4-9-2008

Mammalian DCN-1-like Proteins Enhance Neddylation of CUL3, and Vary in Localization and Tissue-Specific Expression

Russell Ryan

Mammalian DCN-1-like Proteins Enhance Neddylation of CUL3, and Vary in Localization and Tissue-Specific Expression

A Thesis Submitted to the
Yale University School of Medicine
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Medicine

by

Russell J.H. Ryan

2007

MAMMALIAN DCN-1-LIKE PROTEINS ENHANCE NEDDYLATION OF CUL3, AND VARY IN LOCALIZATION AND TISSUE-SPECIFIC EXPRESSION.

Russell Ryan, Alex Kim, Yoshi Yonekawa, Lydia Choi, Andrew Kaufman, Y. Ramanathan, Lyann Mitchell, KeumSil Hwang, Patricia Morris, and Bhuvanesh Singh. Section of Head and Neck Surgery, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY. (Sponsored by Jeffery Sklar, Department of Pathology, Yale University School of Medicine)

The Squamous Cell Carcinoma-Related Oncogene (*SCCRO/DCUN1D1*), located at 3q26.3, is a candidate oncogene that is gene-amplified and overexpressed in squamous-cell carcinomas of the upper aerodigestive tract and lung. The SCCRO protein is a member of a phylogenically conserved protein family, the defective in cullin neddylation – 1 – like domain – containing ("Dcun1d") proteins, which have been shown to promote the covalent modification of cullin proteins by the ubiquitin-like protein Nedd8.

In this study, we sought to characterize the expression and molecular interactions of vertebrate Dcun1d proteins, in order to better understand the normal function of genes in this family. We demonstrate that the majority of SCCRO protein in cultured human epithelial cell lines is not found in stable complexes with known interactors, including cullins, UBC12, and CAND1, a regulator of cullin–RING-finger ubiquitin ligases (CRL's). We show that SCCRO expression is not cell-cycle regulated in HeLa-S cells. We describe five highly conserved vertebrate Dcun1d genes, and show that the products of all five genes interact with cullins and with CAND1. We show that four of the five proteins enhance the neddylation of CUL3 in a dose-dependant fashion *in-vitro*. We demonstrate that the UBA domain-containing SCCRO and Dcun1d2 proteins are pan-cellular in localization, while other family members contain alternate N-terminal domains which mediate their exclusive localization to the nucleus (Dcun1d4 and Dcun1d5) or to cellular membranes (Dcun1d3). We demonstrate that expression of alternate *Dcun1d1* transcipts is specifically regulated in mouse germ cells during spermatogenesis, and we present evidence for specific expression of *Dcun1d2* in brain and muscle tissue. We conclude that spermatogenesis and myogenesis could serve as valuable model systems for future studies of the interaction between expression levels of vertebrate Dcun1d proteins and the activity of CRL complexes. Such work could be valuable in understanding how SCCRO overexpression contributes to cancer.

Acknowledgements

Advisor and Principal Investigator	Bhuvanesh Singh, M.D., Ph.D. (Department of Surgery – Head and Neck, Memorial Sloan-Kettering Cancer Center)
Yale-Sponsored One-Year Medical Student Research Fellowship Advisor	John K. Joe, M.D. (Dept. of Surgery – Otolaryngology, Yale University School of Medicine)
Yale Thesis Advisor	Jeffery Sklar, M.D. PhD (Dept. of Pathology, Yale University School of Medicine)
Funding & Support	Department of Student Research, Yale University School of Medicine

Co-authors: Alex Kim, M.D.; Yoshi Yonekawa; Lydia Choi, M.D.; Y. Ramanathan, Ph.D.; Andrew Kaufman, M.D.; Patricia Morris, Ph.D.; KeumSil Hwang, Ph.D.; Bhuvanesh Singh, M.D., Ph.D.

The author wishes to thank Dr. Andrew Koff for generously providing technical assistance and the use of his equipment. Valuable assistance was also provided by members of the laboratories of Drs. Jaqueline Bromberg, Lorenz Studer, Samuel Singer, Jeffery Boyd, Phillip Paty, Andrew Koff, and Patricia Morris.

This thesis is dedicated to the memory of Dr. John K. Joe, who passed away prior to completion of the manuscript. His humanitarianism, his clinical skill, and his enthusiastic mentorship of aspiring physicians will be greatly missed.

Table of Contents

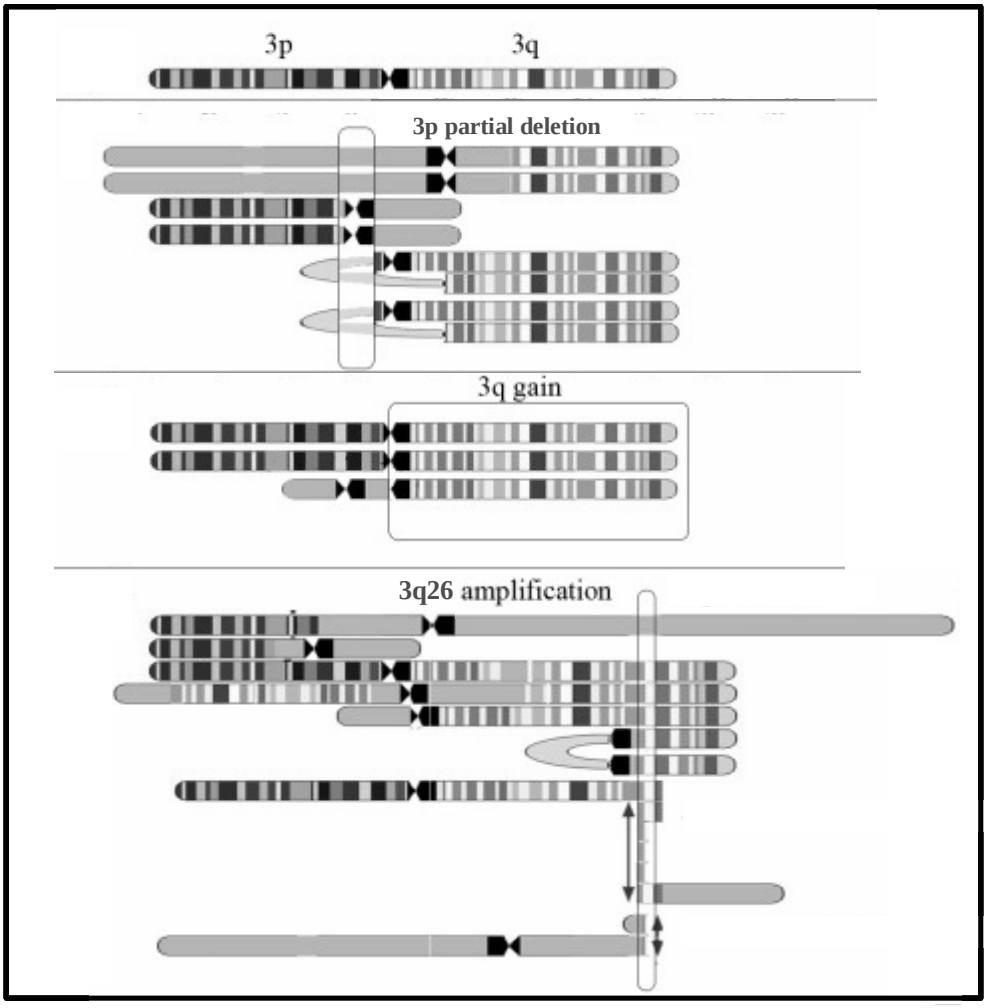
Introduction

A. The role of oncogenes in carcinogenesis

An oncogene is defined as "a gene whose protein product contributes to the development and / or progression of cancer."(1) The first oncogenes were identified as genes present in certain cancer-causing retroviruses. Many of these viral oncogenes, such as v-Src and v-Myc, were found to have originated through viral transduction of host genetic material, and the host genes from which they were derived were designated "proto-oncogenes." Viral oncogenes typically differ from their cellular counterparts in one of two general ways. First, in comparison to proto-oncogenes, viral oncogenes often contain alterations in their coding region, such as partial deletions or point mutations, resulting in protein products which possess abnormal activity in comparison with their cellular counterparts. In the case of other viral oncogenes, their juxtaposition with a powerful viral promoter leads to massive overexpression of the protein product in infected cells. Thus, although the protein itself is functionally similar to those normally produced by the cell, overexpression is sufficient to contribute to cellular transformation and tumorgenesis.

Very few human malignancies are thought to be caused by retroviral oncogene-mediated transformation; rather, most carcinogenesis occurs as the result of series of DNA-damage events which alter the tumor cell's genome, causing the activation of cellular proto-oncogenes and the deactivation of tumor-suppressor genes. However, as in retroviral oncogenes, oncogene activity in non-virally induced tumors typically results from either coding-sequence alteration (through mutation or translocation) or overexpression of cellular proto-oncogenes. Genetic alterations leading to proto-oncogene overexpression include promoter translocations, mutations affecting transcriptional control regions or mutations in sequences affecting transcript stability. Perhaps the most common means of proto-oncogene overexpression in epithelial malignancies, however, results from the chromosomal instability and complex aneuploidy seen in these tumors. As genomic rearrangements and abnormal mitoses produce some pre-malignant or malignant cells with more than two copies of certain chromosomal regions, cells that gain additional copies of proto-oncogenes may be further transformed, resulting in tumor progression.

One of the major strategies employed in recent years to better understand the genetic basis of cancer has been an effort to map recurrent regions of chromosomal copy number abnormalities (CNA's) in specific human malignancies. Loss of a chromosomal region is detectable through loss of heterozygosity (LOH) analysis of marker loci, or through copy number determination via comparative genomic hybridization (CGH), and may represent a "second hit" to one or more tumor suppressor genes at that region. Gain of a chromosomal region may result in overexpression of many encompassed genes, and the observation that a certain region is gained in many tumors of a certain type suggests that one or many genes within that region have pro-oncogenic activity. Because of the large number of genes encompassed by



Introduction Figure 1: A schematic diagram of chromosome 3, along with experimentally determined partial karyotypes of three cancer cell lines containing 3p partial deletion, 3q gain, and 3q26 amplification. Adapted with permission from Darai-Ramqvist, E. et al., 2006.(13)

chromosomal gains, it is difficult to use these findings alone to identify candidate oncogenes. Genomic amplification (Introduction Figure 1) is a form of CNA in which a small chromosomal segment is duplicated many times (often >10, occasionally >100 times), with copies present in tandem, on different chromosomes, or on extra-chromosomal "double minutes."(2) Powerful selective pressures act on malignant and pre-malignant populations of cells, so despite the instability of these amplified genomic regions, those which contain growth or survival enhancing genes are retained in the tumor. Since genomic amplifications often contain fewer than 20 genes, identification of recurrent amplification peaks can be a fruitful strategy for oncogene identification.

B. The candidate oncogene SCCRO

Several tumor types originating from mucosal squamous-cell epithelium demonstrate a similar recurrent gain of the long arm of chromosome 3, including head and neck squamous cell cancer (HNSCC),(3) squamous-type non-small-cell lung cancer (NSCLC),(4) and cancer of the uterine cervix.(5) Frequent gain of this region has also been reported in ovarian cancer. In the cervix, 3q copy number gains or high-level amplification are found in 90% of invasive cancers, but are absent from nearly all pre-malignant dysplastic lesions.(6) Low-resolution mapping studies utilizing comparative genomic hybridization (CGH) and spectral karyotyping (SKY) refined the region of frequent amplification in several squamous cell carcinoma types to 3q25-27. With the use of fine resolution mapping, several candidate oncogenes have been proposed as contributing to selection for this region in tumor progression. These include *PIK3CA*(7, 8), *PKC$_\iota$*(9, 10), *LAMP3*(11), *eIF-5A2*(12), and most recently, the squamous cell carcinoma-related oncogene (*SCCRO/DCUN1D1*).

Our group identified *SCCRO* as a candidate oncogene through FISH-based fine-resolution mapping of a region of maximal amplification at 3q26.3 in head and neck cancer cell lines.(8, 14, 15) Of the protein-coding genes predicted to exist within this amplification peak, only *SCCRO* was found to be highly expressed in tumor cell lines. An overlapping region of recurrent amplification, containing *SCCRO* and only 6 other genes, was noted in an independent SNP-chip based analysis of 101 lung cancer cell lines and

primary tumors.(16) Both studies noted that *PI3KCA*, a well-characterized 3q26 oncogene, was present in a separate and non-overlapping recurrent amplification peak.

SCCRO gene amplification correlates significantly with mRNA and SCCRO protein overexpression in tissue and cell lines derived from head and neck and lung cancers. No mutations were identified in *SCCRO* coding sequences from a large cohort of HNSCC and squamous NSCLC. In primary human tumors, highly significant overexpression of *SCCRO* mRNA relative to normal tissue controls was noted in 36% of head and neck, 44% of cervical, and 48% of non-small cell lung cancers tested. Of the latter group, 61% of squamous cell cancers but only 9% of adenocarcinomas overexpressed *SCCRO*. Furthermore, *SCCRO* overexpression was shown to correlate negatively with cause-free survival in non-small cell lung cancer.(17)

Functional evidence of *SCCRO*'s oncogenic activity came from experiments in which *SCCRO* was transgenically overexpressed in mouse fibroblasts and human immortalized keratinocytes. *SCCRO*-transfected keratinocytes gain anchorage-independent growth, while stably *SCCRO* overexpressing fibroblasts show dedifferentiated morphology, increased proliferation, serum-independent growth, enhanced invasiveness, and tumorgenicity in nude mice.(14)

The term "oncogene addiction" refers to the phenomenon in which tumor cells are dependent on a specific oncogenic signal in order to maintain cell survival and proliferation. Demonstration that a tumor is "addicted" to the activity of an overexpressed or mutant protein is a strong demonstration of that protein's importance in an oncogenic pathway, and can also provide proof-of-principle for the concept of targeting that pathway therapeutically. Importantly, knockdown of SCCRO protein levels via transfection of anti-*SCCRO* RNAi constructs or antisense oligonucleotides resulted in decreased proliferation and increased apoptosis of tumor cell lines, but only in lines which significantly overexpressed the SCCRO protein.(14) This suggests that *SCCRO* overexpressing cancers are "addicted" to a SCCRO-dependant survival pathway.

C. *SCCRO* orthologs and the *Dcun1d1*$^{-/-}$ mouse.

The murine ortholog of *SCCRO* was first described in 2000 by two independent groups. Pourcel et al. described expression of the gene, which they named *Tes3*, in a number of adult mouse tissues. They noted very high-level expression of a number of alternate *Tes3* transcripts in the testes only.(18) Mas et al. identified the same gene, which they named *Rp42*, as part of an effort to characterize genes involved in brain development.(19) They found that *Rp42* transcripts were expressed in a developmentally regulated manner in proliferating neuroblasts of the telencephalon and mesencephalon. Neither group reported on functional properties of the protein product, although Pourcel et al. reported weak protein sequence homology between TES3 and basic helix-loop-helix leucine zipper transcription factors.

Our group has generated a transgenic knockout mouse for *Tes3*/*Rp42*/*Dcun1d1* (hereafter referred to as *Dcun1d1*). The heterozygous mouse shows no significant abnormalities, while the *Dcun1d1*$^{-/-}$ mouse shows increased perinatal mortality, variable growth retardation, and complete male infertility. The testes of mature *Dcun1d1*$^{-/-}$ males become atrophic, with degeneration of seminiferous tubules and production of few and often morphologically abnormal sperm.(20) In combination with the published patterns of *Dcun1d1* expression, these findings suggest a critical role for *Dcun1d1/SCCRO* in spermatogenesis.

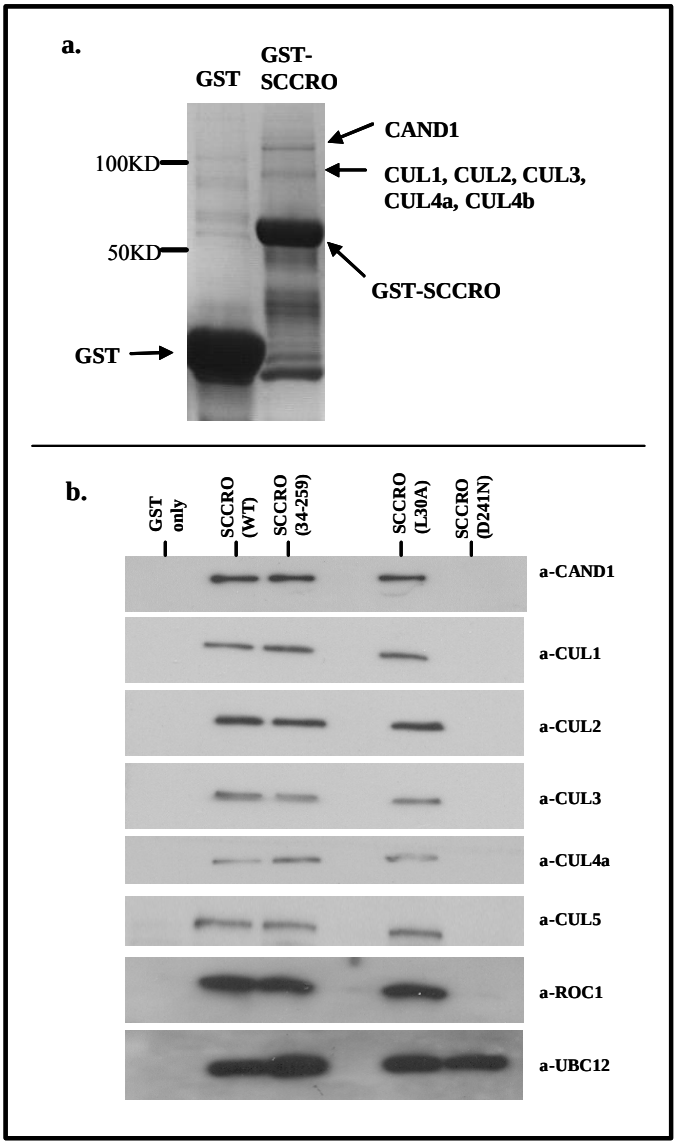
D. SCCRO interacts with CAND1 and cullins

In order to understand the cellular functions of the SCCRO protein, our group generated GST-SCCRO fusion constructs and used GST pulldown to isolate SCCRO binding partners in lysate of the lung cancer cell line SCC15 (Introduction figure 1a). These proteins were then identified using MALDI-TOF mass spectrometry.(21) SCCRO was shown to interact with several proteins in the cullin family, which are subunits of a class of ubiquitin ligase complexes. Also identified was an interaction between SCCRO and CAND1, a well-characterized regulator of cullin activity. These findings, along with subsequent experiments described below, strongly linked SCCRO's function to the cullin–RING-finger family of ubiquitin ligase complexes.

Introduction Figure 2: SCCRO interacts with proteins in the cullin–RING-finger ubiquitin ligase pathway.

(a) Bacterially expressed GST and GST-SCCRO were used for pulldown of proteins from SCC15 lysate. Following SDS-PAGE, the gel was stained for total protein with coomassie blue. Bands uniquely appearing in the GST-SCCRO lane were excised and identified via MALDI-TOF mass spectrometry, which demonstrated the presence of CAND1 and CUL1, CUL2, CUL3, CUL4a, and CUL4b.

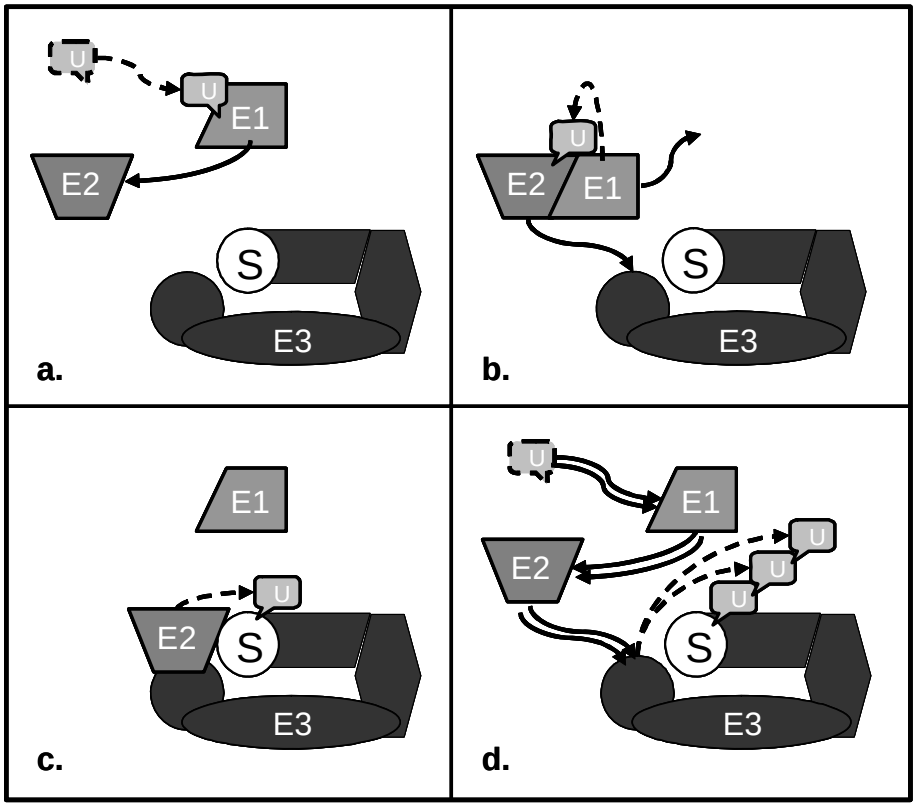
(b) Specific interaction of GST-tagged SCCRO and SCCRO mutants with CDL-associated proteins in HeLa lysate. Note that interactions with CAND1, cullins, ROC1, and UBC12 are retained in the N-terminal truncated SCCRO(34-259) and the UBA domain mutant SCCRO(L20A). However, the SCCRO(D241N) mutant loses binding to the cullins and to the cullin-binding CAND1 and ROC1, but not to the NEDD8 E2 enzyme UBC12.



