Chapter 4
Protein degradation and
Cullins - An overview
Protein degradation and differentiation

Protein degradation is an important mechanism by which various cellular and developmental events are regulated. Often, defective protein-degradation machinery leads to disease conditions, particularly cancer. Many of the morphological and physiological alterations occurring during terminal differentiation also involve protein degradation and clearance of proteins that are no longer required for function in the terminally differentiated cells (Haldeman et al., 1995). Synthesis of a protein and its degradation, both occur simultaneously and continuously in a given cell. Small modifications in the balance between these processes allow cells to rapidly adapt to changes in the extracellular environment. In addition to this continuous turnover of proteins inside cells, abnormally synthesized proteins or proteins incorrectly modified are rapidly eliminated from the cells by the proteolytic systems. Intracellular proteases also participate in many other fundamental cell processes including cell differentiation and cell cycle progression (Lauren et al., 2001).

The proteasome pathway is likely to play a role in this protein modeling, as suggested originally by the high activity of this pathway in reticulocyte lysate (Haldeman et al., 1995). It has been shown that proteasomal degradation of proteins plays an important role in the maturation of erythroblasts (Cheng et al., 2002). It is also suggested that protein degradation plays an important role during differentiation - a fine balance between the molecules of protein degradation machinery is very important for proper development and differentiation. In skeletal muscle, the system is responsible for the breakdown of the major contractile proteins, actin and myosins (Attaix et al., 1998; Lecker et al., 1999; Hasselgren and Fischer, 2001). Ubiquitination of proteins is the major signal for the
degradation of proteins. Marking a specific protein for degradation is an important and complex process, involving three major steps. (1) Synthesis of a polyubiquitin chain (2) covalent attachment of a polyubiquitin chain to the substrate, (3) specific recognition of this signal, and degradation of the tagged protein by the 26S proteasome (Fig.4).

**Ubiquitination**

Ubiquitination is a multiple step process (Hershko and Ciechanover, 1998; Scheffner et al., 1998; Ciechanover et al. 2000). In brief, ubiquitin is initially activated in the presence of ATP to a high-energy thiol ester intermediate by the ubiquitin-activating enzyme (E1). E1 then transfers ubiquitin to one of the ubiquitin conjugating enzymes (E2s), which also forms a thiol ester linkage between the active site cysteine and ubiquitin. E2s and/or ubiquitin-protein ligases (E3s), which play a role in the selection of proteins for conjugation, bind the first ubiquitin molecule to protein substrates via an isopeptide bond between the activated C-terminal glycine residue of ubiquitin and the ε-amino group of a lysine residue of the substrate. The resulting monoubiquitinated protein is usually not targeted for degradation by the proteasome. Typical mono ubiquitinated conjugates are receptors or basic proteins such as histones. Alternatively, E2s and/or E3s catalyze the formation of poly ubiquitinated conjugates. This is usually achieved by transfer of additional activated ubiquitin moieties to Lys48 of the preceding conjugated ubiquitin molecule. Finally, a fourth enzyme called E4, which catalyzes the efficient polymerization of very long polyubiquitin chains, has been characterized in yeast (Koegl et al., 1999). The ubiquitin-conjugating system is hierarchical. In mammals, there is a single E1 (Hershko and Ciechanover, 1998; Pickart, 2001), at least 20-30 E2s (Scheffner et al., 1998), and several dozens of E3s (Wilkinson, 2000; Pickart, 2001).
A Ubiquitin is first activated by the ubiquitin-activating enzyme (E1) and transferred on one ubiquitin-conjugating enzyme (E2). (B). The E2 with or without an ubiquitin-protein ligase (E3), mono- di or triubiquitinates the substrate (Protein-[Ub]1,2,3 ), which is not targeted for breakdown. (C). In contrast, when a polyubiquitin degradation signal is formed (Protein-[Ub]n ), the substrate can be deubiquitinated (D) or is recognized and degraded into peptides by the 26S proteasome (E).
Ubiquitin-activating enzyme (E1)

