pattern shown in Chapter II (see figure 6), while others had various big bright dots (data not shown). DAPI staining and the use of a catalytically inactive mutant ruled out apoptosis (data not shown), suggesting that either the folding of these short species is not very efficient, or that it is cell cycle regulated. Recent reports are showing that the levels of heat

shock proteins Hsp90 and Hsp70 increased during cellular proliferation in a cell cycle regulated manner (34, 35), and that Hsp90 is important for the folding of many cell cycle regulated proteins (90). In either case, from these results overexpression of AUG56 by transfection may not be the correct way to generate active enzyme.

The next question is which form of cathepsin L translocates to the nucleus; is it the proform or the mature enzyme? It is more likely the former, since the mature, active protease could display aberrant processing activity. Reports in the literature revealed that other proteases such as caspase 2 and caspase 1 translocate to the nucleus as proforms (4, 73). The 20S proteosome also localized to the nucleus as a precursor complex (61). Yet, it is not always the case as the prodomain of caspase 7 prevented its nuclear localization (131). If it was the mature enzyme that was translocating, it would need to be associated intermolecularly with an inhibitor other than its prodomain. For example, similarly to MENT, the general cysteine protease inhibitor cystatin B (Stefin B) was found, by electron miscroscopy, in the nucleus of proliferating cells and in the cytoplasm and nucleus of differentiating cells (39, 98). As the enzyme-inhibitor binding is reversible, once in the nucleus, we could hypothesize that binding site competition could remove the inhibitor and allow the substrate to enter in the active site (64). Future studies should also tell us whether the mechanism by which cathepsin L translocates to the nucleus is active or passive. Until now, not much information is available on the nucleocytoplasmic shuttling of proteases. Although calpain activity in the nucleus is widely accepted, the way it migrates to the nucleus is still unknown. To my knowledge, only one paper reported the nuclear localization of one calpain isoform. They showed that mu Calpain was going to the nucleus in an ATP dependent manner and that interestingly, its endogenous inhibitor calpastatin did not prevent this (77). Caspase 2 and the 20S proteasome were shown to

bind the importin alpha/beta complex via their NLS (nuclear localization signal) (4, 61). As pointed out in chapter II, cathepsin L also has a stretch of basic amino acids in its prodomain. However, mutation of 3 lysines did not prevent its nuclear localization, indicating that either the translocation is independent of a typical NLS sequence, or cathepsin L travels by association with another protein, as was proposed for cyclin D (19, 56). Regulated nuclear export also plays an important role in determining protein localization (71). It is therefore not impossible that the nuclear localization of cathepsin L could be regulated in this manner. This could be easily tested by treating cathepsin L transfected cells with leptomycin B an inhibitor of CRM1-mediated nuclear export (127). Finally, the nuclear pore diameter allows proteins up to 50 kDa in size to diffuse freely across the nuclear envelope (75). Even if the molecular masses of the cathepsin L isoforms are between 20 and 33 kDa, the fact that over-expression of cathepsin L in NIH3T3 cells did not process as much CR1CR2 substrate as in the *in vitro* assays (chapter III, figure 6) tends to exclude this possibility and rather suggests the active transport of cathepsin L. Another issue is within which nuclear compartment does cathepsin L resides. As observed in chapter II, transfection of full-length cathepsin L generally gives a specific nuclear signal by indirect immunofluorescence. When the mutant M1 or AUG56 were transfected, a more diffuse pattern was observed. This difference in signal is not yet fully understood.

5.3.2 Cell cycle regulation of cathepsin L

Cathepsin L activity during the cell cycle could be regulated by the rate of its nuclear translocation. It is becoming clear that nuclear localization of many proteins is controlled in a cell cycle-dependent manner at least at two levels. The first one is on the target protein itself where some events such as phosphorylation/dephosphorylation, or proteolytic cleavage can unveil an NLS/NES (nuclear localization signal/nuclear export

signal). The other one is through interaction with the transport machinery (132). It was shown that the capacity of nuclear transport and the size of nuclear pore increased in proliferating cells (25, 26). Furthermore, the rate of nuclear import directly depends on the quantity of cargo proteins available (110, 132). Thus during the cell cycle, cargo protein levels should also be augmented as general protein synthesis increases.

In parallel to its transport into the nucleus, the regulation of cathepsin L activity is also important. Nothing is yet known about post-translational modifications that may affect its activity or localization. RT-PCR analysis using NIH3T3 cells separated at various stages of the cell cycle indicated that the mRNA levels of cathepsin L were constant during all phases (data not shown). However, ^{35}S labeling of synchronized NIH3T3 cells showed an increase in the steady state level of all cathepsin L isoforms at the G1/S transition (data not shown). This result is not really surprising as the general rate of protein synthesis increases as cells are in their growth phase (7). Additionally, during G1/S progression, the translation of specific cell cycle proteins is enhanced (94). The synthesis of potent regulatory proteins, maybe including short cathepsin L, needs to be tightly controlled. This is accomplished not only at the level of transcription (as is the case for some proteins), but also of translation, by regulating the translation initiation factors.

Secondary structures at the 5'UTR of certain mRNAs can affect the efficiency of initiation (55). As mentioned in chapter II, human cathepsin L mRNA can be alternatively spliced into species with shorter 5'UTR that are lacking secondary loops and are translated more effectively (2). Translation is a way to generate amino-terminally truncated proteins, which may then have different localization and/or biological properties. For example, in certain situations ribosomal reinitiation (or shunting) occurs: following a terminator codon, the 40S subunit stays on the mRNA and reinitiates at downstream AUGs. This mechanism

depends on the size of the upstream ORF, and as reinitiation is generally very inefficient, these uORF are used to decrease translation efficiency (54). Leaky scanning is an event that happens when the context surrounding the first AUG codon is suboptimal, or when the initating codon is not an AUG, so the ribosome continues scanning until the next AUG. Usually, a small amount of protein can be generated from leaky scanning (see chapter II). In addition, cap-independent translation takes place via the use of an internal ribosome entry site (IRES). In this case, the ribosome starts translation from an internal sequence. IRES were initially identified in viruses and have since been identified in mammalian cells. The experimental evidence in favor of IRES has remained somewhat controversial (54, 102). Some groups actually found cap-independent translation to be active during the G2/M phase of the cell cycle, when protein synthesis is not as prominent (95). I propose that leaky scanning is the mechanism generating the truncated cathepsin L isoforms. I excluded the presence of an IRES for two reasons. First of all, cloning of the first 162 nucleotides of the cathepsin L mRNA (i.e. from AUG 1 to AUG 56) between the firefly and renilla luciferases genes did not allow the synthesis of the renilla luciferase. Moreover, rapamycin decreased the synthesis of all the cathepsin L isoforms, thus excluding cap independent translation for the shorter isoforms (data not shown). Importantly, rapamycin also prevented proteolytic processing of CDP/Cux. Knowing the effect of p110, it is tempting to suggest that the cell cycle inhibitory action of rapamycin could also be mediated in part by preventing the synthesis of cathepsin L and therefore the processing of CDP/Cux.

A proof for the role of short cathepsin L in the cell cycle would be provided if ectopic expression of cathepsin L would be able to alter cell cycle progression and/or profile. We attempted to generate cell lines using retroviral infections, which permit lower

expression of the protein. The tagged AUG56 isoforms were barely detected by indirect immunofluorescence and Western blotting. The cell cycle analysis of these populations has not been studied yet. However, in light of the potential role for cathepsin L in the cell cycle progression, and since it is a protease, we suspect that an excess amount of it could be deleterious for the cells. Inducible expression systems might circumvent this problem and allow cell cycle studies.