E. An overview of cullin–RING-finger ubiquitin ligase complexes

The ubiquitin-proteasome system is a complex molecular system for regulating the degradation of a wide variety of cellular proteins. Factors whose degradation is controlled by this system include cell-cycle regulatory proteins, transcription factors, transmembrane growth-factor receptors, and components of signal transduction pathways and other cellular processes. Ubiquitin itself is a small, phylogenically conserved peptide, which can be covalently attached to a lysine residue on a target protein via an isopeptide bond with its C-terminal glycine. This reaction is mediated by a three-enzyme cascade involving proteins or

complexes with ubiquitin activating (E1), ubiquitin conjugating (E2), and ubiquitin ligase (E3) activities (Introduction Figure 3). Poly-ubiquitin chains can be created through repeated activity of this cascade on a target protein, leading to recruitment of that protein to the 26S proteosome for ATP-coupled hydrolysis.(22)



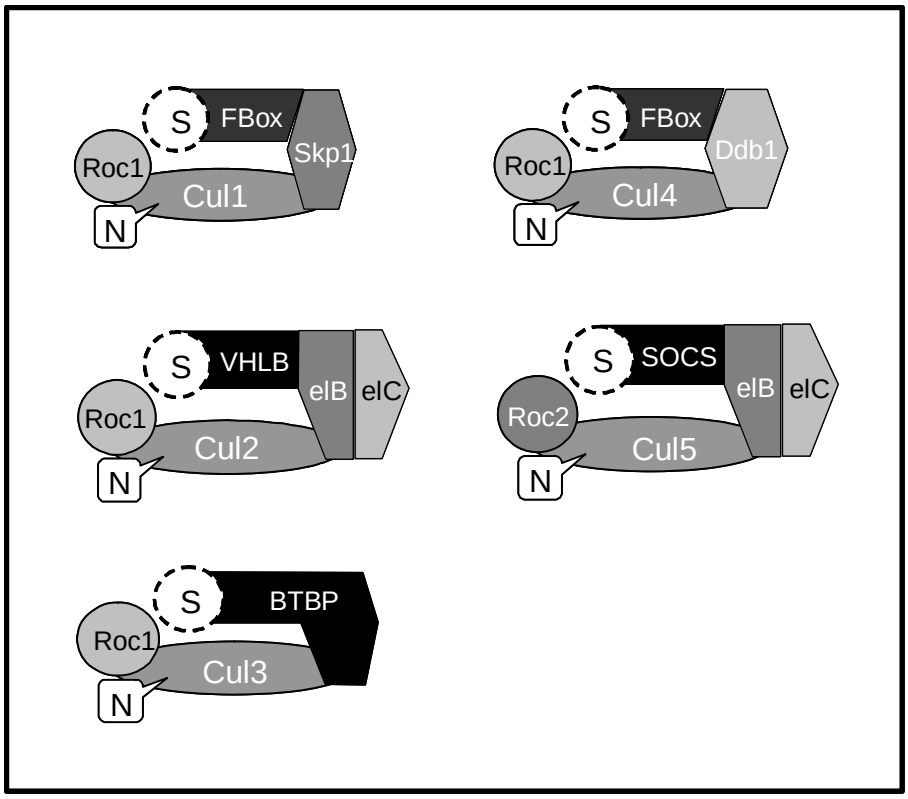
Introduction Figure 3: A schematic diagram of the ubiquitination cascade. **(a)** Free ubiquitin (U) is adenylated by an E1 enzyme, which allows for the formation of a thioester bond between ubiquitin and the catalytic cysteine of the E1. E1 then associates with an E2 enzyme. **(b)** Ubiquitin is covalently transferred from E1 to the reactive cysteine of E2, which then dissociates from E1 and binds to a substrate-bearing E3 enzyme. **(c)** The E3 mediates covalent transfer of ubiquitin to a target lysine on the substrate (S). **(d)** Repeated cycles of the E1-E2-E3 cascade are required for the formation of a poly-ubiquitin chain on the substrate, which is then recruited to the proteasome for degradation.

Specificity and regulation of the ubiquitination reaction are critical to prevent wasteful degradation of proteins, and targeting of a specific protein for proteosomal degradation often occurs in response to a precise cellular signal or event. This specificity is achieved through two general mechanisms. First, the protein target is must be available and in a form that facilitates its recruitment to the ubiquitination machinery. Many proteins are marked for ubiquitination by certain phosphorylation states, which in turn are dependant on the regulated activity of kinases and phosphorylases. Binding to other proteins or non-

peptide molecules, as well as localization to a specific subcellular compartment, may also affect the availability of a given protein for ubiquitination. The second source of specificity and regulation in ubiquitination lies in the E3 component of the cascade, which actually consists of hundreds of different protein complexes, each of which may be specific for a small subset of protein targets. While the identities of some of these complexes have been revealed in the past ten years, this is rapidly evolving field, and many details of this system remain unexplored.

Cullin–RING-finger ubiquitin ligase complexes (CRL's) represent one class of ubiquitin E3 complexes. Unlike other E3's, such as the HECT domain and U-box E3's, CRL's do not appear to contain specific catalytic residues which mediate the transfer of ubiquitin from the E2 to the target protein. Rather, structural models of CRLs suggest that they facilitate this reaction by bringing into proximity separate subunits that bind the ubiquitin-E2 conjugate and the target protein. In all CRL's studied to date, a cullin family protein forms an elongated backbone, with the RING-finger protein Roc1 (or a close homolog) stably bound near the c-terminal end of the cullin protein.(23, 24) It is the Roc1 RING-finger domain which appears to recruit the E2-ubiquitin conjugate. At the N-terminal end of the cullin protein is a domain which recruits the substrate-binding protein, either directly or through an adaptor protein. It is the substrate-binding protein which makes each CRL complex specific for a certain ubiquitination target.(25)

Six human cullin proteins; CUL1, CUL2, CUL3, CUL4a, CUL4b, and CUL5 are representatives of five highly conserved metazoan cullin classes. In fact, four of the five classes appear to be functionally conserved in organisms throughout the major eukaryotic taxa (Cul5 orthologs are restricted to metazoans), underscoring the ancient role of the CRL system in the function of eukaryotic cells. The six human proteins range in size from 87 to 104 kD. A class of larger, more divergent cullin homologs appears to be conserved in terrestrial vertebrates, and is represented by two genes in mammals, Cul7 and Parc (p53-associated parkin-like cytoplasmic protein). Another distant cullin homolog, Anapc2, is a component of a conserved multimeric complex with E3 ligase activity, the *ana*phase-*p*romoting *c*omplex / cyclosome. The following discussion will focus on complexes formed by the five core classes of cullin proteins (Introduction Figure 4), with an emphasis on their known roles in carcinogenesis.



Introduction Figure 4: Schematic model diagrams of CRL complexes formed by the five highly conserved classes of cullin proteins. Each complex contains a characteristic cullin protein, a Roc protein, 0-2 adaptor proteins, and one member of a family of characteristic substrate-binding proteins (S = substrate). The Nedd8 moiety (N) is covalently attached to the cullin protein in at least some active complexes from all five classes. Substrate-binding protein classes are colored black and labeled as follows: FBox = F-box protein, VHLB = VHL-box protein, SOCS = SOCS-box protein, BTBP = BTB protein. Adapted from solved crystal structures for the SCF,(23) VBC,(26) and DDB1-Cul4a-Roc1 complexes,(24) and other sources.

F. Cul1 CRL complexes

The first, and most extensively characterized class of CRL complexes are those formed by Cul1 orthologs. These complexes contain Roc1 bound at the Cul1 C-terminal,(27) while the N-terminal of Cul1 binds to the adaptor protein Skp1. Skp1, in turn, binds to a substrate-recruiting F-box protein.(28, 29) A bioinformatic analysis of predicted proteins in the human genome demonstrated the presence of 68 F-box proteins in three broad classes (Fbxl, Fbxw, and Fbxo)(30), although only a handful are proven members of

active Skp1-Cul1-F-box (SCF) complexes. Three of these, Skp2, Fbxw7, and β-Trcp, are particularly noteworthy as regulators of critical pathways implicated in oncogenesis.(31)

Skp2 is the prototypical "F-box with leucine-rich repeat" (Fbxl) protein and serves as the substrate-binding component of the SCF^{Skp2} complex. SCF^{Skp2} is responsible for the degradation of several critical cell-cycle control regulators, including the CDK inhibitors $p27^{Kip1}$(32), $p57^{Kip2}$(33), and $p21^{Cip1}$(34), as well as the Rb family protein p130(35). Since these proteins are all known to have tumor-suppressing activity, increased SCF^{Skp2} activity might be expected to be pro-oncogenic. In fact, SKP2 overexpression has been observed in a wide range of human malignancies, including HNSCC(36, 37) and NSCLC(38), often in concert with underexpression of a known SCF^{SKP2} target such as $p27^{KIP1}$. It has also been shown to contribute to oncogenesis in a variety of experimental systems.(37, 39-41)

In contrast to SCF^{Skp2}, SCF complexes containing the F-box protein Fbxw7 (orthologs cdc4p and archipelago) seem to target a number of growth-promoting proteins for ubiquitination, including phosphorylated cyclin E,(42) certain members of the notch family of signaling proteins,(43) and c-Myc.(44, 45) Fbxw7 inhibition leads to chromosomal instability,(46) possibly through mechanisms involving Cyclin E or Aurora A. Support for the status of *FBXW7* as a tumor suppressor gene comes from the observation of *FBXW7* mutation or deletion in several human tumors.(47, 48) Also, *Fbxw7* haploinsufficiency cooperates with p53 haploinsufficiency to produce epithelial tumors in an irradiated mouse model.(49)

β-Trcp (SLIMB in *Drosophila*) is a third well-characterized F-box protein, and the prototype of the F-box with WD-40 repeat (Fbxw) family. $SCF^{\beta\text{-}Trcp}$ has activities which could potentially be pro- or anti-oncogenic, depending on the cellular context. $SCF^{\beta\text{-}Trcp}$ activity is necessary for degradation of the growth-promoting Wnt pathway mediator β-catenin,(50, 51) which is activated in many human tumors. *BTRC* is not thought to be a frequently altered tumor suppressor in human carcinogenesis, however, since loss of expression of two paralogous β-Trcp genes (*BTRC* and *BTRC2*) would have to occur in order for this activity to be lost in an evolving tumor(52). In addition, $SCF^{\beta\text{-}Trcp}$ positively regulates the pro-survival activity of the NF-κB pathway by targeting IκB for ubiquitination.(53, 54).

G. Cul4 CRL complexes

All Cul4-containing complexes studied to date, (including both Cul4a and Cul4b complexes in humans) contain Roc1 and the Cul4 C-terminal binding protein Ddb1. Recently it has been shown that Cul4 orthologs share several targets with Cul1, and may even utilize some of the same F-box proteins for adaptor recruitment. For example, mammalian Cul4a-Roc1-DDB1 has been shown to bind Skp2 and mediate ubiquitination of $p27^{Kip1}$.(55) Another apparent target of a Cul4a CRL which is shared with a Cul1 CRL is the replication licensing factor Cdt1(56, 57). Since Cdt1 expression in S phase is an important regulator of chromosomal stability, overexpression of Cul4a could theoretically promote the chromosomal instability phenotype common in epithelial malignancies. Cul4a also mediates ubiquitination and degradation of two proteins involved in the nucleotide excision repair pathway, Ddb2 and Csa(58). These findings are important, because unlike *CUL1*, there is direct evidence that *CUL4A* may act as a true oncogene, with transcript overexpression seen in 47% of primary breast cancers(59) and amplification of the *CUL4A* chromosomal locus seen in breast cancers, hepatocellular carcinomas, esophageal squamous cell carcinomas, and other tumor types(60).

H. Cul2 and Cul5 CRL complexes

All known Cul2 CRL complexes contain two common adaptor proteins, elongin B and elongin C, which bind at the Cul2 N-terminal in a position analogous to that of Skp1 in the SCF complex.(26) The substrate-binding adaptors in Cul2 complexes, called elongin B and C binding proteins (BCBP's), generally contain a short conserved motif known as a BC box.

The most extensively studied of the Cul2 complex adaptor proteins is the von Hippel-Lindau tumor suppressor protein (vHL). Under normoxemic conditions, vHL recruits the transcription factor Hif-1α to the so-called VBC complex, where it is ubiquitinated and rapidly degraded.(61-63) Low oxygen tension causes a decrease in the affinity of Hif-1α for vHL, resulting in Hif-1α accumulation and formation of a

Hif-1α - Hif-1β heterodimer. The heterodimer activates transcription of hypoxia-inducible genes, which may enhance cellular survival and proliferation. Thus, genetic loss of vHL, or other disruption of the Cul2 - vHL - elongin B - elongin C (VBC) complex leads to constitutive activation of an oncogenic pathway. A "second hit" loss of the remaining *VHL* gene is responsible for the development of vascular malformations and tumors in the von Hippel-Lindau familial tumor syndrome in humans, and homozygous inactivation of *VHL* also contributes to the development of sporadic renal cell carcinoma.(64)

Interestingly, Cul5 also forms complexes containing elongin B and elongin C, along with a repertoire of BCBP's which seems to partially overlap those involved in Cul2 complexes. Many Cul5-binding BCBP's contain a second conserved sequence downstream of the BC box, with the combination of those two elements defining a so-called SOCS box.(65) Cul5-containing CRL's are also unique in containing Roc2, rather than the Roc1 found in other cullin complexes.(65, 66)

I. Cul3 CRL complexes

Cul3-containing CRL's differ from those of other classes in that they contain a substrate-binding protein from the BTB family, which binds directly to the Cul3 c-terminal, rather than to an adaptor protein. Thus, a single BTB protein serves the functions played by Skp1 and an F-box protein in the SCF complex.(67) Based on gene sequence analysis, as many as 183 BTB proteins may exist in the human genome,(68) representing one of the most diverse classes of human CRL adaptors, although it is likely that not all proteins containing a BTB domain are true components of Cul3-based E3's. Interestingly, several Cul3-interacting BTB proteins seem to exist primarily as homodimers with multiple distinct Cul3 interacting domains, supporting the notion that active Cul3 CRL's may contain more than one Cul3 molecule.(68)

Several noteworthy Cul3-BTB complexes have been described. In the early *C. elegans* embryo, a CUL-3 CRL containing the BTB protein MEL-26 appears to be responsible for the degradation of MEI-1, a microtubule-severing protein expressed during meiosis. Interference with MEL-26 or CUL-3 function leads

to abnormal mei-1 accumulation in the mitotic cycle which follows meiosis, resulting in abnormal mitotic spindle formation.(69) It has also been suggested that Cul3 plays a major role in mammalian reproduction, based on the greatly increased expression of *Cul3* in adult mouse testes, along with *Klhl-10*, the gene coding for a BTB protein binding partner.(70) Male *Klhl-10*$^{-/-}$ mice are infertile due to disrupted spermatogenesis.

Another important BTB protein, the *Drosophila* roadkill (RDX) protein, has recently been shown to control the Cul3-dependant degradation of the important hedgehog pathway mediator cubitus interruptus (CI) in several developmental contexts.(71, 72) The vertebrate Rdx ortholog, Speckle-type POZ Protein (Spop), appears functionally similar to Rdx, although it is not yet clear that it has a homologous role in controlling the vertebrate hedgehog pathway. Interestingly, in other developmental contexts, CI is degraded by a CUL1-dependant mechanism.(73)

The hedgehog pathway is not the only example of a developmental signaling pathway controlled by both Cul1 and Cul3 SCL's. An SCL containing Cul3 and the BTB protein Klhl-12 appears to be responsible for negatively regulating the β-catenin pathway by mediating ubiquitination of the WNT pathway intermediate Dishevelled.(74) Both the hedgehog and Wnt-β–catenin pathways have been referred to as "onco-developmental" pathways due to their involvement in a range of human malignancies, raising the possibility that Cul3 complexes could have undiscovered tumor-suppressing or tumor-promoting roles.

J. Parc and Cul7 complexes

Parc and Cul7 are homologous cullin-like proteins which are conserved in mammals, and are similar to proteins found in *Gallus gallus* and *Xenopus tropicalis*. Human CUL7 has been shown to form SCF-like complexes containing ROC1, SKP1, and the F-box protein FBXW8. Unlike other cullins, however, Cul7 does not appear to undergo modification with the ubiquitin-like protein Nedd8. It has also been shown to form heterodimeric complexes with Cul1. Parc, on the other hand, does undergo Nedd8 modification, but does not interact with Skp1 or known F-box proteins.(75) It forms heterodimeric

complexes with Cul7 which appear to possess ubiquitin ligase activity(76). Both Cul7 and Parc are notable in that they possess conserved CPH domains which bind the tetramerization domain of the important tumor suppressor protein p53,(77) however, their effect on p53 function is poorly understood.

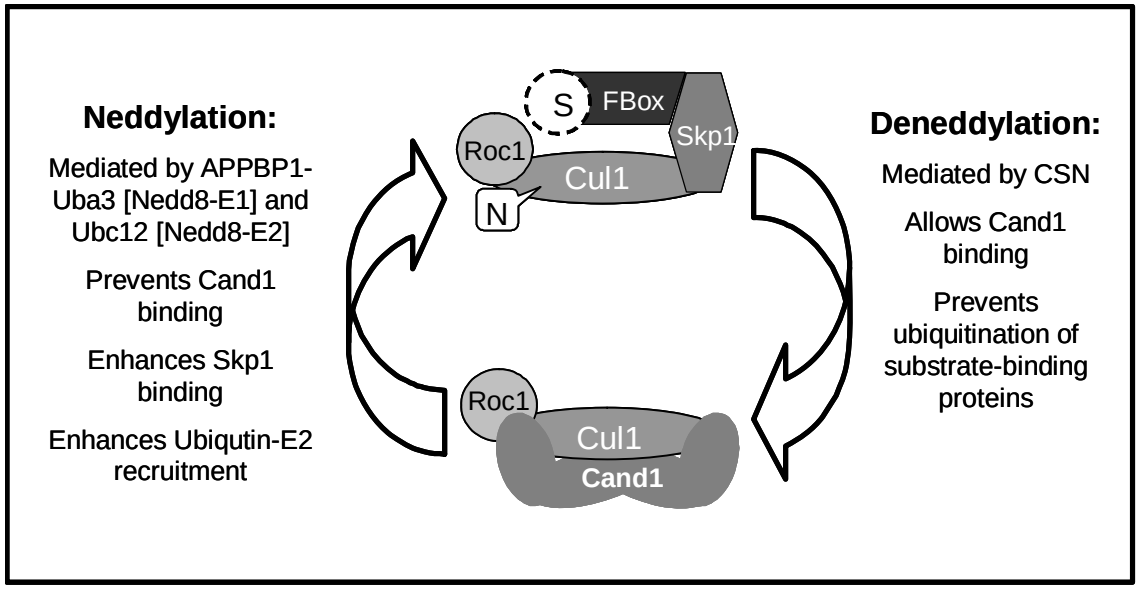
K. Cullin neddylation and regulation of CRL activity

Members of the five conserved cullin classes can be isolated from most cells in two forms: a monomeric form, and a higher molecular weight form which has been shown to represent a covelent conjugate of the cullin with a small ubiquitin-like protein called Nedd8.(78-80) Unlike ubiquitin, Nedd8 does not form long, polymeric chains on the target proteins, and cullin proteins appear to be the primary, if not exclusive target of "neddylation." Cullin neddylation occurs through a process similar to that of ubiquitination. The vertebrate enzymes specific for neddylation are the heterodimeric complex Appbp1-Uba3 (E1) and Ubc12 (E2).(81) The existence or identity of an "E3" for neddylation has been controversial, and will be further discussed below.

In contrast to ubiquitination, neddylation does not appear to mark proteins for degradation. Instead, neddylation has been shown to facilitate activity of CRL complexes in several contexts. The neddylation system was first studied in *S. cerevisiae*, where deletion of the Nedd8 ortholog or its E1 enzyme was found to have no effect on cell growth or viability. Effects were only seen in synthetic mutants: when elimination of the neddylation pathway was combined with a temperature-sensitive mutant of an SCF-related protein, such as the Cul1 ortholog cdc53p, the ubiquitin E2 cdc34p, the Fbxw7 homolog cdc4p, or skp1p, lowering of the temperature restrictive for growth was observed.(79) This suggested that cullin neddylation helps to stabilize the CRL complex, and that in the absence of this stabilizing component, structurally unstable mutant subunits were no longer able to maintain complex activity.

In contrast to the findings in *S. cerevisiae*, interference with the neddylation pathway alone results in non-viability in *S. pombe*(82), and defective embryogenesis in *C. elegans*(83, 84) and in vertebrates(85). Interference with the neddylation pathway prevents ubiquitination of specific mammalian SCF targets *in vitro*(86), and *in vivo*(87, 88), as well as certain VBC targets *in vitro* and *in vivo*.(88)

There have been several proposed explanations for why cullin neddylation is required to maintain CRL activity in many organisms (Introduction Figure 5). One explanation comes from the finding that neddylation of cullin proteins removes them from inhibition by Cand1/Tip120a, a 120 kD protein which binds only to un-neddylated cullins.(89, 90) Structural studies have shown that Cand1 binding sites on Cul1 overlap those utilized by Skp1; thus, Cul1 cannot form an active SCF complex when in complex with Cand1.(91) The absence of Cand1 homologs in *S. cerevisiae* could explain the apparently reduced importance of neddylation in that species. Other proposed mechanisms by which neddylation might enhance CRL activity include improved recruitment of the ubiquitin-E2 conjugate to the CRL complex(89), and facilitation of cullin heterodimer formation.(93)



Introduction Figure 5: Diagram showing the role of neddylation and deneddylation in cycling Cul1 between the active SCF complex and the inactive Cand1-Cullin complex. Both neddylation and deneddylation activities are required for continued activity of CRL complexes *in vivo*.

Several isopeptidases have been identified which are capable of removing Nedd8 from cullin proteins.(94-96) The most physiologically important of these appears to be the COP9 signalsome (CSN), a multi-protein complex with homology to the proteosome lid. One of the subunits of the CSN, Csn5/Jab1, contains a metalloprotease domain which catalyzes cleavage of the cullin-nedd8 isopeptide bond.(97) This function appears to be conserved in *S. cerevisiae*, despite the absence of orthologs to CSN subunits other

than Csn5.(96) Model organisms with defective CSN activity accumulate neddylated cullins,(96, 98). Addition of purified CSN to *in vitro* reactions of SCF^{Skp2}-dependant ubiquitination result in deneddylation of Cul1 and decreased ubiquitination of $p27^{Kip1}$.(99)

While the deneddylating activity and *in-vitro* effects of CSN suggest that it might act as an inhibitor of CRL-dependant ubiquitination, *in-vivo* studies show a more complex picture. Genetic studies in model organisms have demonstrated that CSN is actually necessary for the ubiquitination activity of certain CRL's, and deletion of CSN results in the accumulation of CRL substrates.(97, 98). These findings have since been extended to mammalian systems – for example, heterozygous deletion of Csn5 in mice leads to reduced p27 degradation.(100) The phenomenon of CSN inhibiting CRL activity *in vitro*, but facilitating ubiquitination *in vivo*, is sometimes referred to as the "CSN paradox." Interestingly, a similar phenomenon is observed with Cand1, since knockout of that known CRL inhibitor also seems to reduce, rather than increase CRL activity.(90)

The CSN paradox could be explained if continuous cycles of neddylation and deneddylation were required in order to maintain CRL activity.(101) Some studies have indicated that continuous activation of individual CRL complexes occurs in the absence of deneddylating activity, and that this results in auto-ubiquitination and degradation of CRL components such as F-box proteins, resulting in the observed decrease in global CRL activity.(100, 102) Recently, a new model was proposed, based on experimental evidence that at least some CRL complexes contain a heterodimer of one neddylated and one un-neddylated cullin subunit. In this model, neddylation pathway aberrations which result in either all neddylated or all un-neddylated cullins would prevent heterodimer formation, and thus, CRL activity.(93) These two models need not be mutually exclusive, and regardless of the mechanism, it is empirically clear that there is a critical but non-linear relationship between cullin neddylation and the activity of CRL-dependant ubiquitination pathways.