5.4 Cathepsin L in Cancer: Novel Way to Inhibit Cell Proliferation?

The involvement of cathepsin L in tumorigenesis and metastasis has been studied for many years. First of all, cathepsin L was independently identified as the major excreted protein (MEP) in transformed cells (44, 114, 116). Cathepsin L was then shown to be overexpressed in many malignant human tumors, and many reports correlated this increased expression with the invasive behavior of tumors (5, 12, 20, 65, 89, 100, 108). In tissue culture models, phorbol esters or certain oncogenes such as ras, v-src, SV-40 Large T and raf, cytokines (IL-1, IL-6, TNF α) were shown to also induce cathepsin L expression and secretion (29, 33, 46, 62, 63, 115). Moreover, cathepsin L expression correlated with the metastatic potential of ras-transformed cells (16). Therefore most of the research on cathepsin L and its relationship with cancer focused on the secreted isoform and its involvement in invasion and metastasis. In addition, an imbalance between the endogenous cysteine protease inhibitors (stefin and cystatins) and their enzyme is observed in some cancer patients (52, 53). Only a short paper reported the localization of cathepsin L in the nucleus of some transformed cells by cell fractionation and Western blotting (36). Interestingly, the authors commented on an unrelated 33 kDa isoform observed in the

nuclear fractions of transformed cells. It is possible that this band corresponds to the truncated cathepsin L isoform that I identified in chapter II.

It is believed that secreted cathepsin L mainly cleaves extracellular matrix proteins (47). Cancer cells treated with various inhibitors of cathepsin L, such as chemical compounds, monoclonal antibodies or anti-sense RNA showed not only a reduction in their metastatic potential, but in some cases, in cellular proliferation (48, 83, 123, 126). Some cysteine protease inhibitors were also shown to selectively inhibit proliferation of certain tumor cell lines, such as SkBr3 and MCF7 (129). Expression of antisense cathepsin L mRNA in malignant cells significantly reduced their tumor growth potential in nude mice (50). My own results (chapter III) along with the ones from Mellgreen et al, confirmed that the inhibitor E-64d could cause a block in the G1 phase of the cell cycle (76). Altogether these results suggest that cathepsin L might also control cellular proliferation. Therefore in cancer cells, I suspect that both the full length and the truncated nuclear isoforms are expressed at higher levels and contribute to tumorigenesis in different ways: the full-length isoform would promote metastasis, whereas the truncated nuclear isoform would stimulate proliferation. The novel spliced variant of human cathepsin L, hCATL AIII was expressed more highly in transformed cells and was translated at higher rates than the full length mRNA. Preliminary data from Dr. Chauhan seems to indicate that hCATL AIII also produces more of the short cathepsin L. In light of my results in chapter II and III, it is tempting to speculate that proliferation of cancer cells could be reduced by inhibiting the processing of CDP/Cux into the p110 isoform, via the targeting of cathepsin L.

One question will be to determine whether transgenic mice overexpressing cathepsin L are prone to tumors, and if p110 steady state levels will be increased in these

mice. In addition, I suspect that mice produced from a cross between p200 and cathepsin L animals should develop malignancies at a higher rate.

My results raise the idea that novel cell-permeable cathepsin L inhibitors specific for cathepsin L (that are able to function at neutral pH) or RNAi technology may be useful in preventing the action of intracellular cathepsin L in cancer cells. Most known inhibitors are used in *in vitro* assays at pH 5.5 on synthetic peptide substrates. However, CDP/Cux processing assays *in vitro* using cathepsin L isolated from *P.pastori* were done at pH 7, because CDP/Cux was barely processed at pH 5.5 (Goulet B. and Leduy L. unpublished observation). It is possible that the interaction between the enzyme and CDP/Cux is more stable at pH 7. The entry of CDP/Cux into the active cleft may be affected by the charges of the amino acids in the cleft resulting from the change in pH. It has been described for cathepsin B that inhibitors targeted to the occluding loop were of poor inhibitory design if the enzyme environment had a pH greater then pH 5.5, because of charge hindrance (10). This would suggest that in order to accurately target the nuclear enzymes, the inhibitors should be efficient at neutral pH. If this idea is correct, then we could envisage a situation whereby a tumor would be treated with a mixture of a non-permeable inhibitor that is active at low pH, and a permeable inhibitor that is active at neutral pH.

5.5 Protease Activity in Cancer Cells

Proteolysis is an important event in cell cycle progression. At each stage, key players are degraded by the ubiquitin/proteasome pathway. Studies and clinical trials have focused on the proteasome, as its inhibition resulted in cell cycle arrest and apoptosis in many tumors (1).

It is now becoming clear that site-directed proteolysis plays an important role in cancer. More and more examples of the activation of specific proteases, located in various cellular compartments, are emerging in cancer. As previously mentioned, proteolytic cleavage can generate truncated variants of a protein that either possess a more potent activity than the wild type, or on the contrary, act as dominant negatives. Alternatively, these shorter isoforms may have a different cellular localization. Various transmembrane receptors and some ligands, such as Notch, the Notch ligands Delta1 and jagged2, ErbB4, HER2/neu, E-cadherin, CD44 were shown to be processed (13, 38, 45, 57, 60, 68, 74, 81, 84, 88, 97, 99, 124). Following cleavage of Notch and erbB4, the intracellular domains translocate to the nucleus and function as transcription factors. The processing of E-cadherin can play an important role in cancer in two manners: via the disruption of cell-cell adhesion, and via the activation of β-catenin, which is imported in the nucleus and modulates transcription (86). When the cell surface receptor for hyaluronate, CD44, is cleaved at its intramembrane cleavage site, its intracellular domain is liberated and the truncated receptor loses its ability to interact with its substrate, thereby promoting cell migration. The intracellular fragment is a transcriptional activator (87). Regarding the change in activity of a given protein following cleavage, we showed the example of the transcription factor CDP/Cux. As mentioned in chapter III, cyclin E is proteolytically cleaved into more active molecules (92, 122). These low molecular weight cyclin E forms strongly correlated with a worse survival rate in breast cancer patients (49).

Many over-expressed proteases are active in cancer. Calpain was shown to be up-regulated in transformed cells and in human cancers (8, 9, 99). Metalloproteases were shown to be up-regulated in tumor cells and are involved in metastasis via the degradation of extracellular matrix proteins (121). As mentioned previously, this is the case for

cysteine proteases. Finally, some members of the family of prohormone convertases are expressed in various tumor lines and human primary tumors (6). Active proteases are becoming key targets in the search for inhibitors of cancer progression and metastasis. As more and more proteases have different substrates located in various compartments in the cell, it may be important to find specific cell permeable compounds for each substrate. Indeed, it was shown that secreted cathepsin L could also generate endostatin, an inhibitor of angiogenesis (24).

5.6 Conclusion

At last, these results have revealed that active short CDP/Cux isoforms can be generated by two mechanisms in cancer. Most importantly, the novelty of this thesis comes from the proposed new role for the cysteine protease cathepsin L. Not only is it targeted to the nucleus via alternative translation initiation but it processes a novel substrate: a transcription factor involved in cell cycle regulation. Moreover, my results indicate that this nuclear truncated cathepsin L may play a role in cancer, by stimulating the processing of CDP/Cux and perhaps other nuclear proteins. This suggests that the aberrant expression and activity of certain proteases within the cell could induce cellular transformation and cancer.

References

1. **Adams, J.** 2003. The proteasome: structure, function, and role in the cell. Cancer Treat Rev **29 Suppl 1**:3-9.

2. **Arora, S., and S. S. Chauhan.** 2002. Identification and characterization of a novel human cathepsin L splice variant. Gene **293:**123-31.

3. **Aurrekoetxea-Hernandez, K., and E. Buetti.** 2000. Synergistic action of GA-binding protein and glucocorticoid receptor in transcription from the mouse mammary tumor virus promoter. J Virol **74:**4988-98.

4. **Baliga, B. C., P. A. Colussi, S. H. Read, M. M. Dias, D. A. Jans, and S. Kumar.** 2003. Role of prodomain in importin-mediated nuclear localization and activation of caspase-2. J Biol Chem **278:**4899-905.

5. **Baricos, W. H., S. L. Cortez, Q. C. Le, L. T. Wu, E. Shaw, K. Hanada, and S. V. Shah.** 1991. Evidence suggesting a role for cathepsin L in an experimental model of glomerulonephritis. Arch Biochem Biophys **288:**468-72.

6. **Bassi, D. E., H. Mahloogi, and A. J. Klein-Szanto.** 2000. The proprotein convertases furin and PACE4 play a significant role in tumor progression. Mol Carcinog **28:**63-9.

7. **Brooks, R. F.** 1977. Continuous protein synthesis is required to maintain the probability of entry into S phase. Cell **12:**311-7.

8. **Carragher, N. O., and M. C. Frame.** 2002. Calpain: a role in cell transformation and migration. Int J Biochem Cell Biol **34:**1539-43.

9. **Carragher, N. O., M. A. Westhoff, D. Riley, D. A. Potter, P. Dutt, J. S. Elce, P. A. Greer, and M. C. Frame.** 2002. v-Src-induced modulation of the calpain-calpastatin proteolytic system regulates transformation. Mol Cell Biol **22:**257-69.

10. **Cathers, B. E., C. Barrett, J. T. Palmer, and R. M. Rydzewski.** 2002. pH Dependence of inhibitors targeting the occluding loop of cathepsin B. Bioorg Chem **30:**264-75.

11. **Chapman, R. L., S. E. Kane, and A. H. Erickson.** 1997. Abnormal glycosylation of procathepsin L due to N-terminal point mutations correlates with failure to sort to lysosomes. J Biol Chem **272:**8808-16.

12. **Chauhan, S. S., L. J. Goldstein, and M. M. Gottesman.** 1991. Expression of cathepsin L in human tumors. Cancer Res **51:**1478-81.

13. **Codony-Servat, J., J. Albanell, J. C. Lopez-Talavera, J. Arribas, and J. Baselga.** 1999. Cleavage of the HER2 ectodomain is a pervanadate-activable process that is inhibited by the tissue inhibitor of metalloproteases-1 in breast cancer cells. Cancer Res **59:**1196-201.

14. **Cotran, R. S., Kumar, V., Robbins, S.L.** 1989. Robbins Pathologic Basis of Disease, 4th Edition ed. W.B. Saunders Company.

15. **Cox, D. M., M. Du, X. Guo, K. W. Siu, and J. C. McDermott.** 2002. Tandem affinity purification of protein complexes from mammalian cells. Biotechniques **33:**267-8, 270.

16. **Denhardt, D. T., A. H. Greenberg, S. E. Egan, R. T. Hamilton, and J. A. Wright.** 1987. Cysteine proteinase cathepsin L expression correlates closely with the metastatic potential of H-ras-transformed murine fibroblasts. Oncogene **2:**55-9.

17. **DeRocco, S. E., R. Iozzo, X. P. Ma, R. Schwarting, D. Peterson, and B. Calabretta.** 1997. Ectopic expression of A-myb in transgenic mice causes follicular hyperplasia and enhanced B lymphocyte proliferation. Proc Natl Acad Sci U S A **94:**3240-4.

18. **Descombes, P., and U. Schibler.** 1991. A liver-enriched transcriptional activator protein, LAP, and a transcriptional inhibitory protein, LIP, are translated from the same mRNA. Cell **67:**569-79.

19. **Diehl, J. A., and C. J. Sherr.** 1997. A dominant-negative cyclin D1 mutant prevents nuclear import of cyclin-dependent kinase 4 (CDK4) and its phosphorylation by CDK-activating kinase. Mol Cell Biol **17:**7362-74.

20. **Dohchin, A., J. I. Suzuki, H. Seki, M. Masutani, H. Shiroto, and Y. Kawakami.** 2000. Immunostained cathepsins B and L correlate with depth of invasion and different metastatic pathways in early stage gastric carcinoma. Cancer **89:**482-7.

21. **Dohn, M., S. Zhang, and X. Chen.** 2001. p63alpha and DeltaNp63alpha can induce cell cycle arrest and apoptosis and differentially regulate p53 target genes. Oncogene **20:**3193-205.

22. **Ellgaard, L., and A. Helenius.** 2003. Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol **4:**181-91.

23. **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.

24. **Felbor, U., L. Dreier, R. A. Bryant, H. L. Ploegh, B. R. Olsen, and W. Mothes.** 2000. Secreted cathepsin L generates endostatin from collagen XVIII. Embo J **19:**1187-94.

25. **Feldherr, C. M., and D. Akin.** 1993. Regulation of nuclear transport in proliferating and quiescent cells. Exp Cell Res **205:**179-86.

26. **Feldherr, C. M., and D. Akin.** 1991. Signal-mediated nuclear transport in proliferating and growth-arrested BALB/c 3T3 cells. J Cell Biol **115:**933-9.

27. **Fischer, K., J. Brown, S. W. Scherer, P. Schramm, J. Stewart, G. Fugazza, U. Pascheberg, W. Peter, L. C. Tsui, P. Lichter, and H. Dohner.** 1998. Delineation of genomic regions in chromosome band 7q22 commonly deleted in myeloid leukemias. Recent Results in Cancer Research **144:**46-52.

28. **Frydman, J.** 2001. Folding of newly translated proteins in vivo: the role of molecular chaperones. Annu Rev Biochem **70:**603-47.

29. **Gerber, A., A. Wille, T. Welte, S. Ansorge, and F. Buhling.** 2001. Interleukin-6 and transforming growth factor-beta 1 control expression of cathepsins B and L in human lung epithelial cells. J Interferon Cytokine Res **21:**11-9.

30. **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.

31. **Graumann, J., H. Lilie, X. Tang, K. A. Tucker, J. H. Hoffmann, J. Vijayalakshmi, M. Saper, J. C. Bardwell, and U. Jakob.** 2001. Activation of the redox-regulated molecular chaperone Hsp33--a two-step mechanism. Structure (Camb) **9:**377-87.

32. **Hatada, S., W. Kuziel, O. Smithies, and N. Maeda.** 1999. The influence of chromosomal location on the expression of two transgenes in mice. Journal of Biological Chemistry **274:**948-55.

33. **Heinrich, J., M. Bosse, H. Eickhoff, W. Nietfeld, R. Reinhardt, H. Lehrach, and K. Moelling.** 2000. Induction of putative tumor-suppressing genes in Rat-1 fibroblasts by oncogenic Raf-1 as evidenced by robot-assisted complex hybridization. J Mol Med **78:**380-8.

34. **Helmbrecht, K., and L. Rensing.** 1999. Different constitutive heat shock protein 70 expression during proliferation and differentiation of rat C6 glioma cells. Neurochem Res **24:**1293-9.

35. **Helmbrecht, K., E. Zeise, and L. Rensing.** 2000. Chaperones in cell cycle regulation and mitogenic signal transduction: a review. Cell Prolif **33:**341-65.

36. **Hiwasa, T., and S. Sakiyama.** 1996. Nuclear localization of procathepsin L/MEP in ras-transformed mouse fibroblasts. Cancer Lett **99:**87-91.

37. **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, J. L. Stein, and G. S. Stein.** 1990. Tumor cells exhibit deregulation of the cell cycle histone gene promoter factor HiNF-D. Science **247:**1454-7.

38. **Ikeuchi, T., and S. S. Sisodia.** 2003. The Notch ligands, Delta1 and Jagged2, are substrates for presenilin-dependent "gamma-secretase" cleavage. J Biol Chem **278:**7751-4.

39. **Irving, J. A., S. S. Shushanov, R. N. Pike, E. Y. Popova, D. Bromme, T. H. Coetzer, S. P. Bottomley, I. A. Boulynko, S. A. Grigoryev, and J. C. Whisstock.** 2002. Inhibitory activity of a heterochromatin-associated serpin (MENT) against papain-like cysteine proteinases affects chromatin structure and blocks cell proliferation. J Biol Chem **277:**13192-201.