L. DCN-1 and SCCRO have Nedd8 E3-like activity

Our group's observation that SCCRO protein binds to all of the classic human cullins, as well as to CAND1, suggested that the SCCRO proto-oncogene might play a role in the regulation of cullin activity. Independently, Kurz et. al. identified the SCCRO ortholog *dcn-1* (*d*efective in *c*ullin *n*eddylation) while screening *C. elegans* embryos for mutants which recapitulated phenotypic defects seen in Cul-3 and neddylation pathway mutants.(103) They found that protein extracts from *C. elegans dcn-1*(-) mutants show a significant decrease in the fraction of neddylated cul-3 compared to wild type embryos.

Kurtz et. al. also studied the orthologous gene *dcn1* in *S. cerevisiae*, and found that deletion of *dcn1* led to a decrease in the fraction of neddylated cdc53p. Conversely, overexpression of dcn1p in yeast led to increased neddylation of cdc53p. Through analysis of double mutants of *dcn1* and other genes involved in neddylation or deneddylation, they demonstrated that dcn1p functions by enhancing neddylation, rather than by interfering with deneddylation. These changes in the neddylation state of cdc53p were functionally significant because either deletion or overexpression of *dcn1* lowered the restrictive temperature of temperature-sensitive *cdc53-1* mutants, suggesting that either an excess or a deficit of dcn1p-mediated neddylation activity could interfere with cullin function. Finally, they showed that addition of yeast extract from transgenic *dcn1* overexpressing strains enhanced cdc53p neddylation *in vitro* to a greater extent than did extract from *dcn1* defective strains. They speculated that dcn1p may act as part of the E3 ligase for neddylation, although they found that cdc53p neddylation was not enhanced by the addition of bacterially expressed dcn1p to an in vitro reaction containing minimal purified components, suggesting that additional factors were needed for this activity.

Investigations by other members of our group, performed contemporaneously with the current work, have produced significant additional evidence for the importance of SCCRO in the neddylation of human cullin proteins. The addition of purified, bacterially expressed SCCRO to *in vitro* reactions utilizing purified Nedd8, E1 and E2 enzymes, ATP, and mammalian cell lysate results in dose-dependant enhancement of the rate of CUL3 neddylation. This effect is abrogated in c-terminal truncated SCCRO

proteins or the mutant SCCRO(D241A), both of which lack cullin-binding capability. Conversely, *in vitro* neddylation reactions performed with lysate from *Dcun1d1*$^{-/-}$ mouse tissues produce significantly less CUL3 neddylation compared to reactions performed with wild-type lysate. These findings are analogous to those made for Dcn1p, and strongly suggest that SCCRO acts as a molecular activator of cullin neddylation. However, attempts to extend these findings *in vivo* by overexpressing *SCCRO* or performing RNAi-mediated *SCCRO* knockdown in cultured mammalian cells have not, to date, yielded consistent changes in the level of neddylated cullins or ubiquitination activity.

M. Understanding the function of SCCRO – implications for cancer

Due to the complexity of the CRL system, and our incomplete understanding of the function of the neddylation pathway, it is difficult to make *a priori* hypotheses about the mechanism by which SCCRO overexpression contributes to tumor development. As described in this introduction, there are a wide variety of potentially oncogenic and tumor-suppressing pathways which are known to be affected by specific CRL E3 complexes, and an even greater number likely remain uncharacterized. One strategy for approaching the mechanism of SCCRO-mediated oncogenicity would be to individually assay each of the many "candidate" CRL complexes and ubiquitination targets for altered activity in SCCRO-overexpressing tumors or SCCRO-transformed cell lines. Another approach, and the one which inspired the current work, would be to identify specific cellular processes or tissues in vertebrates which are dependant on SCCRO or SCCRO-like proteins for their normal function. By learning more about the endogenous role of SCCRO, it might be possible to more rationally focus future studies of SCCRO's role in carcinogenesis.

Statement of purpose

The central hypothesis of our research program has two parts. First, we believe that SCCRO acts as an oncogene in carcinoma of the human lung and upper aerodigestive tract, contributes to selection for chromosomal amplification at 3q26 during tumor evolution, and leads to a more aggressive tumor phenotype when overexpressed. Second we hypothesize that the oncogenic effects of SCCRO are mediated through an alteration in the activity of one or more cullin–RING-finger ubiquitin ligase complexes, thus altering the ubiquitination-dependant degradation of one or more critical protein targets within the tumor cell.

While we have published substantial evidence to support the first statement, demonstration of the second poses a greater technical and theoretical challenge, given the multitude of CRL's and potential CRL ubiquitination targets. The purpose of this work is to characterize features of the normal expression and molecular function of Dcun1d proteins in both cellular and tissue-based contexts, with the goal of prioritizing specific endogenous processes which might be regulated by high expression of SCCRO or related proteins. We expect that these endogenous functions of SCCRO would involve regulation of CRL complexes, and could serve as valuable model systems for understanding the role of SCCRO overexpression in epithelial malignancies.

Methods

Gel filtration analysis of HaCaT, HeLa, and SCC15 lysates

HaCaT and HeLa cell lines, obtained from ATCC, were grown in minimal essential media (MEM) supplemented with 10% fetal calf serum (FCS), using standard mammalian cell culture conditions and techniques. Intact cells were removed from the plate via incubation in 0.25% trypsin + 1mM EDTA in Hank's Buffered Saline Solution without Mg^{2+} or Ca^{2+}. Following trypsin neutralization with media, cells were washed 2x in cold PBS, pelleted, and lysed via sonication in 2-3 volumes of FPLC buffer, consisting of 20 mM HEPES*KOH, pH 7.5, 5 mM KCl, 0.75 mM $MgCl_2$, 100 mM NaCl, 0.5 mM DTT, 1 mM PMSF, and 10% glycerol by volume. Lysates were centrifuged at maximum speed in a tabletop microcentrifuge, followed by ultracentrifugation at 50,000 g for 30 minutes to remove particulate matter.

Gel filtration was performed via FPLC on a Superose 6 column. 200 mcl of molecular weight standard, purified protein solution, or cell lysate were loaded onto the system in FPLC buffer running at 0.35 ml/min, with fractions collected continuously over 0.72 minute intervals. The void volume (V_0) was defined as the peak elution volume of blue dextran, which was measured through colorimetry. Elution volume for standards and purified proteins was determined by SDS-PAGE and Coomassie blue staining of successive fractions. A standard curve of K_{av} versus apparent molecular mass was determined by measuring the elution volume V_r for low and high molecular mass protein standards (Amersham), calculating $K_{av} = V_r–V_0 / V_c–V_0$ for each standard, and plotting K_{av} vs. $\log[M_r]$ where M_r is the known apparent molecular mass of the standard. The curve was approximately linear at $R^2 = 0.95$ over the range of standards from $M_r = 13.7 – 669$ kD.

For detection of specific proteins by western blot, aliquots of appropriate fractions were denatured for 3 minutes at 90ºC in Laemmli buffer, loaded onto an appropriate gel for SDS-PAGE (Criterion or Ready-gel systems, Bio-Rad) transferred to PVDF membrane (Immobilon-P, Millipore), blocked for at least 30 minutes in PBST (phosphate buffered saline + 0.05% Tween 20) + 5% powdered milk, incubated with

appropriate primary antibody for 1 hour in PBST + 1% milk, washed for 15 minutes in PBST, incubated for 30 minutes with an appropriate HRP-conjugated secondary antibody in PBST + 1% milk, washed for 15 minutes in PBST, and imaged using the EC chemiluminescent kit (Millipore) and darkroom film exposure. The polyclonal antibody used for detection of SCCRO was previously described(14) and recognizes an epitope in the c-terminus of the SCCRO protein. Other utilized antibodies include anti-CUL1 (Zymed), anti-CUL2 (Abcam), anti-CUL3 (BD), anti-CUL4 (Santa Cruz), anti-CUL5 (Santa Cruz), anti-Tip120a (BD), anti-UBC12 (Rockland), anti-cyclin A (Santa Cruz), anti-p21 (Pharmigen), anti-α-tubulin (CalBiochem), and anti-HA (Convance).

HeLa-S elutriation and FACS analysis

HeLa-S cells were grown in a spinner flask in Dulbecco's modified Eagle's medium (high glucose, with L-glutamine) plus 10% fetal calf serum, penicillin, and streptomycin. Cells were maintained under standard mammalian cell culture conditions at a density of 2 – 8 x 10^5 cells per ml. Centrifugal elutriation was performed using a Beckman J2-21 M centrifuge and a JE-6B rotor with a large (40 ml) separation chamber. Four liters of media containing 6 x 10^5 cells per ml were concentrated via centrifugation into a total volume of 20 ml of cells and ice-cold fresh media. The cells were loaded into the centrifuge spinning at 1600 rpm in a continuously pumped stream of ice-cold PBS without Mg^{2+} or Ca^{2+} and supplemented with 1% FCS. A total of 10 fractions were collected from the centrifuge outflow according to a previously validated algorithm of stepwise increases in PBS flow or decrease in rotational velocity of the centrifuge. 10 mcl of the load cell mixture and 10 ml of each fraction were immediately spun down and frozen at -80ºC for later FACS analysis. Simultaneously, the bulk of cells from each fraction were spun down and frozen at -80ºC for subsequent lysate preparation. Lysates of thawed cell fractions were prepared in FPLC buffer as above.

FACS analysis was performed by fixing aliquots of thawed cells in ethanol and incubating for 30 min at 37°C in 0.5 ml of staining solution (25 µg/ml propidium iodide and 10 µg/ml RNase A in PBS). Stained cells were analyzed for DNA content using a Becton-Dickinson FACScan flow cytometer.

Gel filtration of fractions F3, F6, and F8 was performed as above. SDS-PAGE of gel filtration fractions was performed on two separate gels. These were then cut horizontally and the halves rearranged prior to transfer such that all regions to be probed with the same antibody were transferred to a single membrane. The remainder of the western blot protocol was performed as above.

Detection, phylogenetic, and sequence analysis of Dcun1d genes

Dcn1-like domain-containing genes were identified in *H. sapiens, M. musculus, R. norvegicus, G. gallus, D. rerio, D. melanogaster, C. elegans, S. cerevisiae, S. pombe, A. thaliana, and O. sativa* using publically available sequences in Genbank and analysis tools on the NCBI website, including protein-to-nucleotide translated BLAST against expressed sequences, predicted transcripts, and genomic sequences (See Appendices A and B for representative sequences). Expressed sequence evidence for predicted genes was inspected against genomic loci using Map Viewer and the UCSC genome browser (http://genome.ucsc.edu). Where incomplete or inconsistent expressed sequence evidence was present, genomic sequences and homology to genes in other organisms were used to choose the best full-length coding sequences. Due to ongoing addition of expressed sequence data, and automated and curated gene prediction, nearly all genes identified over the course of this work have since been independently annotated in NCBI and ENSEMBL databases in a manner consistent with our predictions, and the current NCBI gene nomenclature was adopted in writing this manuscript.

Multi-alignment of nucleotide and protein sequences were performed and graphically plotted with the ClustalX program(104) using the Gonnet matrix for amino acids. Refinements (primarily gap-reduction) were performed manually using the GenDoc program (available at http://www.psc.edu/biomed/genedoc/). Sequences for the human Dcn1 domain-containing genes were submitted to the PredictProtein server(105) for identification of functional domains, non-regular secondary structure regions, localization signals and other conserved domains. The “MYR predictor” was used for secondary analysis of putative N-myristoylation signals (http://mendel.imp.ac.at/myristate/ SUPLpredictor.htm).(106)

Acquisition and subcloning of murine Dcun1d genes

cDNA sequences corresponding to intact murine dcn1-domain containing transcripts were identified in the database of the RIKEN consortium's Functional Annotation of the Mouse 3 (FANTOM3) transcriptome sequencing project(107), and sequence-verified cDNA clones were obtained (DNAform, Tokyo Japan – see Appendix A for a complete list). The coding sequences of the five genes were then subcloned, using standard PCR-based directional cloning techniques, into pEGFP-N2, pEGFP-C2 (Clontech), pGex-4T3, and phCMV3 (Genlantis). Site-directed mutagenesis of Dcn1d3 was accomplished via PCR cloning with a single-nucleotide mismatch encoding the mutation contained within the forward primer. All constructs were sequenced prior to use to verify proper insertion and PCR fidelity.

Expression of Dcun1d constructs and fluorescent microscopy

HeLa cells were grown as above and transfected with appropriate gene-bearing or control plasmids using Lipofectamine 2000 (Invitrogen) per the manufacturer's specifications. For western blot, cells were lysed on the plate 24 hours after transfection in Mammalian Cell Lysis Buffer (Cell Signalling) containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, and 1 ug/ml leupeptin.

Cells transfected with EGFP vectors were digitally imaged 24 hours after transfection with a Nikon IX51 fluorescent microscope using standard excitation and filter settings for eGFP fluorescence.

Expression and purification of Dcun1d proteins in E. coli

Cultures of protease-deficient *E. coli* (strain BL21, Novagen) bearing the appropriate pGEX-4T3 vector were grown from starter culture in LB + ampicillin in a 37ºC shaker. Cultures were grown for 3 hours pre-induction, followed by addition of IPTG to 1 mM and 3 hours of post-induction growth. Bacteria were harvested by centrifugation and resuspended in PBS containing 1 mg/ml lysozyme, 5 mcg/ml

DNAase, 5 mcg/ml RNAase, and protease inhibitor cocktail. The bacteria were lysed on ice through repeated sonication over 30 min, and insoluble material was removed via centrifugation. The supernatant was inverted with glutathione-sepharose resin at 4ºC for 40 minutes. The beads were washed 3 times with bacterial bead-wash buffer containing 50mM Tris-Cl, pH 8.0, 150mM NaCl, 2.5mM MgCl2, 0.5% NP40, 1 tablet Complete Mini Protease Inhibitor (Roche) / 50 ml, 2.5 mM DTT, 0.2 mM PMSF, and 0.03% SDS, followed by 2 washes and resuspension in PBS. The beads were incubated with thrombin overnight at 4ºC to cleave the gene product from the GST moity. The supernatant was removed, protein concentrations measured by colormetric assay (Bio-Rad protein assay, Bio-Rad) and sample purity and relative concentration were verified via SDS-PAGE electrophoresis and Coomassie blue staining. Purified proteins were diluted to appropriate concentrations in PBS + 5% glycerol and stored at -80ºC in single-use aliquots.

GST pulldown of Dcun1d protein binding partners

Lysates of bacteria expressing GST or GST-tagged Dcun1d proteins were prepared as above and were incubated with equal aliquots of glutathione-sepharose beads, which were washed 3 times in bacterial bead-wash buffer, followed by 2 washes with PBS. Volumes of HeLa lysate containing 500 mcg of protein were added to each aliquot of beads and inverted at 4ºC for 1 hour. Beads were then washed 5 times in a buffer containing 20 mM Tris-Cl, ph 8.0, 150 mM NaCl, 0.5% Np40, 1 tab complete mini PI / 50 ml, 2.5 mM DTT and 0.2 mM PMSF. The beads were then boiled for 2 minutes in 1x Laemmli buffer, and equal volumes of supernatant were loaded for SDS-PAGE and Western Blot as above. A second SDS-PAGE gel was stained with Coomassie blue to verify that equivalent concentrations of GST or GST-Dcun1d proteins were present in each sample.

In-vitro neddylation reactions

In-vitro neddylation reactions were performed in 50 mcl reactions containing 10 ng APPBP1 (Boston Biochem), 20 ng UBC12 (Boston Biochem), and 5 mcg Nedd8 protein (Boston Biochem), 4 mcM ATP, 50 mcg HeLa-S lysate, 100 mM Tris HCl pH 7.45, 4.4 mM Na F, 110 mM NaCl, 24 mM $MgCl_2$, 4 mM $CaCl_2$

and 20 to 1000ng of the appropriate bacterially expressed, purified Dcun1d protein. Reactions were prepared on ice, incubated simultaneously at 37°C for 10-15 minutes, and stopped by addition of 10 mcl of 6x denaturing Laemmli buffer. 20-30 mcl of each reaction was then run on a 7.5% Criterion SDS-PAGE gel (Bio-Rad), transferred to PVDF membrane, and probed with appropriate antibodies.

RNA extraction and Northern Blot

Total RNA was extracted from 100 mg of flash-frozen mouse tissue, using the Qia-shredder system (Qiagen) for homogenization of tissue in Trizol reagent (Invitrogen). RNA was extracted using the Trizol manufacturer's protocol and was then further purified using the RNeasy kit with on-column DNAase digestion (Qiagen). RNA was frozen in single-use aliquots and samples were submitted for quality assessment with an Agilent Bioanalyzer system through the Sloan-Kettering Institute Microarray core facility.

Northern blot was performed using the Northern Max Gly kit (Ambion) according to the manufacturer's instructions, with 15 mcg total RNA loaded per lane and transferred to a positively charged nylon membrane (Roche). ^{32}P – labeled probe was prepared using the Strip-EZ PCR kit (Ambion), appropriate PCR primers, and a plasmid bearing murine *Dcun1d1* cDNA (RIKEN clone F630119O03) as a template. Excess nucleotide was removed from the ^{32}P-labeled PCR product using Probequant G-50 microcolumns (Amersham), and specific radioactivity of the probe was quantified using a scintillation counter. Following hybridization and washing, membranes were wrapped in plastic wrap and visualized with a phosphorimager screen and a Storm phosphorimager reader (GE).

RT-PCR and quantitative RT-PCR

RT-PCR for *Dcun1d* gene transcripts was performed with primer pairs designed using the Primer3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www_slow.cgi).(108) All primer pairs except for mDcun1D1_RT2 crossed at least one splice junction to minimize mispriming on contaminating

genomic DNA, and were designed to minimize homology with related genes. Primer pairs encompassing the full predicted coding sequence of each of the murine Dcun1d paralogs were constructed to verify the existence of full-length transcripts for each gene in mouse tissues.

Primer pairs for SYBER Green quantitative RT-PCR were designed to produce products from 75-150 bp in length. Where Genbank expressed sequence data or non-quantitative RT-PCR of the full coding sequence suggested the existence of alternate transcripts with truncated coding sequences, primer pairs were designed to straddle splice junctions which existed only in the full-length transcripts. Locations of primer pairs on murine Dcun1d transcripts are diagrammed in Figure 7 (Results).

Primer pairs used for SYBER green quantitative RT-PCR:

mDcun1D1_RT2_F:	GAAGGGCACACTCACTCACA
mDcun1D1_RT2_R:	GAGACCTGACCCACACGAAG
mDcun1D1_RT3_F:	TGGAATCTTCTGTTAGACTTCAGTTC
mDcun1D1_RT3_R:	TGTTGTACTTTTTGTCCCAGCA
mDcun1D1ex1_RT1_F:	GGAGGAGGAGGGGAGAGG
mDcun1D1ex1_RT1_R:	TTGTGTGAAGATCATAAACTGACG
mDcun1D1ex1AB_RT1_F:	TGACAGACTGGCTCTGAACAA
mDcun1D1ex1AB_RT1_R:	CATTTTGAGAAAGACAACTTACTGC
mDcun1D1ex1C_RT1_F:	GCAGTCCTTCCTTCAGCTTC
mDcun1D1ex1C_RT1_R:	TGCAGTTTTCTCACTAGATTGTGT
mDcun1D2_RT1_F:	CTAAGAACCCTGGGCAGAAG
mDcun1D2_RT1_R:	TGATGCTCCAGCAGAAATGT
mDcun1D3_RT1_F:	GAATTTCGAGTGCTGCTCTT
mDcun1D3_RT1_R:	ACAGATCCCATCAATGCTGT
mDcun1D4_RT2_F:	CAACACTGGCAAGCATTCAT
mDcun1D4_RT2_R:	AGTCCGAAGAACTGCAAGAG
mDcun1D5_RT1_F:	AGAACAGTTCATGCCGATCTT
mDcun1D5_RT1_R:	AGGTTTGCATTGTCTTCACG
m18s_F:	GTAACCCGTTGAACCCCATT
m18s_R:	CCATCCAATCGGTAGTAGCG
mGAPDH:	TGCACCACCAACTGCTTAGC
mGAPDH:	GGCATGGACTGTGGTCATGAG

cDNA was prepared from mouse RNA using the iScript reverse transcription kit (Bio-Rad) per the manufacturer's instructions, with 1mcg RNA per 20 mcl reaction. For non-quantitative RT-PCR, PCR reactions were run with Platinum Taq PCR supermix (Invitrogen) for 30-40 cycles and run on a 2% agarose/TAE gel.

For quantitative RT-PCR, reactions were prepared using iQ SYBER green Supermix (Bio-Rad) according to the manufacturer's instructions and were run on a MyQ single-color real-time PCR detection system (Bio-Rad). PCR settings were the same for all primer sets: following a 5-minute denaturation step at 95 degrees, each PCR cycle consisted of 30 seconds denaturation at 95°C, 30 seconds annealing at 58°C, and 30 seconds elongation at 72°C for 40 cycles. After cycling was complete, melting curves were determined for the products of each sample at 0.5°C increments from 55-95°C. All primer sets were shown to have high amplification efficiency (>90%) under experimental conditions, and produced a single melting peak and a single product of the expected length on agarose gel electrophoresis. Primer sets which did not meet these criteria were redesigned as needed. Primer pairs were verified for specificity by running each pair under standard experimental conditions against 5 pM of cDNA plasmid containing each of the other genes. The threshold was reached at 15-18 cycles for specific gene / primer combinations, while no threshold was reached at 30 cycles for nonspecific combinations.

Two primer sets were used for each experimental run on a 96-well plate. For each primer set, 20 mcl reactions were set up in triplicate for each of nine tissues, with each reaction containing $1/20^{th}$ of the product of one RT reaction. One control well was included for each cDNA/primer set combination, containing the product of a mock RT reaction in which reverse transcriptase was omitted. Also included on the same plate were duplicate sets of five 10x dilutions of testes or (for Dcun1d2) heart cDNA prepared at a higher concentration than the cDNA used in the experimental wells, such that all experimental data fell within the standard curve. Fluorescence threshold level and data analysis mode were predetermined and uniform for all runs.

Analysis of quantitative RT-PCR data was performed via an adaptation of the method of Pfaffl,(109) in which the efficiency of the PCR reaction was calculated using the dilution curves on each experimental plate. The primer-pair specific efficiency was then used to calculate the fold-difference in expression (averaged across three replicates) for each tissue relative to brain, using the experimentally obtained Ct values. The values for each tissue were then normalized to the relative fold-difference in expression of 18s ribosomal RNA. During pilot experiments, *GAPDH* and *HPRT* were also evaluated as normalization genes and were found to produce equivalent results. Error bars reflect the range of expression calculated using the average Ct for that gene plus or minus the root-mean-square of the variances for the three replicates.

Germ cell elutriation

Elutriation of mouse and rat germ cells was performed according to previously published protocols(110-112) Briefly, testes were removed from normal Sprague-Dawley rats, minced and trypsinized with 1 mg/ml trypsin and 19 Fg/ml DNase in phosphate buffered saline containing 0.1% glucose, 0.06% lactate, and 1 mM pyruvate at 32°C for 30 min. The suspension was filtered through 300~km and 80-pm nylon mesh membranes, and then through glass wool. The suspension was centrifuged and the pellets were resuspended in 100 ml of PBS with 250 Fg/ml DNase, 0.02% soybean trypsin inhibitor, and 5 mM naphthol disulfonic acid. The cells were cooled, filtered through a 25pm nylon mesh membrane, and loaded into a JE-6B elutriation rotor. Spermatogonia, mid-late pachytene spermatocyte, round spermatid, elongating spermatid, and late spermatid-enriched fractions were collected according to a validated elutriation sequence and were checked morphologically for purity. RNA was extracted from purified cell populations using standard techniques.