40. **Jhappan, C., D. Gallahan, C. Stahle, E. Chu, G. H. Smith, G. Merlino, and R. Callahan.** 1992. Expression of an activated Notch-related int-3 transgene interferes with cell differentiation and induces neoplastic transformation in mammary and salivary glands. Genes Dev **6:**345-55.

41. **Johnson, D. G.** 2000. The paradox of E2F1: oncogene and tumor suppressor gene. Mol Carcinog **27:**151-7.

42. **Johnson, D. G., W. D. Cress, L. Jakoi, and J. R. Nevins.** 1994. Oncogenic capacity of the E2F1 gene. Proc Natl Acad Sci U S A **91:**12823-7.

43. **Johnson, E. J., S. W. Scherer, L. Osborne, L. C. Tsui, D. Oscier, S. Mould, and F. E. Cotter.** 1996. Molecular definition of a narrow interval at 7q22.1 associated with myelodysplasia. Blood **87:**3579-86.

44. **Joseph, L., S. Lapid, and V. Sukhatme.** 1987. The major ras induced protein in NIH3T3 cells is cathepsin L. Nucleic Acids Res **15:**3186.

45. **Kajita, M., Y. Itoh, T. Chiba, H. Mori, A. Okada, H. Kinoh, and M. Seiki.** 2001. Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. J Cell Biol **153:**893-904.

46. **Kakegawa, H., K. Tagami, Y. Ohba, K. Sumitani, T. Kawata, and N. Katunuma.** 1995. Secretion and processing mechanisms of procathepsin L in bone resorption. FEBS Lett **370:**78-82.

47. **Kane, S. E., and M. M. Gottesman.** 1990. The role of cathepsin L in malignant transformation. Semin Cancer Biol **1:**127-36.

48. **Katunuma, N., H. Tsuge, M. Nukatsuka, and M. Fukushima.** 2002. Structure-based development of cathepsin L inhibitors and therapeutic applications for prevention of cancer metastasis and cancer-induced osteoporosis. Adv Enzyme Regul **42:**159-72.

49. **Keyomarsi, K., S. L. Tucker, T. A. Buchholz, M. Callister, Y. Ding, G. N. Hortobagyi, I. Bedrosian, C. Knickerbocker, W. Toyofuku, M. Lowe, T. W. Herliczek, and S. S. Bacus.** 2002. Cyclin E and survival in patients with breast cancer. N Engl J Med **347:**1566-75.

50. **Kirschke, H., R. Eerola, V. K. Hopsu-Havu, D. Bromme, and E. Vuorio.** 2000. Antisense RNA inhibition of cathepsin L expression reduces tumorigenicity of malignant cells. Eur J Cancer **36:**787-95.

51. **Kitanaka, C., K. Kato, R. Ijiri, K. Sakurada, A. Tomiyama, K. Noguchi, Y. Nagashima, A. Nakagawara, T. Momoi, Y. Toyoda, H. Kigasawa, T. Nishi, M. Shirouzu, S. Yokoyama, Y. Tanaka, and Y. Kuchino.** 2002. Increased Ras expression and caspase-independent neuroblastoma cell death: possible mechanism of spontaneous neuroblastoma regression. J Natl Cancer Inst **94:**358-68.

52. **Kos, J., and T. T. Lah.** 1998. Cysteine proteinases and their endogenous inhibitors: target proteins for prognosis, diagnosis and therapy in cancer (review). Oncol Rep **5:**1349-61.

53. **Kos, J., B. Werle, T. Lah, and N. Brunner.** 2000. Cysteine proteinases and their inhibitors in extracellular fluids: markers for diagnosis and prognosis in cancer. Int J Biol Markers **15:**84-9.

54. **Kozak, M.** 2001. New ways of initiating translation in eukaryotes? Mol Cell Biol **21:**1899-907.

55. **Kozak, M.** 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene **299:**1-34.

56. **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.

57. **Lammich, S., M. Okochi, M. Takeda, C. Kaether, A. Capell, A. K. Zimmer, D. Edbauer, J. Walter, H. Steiner, and C. Haass.** 2002. Presenilin-dependent intramembrane proteolysis of CD44 leads to the liberation of its intracellular domain and the secretion of an Abeta-like peptide. J Biol Chem **277:**44754-9.

58. **Leder, A., P. K. Pattengale, A. Kuo, T. A. Stewart, and P. Leder.** 1986. Consequences of widespread deregulation of the c-myc gene in transgenic mice: multiple neoplasms and normal development. Cell **45:**485-95.

59. **Ledford, A. W., J. G. Brantley, G. Kemeny, T. L. Foreman, S. E. Quaggin, P. Igarashi, S. M. Oberhaus, M. Rodova, J. P. Calvet, and G. B. Vanden Heuvel.** 2002. Deregulated Expression of the Homeobox Gene Cux-1 in Transgenic Mice Results in Downregulation of p27(kip1) Expression during Nephrogenesis, Glomerular Abnormalities, and Multiorgan Hyperplasia. Dev Biol **245:**157-71.

60. **Lee, H., R. W. Akita, M. X. Sliwkowski, and N. J. Maihle.** 2001. A naturally occurring secreted human ErbB3 receptor isoform inhibits heregulin-stimulated activation of ErbB2, ErbB3, and ErbB4. Cancer Res **61:**4467-73.

61. **Lehmann, A., K. Janek, B. Braun, P. M. Kloetzel, and C. Enenkel.** 2002. 20 S proteasomes are imported as precursor complexes into the nucleus of yeast. J Mol Biol **317:**401-13.

62. **Lemaire, R., R. M. Flipo, D. Monte, T. Dupressoir, B. Duquesnoy, J. Y. Cesbron, A. Janin, A. Capron, and R. Lafyatis.** 1994. Synovial fibroblast-like cell transfection with the SV40 large T antigen induces a transformed phenotype and permits transient tumor formation in immunodeficient mice. J Rheumatol **21:**1409-19.

63. **Lemaire, R., G. Huet, F. Zerimech, G. Grard, C. Fontaine, B. Duquesnoy, and R. M. Flipo.** 1997. Selective induction of the secretion of cathepsins B and L by cytokines in synovial fibroblast-like cells. Br J Rheumatol **36:**735-43.

64. **Leonardi, A., B. Turk, and V. Turk.** 1996. Inhibition of bovine cathepsins L and S by stefins and cystatins. Biol Chem Hoppe Seyler **377:**319-21.

65. **Levicar, N., T. Strojnik, J. Kos, R. A. Dewey, G. J. Pilkington, and T. T. Lah.** 2002. Lysosomal enzymes, cathepsins in brain tumour invasion. J Neurooncol **58:**21-32.

66. **Lewis, S., G. Abrahamson, J. Boultwood, C. Fidler, A. Potter, and J. S. Wainscoat.** 1996. Molecular characterization of the 7q deletion in myeloid disorders. British Journal of Haematology **93:**75-80.

67. **Liang, H., J. Fairman, D. F. Claxton, P. C. Nowell, E. D. Green, and L. Nagarajan.** 1998. Molecular anatomy of chromosome 7q deletions in myeloid neoplasms: evidence for multiple critical loci. Proc Natl Acad Sci U S A **95:**3781-5.

68. **Lin, Y. Z., and G. M. Clinton.** 1991. A soluble protein related to the HER-2 proto-oncogene product is released from human breast carcinoma cells. Oncogene **6:**639-43.

69. **Lund, F. E., and R. B. Corley.** 1991. Regulated expression of mouse mammary tumor proviral genes in cells of the B lineage. J Exp Med **174:**1439-50.

70. **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.

71. **Macara, I. G.** 2001. Transport into and out of the nucleus. Microbiol Mol Biol Rev **65:**570-94, table of contents.

72. **Maillard, I., and W. S. Pear.** 2003. Notch and cancer: Best to avoid the ups and downs. Cancer Cell **3:**203-5.

73. **Mao, P. L., Y. Jiang, B. Y. Wee, and A. G. Porter.** 1998. Activation of caspase-1 in the nucleus requires nuclear translocation of pro-caspase-1 mediated by its prodomain. J Biol Chem **273:**23621-4.

74. **Marambaud, P., J. Shioi, G. Serban, A. Georgakopoulos, S. Sarner, V. Nagy, L. Baki, P. Wen, S. Efthimiopoulos, Z. Shao, T. Wisniewski, and N. K. Robakis.** 2002. A presenilin-1/gamma-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. Embo J **21:**1948-56.

75. **Mattaj, I. W., and L. Englmeier.** 1998. Nucleocytoplasmic transport: the soluble phase. Annu Rev Biochem **67:**265-306.

76. **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.

77. **Mellgren, R. L., and Q. Lu.** 1994. Selective nuclear transport of mu-calpain. Biochem Biophys Res Commun **204:**544-50.

78. **Menon, S. G., E. H. Sarsour, D. R. Spitz, R. Higashikubo, M. Sturm, H. Zhang, and P. C. Goswami.** 2003. Redox regulation of the G1 to S phase transition in the mouse embryo fibroblast cell cycle. Cancer Res **63:**2109-17.

79. **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. Molecular & Cellular Biology **21:**6332-6345.

80. **Moon, N. S., W. Rong Zeng, P. Premdas, M. Santaguida, G. Berube, and A. Nepveu.** 2002. Expression of N-terminally truncated isoforms of CDP/CUX is increased in human uterine leiomyomas. Int J Cancer **100:**429-32.

81. **Murakami, D., I. Okamoto, O. Nagano, Y. Kawano, T. Tomita, T. Iwatsubo, B. De Strooper, E. Yumoto, and H. Saya.** 2003. Presenilin-dependent gamma-secretase activity mediates the intramembranous cleavage of CD44. Oncogene **22:**1511-6.

82. **Nardai, G., B. Sass, J. Eber, G. Orosz, and P. Csermely.** 2000. Reactive cysteines of the 90-kDa heat shock protein, Hsp90. Arch Biochem Biophys **384:**59-67.

83. **Navab, R., J. S. Mort, and P. Brodt.** 1997. Inhibition of carcinoma cell invasion and liver metastases formation by the cysteine proteinase inhibitor E-64. Clin Exp Metastasis **15:**121-9.

84. **Ni, C. Y., H. Yuan, and G. Carpenter.** 2003. Role of the ErbB-4 carboxyl terminus in gamma-secretase cleavage. J Biol Chem **278:**4561-5.

85. **Nicolas, M., A. Wolfer, K. Raj, J. A. Kummer, P. Mill, M. van Noort, C. C. Hui, H. Clevers, G. P. Dotto, and F. Radtke.** 2003. Notch1 functions as a tumor suppressor in mouse skin. Nat Genet **33:**416-21.

86. **Nollet, F., G. Berx, and F. van Roy.** 1999. The role of the E-cadherin/catenin adhesion complex in the development and progression of cancer. Mol Cell Biol Res Commun **2:**77-85.

87. **Okamoto, I., Y. Kawano, D. Murakami, T. Sasayama, N. Araki, T. Miki, A. J. Wong, and H. Saya.** 2001. Proteolytic release of CD44 intracellular domain and its role in the CD44 signaling pathway. J Cell Biol **155:**755-62.

88. **Okamoto, I., Y. Kawano, H. Tsuiki, J. Sasaki, M. Nakao, M. Matsumoto, M. Suga, M. Ando, M. Nakajima, and H. Saya.** 1999. CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumor cell migration. Oncogene **18:**1435-46.

89. **Park, I. C., S. Y. Lee, D. G. Jeon, J. S. Lee, C. S. Hwang, B. G. Hwang, S. H. Lee, W. S. Hong, and S. I. Hong.** 1996. Enhanced expression of cathepsin L in metastatic bone tumors. J Korean Med Sci **11:**144-8.

90. **Picard, D.** 2002. Heat-shock protein 90, a chaperone for folding and regulation. Cell Mol Life Sci **59:**1640-8.

91. **Pierce, A. M., R. Schneider-Broussard, I. B. Gimenez-Conti, J. L. Russell, C. J. Conti, and D. G. Johnson.** 1999. E2F1 has both oncogenic and tumor-suppressive properties in a transgenic model. Mol Cell Biol **19:**6408-14.

92. **Porter, D. C., N. Zhang, C. Danes, M. J. McGahren, R. M. Harwell, S. Faruki, and K. Keyomarsi.** 2001. Tumor-specific proteolytic processing of cyclin E generates hyperactive lower-molecular-weight forms. Mol Cell Biol **21:**6254-69.

93. **Puig, O., F. Caspary, G. Rigaut, B. Rutz, E. Bouveret, E. Bragado-Nilsson, M. Wilm, and B. Seraphin.** 2001. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. Methods **24:**218-29.

94. **Pyronnet, S., J. Dostie, and N. Sonenberg.** 2001. Suppression of cap-dependent translation in mitosis. Genes Dev **15:**2083-93.

95. **Pyronnet, S., L. Pradayrol, and N. Sonenberg.** 2000. A cell cycle-dependent internal ribosome entry site. Mol Cell **5:**607-16.

96. **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:**22624-22634.

97. **Rashid, M. G., M. G. Sanda, C. J. Vallorosi, J. Rios-Doria, M. A. Rubin, and M. L. Day.** 2001. Posttranslational truncation and inactivation of human E-cadherin distinguishes prostate cancer from matched normal prostate. Cancer Res **61:**489-92.

98. **Riccio, M., R. Di Giaimo, S. Pianetti, P. P. Palmieri, M. Melli, and S. Santi.** 2001. Nuclear localization of cystatin B, the cathepsin inhibitor implicated in myoclonus epilepsy (EPM1). Exp Cell Res **262:**84-94.

99. **Rios-Doria, J., K. C. Day, R. Kuefer, M. G. Rashid, A. M. Chinnaiyan, M. A. Rubin, and M. L. Day.** 2003. The role of calpain in the proteolytic cleavage of E-cadherin in prostate and mammary epithelial cells. J Biol Chem **278:**1372-9.

100. **Rozhin, J., R. L. Wade, K. V. Honn, and B. F. Sloane.** 1989. Membrane-associated cathepsin L: a role in metastasis of melanomas. Biochem Biophys Res Commun **164:**556-61.

101. **Schilling, K., S. Pietschmann, M. Fehn, I. Wenz, and B. Wiederanders.** 2001. Folding incompetence of cathepsin L-like cysteine proteases may be compensated by the highly conserved, domain-building N-terminal extension of the proregion. Biol Chem **382:**859-65.

102. **Schneider, R., V. I. Agol, R. Andino, F. Bayard, D. R. Cavener, S. A. Chappell, J. J. Chen, J. L. Darlix, A. Dasgupta, O. Donze, R. Duncan, O. Elroy-Stein, P. J. Farabaugh, W. Filipowicz, M. Gale, Jr., L. Gehrke, E. Goldman, Y. Groner, J. B. Harford, M. Hatzglou, B. He, C. U. Hellen, M. W. Hentze, J. Hershey, P. Hershey, T. Hohn, M. Holcik, C. P. Hunter, K. Igarashi, R. Jackson, R. Jagus, L. S. Jefferson, B. Joshi, R. Kaempfer, M. Katze, R. J. Kaufman, M. Kiledjian, S. R. Kimball, A. Kimchi, K. Kirkegaard, A. E. Koromilas, R. M. Krug, V. Kruys, B. J. Lamphear, S. Lemon, R. E. Lloyd, L. E. Maquat, E. Martinez-Salas, M. B. Mathews, V. P. Mauro, S. Miyamoto, I. Mohr, D. R. Morris, E. G. Moss, N. Nakashima, A. Palmenberg, N. T. Parkin, T. Pe'ery, J. Pelletier, S. Peltz, T. V. Pestova, E. V. Pilipenko, A. C. Prats, V. Racaniello, G. S. Read, R. E. Rhoads, J. D. Richter, R. Rivera-Pomar, T. Rouault, A. Sachs, P. Sarnow, G. C. Scheper, L. Schiff, D. R. Schoenberg, B. L. Semler, A. Siddiqui, T. Skern, N. Sonenberg, S. M. Tahara, A. A. Thomas, J. J. Toulme, J. Wilusz, E. Wimmer, G. Witherell, and M. Wormington.** 2001. New ways of initiating translation in eukaryotes. Mol Cell Biol **21:**8238-46.