Quantitative RT-PCR for these samples was performed with the previously described primer sets, using a PE Applied Biosystems model 7700 Sequence Detection System. Expression relative to whole testes was calculated by the standard curve method, and normalized relative to 18s ribosomal RNA.

Contributions of co-authors and other collaborators

Except as stated below, all experiments presented in the "Results" section were planned and executed by R. Ryan. As principal investigator of the Laboratory of Epithelial Cancer Biology, B. Singh provided guidance and approved all experiments. Senior Research Scientist Y. Ramanathan assisted in the planning of all experiments. Gel filtration and HeLa-S elutriation experiments were conducted by R. Ryan in the laboratory of Andrew Koff, who generously provided his equipment, time, and expertise. L. Choi assisted with the cloning of eGFP fusion constructs and with fluorescent microscopy. Y. Yonekawa performed GST pulldown experiments, using fusion constructs created by R. Ryan. In-vitro neddylation reactions were performed by R. Ryan, and were adapted from assays optimized by A. Kim and Y. Yonekawa, who also assisted in preparation of materials. Mouse tissues for RNA extraction were harvested by A. Kaufman, who also provided helpful insights regarding his Dcun1d1$^{-/-}$ mouse work. Mouse germ cell elutriation was performed by Lyann Mitchell in the laboratory of P. Morris, who contributed her equipment, time and expertise. RNA extraction and quantitative RT-PCR for germ cell elutriation experiments were performed and analyzed by K. Hwang, using primers sets designed by R. Ryan. All other quantitative RT-PCR experiments were performed by R. Ryan. The full text of this manuscript was written by R. Ryan, in consultation with the co-authors.

Results

A. Gel filtration analysis of SCCRO and CDL complexes

Our group has previously demonstrated that SCCRO participates in protein-protein interactions with several cullin proteins, as well as with ROC1 and CAND1, which are well-established as members of cullin-containing complexes. Interactions with cullin proteins have been independently reported for SCCRO orthologs in *S. cerevisiae* and *C. elegans*. The observation that SCCRO interacts with CAND1 but not with SKP1(21) suggested that SCCRO may directly interact with the inhibitory CAND1-cullin-ROC1 complex, but not with active E3's such as the SKP1-CUL1-F-box (SCF) complex. We had also demonstrated an interaction between SCCRO and the E2 for neddylation, UBC12. There has been no previous evidence as to whether SCCRO's interactions with cullin-containing complexes, or with UBC12, are transient, or if in fact it forms stable complexes with any of its known interactors.

To approach this question, we performed size exclusion chromatography on lysates prepared from several human cell lines, including an immortalized human keratinocyte line (HaCaT), and the epithelial tumor lines HeLa and SCC15. We then used western blot analysis to determine the apparent molecular weight of complexes containing SCCRO and SCCRO-associated proteins, based on a previously determined standard curve (Figure 1). Similar data was obtained for all three cell lines.

Western blot analysis demonstrated similar distributions for CUL1 and ROC1, which co-migrated in a broad range of fractions corresponding to complexes from 115 to greater than 1000 kD in size (Figure 1). This is consistent with the known, stable association of ROC1 in a range of active and inactive cullin-containing complexes. A small amount of ROC1 immunoreactivity was also noted in fractions with an apparent molecular mass of 9-12 kD, consistent with the size of monomeric ROC1 (13kD). CAND1 co-migrated with CUL1 and ROC1 in a subset of fractions representing complexes between 115 and 575 kD. These findings are consistent with the presence of a stable CAND1-CUL1-ROC1 complex.

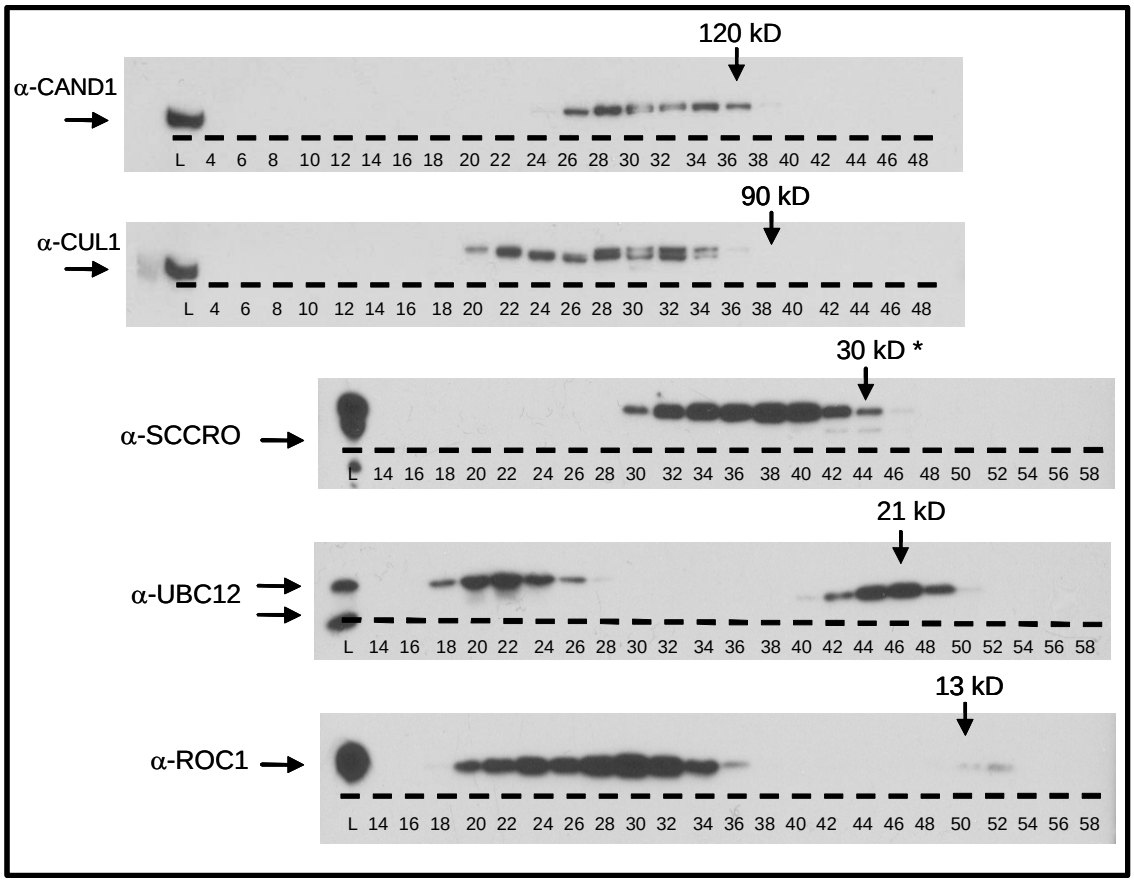


Figure 1: Gel filtration analysis followed by western blot for SCCRO and associated proteins. Arrows indicate fractions in which monomeric (non-complexed) forms of each protein would be predicted to appear, based on MW and a previously determined standard curve. The actual migration of monomeric SCCRO at an apparent MW of 32 kD was separately confirmed using purified protein (*). Note similarity in distribution of complexes containing CUL1 and ROC1.

We next looked at the distribution of UBC12, the E2 for cullin neddylation, which is known to transiently interact with cullin-containing complexes. UBC12 immunoreactivity was noted in two separate, distinct peaks, with one peak corresponding to an M_r of 27 kD (close to the actual molecular weight of monomeric UBC12, 21 kD), and a second peak within the upper range of CUL1 containing complexes.

Peak SCCRO immunoreactivity was not found in fractions containing significant amounts of CAND1, CUL1, or ROC1. SCCRO was present in a broad range of fractions, with a peak at an apparent molecular weight of 72 kD and an extended "tail" of immunoreactivity into fractions representing M_r >200 kD. This contrasts with the migration pattern of purified, bacterially expressed SCCRO, which migrated at an M_r of

32 kD (consistent with the true MW of monomeric SCCRO, 30 kD). The migration of lysate-derived SCCRO at a higher molecular weight than that of purified SCCRO is likely due to interactions with other proteins, which were not specifically identified in this experiment. We concluded that the majority of endogenous cellular SCCRO is not stably associated with CAND1 or cullin-containing complexes, nor with UBC12. It is likely that SCCRO interactions with those complexes are either transient or involve only a small fraction of endogenous SCCRO molecules.

B. Expression of SCCRO across the cell cycle

S. cerevisiae dcn1p has been shown to affect the neddylation state and function of the Cul1 homolog cdc53p.(106) One of the best-established roles for Cul1-containing SCF complexes is in control of the cell cycle. While cellular levels of Cul1 protein are constant across the cell cycle, other factors involved in SCF function, such as the ubiquitin E2 enzymes Ubc10 and Ubc7, do vary at specific stages.(113) These variations allow for increased SCF-dependant ubiquitination activity towards specific targets, such as cyclins or cdk inhibitors, at appropriate transition points in the cycle. We reasoned that SCCRO expression might also be differentially regulated throughout the cell cycle, and that such a pattern of expression might provide clues to SCCRO's normal function.

We first evaluated the expression of SCCRO protein in three well-characterized human suspension culture cell lines, HeLa-S, HeLa-S3, MANCA, and TK-6, and found them to be equivalent (data not shown). We chose to continue our analysis in the HeLa-S cells. Elutriation is a method of fractionating cells according to cell volume, with smaller cells appearing in earlier fractions, and larger cells in later fractions. Because suspension culture cells increase in size continuously throughout the cell cycle, the first fractions should contain post-mitotic, early G_1 cells, while successive fractions should be enriched for later cell-cycle stages. We collected a total of 10 separate elutriated fractions from a single log-phase culture of HeLa-S cells. Small aliquots of each fraction were frozen for FACS analysis of nuclear DNA content, while a cell lysate was prepared from the bulk of the cells in each fraction.

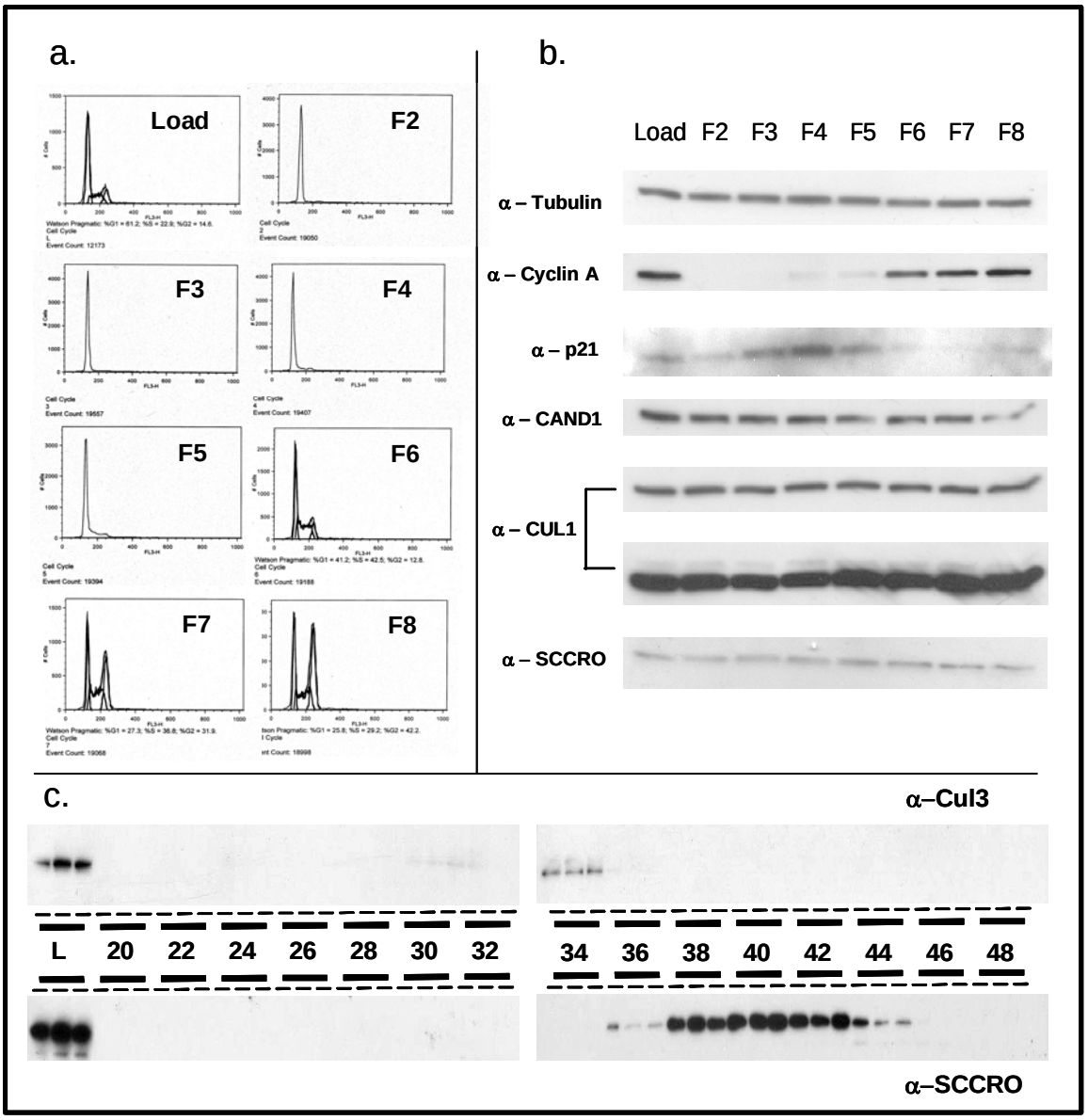


Figure 2: (a) FACS analysis of nuclear DNA content showing cell-cycle stage distribution of elutriated HeLa-S fractions. **(b)** Western Blot analysis of lysate from elutriated cell fractions shows expected cell-cycle specific expression of cyclin A and p21, but no variation in levels of CAND1, Cul1, high molecular weight (neddylated) cul1, or SCCRO. **(c)** Western blot following gel filtration of elutriated HeLa fractions. The three lanes above and below each number represent equivalent gel filtration fractions from three separate runs, each with a different elutriated cell lysate input. For each trio, the first lane represents lysate from elutriated fraction F3 (G_1-enriched), the second F6 (S-enriched), and the third F8 (G_2-enriched).

Cell cycle analysis based on DNA content revealed that unfractionated cells were approximately 61% G_1 (1n), 23% S (1n < x < 2n), and 15% G_2 (2n). DNA content analysis of each fraction revealed that fractions 2-8 represented enrichment of successively later phases of the cell cycle, with fractions 2 and 3 representing essentially 100% G_1 cells, while fraction 6 contained 41% G_1, 43% S, and 13% G_2 cells, and

fraction 8 contained 26% G_1, 29% S, and 42% G_2 (Figure 2a). Pure fractions representing cells in later stages of the cell cycle were not achieved, likely due to co-elutriation of clumps of small G_1 cells along with the larger, late-phase cells. However, the enrichments achieved were sufficient to detect significant variations in protein expression across the cell cycle, as demonstrated in the next phase of the experiment.

To further validate the cell-cycle enrichment of our fractions, we used western blot to analyze the expression of two proteins whose levels are known to vary across the cell cycle. Cyclin A is essentially absent during G_1, but is increasingly expressed in S and G_2 phases(114). $p21^{CIP1}$ is a cyclin-dependant kinase inhibitor whose cellular levels rise during G_1, fall precipitously at the G_1-S transition, remain low throughout S, then rise again moderately in G_2(115). As expected, western blot analysis of elutriated HeLa-S fractions showed that Cyclin A was absent from pure G_1 fractions, but was increasingly expressed in S and G_2 – enriched fractions (Figure 2b). Also as expected, $p21^{CIP1}$ revealed a strong expression peak in fractions 3-5 (late G_1), with low levels detected in later fractions. In contrast to the clear variations seen in the levels of these known cell-cycle regulated proteins across our fractions, we could detect no significant differences in the expression of CAND1, CUL1, high molecular weight (neddylated) CUL1, or SCCRO.

While overall levels of SCCRO did not vary across the cell cycle in HeLa-S, it remained possible that the levels or binding properties of unknown SCCRO-interacting proteins might vary, resulting in differences in the size of SCCRO-containing complexes across the cell cycle. We tested this possibility by performing size exclusion chromatography on lysate from elutriated HeLa-S cell fractions 3, 6, and 8, and evaluating the apparent molecular weight of SCCRO and CUL3 complexes by western blot (Figure 2c). There was no difference between the three cell fractions in the M_r of complexes containing SCCRO or CUL3.

C. Genomic / Bioinformatic / Phylogenetic analysis of SCCRO and Dcn1d genes

Analysis of Genbank genomic and transcript databases revealed a variety of vertebrate genes with coding sequence homology to *SCCRO* and *dcn-1* (hereafter referred to as *d*efective in *cu*llin *n*eddylation –

1 – like *d*omain – containing, or "*Dcun1d*" genes). While transcripts of the murine ortholog of SCCRO have been described, there has been no published characterization of any paralogous Dcun1d genes or gene products.[1]

Human Sequence	Entrez Gene name	Unigene Identifier	Genomic Locus	ORF length (bp)	Predicted protein product (kD)	Mouse Ortholog
DCUN1D1 /SCCRO (gene)	*DCUN1D1*	Hs.104613	3q26.3	777	30.0	*Dcun1d1*/ Mm.379305
DCUN1D2 (gene)	*DCUN1D2*	Hs.369453	13q34	777	21.3	*Dcun1d2*/ Mm.368890
DCUN1D3 (gene)	*DCUN1D3*	Hs.101007	16p12.3	912	34.3	*Dcun1d3*/ Mm.31539
DCUN1D4 (gene)	*DCUN1D4*	Hs.605388	4q11	876	34.1	*Dcun1d4*/ Mm.220312
DCUN1D5 (gene)	*DCUN1D5*	Hs.503716	11q22.3	711	27.4	*Dcun1d5*/ Mm.27293
RP42Ψ (pseudogene)	*LOC153893*	N/A	6q16.3	N/A	N/A	N/A
Dcun1d-like fragment	N/A	N/A	9p24	N/A	N/A	N/A

Table 1: Genes, pseudogenes, and other sequences in the human genome with coding homology to *Dcn1*

We used Genomic BLAST to search the human genome for all DNA sequences with coding homology to *SCCRO*. We identified sequences at seven unique loci in the human genome, including the known SCCRO gene at 3q26.3 (Table 1). A sequence at 6q16, previously reported as the locus of a human "RP42 homolog,"(19) proved in fact to be an intronless pseudogene that lacks a significant open reading frame. Another non-coding sequence homologous to SCCRO was found at 9p24. The remaining four loci were each predicted to represent a true protein-coding gene containing a dcn-1-like domain. These genes had not been previously characterized, and at the time of our initial analysis lacked consistent annotation in NCBI databases. We verified these findings by performing translated BLAST of SCCRO against real and predicted human expressed sequences, and recovered a corresponding set of transcripts for all five genes.

[1] For a thorough discussion of how gene orthology and paralogy are determined, see (116)

Using a similar methodology, we identified highly conserved orthologs to each of the five human *DCUN1D* genes in *Mus musculus* and *Danio rerio.* All five identified genes share common properties of protein-coding genes, including multiple exons, strong predicted translation initiation (Kozak) sequences, and significant (>700 bp) open reading frames. Since our original analysis, these genes have been independently assigned a uniform NCBI nomenclature: *DCUN1D1*, *DCUN1D2*, *DCUN1D3*, *DCUN1D4*, and *DCUN1D5*, which will be used throughout the remainder of this manuscript.

Because Genbank expressed sequences suggested the existence of alternate spliceforms for human *DCUN1D2* and *DCUN1D4,* as well as for their mouse orthologs, we wanted to confirm that full-length coding sequence transcripts were expressed before further analysis. Due to our ongoing work with transgenic mouse models of *Dcun1d1* overexpression and deletion, as well as the comprehensive availability of murine cDNA's, we chose to perform our characterization of the vertebrate *Dcun1d* genes in the mouse. We designed RT-PCR primers to amplify the coding sequences of murine *Dcun1d1*, *Dcun1d2*, *Dcun1d4*, and *Dcun1d5* (the coding sequence of *Dcun1d3* is contained in only two exons, which are present in all *Dcun1d3* mRNA sequences in Genbank). We extracted RNA from murine testes, lung, and heart, and performed RT-PCR with the appropriate primer sets. Products of the expected length were the dominant species produced for all four genes – three faint, lower MW bands were produced for Dcun1d2 in heart RNA, suggesting that truncated spliceforms are expressed in that tissue, but at a much lower level than the full-length CDS transcript (data not shown). No PCR products suggestive of alternately spliced transcripts were seen in mouse for *Dcun1d4*.

We next used Genbank expressed sequences and data from completed genome projects to perform a phylogenic analysis of the five paralogous *Dcun1d* genes. All five genes appear to be conserved among vertebrates for which sufficient genomic or transcriptomic information is available. The amino acid sequences of DCUN1D1 and DCUN1D2 in mice are identical in length and show 80% sequence identity to one another, suggesting that these proteins serve very similar or identical molecular functions. They appear to have resulted from a gene duplication event early in the deuterostome / chordate lineage, since they are

both present in *Danio rerio,* but are co-orthologous to a single gene, "*dcun1d1l*," found in the invertebrates *Drosophila melanogaster* and *Anopheles gambiae*. A similar pattern is seen for *Dcun1d4* and *Dcun1d5*, which share co-orthology to a single insect gene, "*dcun1d4l*." *Dcun1d3* is well conserved in the purple sea urchin *Strongyloides purpuratus,* and is more divergent but recognizable as an ortholog in *Drosophila* and *Anopheles*. Thus, the five vertebrate genes can be sub-classified into three orthologous groups which are conserved in diverse metazoan clades. While all these genes show strong conservation of the central and C-terminal sequences (the Dcn1-like domain), each of the orthologous groups possesses a unique and characteristic N-terminal motif (Figure 3, 4a).

Proteins of the first group, the orthologs of Dcun1d1 and Dcun1d2, possess an N-terminal domain with strong homology to the UBA-like domain found in *S. cerevisiae* Dcn1 and *C. elegans* Dcn-1.(103) In fact, a BLAST search for similar expressed sequences in the "taxonomically broad EST database" (TBestDB)(117) reveals that "UBA-Dcun1d" orthologs are conserved in species from five of the six phylogenic "supergroups" currently proposed to encompass most or all of eukaryotic diversity, including Opisthokonta, Plantae (*Arabidopsis*), Amoebozoa (*Dictyostelium*), Chromalveolata (*Tetrahymena*), and Excavata (*Reclinomonas, Malawimonas*) strongly suggesting that a UBA-Dcun1d gene was present in the last eukaryotic common ancestor. This underscores the importance of Dcun1d proteins in a fundamental process of eukaryotic cells.

Proteins of the second class are represented by Dcun1d4 and Dcun1d5 in vertebrates. Their N-terminals contain a conserved motif consisting of proline followed one or two residues later by 4-5 basic amino acids, a sequence similar to the canonical nuclear localization signal (NLS). We henceforth refer to this class of predicted Dcun1d proteins as the "NLS-Dcun1d" proteins. The nucleotide sequence coding for the NLS-like motif of both *Dcun1d4* and *Dcun1d5* is preceded by an AUG codon. However, Dcun1d4 proteins are predicted to have an extended N-terminal prior to the NLS-like sequence, due to a conserved upstream translation initiation sequence which is not shared with Dcun1d5. The murine Dcun1d5 protein sequence is 65% identical to the corresponding sequence in the Dcun1d4 protein. We found no evidence for

NLS-Dcun1d genes in non-metazoans, suggesting that the first *NLS-Dcun1d* gene may have evolved early in animal evolution.

The third orthologous group consists of the single gene *Dcun1d3*, which appears in the genomes of most fully sequenced metazoans, although like the *NLS-Dcun1d* genes, *Dcun1d3* is not present in *C. elegans*. Dcun1d3 has a slightly more divergent dcn1-like domains compared to the other classes, and is distinguished by a long N-terminal domain consisting of a predicted non-ordered secondary structure (NORS) domain, and a potential N-myristoylation sequence at the N-terminal. The sequences of *Homo*, *Mus*, *Tetraodon*, and *Strongylocentrotus* Dcun1d3 proteins classify as "reliable" substrates for N-myristoyltransferase according to the "MYR predictor" web server. These findings suggest that the non-ordered N-terminal portion of the Dcun1d3 protein might function as a flexible tether, linking the protein to a plasma or organelle membrane via a fatty acid anchor. We tentatively refer to this orthologous group as the "*MYR-Dcun1d*" genes. However, it should be noted that the divergent N-terminal of insect Dcun1d3-like proteins are not predicted to be likely N-myristoylation substrates by the MYR algorithm.

D. Cloning and expression of murine *Dcun1d* genes

The 5' untranslated regions of transcripts for all five genes are well-represented by expressed sequences in Genbank, and translation initiation sites can be predicted for all five genes, based on the lack of upstream AUG codons and conservation between organisms. While initiation sequences for the five paralogs display varying degrees of homology to the Kozak consensus sequence, all five possess an adenine at the -3 position, which is known to be the most important upstream determinant of an efficient initiation sequence.(118)

We obtained full-length, sequence-verified cDNA clones for each of the five murine dcun1d genes from the RIKEN consortium. The five paralogs were cloned into the mammalian expression vector phCMV3 to produce c-terminal HA-tagged coding sequences, preceded by the endogenous initiation sequence to the -9 position. Expression of these plasmids in HeLa cells produced proteins of the expected

molecular weight on HA western blot (Figure 4b).

Interestingly, we were unable to produce strong expression of Dcun1d4 protein using the phCMV3 vector. This was true even after replacing the predicted endogenous initiation sequence with the Kozak consensus sequence. We also found that expression of Dcun1d4 in bacteria via an N-terminal GST-fusion vector (pGex3T4) resulted in very weak expression compared to similar constructs for the other four genes under a variety of conditions. Thus, low expression of the Dcun1d4 protein is independent of promoter, initiation sequence or translation in a eukaryotic or prokaryotic system. A likely explanation for this is that translation is inhibited by the secondary structure of the *Dcun1d4* mRNA sequence which was predicted to code for a unique N-terminal not shared with *Dcun1d5*.

It is possible that endogenous *Dcun1d4* transcripts might be translated from the downstream AUG that corresponds to the initiation site of *Dcun1d5*, thus producing a shorter protein with full-length homology to Dcun1d5 and *Drosophila* Dcun1d4L. However, this would require an unusual 5'UTR regulatory mechanism to prevent the ribosome from initiating translation on what appears to be a favorable, in-frame upstream site. It is interesting to note that of the eight vertebrate species for which *Dcun1d4* transcript or genomic sequences are available, the upstream AUG start codon, though conserved, is shifted out of frame with the downstream coding sequence in two species (*Bos taurus* and *Canis familiaris)*. The downstream AUG is conserved in all eight species. A definitive answer to this question could only be produced through the purification and sequencing of endogenous Dcun1d4 proteins, which was not attempted in this study.

We cannot completely exclude the possibility that the transgenically expressed Dcun1d4 protein is more rapidly degraded in both bacteria and eukaryotic cells than are the other Dcun1d proteins. However it is unlikely that the surviving Dcun1d4 protein is misfolded, given the similarities in subcellular localization, protein-protein interactions, and *in-vitro* functional activity seen between Dcun1d4 and Dcun1d5 in subsequent experiments.

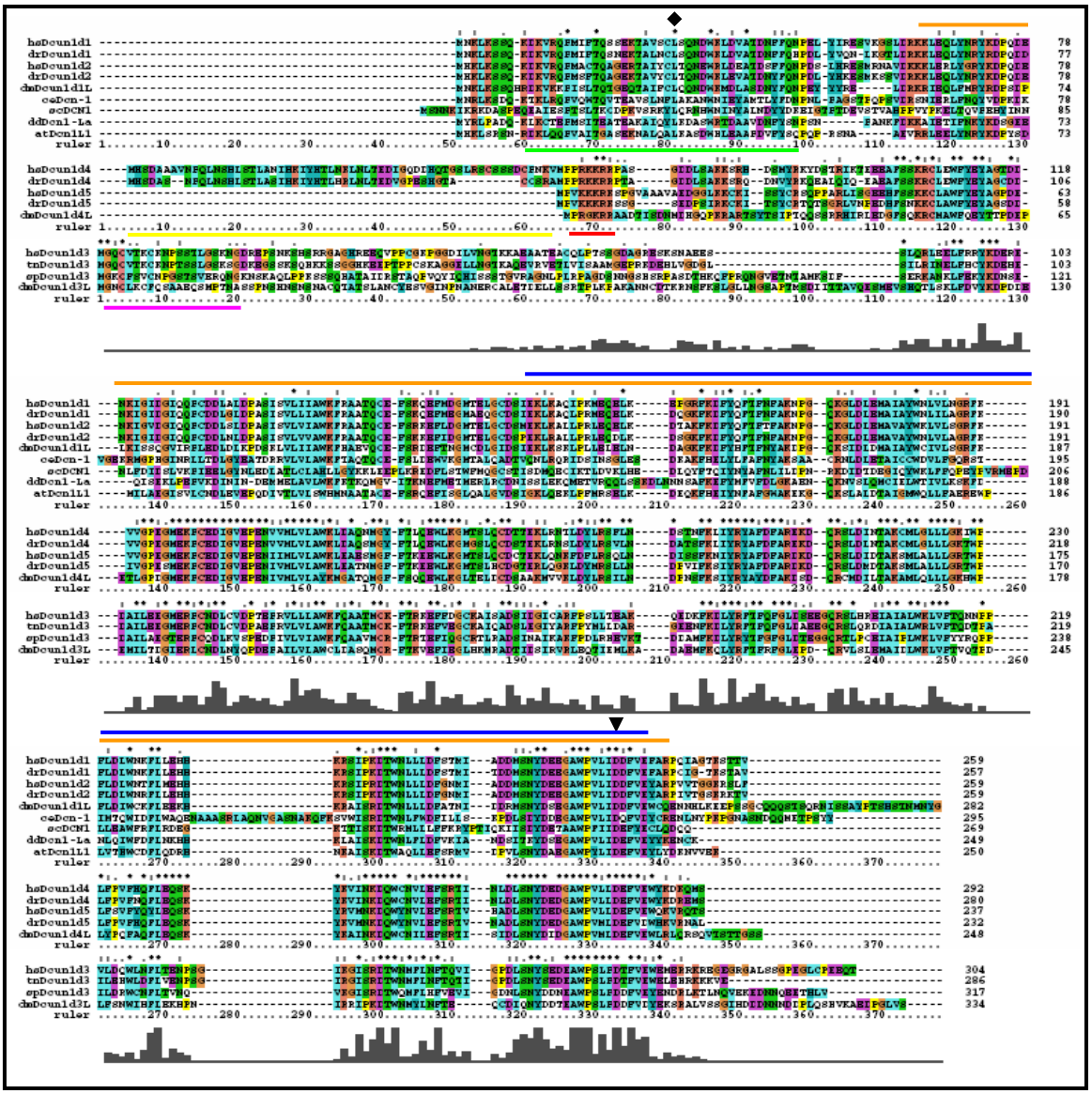


Figure 3: Multi-alignment of vertebrate Dcun1d proteins and orthologs in other organisms. Sequences are grouped by orthologous group, with proteins containing the UBA domain (green underline) at top, basic nuclear localization signal (red underline) in the middle, and N-myristoylation signal (purple underline) at bottom (note poor conservation of the MYR domain in *Drosophila* Dcun1d3L). The yellow underline marks the region of the predicted vertebrate Dcun1d4 protein which may not be endogenously translated (see further discussion in text).

All Dcun1d proteins show conservation of the Dcn-1-like domain (orange overline), which includes the more strongly conserved PFAM DUF298 domain (blue overline). The black arrowhead marks the universally conserved aspartate 241 (numbered with reference to *Homo* Dcun1d1) which is required for cullin binding and lies in the most strictly conserved region of the Dcn1-like domain.

Species names are abbreviated as follows: Hs = *Homo sapiens*, Dr = *Danio rerio*, Tn = *Tetraodon nigroviridis*, Sp = *Strongylocentrus purpuratus*, Dm = *Drosophila melanogaster*, Ce = *Caenorhabditis elegans*, Sc = *Saccharomyces cerevisiae*, Dd = *Dictyostelium discoideum*, At = *Arabidopsis thaliana*. Genbank ID's for all sequences are listed in Appendix B.

E. Subcellular localization of Dcun1d proteins is dependant on N-terminal functional domains

As an initial approach to distinguishing the function of mammalian Dcun1d genes, we decided to test the function of predicted N-terminal localization motifs by comparing the localization of N-and C-terminal eGFP-tagged SCCRO (human DCUN1D1) protein to equivalent fusion proteins containing murine Dcun1d2, Dcun1d3, Dcun1d4, and Dcun1d5 (Figure 4c).

The localization of N- and C- terminal tagged SCCRO and Dcun1d2 is diffusely nuclear and cytoplasmic, with slightly more intense fluorescence in the nucleus compared to EGFP alone. No significant localization heterogeneity was seen between interphase cells in asynchronous cultures, suggesting that there is no cell-cycle specific localization of these proteins.

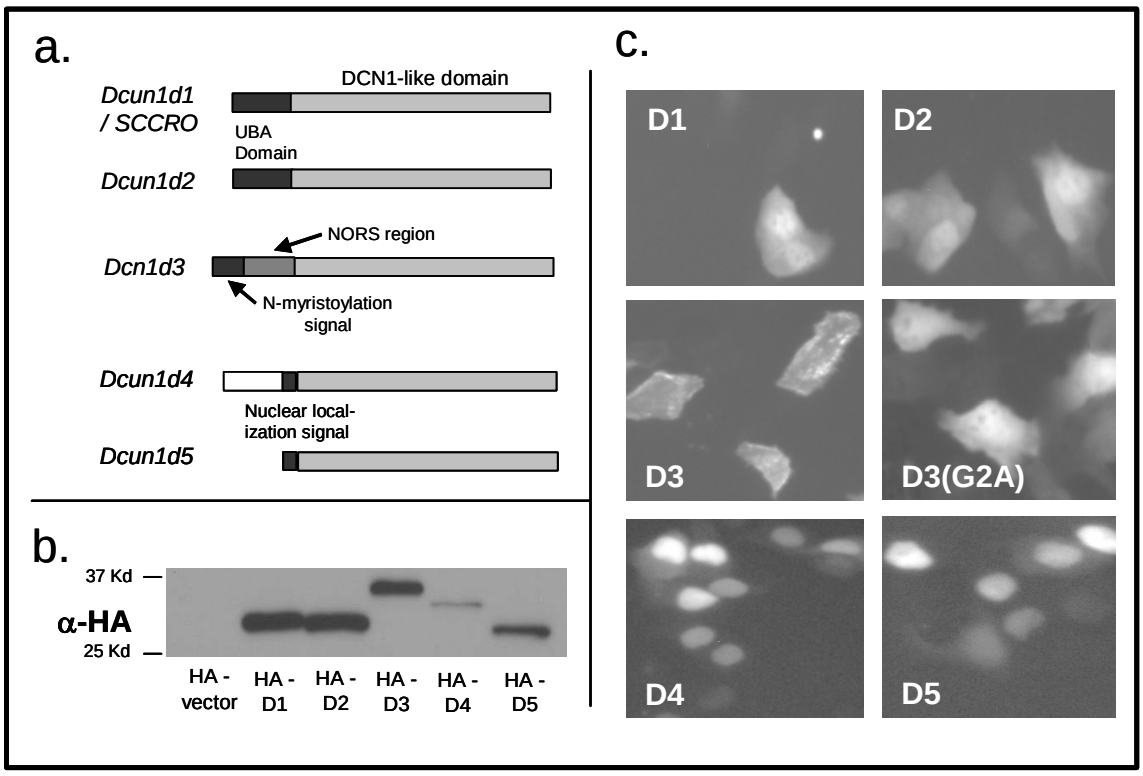


Figure 4: a. Schematic representation of domains in the five conserved vertebrate Dcun1d proteins. **b.** Western blot showing transgenic expression of C-terminal HA-tagged murine Dcun1d proteins in HeLa cells. **c.** Localization of c-terminal eGFP tagged SCCRO and murine Dcun1d proteins in transfected HeLa cells. D1 = SCCRO, D2 = Dcun1d2, D3 = Dcun1d3, D3(G2A) = Dcun1d3(G2A), D4 = Dcun1d4, D5 = Dcun1d5.

In contrast to the UBA-Dcun1d proteins, the murine NLS-Dcun1d proteins, Dcun1d4 and Dcun1d5, localized almost exclusively to the nucleus, suggesting that the basic N-terminal sequence does indeed function as a nuclear localization signal.

The C-terminal tagged Dcun1d3 protein shows a strikingly different localization pattern from the other four paralogs, with florescence seen outlining the plasma membrane as well as peri-nuclear organelles. Fluorescence was notably absent from the nucleus. Interestingly, the N-terminal tagged fusion protein (not shown) had a diffuse localization pattern throughout the cell. This would be expected if membrane localization of Dcun1d3 were dependent on N-myristoylation, a modification which can only occur at a sequence located on a free N-terminal. This was further confirmed by loss of membrane localization in a C-terminal GFP-tagged Dcun1d3(G2A) mutant, since N-terminal fatty acid conjugation can only occur on a glycine residue (the methionine in position 1 is cleaved from N-lipidated proteins).

F. *In-vitro* interaction of bacterially expressed Dcun1d proteins and CDL components.

The high degree of homology between the dcn-1-like domains of the five mammalian Dcun1d proteins suggests that they likely carry out similar molecular functions. We tested the *in vitro* protein binding

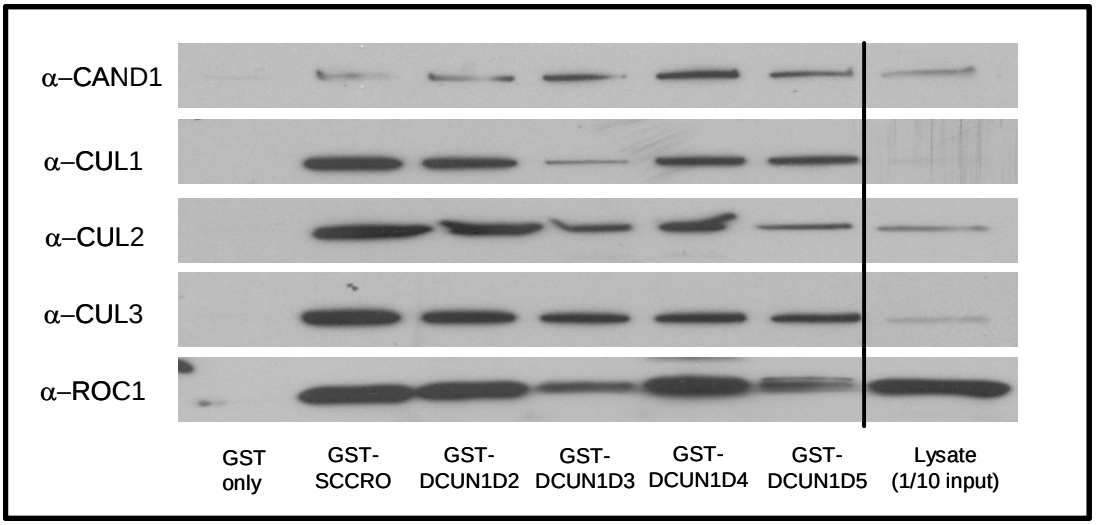


Figure 5. All five mouse Dcun1d genes interact with CAND1, cullins, and ROC1, as measured by pulldown of bacterially expressed GST-Dcun1d fusion proteins in HeLa lysate. 1/10 the original volume of lysate was run in the right-hand lane for comparison.

properties of the five murine Dcun1d genes in a GST pulldown assay, and confirmed that all five paralogs showed specific interactions with CUL1, CUL2, CUL3, CAND1, and ROC1 (Figure 5). The similar binding properties of the paralogs are not surprising, given the high degree of conservation of the Dcn-1-like domain.

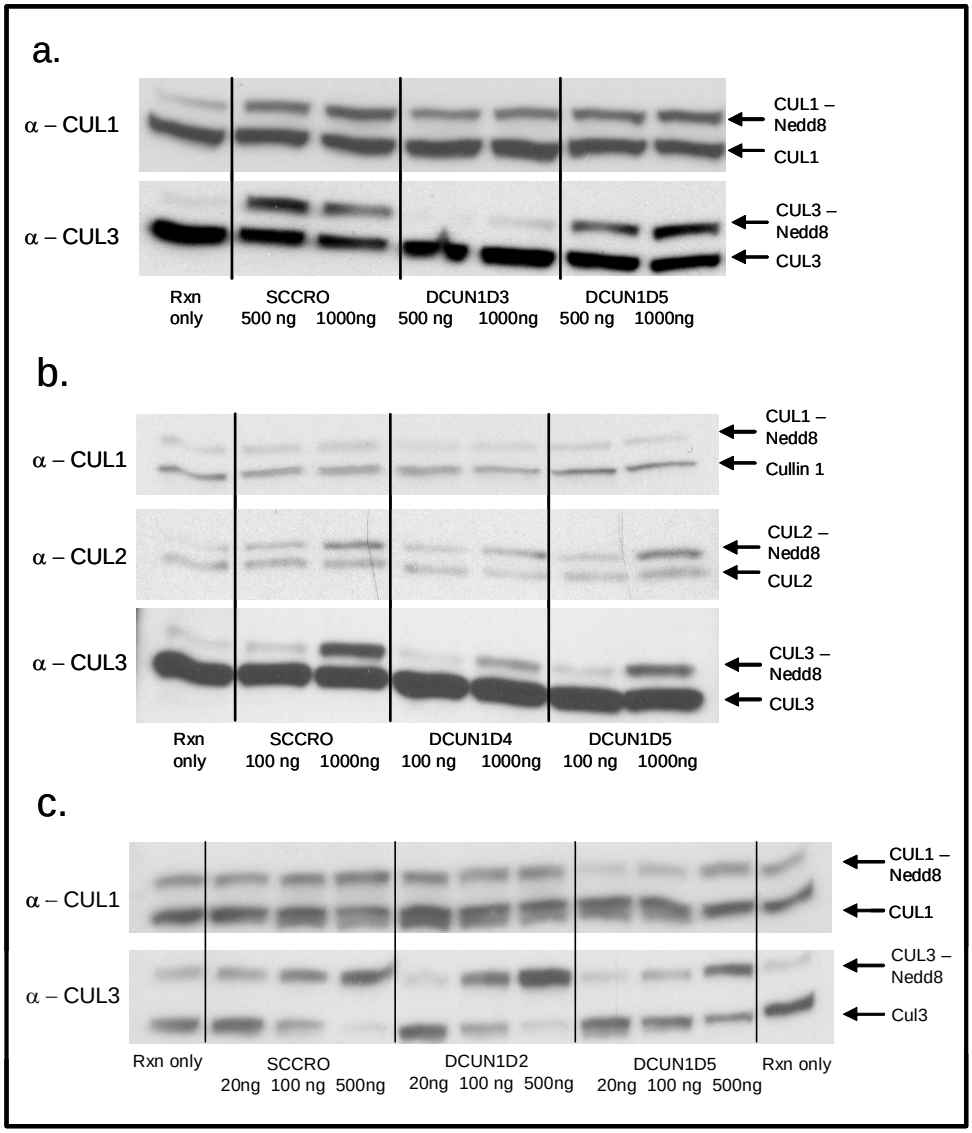


Figure 6. Dose-dependant effects of bacterially expressed SCCRO or murine Dcun1d proteins in three separate sets of 10-minute *in-vitro* neddylation reactions. Products of the same reactions were used for all blots within each section. **a.** Significantly increased CUL3 neddylation is seen with addition of SCCRO and Dcun1d5, but not with Dcun1d3. **b.** SCCRO, Dcun1d4, and Dcun1d5 increase neddylation of CUL3 in a dose-dependant manner. **c.** SCCRO, Dcun1d2, and Dcun1d5 increase neddylation of CUL3 in a dose-dependant manner. Note the lack of a significant effect on CUL1 neddylation in **a, b,** and **c,** and the modest but consistent enhancement of CUL2 neddylation in **b.**

G. Specific enhancement of *in vitro* cullin neddylation by bacterially expressed DCUN1D proteins

To test the functional properties of Dcn-1 domain-containing proteins, we have established an in-vitro assay for the neddylation of cullins in mammalian cell lysate. In previous experiments, addition of bacterially expressed SCCRO protein, but not SCCRO(D241A), which loses binding affinity for cullins, significantly enhanced the neddylation of cullin 3 in a dose-dependant manner.(21)

Under similar conditions, we found that SCCRO and murine Dcun1d2, Dcun1d4, and Dcun1d5 all enhance neddylation of CUL3 in a dose-dependant manner (Figure 6a, 6b, 6c). Enhancement of CUL2 neddylation by these four proteins was far more modest (Figure 6b), and little or no enhancement of CUL1 neddylation was observed (Figure 6a, 6b, 6c). This suggests that the neddylation-enhancing effects of Dcun1d proteins may be selective within the cullin family. Interestingly, we observed no significant enhancement of cullin neddylation with any amount of added Dcun1d3 (Figure 6a).