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

104. **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.

105. **Singh, P., S. H. Wong, and W. Hong.** 1994. Overexpression of E2F-1 in rat embryo fibroblasts leads to neoplastic transformation. Embo J **13:**3329-38.

106. **Spotts, G. D., S. V. Patel, Q. Xiao, and S. R. Hann.** 1997. Identification of downstream-initiated c-Myc proteins which are dominant-negative inhibitors of transactivation by full-length c-Myc proteins. Mol Cell Biol **17:**1459-68.

107. **Staudt, L. M., A. L. Dent, A. L. Shaffer, and X. Yu.** 1999. Regulation of lymphocyte cell fate decisions and lymphomagenesis by BCL-6. Int Rev Immunol **18:**381-403.

108. **Strojnik, T., B. Zidanik, J. Kos, and T. T. Lah.** 2001. Cathepsins B and L are markers for clinically invasive types of meningiomas. Neurosurgery **48:**598-605.

109. **Suda, Y., S. Aizawa, Y. Furuta, T. Yagi, Y. Ikawa, K. Saitoh, Y. Yamada, K. Toyoshima, and T. Yamamoto.** 1990. Induction of a variety of tumors by c-erbB2 and clonal nature of lymphomas even with the mutated gene (Val659----Glu659). Embo J **9:**181-90.

110. **Tachibana, T., M. Hieda, and Y. Yoneda.** 1999. Up-regulation of nuclear protein import by nuclear localization signal sequences in living cells. FEBS Lett **442:**235-40.

111. **Tao, K., N. A. Stearns, J. Dong, Q. L. Wu, and G. G. Sahagian.** 1994. The proregion of cathepsin L is required for proper folding, stability, and ER exit. Archives of Biochemistry & Biophysics **311:**19-27.

112. **Tavares, A. T., T. Tsukui, and J. C. Izpisua Belmonte.** 2000. Evidence that members of the Cut/Cux/CDP family may be involved in AER positioning and polarizing activity during chick limb development. Development **127:**5133-5144.

113. **Tosi, S., S. W. Scherer, G. Giudici, B. Czepulkowski, A. Biondi, and L. Kearney.** 1999. Delineation of multiple deleted regions in 7q in myeloid disorders. Genes Chromosomes Cancer **25:**384-92.

114. **Troen, B. R., D. Ascherman, D. Atlas, and M. M. Gottesman.** 1988. Cloning and expression of the gene for the major excreted protein of transformed mouse fibroblasts. A secreted lysosomal protease regulated by transformation. J Biol Chem **263:**254-61.

115. **Troen, B. R., S. S. Chauhan, D. Ray, and M. M. Gottesman.** 1991. Downstream sequences mediate induction of the mouse cathepsin L promoter by phorbol esters. Cell Growth Differ **2:**23-31.

116. **Troen, B. R., S. Gal, and M. M. Gottesman.** 1987. Sequence and expression of the cDNA for MEP (major excreted protein), a transformation-regulated secreted cathepsin. Biochem J **246:**731-5.

117. **Turano, C., S. Coppari, F. Altieri, and A. Ferraro.** 2002. Proteins of the PDI family: unpredicted non-ER locations and functions. J Cell Physiol **193:**154-63.

118. **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.

119. **van Wijnen, A. J., F. M. van den Ent, J. B. Lian, J. L. Stein, and G. S. Stein.** 1992. Overlapping and CpG methylation-sensitive protein-DNA interactions at the histone H4 transcriptional cell cycle domain: distinctions between two human H4 gene promoters. Mol Cell Biol **12:**3273-87.

120. **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. Proceedings of the National Academy of Sciences of the United States of America **93:**11516-21.

121. **Visse, R., and H. Nagase.** 2003. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. Circ Res **92:**827-39.

122. **Wang, X. D., J. L. Rosales, A. Magliocco, R. Gnanakumar, and K. Y. Lee.** 2003. Cyclin E in breast tumors is cleaved into its low molecular weight forms by calpain. Oncogene **22:**769-74.

123. **Weber, E., D. Gunther, F. Laube, B. Wiederanders, and H. Kirschke.** 1994. Hybridoma cells producing antibodies to cathepsin L have greatly reduced potential for tumour growth. J Cancer Res Clin Oncol **120:**564-7.

124. **Weijzen, S., P. Rizzo, M. Braid, R. Vaishnav, S. M. Jonkheer, A. Zlobin, B. A. Osborne, S. Gottipati, J. C. Aster, W. C. Hahn, M. Rudolf, K. Siziopikou, W. M. Kast, and L. Miele.** 2002. Activation of Notch-1 signaling maintains the neoplastic phenotype in human Ras-transformed cells. Nat Med **8:**979-86.

125. **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.

126. **Wille, A., A. Heimburg, A. Gerber, A. Reisenauer, T. Welte, and F. Buhling.** 2002. Functional consequences of cathepsin L deficiency in human lung epithelial cells. Biol Chem **383:**1291-6.

127. **Wolff, B., J. J. Sanglier, and Y. Wang.** 1997. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human

immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. Chem Biol **4:**139-47.

128. **Wu, G., S. Nomoto, M. O. Hoque, T. Dracheva, M. Osada, C. C. Lee, S. M. Dong, Z. Guo, N. Benoit, Y. Cohen, P. Rechthand, J. Califano, C. S. Moon, E. Ratovitski, J. Jen, D. Sidransky, and B. Trink.** 2003. DeltaNp63alpha and TAp63alpha regulate transcription of genes with distinct biological functions in cancer and development. Cancer Res **63:**2351-7.

129. **Xing, R., F. Wu, and R. W. Mason.** 1998. Control of breast tumor cell growth using a targeted cysteine protease inhibitor. Cancer Res **58:**904-9.

130. **Xu, G., D. M. Livingston, and W. Krek.** 1995. Multiple members of the E2F transcription factor family are the products of oncogenes. Proc Natl Acad Sci U S A **92:**1357-61.

131. **Yaoita, Y.** 2002. Inhibition of nuclear transport of caspase-7 by its prodomain. Biochem Biophys Res Commun **291:**79-84.

132. **Yoneda, Y.** 2000. Nucleocytoplasmic protein traffic and its significance to cell function. Genes Cells **5:**777-87.

133. **Zagouras, P., S. Stifani, C. M. Blaumueller, M. L. Carcangiu, and S. Artavanis-Tsakonas.** 1995. Alterations in Notch signaling in neoplastic lesions of the human cervix. Proc Natl Acad Sci U S A **92:**6414-8.

134. **Zang, Z., J. Wang, Y. Zuo, Y. Hu, Y. Shi, and H. Xu.** 2001. Squamous cell carcinoma in an ileal neobladder: a case report and review of the literature. Chin Med J (Engl) **114:**1105-6.

135. **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:**2355-2365.

136. **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:**2015-2021.

Figure 1: Shorter cathepsin L isoforms localize to the nucleus in S phase and cleave CDP/Cux.

A fraction of cathepsin L is translated from downstream AUGs. These shorter isoforms devoid of signal peptide localize to the nucleus in a cell cycle dependent manner. It is currently unclear which isoform of cathepsin L translocates to the nucleus, and the mechanism by which it translocates. As shown in this schematic representation, short cathepsin L (including part of the prodomain) could be matured in the cytoplasm, therefore cathepsin L would translocates as a mature enzyme. Alternatively, short cathepsin L could translocate as is in the nucleus where it would then become mature and process nuclear substrates such as CDP/Cux.