H. Tissue-specific expression of *Dcun1d* genes in the mouse

One of the more puzzling features of our initial work on murine *Dcun1d1* was the limited phenotypic abnormalities of the *Dcun1d1*$^{-/-}$ mouse. The existence of a second conserved protein in vertebrates, Dcun1d2, with nearly identical structural and functional properties to Dcun1d1, suggested that the two corresponding genes might be distinguished primarily by their patterns of expression. In other words, Dcun1d1 and Dcun1d2 may perform the same cellular function, but in different tissues or stages in development. According to this hypothesis, we would only expect to see consistent defects in the *Dcun1d1*$^{-/-}$ mouse in tissues where *Dcun1d2* shows little or no expression. Given the similar functional properties of their protein products *in vitro*, *Dcun1d4* and *Dcun1d5* could potentially compensate for loss of *Dcun1d1* expression as well. However, the distinct pattern of sub-cellular localization and absence of the UBA domain in the NLS-Dcun1d proteins suggests that they are not likely to replicate all the functions of Dcun1d1 *in vivo*. To test these hypotheses, we decided to further investigate patterns of *Dcun1d* gene expression in the mouse.

While previous authors have reported the expression of murine *Dcun1d1* by northern blot,(18, 19) this work was done with probes complementary to the coding sequence of *Dcun1d1*, which contains regions nearly identical to the corresponding sequences in *Dcun1d2*. This raised the possibility that cross-hybridization to *Dcun1d2* transcripts could have confounded the earlier data. To rule this out, we performed northern blot using a short, PCR-generated probe against a 100 BP region of murine *Dcun1d1* (diagrammed in Figure 7) which shares no homology with *Dcun1D2*. Our results confirmed the previously reported pattern of two long (~3.5 and 4.0 kD), ubiquitously expressed *Dcun1d1* transcripts as well as several shorter, highly expressed, testis–specific transcripts (Figure 8a).

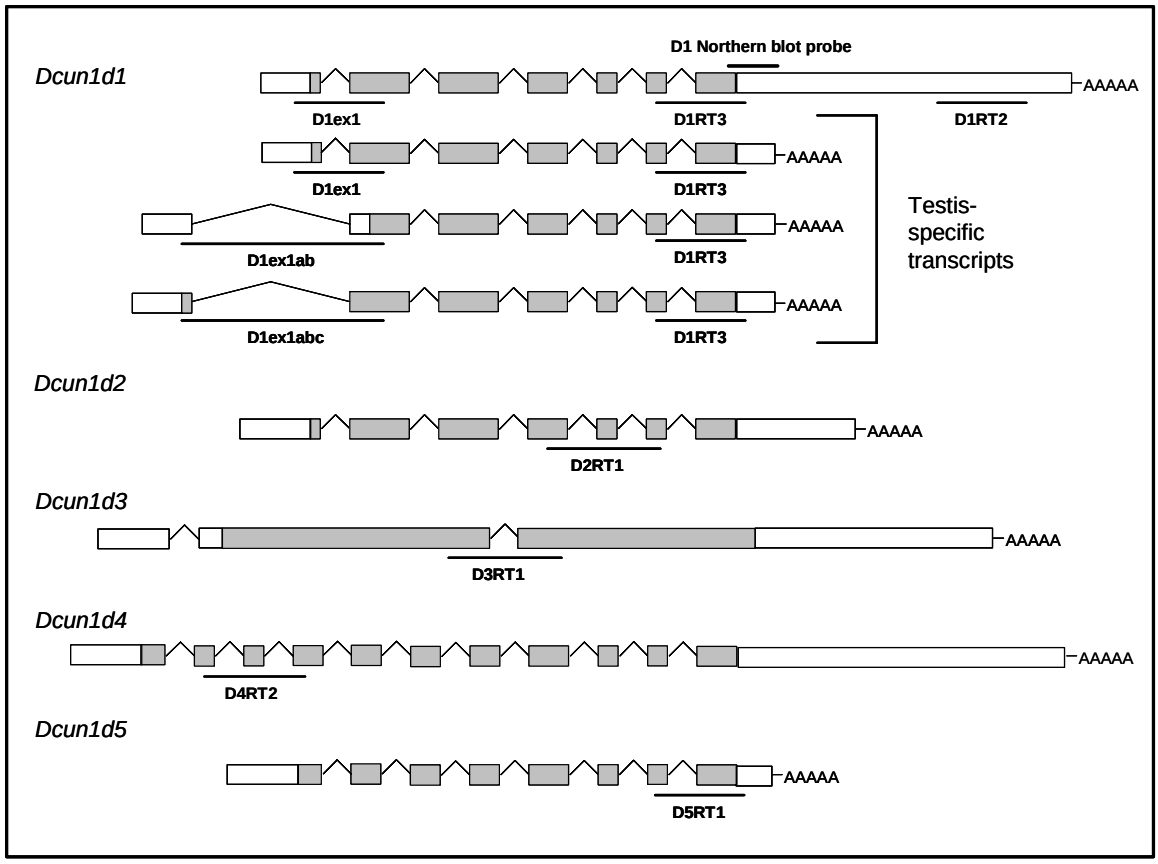


Figure 7: Diagram of murine *Dcun1d* gene transcripts showing the location of PCR products amplified for quantitative real-time RT-PCR.

Pourcel et al. cloned several of the testis-specific short transcripts of *Dcun1d1* and showed that they appear to be produced through the use of alternate polyadenylation sites, which result in truncation of the 3'UTR. While some of the testis-specific transcripts sequenced by Pourcel et al. contained the usual first

exon, others contained alternate 5' sequences, which they speculated were produced through transcription of the *Dcun1d1* gene from an unidentified alternate promoter, followed by the splicing of several new upstream exons, dubbed 1a, 1b, and 1c. Our examination of public genome project data revealed a

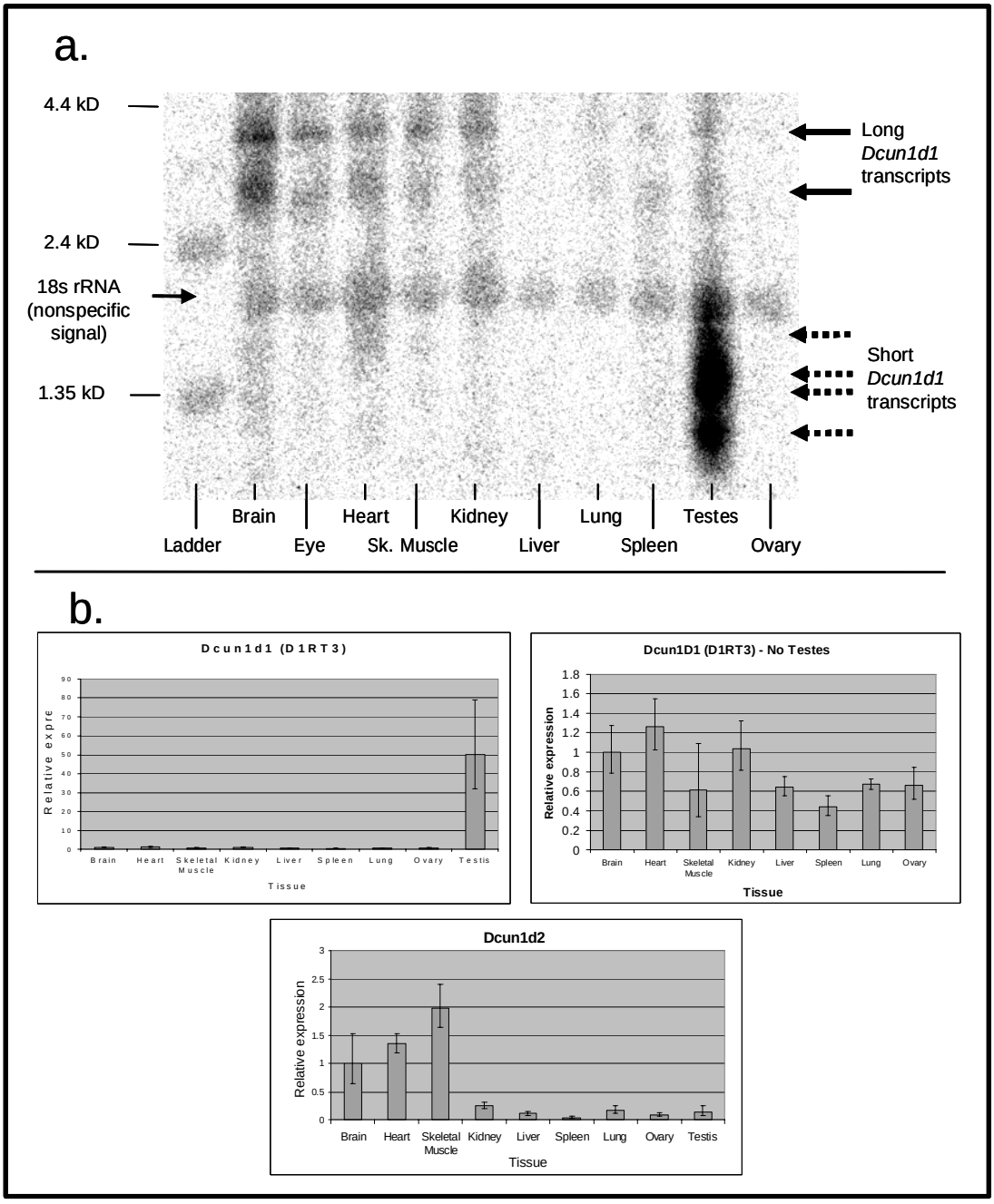


Figure 8. (a) Northern blot analysis of *Dcun1d1* transcripts using a short probe lacking homology to other Dcn1-domain containing genes. Note expression of long *Dcun1d1* transcripts in many tissues, while short transcripts are restricted to the testes. **(b)** qRT-PCR expression analysis of *Dcun1d1* and *Dcun1d2*. The second *Dcun1d1* graph omits data from the testes to highlight differences in expression of other genes, but the data otherwise identical. Tissue-specific expression was normalized to 18s rRNA. Error bars show +/- SEM of fold-difference for three replicate wells.

continuous sequence exactly matching the sequence of all three alternate "exons" 6.5 kb upstream of the transcriptional start site for the ubiquitous transcript. Thus, this sequence actually seems to represent a single alternate first exon (renamed "exon 1abc"), which is expressed from a second upstream promoter and contains two alternate donor sites for splicing to the ubiquitous exon 2.

A comparison of the *Dcun1d1* exon 1abc sequence to genomic sequences from human, rat, cow, and dog shows that only the shorter of the two spliceforms (containing the "exon1ab" sequence) seems to be phylogenically conserved. All Genbank EST's containing the alternate first exon were cloned from testis or testis-derived cells, supporting the idea that these transcripts are testis-specific.

To further examine the relative expression of Dcun1d1 transcripts in murine testes and other tissues, we designed sets of quantitative RT-PCR primer pairs to detect features of specific Dcun1d1 transcripts, including those containing the ubiquitous first exon (D1ex1), or exons 1ab or 1abc (D1ex1ab and D1ex1abc respectively). We designed a primer set to amplify a region of the *Dcun1d1* 3' UTR present in the ubiquitously expressed transcripts, but absent from the short, testes-specific transcripts (D1RT2). We also designed a primer pair specific for the full-length *Dcun1d2* transcript, taking care to minimize homology with *Dcun1d1*.

The primer pair D1RT3, which amplifies a region common to all known *Dcun1d1* transcripts, showed a multi-tissue expression profile consistent with our northern blot findings, with moderately increased expression in brain, heart, and kidney relative to liver, lung and spleen, and approximately 50-fold higher expression in the testes than in all other tissues (Figure 8b). However, expression of the long-transcript specific primer pair D1RT2 was similar in testes and other tissues (data not shown), consistent with the northern blot finding that only short Dcun1d1 transcripts are highly expressed in testes. In contrast, Dcun1d2 showed significantly higher expression in brain, heart, and skeletal muscle than in testis and other tissues (Figure 8b). Primer pairs specific for *Dcun1d1* transcripts containing exon 1ab or exon 1abc produced little or no signal in tissues other than the testes (data not shown).

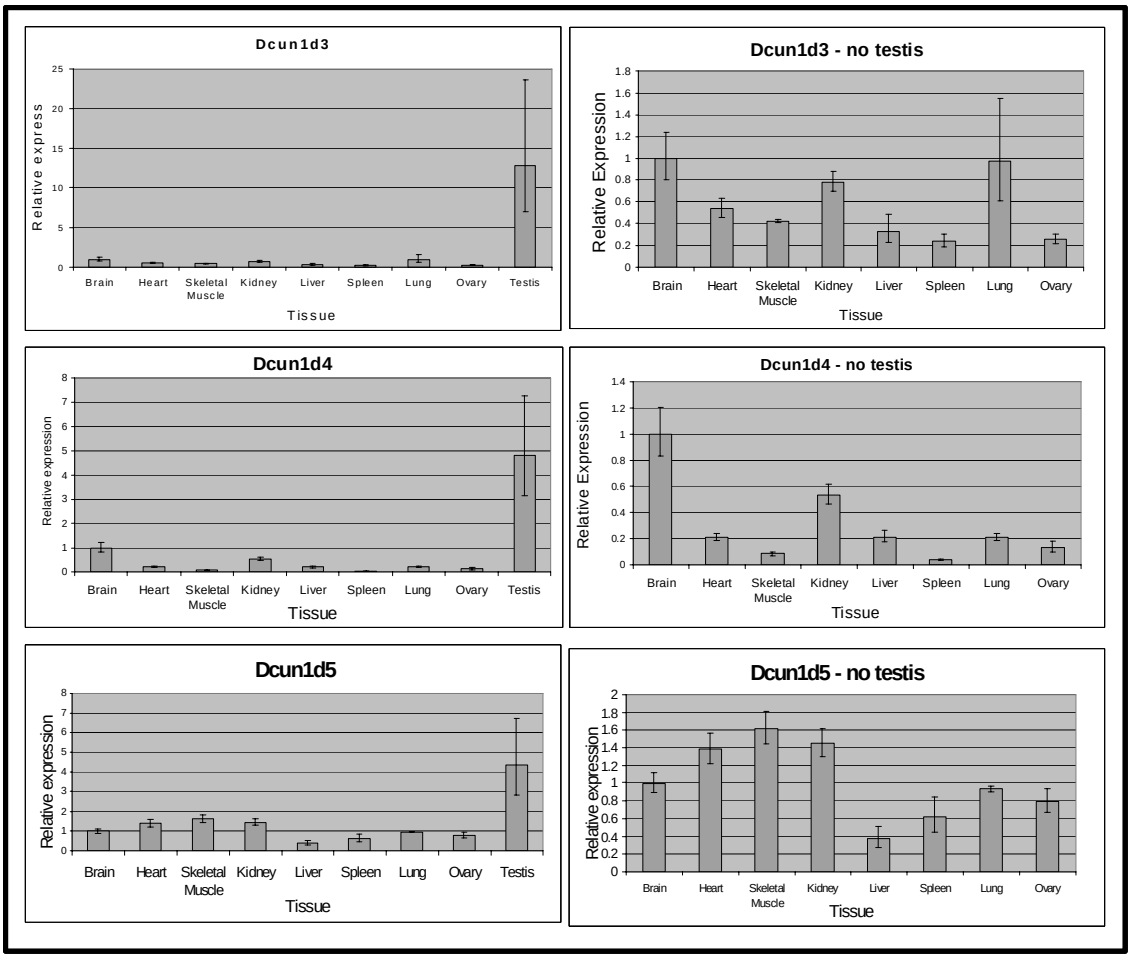


Figure 9: Expression of murine *Dcun1d* genes by RT-PCR. All three are strongly overexpressed in testis. The testis data is omitted in right-hand charts to highlight variation in expression between other tissues. Tissue-specific expression was normalized to 18s rRNA. Error bars show +/- SEM of fold-difference for three replicate wells.

Quantitative RT-PCR was also used to compare the expression of *Dcun1d3*, *Dcun1d4*, and *Dcun1d5* across nine mouse tissues (Figure 9). All three genes had their highest expression in the testis, though the difference was not as dramatic as with *Dcun1d1*. *Dcun1d4* and *Dcun1d5* otherwise showed patterns of expression which were quite distinct from one another, which may explain their conservation as separate genes, given their similar sub-cellular localization and functional properties.

I. Germ cell elutriation

The confirmation of high *Dcun1d1* expression in the testes, in light of the male sterility phenotype of the *Dcun1d1*$^{-/-}$ mouse, strongly suggests that *Dcun1d1* plays a specific role in spermatogenesis. To further explore this, we took advantage of a robust elutriation method for the isolation of mouse germ cells in

specific stages of spermatogenesis. We then isolated RNA from whole testis and enriched germ cell fractions, and performed quantitative RT-PCR using the previously described primer sets (Figure 10).

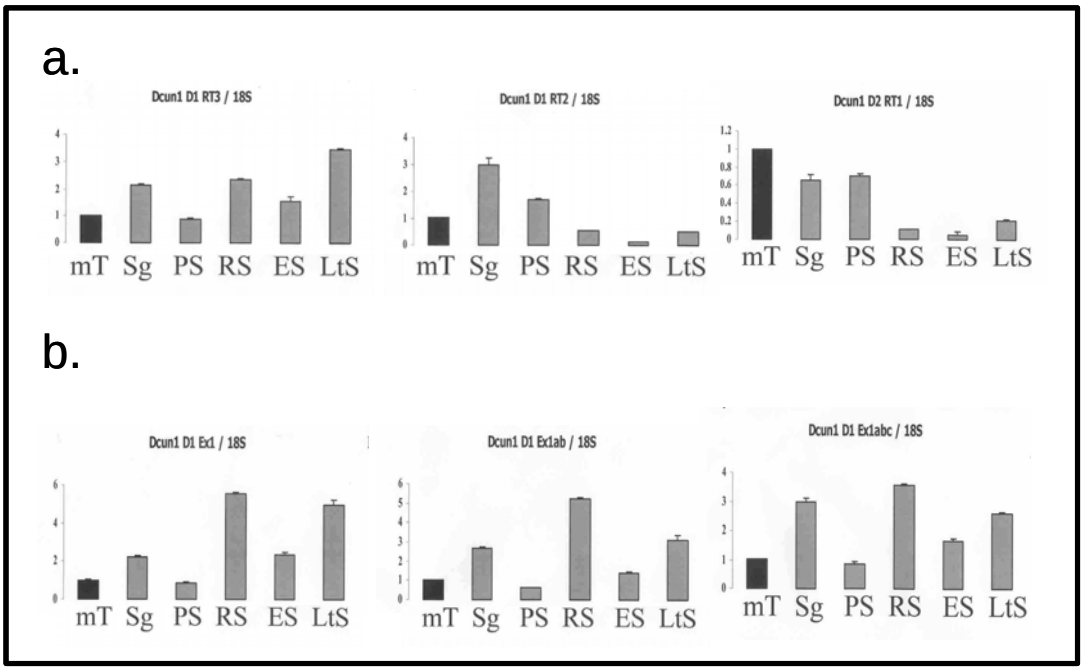


Figure 10. (a) Overall expression of transcripts containing the *Dcun1d1* coding sequence increase during male germ cell maturation, as measured by the D1RT3 primer pair. In contrast, expression of long transcripts of *Dcun1d1* decrease during germ cell maturation, as measured by the primer set D1RT2, located on the 3' UTR of the long Dcun1d1 transcript. *Dcun1d2* expression also decreases during spermatogenesis. **(b)** Expression of alternate first exons of *Dcun1d1* in elutriated mouse germ cells.
mT = whole mouse testes, Sg = spermatogonia, PS = pachytene spermatocytes, RS = round spermatids, ES = elongating spermatids, LtS = late spermatids. Expression of each gene is quantified relative to expression in whole testes. Error bars show SEM of fold-difference for three replicate wells.

Several interesting observations emerge from this data. First, levels of *Dcun1d1* transcripts as a whole (primer set D1RT3) are equal or higher in germ cells than in whole testis (recalling that whole testis expression of *Dcun1d1* is 50-100 fold higher than in other tissues), suggesting that germ cells specifically up-regulate *Dcun1d1*. While *Dcun1d1* transcripts as a whole increase during germ cell development, the level of transcripts containing the 3' UTR primer set D1RT2 show a steady decline, demonstrating that the alternate polyadenylation mechanisms which produce short *Dcun1d1* transcripts are up-regulated during spermatogenesis. Thus, the "testes-specific" short *Dcun1d1* transcripts noted on northern blot appear to be

enriched in maturing germ cells. Conversely, levels of *Dcun1D2*, which are very low in whole testes compared to brain and muscle tissue, decline even further in developing germ cells.

The patterns of expression of the three different first exons found in testis *Dcun1d1* transcripts are nearly identical across the stages of spermatogenesis. This argues that while the testes-specific alternate promoter is specifically turned on in germ cells, activity of the ubiquitous promoter also increases, resulting in a mix of the three alternate transcripts. There is evidence of developmentally precise regulation of the *Dcun1d1* gene across stages of spermatogenesis, with a sharp decrease in expression of all short *Dcun1d1* transcripts seen in the pachytene spermatocyte stage, and again in elongating spermatids. Elutriation of early germ cells in rats (not shown) also demonstrated distinct expression peaks in spermatogonia and round spermatids.

Discussion

Little is known about the significance and function of DCN-1-like proteins in vertebrates. To our knowledge, our group is the first to provide any functional data about the protein products of these genes in humans and mice, and four of the five genes characterized in this work have not been studied previously.

Core elements of the cullin-dependant ubiquitin ligase system are ancient in origin, and have been strongly conserved over more than a billion years of divergent evolution between plants and metazoans. Four of the five cullin classes (all but Cul5), Cand proteins, and the neddylation system are all present in *Arabidopsis.* We have shown that DCN-1-like proteins are similarly ancient, and thus are likely to be involved in basic processes common to most or all eukaryotic cells.

SCCRO expression levels are not a cell-cycle-specific influence on CRL activity

Because of the many cell-cycle specific proteins regulated by CRL complexes, and the importance of these pathways in cancer, regulation of CRL complexes in specific stages of the cell cycle has been of great interest in previous studies. Neither Cul-1 or Skp1, nor the interaction between the two, vary across stages of the cell cycle in HeLa cells.(119) However, cell-cycle specific variations are seen in the levels of some ubiquitin E2 enzymes as well as in Cul5, which is expressed in a cell-cycle dependant fashion in rat adrenal endothelial cells.(120)

In contrast, we found no evidence for cell-cycle stage-specific variations in SCCRO or SCCRO complex formation. Several caveats are necessary in interpreting these experiments. First, we have shown that our polyclonal SCCRO antibody cross-reacts equally with other Dcun1d proteins, including murine Dcun1d2, Dcun1d4, and Dcun1d5, so it is possible that variations in SCCRO expression might have been masked by constant expression of another Dcun1d protein of similar molecular weight (i.e. Dcun1d2). However, we have previously shown that RNAi constructs specific for SCCRO are capable of knocking down >80% of anti-SCCRO immunoreactivity at 30 kD in many epithelial cell lines (including HeLa), so

SCCRO is likely the major if not sole component of that band in HeLa-S. Second, HeLa-S cells are derived from an epithelial malignancy, and it is possible that SCCRO levels are abnormally regulated in this cell line. This is a reasonable concern, although expression of SCCRO in HeLa-S cells was not increased relative to other human suspension cell lines we tested. If future studies suggest a cell-cycle regulatory role for SCCRO, it might be useful to confirm these findings in cultured primary cells, or another non-transformed system.

Dcun1d genes diversified in metazoans to support cullin-related functions in specific subcellular compartments

Analysis of model organisms with completely sequenced genomes reveals that a single ancestral Dcun1d gene exists in plants, fungi, and many metazoans. This ancestral gene contains a UBA-like domain, which was shown by our group and others to facilitate ubiquitin binding. As noted by Kurz et al., this domain is lost in one of two isoforms of *C. elegans* DCN-1, and does not appear to be necessary for the neddylation-enhancing function of Dcun1d proteins, since the UBA-lacking proteins Dcun1d4 and Dcun1d5 enhance *in-vitro* neddylation reactions with roughly the same efficiently as the UBA-containing Dcun1d1. There is little positive evidence to date regarding what the function of the Dcun1d UBA domain might be.

The vertebrate Dcun1d4 and Dcun1d5 genes appear to have their origins in two gene duplication events. The first of these, and the replacement of the new gene's UBA domain with a basic nuclear localization signal, must have occurred in a common ancestor of arthropods and chordates, since one gene in this class is present in *Drosophila.* Because the murine UBA-Dcun1d genes are also present in the nucleus, and because at least one of the NLS-Dcun1d genes (Dcun1d5) does not posses any other novel domains in comparison to the ancestral Dcun1d genes, it is tempting to infer that these genes perform a restricted subset of the functions performed by the UBA-Dcun1d genes – those which take place in the nucleus and do not require certain UBA-mediated interactions. We did not create deletion mutants to directly test whether the basic N-terminal motif of Dcun1d4 and Dcun1d5 was responsible for nuclear

localization of these proteins. However, given its homology to known nuclear localization signals, and the fact that it is the only conserved motif in Dcun1d4 and Dcun1d5 which is not present in the UBA-Dcun1d genes, we believe this interpretation to be sound.

It will be important to know whether murine *NLS-Dcun1d* genes display the same transforming properties as *SCCRO*, since a positive answer would focus our oncogenesis research on nuclear functions of *SCCRO*, and could also prompt the investigation of a possible role for the *NLS-DCUN1D* genes in human cancer. A negative answer would suggest that the oncogenic activity of SCCRO protein requires cytoplasmic activity, nuclear-cytoplasmic shuttling, or the function of the UBA domain. Further experiments with chimeric *Dcun1d* genes containing both the UBA and NLS domains, or containing a nuclear export sequence, could help clarify these possibilities. Unfortunately, an initial attempt to test the transforming potential of *Dcun1d4* and *Dcun1d5* in human epithelial cells was technically unsuccessful.

Our experiments with GFP-tagged Dcun1d3 and the MYR sequence mutant Dcun1d3(G2A) strongly support the conclusion that this protein is membrane-bound via modification of its N-terminal, most likely with a hydrophobic myristic acid anchor. It is unclear if this property is shared with all members of this class, given the divergence of the *Drosophila dcun1d3* gene, suggesting that *Dcun1d3* orthologs may have different functions in different metazoan phyla. We are also less confident that the function of murine Dcun1d3 protein involves enhancement of cullin neddylation, since bacterially expressed Dcun1d3 did not enhance our in-vitro neddylation reaction. However, the DCN-1-like domain of Dcun1d3 is relatively well conserved, and this no doubt is responsible for its retained ability to bind CAND1 and cullins in GST pulldown. A number of important ubiquitination pathways are known to operate at cellular membranes, including ubiquitination of growth factor receptors on the plasma membrane, and ER quality-control mechanisms which degrade misfolded proteins after translocation to the cytoplamic side of the ER membrane, but to link any of these specific pathways to Dcun1d3 at this time would be speculative.

An important unresolved issue is whether or not Dcun1d3 binds to UBC12. Initial attempts to test this question via GST-pulldown were technically unsuccessful. Failure of Dcun1d3 to bind UBC12 could

explain its lack of neddylation activity, and could help in the identification of the UBC12 binding region in SCCRO. Deletion analyses have suggested that the UBC12 binding site on SCCRO, like the cullin binding site, is located in the C-terminal DUF298 domain. There are several residues in the highly conserved C-terminal domain which distinguish all Dcun1d3 proteins from SCCRO and other neddylation-competent Dcn1d proteins. These include glycine 134, valine 138, and isoleucine 140 (numbers refer to the SCCRO sequence), for which glutamate, serine, and phenylalanine are substituted in Dcun1d3 sequences. It would be interesting to study the binding and functional properties of SCCRO proteins mutated at one or more of these sites to contain the corresponding Dcun1d3 residues.

Our findings that the neddylation-enhancing activity of murine Dcun1d proteins is specific for CUL3 and, to a lesser extent CUL2, should be interpreted cautiously. As previously mentioned, *in vitro* studies of the neddylation pathway have been useful for determining the basic molecular mechanisms of proteins in this system, but do not fully capture the complexity of the system *in vivo.*

Tissue-specific expression patterns of Dcun1d genes are consistent with a role in neddylation

Despite the basic role that CRL's and neddylation pathways seem to play in most eukaryotic cells, there is significant evidence for specialization of function for many CRL's and for DCN-1-like proteins, in specific metazoan tissues and cell types. Expression levels of the six classic human cullins vary widely across developmental stages and tissue types. CUL1, CUL2, and CUL3 are more highly expressed in 7-day mouse embryos than in later stages, while the peak of CUL4b expression occurs at day 11. Relative to other tissues, increased expression of CUL1 and CUL3, and to a lesser extent CUL2, CUL4a, and CUL4b, are seen in the testes. CUL5 expression was not high in testes, but was elevated in heart and skeletal muscle, which also contained relatively high expression of CUL1, CUL2, CUL3, and CUL4a. Brain represented a peak in expression of CUL4b expression, while both CUL4a and CUL4b are high in spleen.

Patterns of Nedd8 expression are also highly variable across tissues and stages of development. Nedd8 expression is relatively high in early mouse embryos, peaking at about 11 days, before declining at later

fetal ages. In human tissues, Nedd8 expression is highest in the heart and skeletal muscle, with moderate expression in the brain. While Nedd8 itself shows only modestly increased expression in testes, the Nedd8 E1 and E2 enzymes have more pronounced expression peaks in that tissue, and are also high in skeletal muscle, heart, and to a lesser extent, brain.(80, 121)

There are several interesting point to consider when comparing our expression data for the murine *Dcun1d* genes to these previous findings. The first is that all tissues which highly express either *Dcun1d1* or *Dcun1d2* also show high expression of the neddylation pathway components, and conversely, neither *Dcun1d1* nor *Dcun1d2* is highly expressed in a tissue which is low in *Nedd8* or its E1 and E2 enzymes. This further supports our hypothesis that the primary functions of *Dcun1d1* and *Dcun1d2* involve the neddylation pathway. Furthermore, it suggests that certain tissues have a higher need for neddylation activity than others, and that the diversification of *Dcun1d* genes in more complex metazoans may be related to the need for more complex regulation of this system. Regarding the expression of the other three *Dcun1d* genes, they were all relatively highly expressed in the testes, and *Dcun1d4* was significantly higher in brain than in other non-testes tissues. With the exception of our germ cell elutriation findings, most of our data refers to gene expression at the level of tissues which contain a heterogeneous mix of cell types. Expression of *Dcn1d* genes in specific cell types, such as hematopoetic cells or endothelial cells, will have to be evaluated in future experiments.

Specific Dcun1d1 transcript expression patterns during spermatogenesis point to an expression level-dependant role

The concept that *Dcun1d1* specifically facilitates an increased need for neddylation activity in the testes is supported both by expression data and the phenotype of the *Dcun1d1*$^{-/-}$ mouse. In the present study, we confirmed previous findings of short *Dcun1d1* isoforms present only in testes, and extended those findings by showing that these short transcripts are specifically upregulated in germ cells. While truncation of the *Dcun1d1* 3' UTR does not chance its coding sequence, it may play a role in maintaining the high expression level of those transcripts in germ cells, since many mRNA's have been shown to possess

sequences in their 3' UTR which enhance transcript degradation through interaction with RNA binding proteins and microRNA's (122). The existence of testes-specific transcript processing mechanisms is well established, and the testis-specific poly-A polymerase (TPAP) is required for normal spermatogenesis.(123)

The testes-specific alternate promoter and first exon for *Dcun1d1* likely play an important functional role, given their phylogenic conservation. However, the protein product of this alternate transcript is hard to predict, since the AUG initiation codon of the ubiquitous *SCCRO* transcript is replaced by a conserved CUG encoded on exon 1ab. Exon 1ab does contain an upstream AUG in reasonable Kozak context, but it is out of frame with the SCCRO protein sequence and lacks a coding sequence of significant length. An in-frame AUG codon is present downstream in the *Dcun1d1* coding sequence, but it is in very poor Kozak context, and would also require ribosomal leakage over the upstream ATG or re-initiation for translation. A more unusual, but intriguing possibility for translation of *Dcun1d1* exon 1ab-containing transcripts would be through the use of a non-canonical initiation codon, such as the previously mentioned conserved CUG. Initiation of translation at non-AUG codons has been reported in mammalian germ cells for several genes.(124-126) For example, CUG initiation codons are utilized for translation of several FGF-2 isoforms in the testes of mice, rats and humans via a conserved internal ribosomal entry sequence (IRES) at the 5' end of the transcript.(127, 128) Translation at the FGF-2 IRES seems to be activated in mouse germ cells in response to rising testosterone levels in adulthood,(129), which correlates with the temporal pattern of expression for short *Dcun1d1* transcripts in mouse testes.(18)

Endogenous eukaryotic IRES sequences are poorly understood, and reliable prediction algorithms for IRES sequences are currently lacking(130). It would be interesting to test the sequence of *Dcun1d1* exon 1ab for IRES activity by cloning it into an appropriate bicistronic construct and expressing the construct in germ cells. Previously we discussed evidence that the initial portion of the predicted *Dcun1d4* coding sequence is not efficiently translated. Since that gene is also most highly expressed in the testes, it might also be fruitful to evaluate *Dcun1d4* transcripts for the presence of an IRES or other non-canonical means of translation.

Our in-vitro studies provide evidence that two classes of DCN-1-like proteins show a neddylation-enhancing activity which is relatively specific for CUL3, a result which is consistent with the original identification of *C. elegans dcn-1* through its effects on CUL-3–dependant degradation of the microtubule–severing protein MEI-1. Our results suggest that the UBA-Dcun1d proteins may be significant regulators of Cul3–dependent degradation pathways in mammals. The defects in embryogenesis observed in *C. elegans* neddylation pathway mutants appear to occur because MEI-1 accumulates during a meiotic cycle which occurs in the *C. elegans* embryo prior to the first mitotic division. The failure of a CUL-3 CRL to degrade MEI-1 after meiosis is complete results in persistence of MEI-1 into mitosis, which interferes with microtubule polymerization and proper formation of the mitotic spindle. It is therefore interesting to note that spermatogenesis is also a process in which meiosis is followed by a complex, microtubule-dependent event - the assembly of the flagellum and other structures of the mature sperm. The few sperm produced by *Dcun1d1*$^{-/-}$ mice show gross morphological abnormalities, suggesting that disregulation of spermatid structural proteins results from the absence of *Dcun1d1*.

Our work to define the function of *Dcun1d1* in spermatogenesis is still ongoing, but it is clear from the present data that Dcn1-domain containing genes play a highly regulated and vital role in spermatogenesis. Spermatogenesis is difficult to model in cell culture or other easily manipulated systems, so in some regards it is not an ideal candidate system for studying the endogenous function of dcn1 domain-containing proteins. However, one potential member of a specific Cul3 complex, the BTB protein KLHL10(70), is required for spermatogenesis, so this could be an interesting complex to study for modulation of CRL activity by changes in SCCRO expression. Abnormal expression of germ cell-specific proteins and pathways are thought to be important in certain cancers, and the very high levels of SCCRO expression in testes may allow for specific functions which are similar to those occurring in SCCRO overexpressing tumors.

Gene duplications creating Dcun1d2 and Cand2 may have co-evolved to support a muscle-specific, neddylation-dependant function in vertebrates.

In contrast to *Dcun1d1*, expression patterns of *Dcun1d2* suggest that it plays a role specifically in the regulation of the neddylation system in muscle, and possibly nervous tissue. Intriguingly, the tissue-specific expression pattern of *Dcun1d1* closely parallels that of *Cand1*, while *Cand2*, a close homolog present in birds and mammals, is expressed specifically in heart and skeletal muscle.(131) These parallels support the notion that UBA-Dcn1d protein function is closely linked to that of Cand proteins, and that *Dcun1d2* and *Cand2* may have coevolved to support a specialized neddylation-related function in muscle tissue. Murine *Cand2* is specifically upregulated during differentiation of C2C12 myoblasts into myotubules *in vitro*(132). If *Dcun1d2* shows a similar expression pattern, myoblast differentiation might represent an excellent model system in which to explore the effect of varying expression of Cand proteins and UBA-Dcun1d proteins on cullin neddylation and CRL activity in vertebrates.

While this manuscript was undergoing final revision, Shirashi et al. (133) published exciting new data about the role of *Cand2* and cullin neddylation in the regulation of myogenesis. They focused on the effects of *Cand2* expression on the protein levels of myogenin, a transcription factor which is rapidly degraded in undifferentiated myoblasts, but which accumulates along with increasing *Cand2* expression during myoblast differentiation. The authors provide evidence that ubiquitination of myogenin is dependant on the activity of CUL1-containing complexes. Interestingly, transgenic manipulation of the rise of *Cand2* expression in differentiating myoblasts has a direct and logical effect on myogenin levels and the rate of cell differentiation into myotubules. When *Cand2* expression is suppressed with RNAi, myogenin levels fail to rise and differentiation is delayed, while overexpression of *Cand2* conversely stabilizes myogenin and accelerates differentiation, an *in vivo* recapitulation of the simple inhibitory effect of Cand proteins on CRL activity *in vitro*. This finding is exciting because, as previously noted, *in vivo* manipulations of neddylation pathway regulators such as *Cand1* and *Csn5* in other cell types have either had modest effects, or resulted in "paradoxical" inhibition of CRL activity following either RNAi knockdown or overexpression of the target gene.

One possible unifying hypothesis from these observations is that the neddylation pathway and its protein regulators, such as Csn5, Cand proteins, and Dcun1d proteins, serve a "housekeeping" role in most

vertebrate cells, with constant expression levels which facilitate CRL activity but do not dynamically regulate ubiquitination in response to cellular events. Instead, this regulation is carried out by specific, dynamically expressed CRL components, such as F-box proteins or other substrate-binding proteins. In specific contexts, however, such as muscle-lineage cells, or germ cells, the neddylation pathway may become a "limiting factor" for the activity of some CRL complexes. Thus, alternate neddylation-pathway regulating genes with dynamically regulated expression (such as *Cand2*, *Dcun1d2*, or testes-specific alternate *Dcun1d1* transcripts) may have evolved to regulate CRL activity in those cells. The current work provides initial but promising leads for the exploration of this hypothesis, which must by confirmed by functional studies *in vivo*. If further work confirms that *Dcun1d2* is specifically upregulated during myogenesis, it would be interesting to explore whether it specifically enhances Cul3 neddylation (in keeping with our *in vitro* data), while allowing the activity of Cul1 and Cul2 – containing complexes to be suppressed by Cand2. Such a phenomenon would have logical appeal, since many of the known functions of Cul1 and Cul2 complexes involve cell-cycle regulation, which would be unnecessary in terminally differentiated myocytes, while two of the best-characterized functions of Cul3 involve regulation of microtubule polymerization and modulation of hedgehog pathway signaling, both of which are critical to muscle cell differentiation and maintenance. The effect of Dcun1d1 proteins on the hedgehog pathway is of particular interest to us, since our previous work showed that *SCCRO*-transformed cells show highly significant alterations in the expression of several hedgehog pathway genes.

The oncogenic role of Dcun1d genes – future investigations

Our original interest in DCN-1-like proteins arose from evidence that DCUN1D1/SCCRO is a gene-amplified, overexpressed oncogene in human squamous-cell cancers of mucosal origin. While cancer was not explicitly the subject of the current study, our discovery of four additional human genes with many similar molecular properties raises the question of whether they might also be candidate oncogenes. Further investigation of this possibility is warranted.

Returning to the mechanism of SCCRO-mediated oncogenesis, our hypothesis that *dcn-1*-like genes serve a "housekeeping" role in most cells and a more specific, expression-level-dependant function in others, raises an interesting question: do tumor cells represent systems where the ubiquitination of critical factors is controlled by the expression levels of neddylation pathway proteins? Frequent overexpression of another neddylation pathway regulator, CSN5/JAB1, in several human cancers, suggests that this may be the case.(134-137) Since some expression-regulated CRL components (e.g. Skp2) are stably overexpressed in the same tumor types, common mechanisms of CRL regulation may be defective, making neddylation activity a "limiting factor" in CRL activation and cancer maintenance. Such derangements may identify a class of cancer patients who might benefit from trials of existing drugs, such as the proteasome inhibitor bortezomib, or future drugs which specifically target the neddylation pathway.

X Appendices

A: Unique identifiers for *Mus musculus* cDNA clones used in this study

	RIKEN FANTOM3 ID	Genbank transcript ID	Genbank protein ID
Dcun1d1:	F630119O03	AK155081	BAE33033.1
Dcun1d2:	F730211A17	AK155726	BAE33403.1
Dcun1d3:	F630204O15	AK155160	BAE33084.1
Dcun1d4:	9030622J04	AK033559	BAC28359.1
Dcun1d5:	3110001A18	AK013933	BAB29066.1

B: Genbank identifiers for other representative Dcn1-domain containing genes[2]

	Genbank transcript ID	Genbank EST ID	Genbank protein ID
Homo sapiens			
DCUN1D1/SCCRO	NM_020640		NP_065691.2
DCUN1D1(Ex1ab)		DB095571	
DCUN1D2	NM_001014283		NP_001014305.1
DCUN1D3	NM_173475		NP_775746.1
DCUN1D4	NM_001040402.1		NP_001035492.1
DCUN1D5	NM_032299		NP_115675.1
Mus musculus			
Dcun1d1(Ex1ab)		BU946323	
Dcun1d1(Ex1abc)		CN842488	
Macaca fascicularis			
Dcun1d1(Ex1ab)	AB169138.1		BAE01231.1
Bos taurus			
Dcun1d1(Ex1ab)		BC102675.1	
Canis familiaris			
Dcun1d1(Ex1ab)	XM_843562.1		XP_848655.1
Danio rerio			
dcun1d1	BC068381		AAH68381.1
dcun1d2	NM_205538		NP_991101.1
dcun1d4	NM_001032366		NP_001027538.1
dcun1d5	NM_213331		NP_998496.1
Tetraodon nigroviridis			
dcun1d3			CAG04390
Strongylocentrus purpuratus			
Dcun1d3	XM_001192744		XP_001192744.1
Drosophila melanogaster			
dcun1d1l	NM_140520		NP_648777.1
dcun1d3l	NM_165963		NP_725244.1
dcun1d4l	BT012470		AAS93741.1

[2] Gene symbols are those used in Figure 3 and throughout this manuscript, but may not correspond to current NCBI nomenclature.

	Genbank transcript ID	Genbank protein ID
Caenorhabditis elegans		
dcn-1	NM_065465	NP_497866.2
Saccharomyces cerevisiae		
dcn1	Z73300	Q12395
Dictyostelium discoideum		
Dcn1L1	XM_630829	XP_635921.1
Dcn1L2	XM_633496	XP_638588.1
Dcn1L3	XM_640290	XP_645382.1
Arabidopsis thaliana		
Dcn1F1	NM_112112	NP_566436.1
Dcn1F2	NM_101454	NP_563983.1

XI References

1. Oster S, Penn L, Stambolic V. Oncogenes and tumor supressor genes. In: Tannock IF, Hill RP, Bristow RG, Harrington L, editors. The Basic Science of Oncology. 4th ed. United States of America: McGraw-Hill; 2005. p. 123.

2. Albertson DG. Gene amplification in cancer. Trends Genet 2006 Aug;22(8):447-55.

3. Riazimand SH, Welkoborsky HJ, Bernauer HS, Jacob R, Mann WJ. Investigations for fine mapping of amplifications in chromosome 3q26.3-28 frequently occurring in squamous cell carcinomas of the head and neck. Oncology 2002;63(4):385-92.

4. Petersen I, Bujard M, Petersen S, et al. Patterns of chromosomal imbalances in adenocarcinoma and squamous cell carcinoma of the lung. Cancer Res 1997;57(12):2331-5.

5. Heselmeyer K, Macville M, Schrock E, et al. Advanced-stage cervical carcinomas are defined by a recurrent pattern of chromosomal aberrations revealing high genetic instability and a consistent gain of chromosome arm 3q. Genes Chromosomes Cancer 1997;19(4):233-40.

6. Heselmeyer K, Schrock E, du Manoir S, et al. Gain of chromosome 3q defines the transition from severe dysplasia to invasive carcinoma of the uterine cervix. Proc Natl Acad Sci U S A 1996;93(1):479-84.

7. Shayesteh L, Lu Y, Kuo WL, et al. PIK3CA is implicated as an oncogene in ovarian cancer. Nat Genet 1999;21(1):99-102.

8. Singh B, Reddy PG, Goberdhan A, et al. p53 regulates cell survival by inhibiting PIK3CA in squamous cell carcinomas. Genes Dev 2002;16(8):984-93.

9. Regala RP, Weems C, Jamieson L, et al. Atypical protein kinase C{iota} is an oncogene in human non-small cell lung cancer. Cancer Res 2005;65(19):8905-11.

10. Eder AM, Sui X, Rosen DG, et al. Atypical PKC{iota} contributes to poor prognosis through loss of apical-basal polarity and cyclin E overexpression in ovarian cancer. PNAS 2005;102(35):12519-24.

11. Kanao H, Enomoto T, Kimura T, et al. Overexpression of LAMP3/TSC403/DC-LAMP promotes metastasis in uterine cervical cancer. Cancer Res 2005;65(19):8640-5.

12. Brass N, Heckel D, Sahin U, Pfreundschuh M, Sybrecht GW, Meese E. Translation initiation factor eIF-4gamma is encoded by an amplified gene and induces an immune response in squamous cell lung carcinoma. Hum Mol Genet 1997;6(1):33-9.

13. Darai-Ramqvist E, de Stahl TD, Sandlund A, et al. Array-CGH and multipoint FISH to decode complex chromosomal rearrangements. BMC Genomics 2006;7:330.

14. Sarkaria I, O-Charoenrat P, Talbot SG, et al. Squamous cell carcinoma related Oncogene/DCUN1D1 is highly conserved and activated by amplification in squamous cell carcinomas. Cancer Res 2006;66(19):9437-44.

15. Singh B, Gogineni SK, Sacks PG, et al. Molecular cytogenetic characterization of head and neck squamous cell carcinoma and refinement of 3q amplification. Cancer Res 2001;61(11):4506-13.

16. Zhao X, Weir BA, LaFramboise T, et al. Homozygous deletions and chromosome amplifications in human lung carcinomas revealed by single nucleotide polymorphism array analysis. Cancer Res 2005;65(13):5561-70.

17. Sarkaria IS, Pham D, Ghossein RA, et al. SCCRO expression correlates with invasive progression in bronchioloalveolar carcinoma. Ann Thorac Surg 2004;78(5):1734-41.