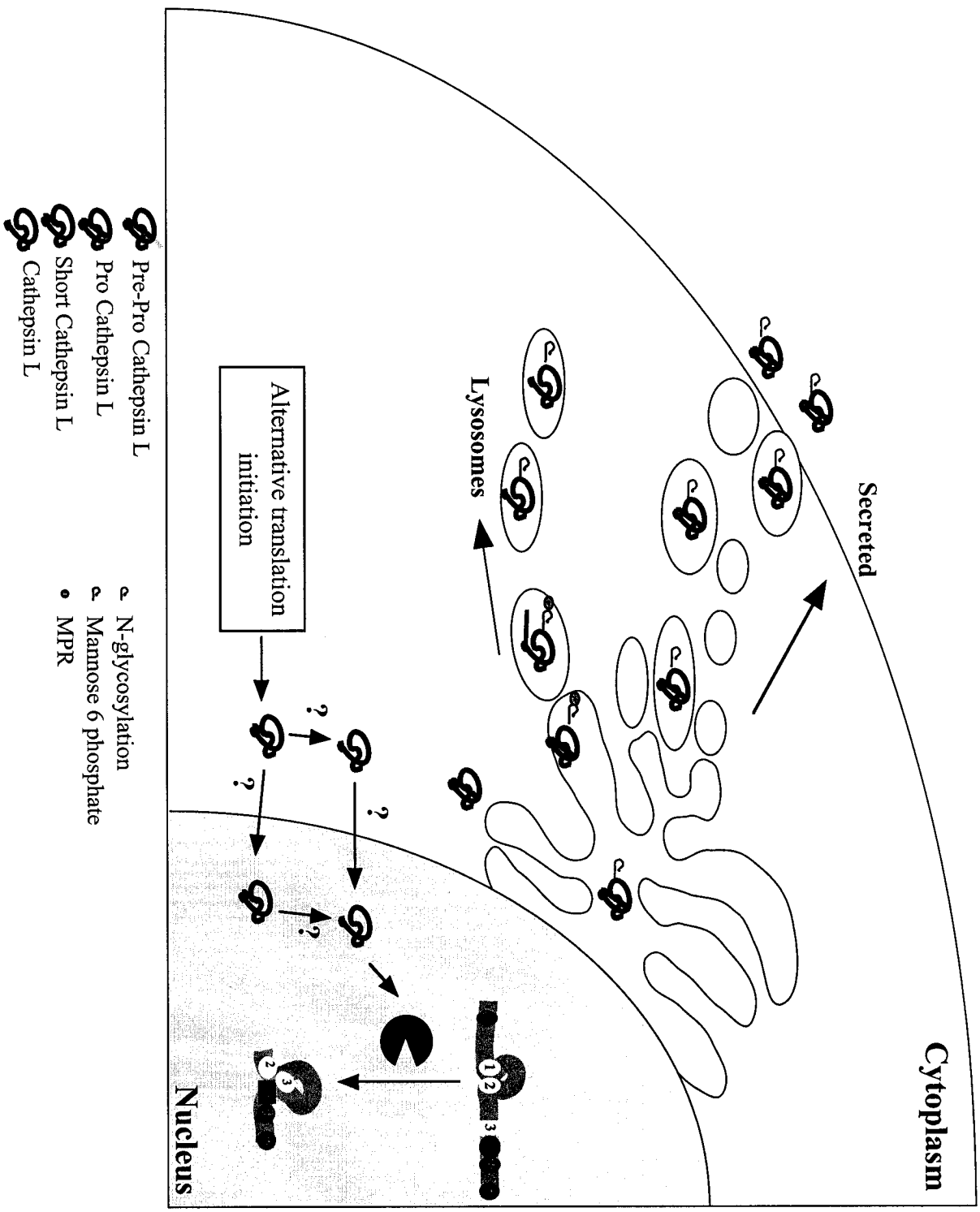
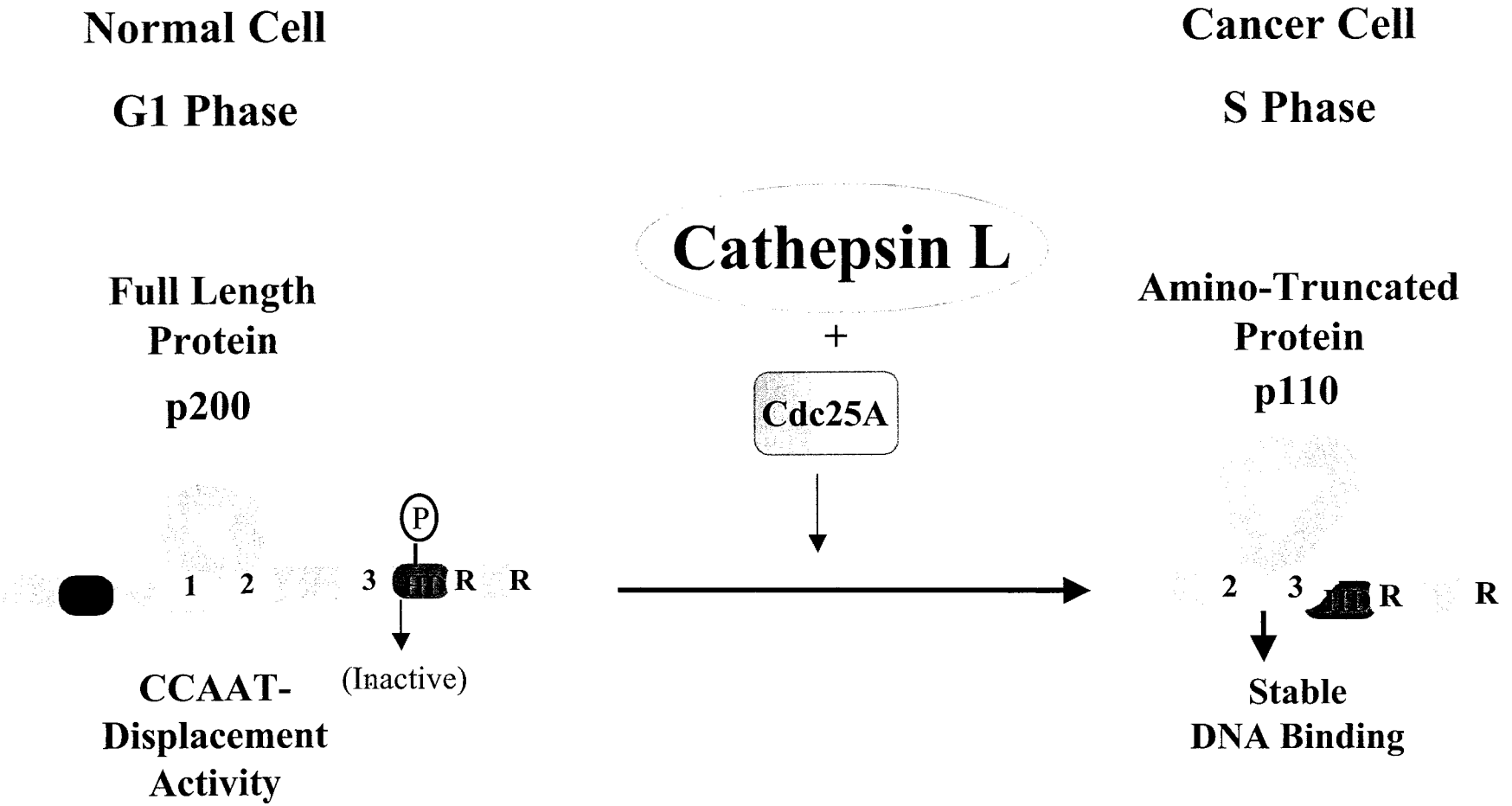

Cytoplasm
Secreted
Lysosomes
Alternative translation initiation
Nucleus
Pre-Pro Cathepsin L
Pro Cathepsin L
Short Cathepsin L
Cathepsin L
N-glycosylation
Mannose 6 phosphate
MPR

Figure 2: Increased CDP/Cux processing in cancer cells

In cancer cells, CDP/Cux is constitutively processed into the active p110 isoform, which correlates with the increased nuclear cathepsin L expression and activity observed in these cells.


Normal Cell
G1 Phase
Cancer Cell
S Phase
Cathepsin L
+
Cdc25A
Full Length Protein
p200
Amino-Truncated Protein
p110
P
1
2
3
R
R
2
3
R
R
CCAAT-Displacement Activity
(Inactive)
Stable DNA Binding

Chapter VI Contribution to Original Research

1 I identified the "lysosomal" cysteine protease cathepsin L as the protease responsible for CDP/Cux processing into the p110 isoform at the G1/S transition in the nucleus. One mechanism that accounts for the nuclear localization of cathepsin L is translation initiation at downstream AUG sites within the cathepsin L mRNA, thereby generating truncated enzymes devoid of a signal peptide. These results demonstrate for the first time the activity of a "lysosomal" enzyme in the nucleus and establish a novel role for cathepsin L in cell cycle regulation.

2 I demonstrated that CDP/Cux processing is increased and constitutive throughout the cell cycle in transformed cells, and that this is caused by the overexpression of cathepsin L. These findings suggest a novel role for cathepsin L in tumorigenesis. More generally, when considered together with the findings reported in a few other articles, my results raise the notion that the increased activity of certain site-specific proteases, within the cell, plays a role in cancer by activating certain proto-oncogenes or inactivating tumor suppressors.

3 I identified and characterized of a novel mRNA initiated within the intron 20 of CUTL1, which codes for a protein of 75 kDa. The intron20 mRNA and p75 are usually expressed in a tissue specific manner, but were found to also be expressed in breast tumor cell lines and in a subset of breast tumors. This finding suggests a novel mechanism to generate short CDP/Cux isoforms in some cancer cells. It also raises the possibility that

shorter CDP/Cux isoforms might be used as prognostic or predictive markers in breast cancer.

Abbreviations

a.a	Amino-acid
ABP	Activity-based probe
ALLN	Calpain inhibitor 1 N-acetyl-Leu-Leu-Met-CHO
APC	Anaphase promoting complex
ATP	Adenosine triphosphate
Bcr-Abl	Break-cell region-Abelson
bp	base pair
BRCA	Breast cancer
C/EBP	CCAAT enhancer binding protein
CASP	Cut alternatively spliced protein
Cat L	Cathepsin L
CC	Coiled-coil
CDC	Cell cycle division
Cdk	Cyclin dependent kinase
CDP/Cux	CCAAT displacement protein/Cut homeobox
ChIP	Chromatin immunoprecipitation
CKI	Casein kinase I
CKI	Cyclin dependent kinase inhibitor
CML	Chronic myelogenous leukemia
CMV	Cytomegalovirus
CPEB	Cytoplasmic polyadenylation element binding protein

CR	Cut repeat
CUTL1	Cut-like 1
DCIS	Ductal carcinoma in situ
DMEM	Delbecco's modified Eagle medium
DNA	Deoxyribonucleic acid
DOC	Deoxycholate
DPE	Downstream promoter element
DTT	Dithiothreitol
E-64d	(2S,3S)-trans-epoxysuccinyl-L-leucylamido-3-methylbutane ethyl ester
EDTA	Ethylenediaminetetra-acetic acid
EGF	Epithelial growth factor
EGTA	Ethylenebis(oxyethylenenitrilo)tetra-acetic acid
EMSA	Electrophoretic mobility shift assay
EMT	Epithelial-mesenchymal transition
ER	Endoplasmic reticulum
ER	Estrogen receptor
FACS	Fluorescence-activated cell sorting
FBS	Fetal bovine serum
Fs	Furless
G1	Gap 1
G2	Gap 2
GAPDH	Glyceraldehyde phosphate dehydrogenase
GSK3	Glycogen synthase kinase 3
GST	Glutatione-S tranferase

HA	Hematoglutinin
HCF-1	Host cell factor 1
HD	Cut homeodomain
HDAC	Histone deacetylase
Hepes	N-2-hydroxyethylpiperazine-N'2-ethanesulfonic acid
HER-2	Human epithelial growth factor receptor 2
HiNF-D	Histone nuclear factor-D
His	Histidine
HMEC	Human mammary epithelial cells
Hsp	Heat-shock protein
I20-mRNA	Intron 20 initiated mRNA
ICE	Interleukin 1β converting enzyme
I-Clip	Intramembrane cleaving protease
InR	Initiator region
IP	Immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactopyranoside
IRES	Internal ribosome entry site
IRF-1/2	Interferon factor 1 or 2
JPM-OET	N-[[L-trans-3-(Ethoxycarbonyl)oxiran-2-yl]carbonyl]-L-leucyl-3-(p-hydroxyphenyl)ethylamide
KCl	Potassium chloride
kDa	Kilo-Dalton
LAP	Liver-enriched transcriptional activated protein

LCIS	Lobular carcinoma in situ
LIP	Liver-enriched transcriptional inhibitory protein
LOH	Loss of heterozygosity
LT	Large T
M	Mitosis
M6P	Mannose-6-phosphate
MAPK	Mitogen-activated protein kinase
MCH	Myc-Cux-HA
MEF	Mouse embryo fibroblasts
MENT	Myeloid and erythroid nuclear termination stage-specific protein
MEP	Major excreted protein
MES	2-(N-morpholino)ethanesulfonic acid
Met (or M)	Methionine
MG132	Carbobenzoxy-L-leucyl-L-leucyl-L-leucinal
$MgCl_2$	Magnesium chloride
MHC II	Major histocompatibility complex class II
MMTV	Mouse mammary tumor virus
MRP	Mannose-6-phosphate receptor
NaCl	Sodium Chloride
NaF	Sodium fluoride
NaVO3	Sodium vanadate
NES	Nuclear export signal
NF-Y	Nuclear factor Y
NF-κB	Nuclear factor κB

NH_4CL	Ammonium Chloride
NLS	Nuclear localization signal
NP-40	Nonidet P40
NT-AT	Nuclear factor
P/CAF	p300/CBP associated factor
P300/CBP	CREB binding protein (cAMP-dependent response element binding protein)
PABP	Poly(A) binding protein
PAGE	Poly-acrylamide gel electrophoresis
PBS	Phosphate buffered saline
PDI	Protein disulfide isomerase
PI3K	phosphatidylinositol-3 phosphate
PKA	Proteine kinase A
PLC	Phospholipase C
PR	Progesterone receptor
PS1/2	Presenilin
Ci	Cubitus interruptus
R	Repression domain
RalGEF	Ras related guanine nucleotide exchange factor
Ras	Rat sarcoma
Rb	Retinoblastoma protein
RIP	Regulated intramembrane proteolysis
RNA	Ribonucleic acid
RTK	Receptors tyrosine kinase
RT-PCR	Reverse-transcriptase polymerase chain reaction

S	Synthesis
S1P	Site 1 protease
S2P	Site 2 protease
SAF-1	Serum ameloid A factor-1
SATB	Special AT-rich binding protein
SCF	Skp-cullin-F-Box complex
SDS	Sodium dodecyl sulfate
Ser	Serine
Skp2	S-phase kinase-associated protein 2
Src	Rous sarcoma-virus transforming protein
STAT	Signal transducer and activator of transcription
TBE	Tris-borate-EDTA
TBP	TATA box binding protein
TBS	Tris buffered saline
TGN	Trans-golgi network
Thr	Threonine
TNT	Transcription and translation
uORF	Upstream open reading frame
UTR	Untranslated region