18. Pourcel C, Jaubert J, Hadchouel M, Wu X, Schweizer J. A new family of genes and pseudogenes potentially expressing testis- and brain-specific leucine zipper proteins in man and mouse. Gene 2000 May;249(1-2):105-13.

19. Mas C, Bourgeois F, Bulfone A, Levacher B, Mugnier C, Simonneau M. Cloning and expression analysis of a novel gene, RP42, mapping to an autism susceptibility locus on 6q16. Genomics 2000;65(1):70-4.

20. Kaufman A, Singh B. Unpublished observations.

21. Kim A, Lee B, Singh B. Unpublished observations.

22. von Arnim AG. A hitchhiker's guide to the proteasome. Sci STKE 2001;2001(97):PE2.

23. Zheng N, Schulman BA, Song L, et al. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. Nature 2002;416(6882):703-9.

24. Angers S, Li T, Yi X, MacCoss MJ, Moon RT, Zheng N. Molecular architecture and assembly of the DDB1-CUL4A ubiquitin ligase machinery. Nature 2006;443(7111):590-3.

25. Willems AR, Schwab M, Tyers M. A hitchhiker's guide to the cullin ubiquitin ligases: SCF and its kin. Biochim Biophys Acta 2004;1695(1-3):133-70.

26. Stebbins CE, Kaelin WG,Jr, Pavletich NP. Structure of the VHL-ElonginC-ElonginB complex: Implications for VHL tumor suppressor function. Science 1999;284(5413):455-61.

27. Seol JH, Feldman RM, Zachariae W, et al. Cdc53/cullin and the essential Hrt1 RING-H2 subunit of SCF define a ubiquitin ligase module that activates the E2 enzyme Cdc34. Genes Dev 1999;13(12):1614-26.

28. Lyapina SA, Correll CC, Kipreos ET, Deshaies RJ. Human CUL1 forms an evolutionarily conserved ubiquitin ligase complex (SCF) with SKP1 and an F-box protein. Proc Natl Acad Sci U S A 1998;95(13):7451-6.

29. Winston JT, Koepp DM, Zhu C, Elledge SJ, Harper JW. A family of mammalian F-box proteins. Curr Biol 1999;9(20):1180-2.

30. Jin J, Cardozo T, Lovering RC, Elledge SJ, Pagano M, Harper JW. Systematic analysis and nomenclature of mammalian F-box proteins. Genes Dev 2004;18(21):2573-80.

31. Guardavaccaro D, Pagano M. Oncogenic aberrations of cullin-dependent ubiquitin ligases. Oncogene 2004;23(11):2037-49.

32. Tsvetkov LM, Yeh KH, Lee SJ, Sun H, Zhang H. p27(Kip1) ubiquitination and degradation is regulated by the SCF(Skp2) complex through phosphorylated Thr187 in p27. Curr Biol 1999;9(12):661-4.

33. Kamura T, Hara T, Kotoshiba S, et al. Degradation of p57Kip2 mediated by SCFSkp2-dependent ubiquitylation. Proc Natl Acad Sci U S A 2003;100(18):10231-6.

34. Yu ZK, Gervais JL, Zhang H. Human CUL-1 associates with the SKP1/SKP2 complex and regulates p21(CIP1/WAF1) and cyclin D proteins. Proc Natl Acad Sci U S A 1998;95(19):11324-9.

35. Tedesco D, Lukas J, Reed SI. The pRb-related protein p130 is regulated by phosphorylation-dependent proteolysis via the protein-ubiquitin ligase SCF(Skp2). Genes Dev 2002;16(22):2946-57.

36. Kudo Y, Kitajima S, Sato S, Miyauchi M, Ogawa I, Takata T. High expression of S-phase kinase-interacting protein 2, human F-box protein, correlates with poor prognosis in oral squamous cell carcinomas. Cancer Res 2001;61(19):7044-7.

37. Gstaiger M, Jordan R, Lim M, et al. Skp2 is oncogenic and overexpressed in human cancers. Proc Natl Acad Sci U S A 2001;98(9):5043-8.

38. Yokoi S, Yasui K, Mori M, Iizasa T, Fujisawa T, Inazawa J. Amplification and overexpression of SKP2 are associated with metastasis of non-small-cell lung cancers to lymph nodes. Am J Pathol 2004;165(1):175-80.

39. Carrano AC, Pagano M. Role of the F-box protein Skp2 in adhesion-dependent cell cycle progression. J Cell Biol 2001;153(7):1381-90.

40. Latres E, Chiarle R, Schulman BA, et al. Role of the F-box protein Skp2 in lymphomagenesis. Proc Natl Acad Sci U S A 2001;98(5):2515-20.

41. Shim EH, Johnson L, Noh HL, et al. Expression of the F-box protein SKP2 induces hyperplasia, dysplasia, and low-grade carcinoma in the mouse prostate. Cancer Res 2003;63(7):1583-8.

42. Koepp DM, Schaefer LK, Ye X, et al. Phosphorylation-dependent ubiquitination of cyclin E by the SCFFbw7 ubiquitin ligase. Science 2001;294(5540):173-7.

43. Tsunematsu R, Nakayama K, Oike Y, et al. Mouse Fbw7/Sel-10/Cdc4 is required for notch degradation during vascular development. J Biol Chem 2004;279(10):9417-23.

44. Welcker M, Orian A, Jin J, et al. The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-myc protein degradation. Proc Natl Acad Sci U S A 2004;101(24):9085-90.

45. Yada M, Hatakeyama S, Kamura T, et al. Phosphorylation-dependent degradation of c-myc is mediated by the F-box protein Fbw7. EMBO J 2004;23(10):2116-25.

46. Rajagopalan H, Jallepalli PV, Rago C, et al. Inactivation of hCDC4 can cause chromosomal instability. Nature 2004;428(6978):77-81.

47. Spruck CH, Strohmaier H, Sangfelt O, et al. hCDC4 gene mutations in endometrial cancer. Cancer Res 2002;62(16):4535-9.

48. Calhoun ES, Jones JB, Ashfaq R, et al. BRAF and FBXW7 (CDC4, FBW7, AGO, SEL10) mutations in distinct subsets of pancreatic cancer: Potential therapeutic targets. Am J Pathol 2003;163(4):1255-60.

49. Mao JH, Perez-Losada J, Wu D, et al. Fbxw7/Cdc4 is a p53-dependent, haploinsufficient tumour suppressor gene. Nature 2004;432(7018):775-9.

50. Fuchs SY, Chen A, Xiong Y, Pan ZQ, Ronai Z. HOS, a human homolog of slimb, forms an SCF complex with Skp1 and Cullin1 and targets the phosphorylation-dependent degradation of IkappaB and beta-catenin. Oncogene 1999;18(12):2039-46.

51. Latres E, Chiaur DS, Pagano M. The human F box protein beta-trcp associates with the Cul1/Skp1 complex and regulates the stability of beta-catenin. Oncogene 1999;18(4):849-54.

52. Guardavaccaro D, Kudo Y, Boulaire J, et al. Control of meiotic and mitotic progression by the F box protein beta-Trcp1 in vivo.see comment. Dev Cell 2003;4(6):799-812.

53. Hatakeyama S, Kitagawa M, Nakayama K, et al. Ubiquitin-dependent degradation of IkappaBalpha is mediated by a ubiquitin ligase Skp1/Cul 1/F-box protein FWD1. Proc Natl Acad Sci U S A 1999;96(7):3859-63.

54. Tan P, Fuchs SY, Chen A, et al. Recruitment of a ROC1-CUL1 ubiquitin ligase by Skp1 and HOS to catalyze the ubiquitination of I kappa B alpha. Mol Cell 1999;3(4):527-33.

55. Bondar T, Kalinina A, Khair L, et al. Cul4A and DDB1 associate with Skp2 to target p27Kip1 for proteolysis involving the COP9 signalosome. Mol Cell Biol 2006;26(7):2531-9.

56. Hu J, McCall CM, Ohta T, Xiong Y. Targeted ubiquitination of CDT1 by the DDB1-CUL4A-ROC1 ligase in response to DNA damage. Nat Cell Biol 2004;6(10):1003-9.

57. Jin J, Arias EE, Chen J, Harper JW, Walter JC. A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. Mol Cell 2006;23(5):709-21.

58. Groisman R, Polanowska J, Kuraoka I, et al. The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. Cell 2003;113(3):357-67.

59. Chen LC, Manjeshwar S, Lu Y, et al. The human homologue for the caenorhabditis elegans cul-4 gene is amplified and overexpressed in primary breast cancers. Cancer Res 1998;58(16):3677-83.

60. Yasui K, Arii S, Zhao C, et al. TFDP1, CUL4A, and CDC16 identified as targets for amplification at 13q34 in hepatocellular carcinomas. Hepatology 2002;35(6):1476-84.

61. Maxwell PH, Wiesener MS, Chang GW, et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature 1999;399(6733):271-5.

62. Lonergan KM, Iliopoulos O, Ohh M, et al. Regulation of hypoxia-inducible mRNAs by the von hippel-lindau tumor suppressor protein requires binding to complexes containing elongins B/C and Cul2. Mol Cell Biol 1998;18(2):732-41.

63. Kamura T, Sato S, Iwai K, Czyzyk-Krzeska M, Conaway RC, Conaway JW. Activation of HIF1alpha ubiquitination by a reconstituted von hippel-lindau (VHL) tumor suppressor complex. Proc Natl Acad Sci U S A 2000;97(19):10430-5.

64. Pantuck AJ, Zeng G, Belldegrun AS, Figlin RA. Pathobiology, prognosis, and targeted therapy for renal cell carcinoma: Exploiting the hypoxia-induced pathway. Clin Cancer Res 2003;9(13):4641-52.

65. Kamura T, Maenaka K, Kotoshiba S, et al. VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. Genes Dev 2004;18(24):3055-65.

66. Kohroki J, Nishiyama T, Nakamura T, Masuho Y. ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes. FEBS Lett 2005;579(30):6796-802.

67. Xu L, Wei Y, Reboul J, et al. BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. Nature 2003;425(6955):316-21.

68. Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG. Sequence and structural analysis of BTB domain proteins. Genome Biol 2005;6(10):R82.

69. Pintard L, Willis JH, Willems A, et al. The BTB protein MEL-26 is a substrate-specific adaptor of the CUL-3 ubiquitin-ligase. Nature 2003;425(6955):311-6.

70. Wang S, Zheng H, Esaki Y, Kelly F, Yan W. Cullin3 is a KLHL10-interacting protein preferentially expressed during late spermiogenesis. Biol Reprod 2006;74(1):102-8.

71. Kent D, Bush EW, Hooper JE. Roadkill attenuates hedgehog responses through degradation of cubitus interruptus. Development 2006;133(10):2001-10.

72. Zhang Q, Zhang L, Wang B, Ou CY, Chien CT, Jiang J. A hedgehog-induced BTB protein modulates hedgehog signaling by degrading Ci/Gli transcription factor. Dev Cell 2006;10(6):719-29.

73. Ou CY, Lin YF, Chen YJ, Chien CT. Distinct protein degradation mechanisms mediated by Cul1 and Cul3 controlling ci stability in drosophila eye development.see comment. Genes Dev 2002;16(18):2403-14.

74. Angers S, Thorpe CJ, Biechele TL, et al. The KLHL12-cullin-3 ubiquitin ligase negatively regulates the wnt-beta-catenin pathway by targeting dishevelled for degradation. Nat Cell Biol 2006;8(4):348-57.

75. Skaar JR, Florens L and Tsutsumi T, et al. PARC and CUL7 Form Atypical Cullin RING Ligase Complexes. Cancer Res 2007;67(5):2006-14.

76. Tsunematsu R, Nishiyama M, Kotoshiba S, Saiga T, Kamura T and Nakayama KI. Fbxw8 is essential for Cul1-Cul7 complex formation and for placental development. Mol Cell Biol 2006;26(16):6157-69.

77. Kaustov L, Lukin J and Lemak A, et al. The conserved CPH domains of Cul7 and PARC are protein-protein interaction modules that bind the tetramerization domain of P53. J Biol Chem 2007.

78. Liakopoulos D, Doenges G, Matuschewski K, Jentsch S. A novel protein modification pathway related to the ubiquitin system. EMBO J 1998;17(8):2208-14.

79. Lammer D, Mathias N, Laplaza JM, et al. Modification of yeast Cdc53p by the ubiquitin-related protein rub1p affects function of the SCFCdc4 complex. Genes Dev 1998;12(7):914-26.

80. Hori T, Osaka F, Chiba T, et al. Covalent modification of all members of human cullin family proteins by NEDD8. Oncogene 1999;18(48):6829-34.

81. Gong L, Yeh ET. Identification of the activating and conjugating enzymes of the NEDD8 conjugation pathway. J Biol Chem 1999;274(17):12036-42.

82. Osaka F, Saeki M, Katayama S, et al. Covalent modifier NEDD8 is essential for SCF ubiquitin-ligase in fission yeast. EMBO J 2000;19(13):3475-84.

83. Jones D, Crowe E, Stevens TA, Candido EP. Functional and phylogenetic analysis of the ubiquitylation system in caenorhabditis elegans: Ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. Genome Biol 2002;3(1):RESEARH0002.

84. Kurz T, Pintard L, Willis JH, et al. Cytoskeletal regulation by the Nedd8 ubiquitin-like protein modification pathway. Science 2002;295(5558):1294-8.

85. Tateishi K, Omata M, Tanaka K, Chiba T. The NEDD8 system is essential for cell cycle progression and morphogenetic pathway in mice. J Cell Biol 2001;155(4):571-9.

86. Podust VN, Brownell JE, Gladysheva TB, et al. A Nedd8 conjugation pathway is essential for proteolytic targeting of p27Kip1 by ubiquitination. Proc Natl Acad Sci U S A 2000;97(9):4579-84.

87. Amir RE, Iwai K, Ciechanover A. The NEDD8 pathway is essential for SCF(beta -TrCP)-mediated ubiquitination and processing of the NF-kappa B precursor p105. J Biol Chem 2002;277(26):23253-9.

88. Ohh M, Kim WY, Moslehi JJ, et al. An intact NEDD8 pathway is required for cullin-dependent ubiquitylation in mammalian cells. EMBO Rep 2002;3(2):177-82.

89. Liu J, Furukawa M, Matsumoto T, Xiong Y. NEDD8 modification of CUL1 dissociates p120(CAND1), an inhibitor of CUL1-SKP1 binding and SCF ligases. Mol Cell 2002;10(6):1511-8.

90. Zheng J, Yang X, Harrell JM, et al. CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. Mol Cell 2002;10(6):1519-26.

91. Goldenberg SJ, Cascio TC, Shumway SD, et al. Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. Cell 2004;119(4):517-28.

92. Kawakami T, Chiba T, Suzuki T, et al. NEDD8 recruits E2-ubiquitin to SCF E3 ligase. EMBO J 2001;20(15):4003-12.

93. Wimuttisuk W, Singer JD. The Cullin3 ubiquitin ligase functions as a Nedd8-bound heterodimer. Mol Biol Cell 2007;18(3):899-909

94. Gong L, Kamitani T, Millas S, Yeh ET. Identification of a novel isopeptidase with dual specificity for ubiquitin- and NEDD8-conjugated proteins. J Biol Chem 2000;275(19):14212-6.

95. Mendoza HM, Shen LN, Botting C, et al. NEDP1, a highly conserved cysteine protease that deNEDDylates cullins. J Biol Chem 2003;278(28):25637-43.

96. Lyapina S, Cope G, Shevchenko A, et al. Promotion of NEDD-CUL1 conjugate cleavage by COP9 signalosome. Science 2001;292(5520):1382-5.

97. Cope GA, Suh GS, Aravind L, et al. Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1.see comment. Science 2002;298(5593):608-11.

98. Schwechheimer C, Serino G, Callis J, et al. Interactions of the COP9 signalosome with the E3 ubiquitin ligase SCFTIRI in mediating auxin response. Science 2001;292(5520):1379-82.

99. Yang X, Menon S, Lykke-Andersen K, et al. The COP9 signalosome inhibits p27(kip1) degradation and impedes G1-S phase progression via deneddylation of SCF Cul1. Curr Biol 2002;12(8):667-72.

100. Denti S, Fernandez-Sanchez ME, Rogge L, Bianchi E. The COP9 signalosome regulates Skp2 levels and proliferation of human cells. J Biol Chem 2006;281(43):32188-96.

101. Wei N, Deng XW. The COP9 signalosome. Annu Rev Cell Dev Biol 2003;19:261-86.

102. Wirbelauer C, Sutterluty H, Blondel M, et al. The F-box protein Skp2 is a ubiquitylation target of a Cul1-based core ubiquitin ligase complex: Evidence for a role of Cul1 in the suppression of Skp2 expression in quiescent fibroblasts. EMBO J 2000;19(20):5362-75.

103. Kurz T, Ozlu N, Rudolf F, et al. The conserved protein DCN-1/Dcn1p is required for cullin neddylation in C. elegans and S. cerevisiae. Nature 2005;435(7046):1257-61.

104. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 1997;25(24):4876-82.

105. Rost B, Yachdav G, Liu J. The PredictProtein server. Nucleic Acids Res 2004 Jul 1;32(Web Server issue):W321-6.

106. Maurer-Stroh S, Eisenhaber B, Eisenhaber F. N-terminal N-myristoylation of proteins: Refinement of the sequence motif and its taxon-specific differences. J Mol Biol 2002;317(4):523-40.

107. Carninci P, Kasukawa T, Katayama S, et al. The transcriptional landscape of the mammalian genome. Science 2005;309(5740):1559-63.

108. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 2000;132:365-86.

109. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29(9):e45.

110. Meistrich ML, Longtin J, Brock WA, Grimes SR,Jr, Mace ML. Purification of rat spermatogenic cells and preliminary biochemical analysis of these cells. Biol Reprod 1981;25(5):1065-77.

111. Fujisawa M, Bardin CW, Morris PL. A germ cell factor(s) modulates preproenkephalin gene expression in rat sertoli cells. Mol Cell Endocrinol 1992;84(1-2):79-88.

112. Jenab S, Morris PL. Testicular leukemia inhibitory factor (LIF) and LIF receptor mediate phosphorylation of signal transducers and activators of transcription (STAT)-3 and STAT-1 and induce c-fos transcription and activator protein-1 activation in rat sertoli but not germ cells. Endocrinology 1998 Apr;139(4):1883-90.

113. Liu Q, Shang F, Guo W, et al. Regulation of the ubiquitin proteasome pathway in human lens epithelial cells during the cell cycle. Exp Eye Res 2004;78(2):197-205.

114. Pines J, Hunter T. Human cyclin A is adenovirus E1A-associated protein p60 and behaves differently from cyclin B. Nature 1990;346(6286):760-3.

115. Li Y, Jenkins CW, Nichols MA, Xiong Y. Cell cycle expression and p53 regulation of the cyclin-dependent kinase inhibitor p21. Oncogene 1994;9(8):2261-8.

116. Koonin EV. Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet 2005;39:309-38.

117. O'Brien EA, Koski LB, Zhang Y, et al. TBestDB: A taxonomically broad database of expressed sequence tags (ESTs). Nucleic Acids Res 2007;35(Database issue):D445-51.

118. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 1987;15(20):8125-48.

119. Michel JJ, Xiong Y. Human CUL-1, but not other cullin family members, selectively interacts with SKP1 to form a complex with SKP2 and cyclin A. Cell Growth Differ 1998;9(6):435-49.

120. Burnatowska-Hledin, Zeneberg A, Roulo A, et al. Expression of VACM-1 protein in cultured rat adrenal endothelial cells is linked to the cell cycle. Endothelium 2001;8(1):49-63.

121. Kamitani T, Kito K, Nguyen HP, Yeh ET. Characterization of NEDD8, a developmentally down-regulated ubiquitin-like protein. J Biol Chem 1997;272(45):28557-62.

122. Wilkie GS, Dickson, KS, Gray, NK. Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. Trends Biochem Sci;28(4)182-8

123. Kashiwabara S, Noguchi J, Zhuang T, et al. Regulation of spermatogenesis by testis-specific, cytoplasmic poly(A) polymerase TPAP. Science 2002;298(5600):1999-2002.

124. Florkiewicz RZ, Sommer A. Human basic fibroblast growth factor gene encodes four polypeptides: Three initiate translation from non-AUG codons. Proc Natl Acad Sci U S A 1989;86(11):3978-81.

125. Taira M, Iizasa T, Shimada H, Kudoh J, Shimizu N, Tatibana M. A human testis-specific mRNA for phosphoribosylpyrophosphate synthetase that initiates from a non-AUG codon. J Biol Chem 1990;265(27):16491-7.

126. Scharnhorst V, Dekker P, van der Eb AJ, Jochemsen AG. Internal translation initiation generates novel WT1 protein isoforms with distinct biological properties. J Biol Chem 1999;274(33):23456-62.

127. Arnaud E, Touriol C, Boutonnet C, et al. A new 34-kilodalton isoform of human fibroblast growth factor 2 is cap dependently synthesized by using a non-AUG start codon and behaves as a survival factor. Mol Cell Biol 1999;19(1):505-14.

128. Bonnal S, Schaeffer C, Creancier L, et al. A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons. J Biol Chem 2003;278(41):39330-6.

129. Gonzalez-Herrera IG, Prado-Lourenco L, Pileur F, et al. Testosterone regulates FGF-2 expression during testis maturation by an IRES-dependent translational mechanism. FASEB J 2006;20(3):476-8.

130. Baird SD, Turcotte M, Korneluk RG, Holcik M. Searching for IRES. RNA 2006;12(10):1755-85.

131. Aoki T, Okada N, Ishida M, Yogosawa S, Makino Y, Tamura TA. TIP120B: A novel TIP120-family protein that is expressed specifically in muscle tissues. Biochem Biophys Res Commun 1999;261(3):911-6.

132. Aoki T, Okada N, Wakamatsu T, Tamura TA. TBP-interacting protein 120B, which is induced in relation to myogenesis, binds to NOT3. Biochem Biophys Res Commun 2002;296(5):1097-103.

133. Shiraishi S, Zhou C, Aoki T, et al. TBP-interacting protein 120b (TIP120B) /cullin-associated and neddylation-dissociated 2 (CAND2) inhibits SCF-dependent ubiquitination of myogenin and accelerates myogenic differentiaton. J Biol Chem 2007 (Epub ahead of print).

134. Sui L, Dong Y, Ohno M, et al. Jab1 expression is associated with inverse expression of p27(kip1) and poor prognosis in epithelial ovarian tumors. Clin Cancer Res 2001;7(12):4130-5.

135. Esteva FJ, Sahin AA, Rassidakis GZ, et al. Jun activation domain binding protein 1 expression is associated with low p27(Kip1)levels in node-negative breast cancer. Clin Cancer Res 2003;9(15):5652-9.

136. Shintani S, Li C, Mihara M, Hino S, Nakashiro K, Hamakawa H. Skp2 and Jab1 expression are associated with inverse expression of p27(KIP1) and poor prognosis in oral squamous cell carcinomas. Oncology 2003;65(4):355-62.

137. Adler AS, Lin M, Horlings H, Nuyten DS, van de Vijver MJ, Chang HY. Genetic regulators of large-scale transcriptional signatures in cancer. Nat Genet 2006;38(4):421-30.