E1 exists as two isoforms of 110- and 117-kDa, which derive from a single gene and are found in both the nucleus and the cytosol (Haas and Siepmann, 1997). The reaction starts with the binding of ATP-Mg2+, and then of ubiquitin, leading to the formation of a ubiquitin adenylate intermediate that serves as the donor of ubiquitin to the critical cysteine residue in the E1 active site. When fully loaded, the E1 carries two molecules of activated ubiquitin (as a thiol ester and as an adenylate, respectively), so that the thiol-linked ubiquitin is transferred to one of the E2s. This reaction is very efficient and allows the production of activated ubiquitin for the entire downstream ubiquitin conjugation pathway. Thus, and not surprisingly, deletion of E1 is lethal in yeast.

Ubiquitin-conjugating enzymes (E2s)

E2s are a super family of related proteins, with a molecular weight range of ~14- to 35-kDa. There are 4 classes of E2s that share a central catalytic domain of ~150 amino acids with the active site cysteine, and some variable N- and/or C-terminal extensions that may play a role in substrate specificity. There are eleven E2s in yeast (Pickart, 2001), and 20-30 E2s in mammals (Scheffner et al., 1998). Despite their structural similarities, E2s are responsible for distinct biological functions so that only a limited number of E2s (e.g. three E2s in yeast) play a role in the formation of the polyubiquitin degradation signal. In general, this signal is catalyzed by the E1, one E2 and one E3. A given E2 can interact with a limited number of E3s (and vice versa), which in turn recognize their specific protein substrates (Pickart, 2001). Moreover, a given protein substrate can be ubiquitinated by different combinations of E2s and E3s (Gonen et al., 1996). This results in a wide range of ubiquitination pathways, which are presumably specific for a given protein or a class of substrates.
Ubiquitin-protein ligases (E3s)

E3s play a key role in the ubiquitin pathway, as they are responsible for the selective recognition of protein substrates. Very few E3s that shared poor apparent structural similarities were first described (Hershko and Ciechanover, 1998). Recently, our knowledge of these enzymes has been rapidly growing. All known E3s are HECT domain E3s or RING finger E3s (Pickart, 2001). The first major group of E3s corresponds to enzymes of the HECT (Homologous to E6-AP CTerminus) domain family. E6-AP (E6-Associated Protein) forms a complex with the papilloma virus E6 oncoprotein (that acts as an adapter protein) to ubiquitinate the tumor suppressor p53 protein. The final third of the E6-AP sequence, called the HECT domain, is approximately 350 amino acids in length (Fig. 5A). The HECT domain itself mediates E2 binding and ubiquitination of the target protein via thiol ester linkage formation with ubiquitin. The N-terminus region of every HECT E3 binds to specific substrate(s) (Fig.5A). Twenty different human HECT E3s that interact with two classes of E2s have been described (Schwarz et al., 1998), but mammalian genome sequencing projects have identified numerous potential uncharacterized HECT E3s (Pickart, 2001).

The RING finger structure is defined by eight cysteine and histidine residues that coordinate two zinc ions (Freemont, 2000; Pickart, 2001) (Fig.5B). There are several hundred cDNAs encoding RING finger proteins in the GeneBank database, and many unrelated RING finger proteins with unknown functions behave in vitro as E3s (Lorick et al., 1999). This suggests that the number of E3s could be much larger than previously believed. The RING finger E3s are either monomeric proteins or multiple subunit complexes (Pickart, 2001).
Fig 5. Different families of E3 ubiquitin Ligases

(A) HECT (Homologous to E6-AP C-terminus) domain E3
(B) RING finger E3
The multiple subunit complexes of RING finger E3s comprise so far, at least three distinct E3 families called the cyclosome or APC (Anaphase Promoting Complex) (Hershko and Ciechanover, 1998), the VCB-like (Von Hippel-Lindau tumor suppressor-ElonginC/B) E3s (Tyers and Willems, 1999; Lisztwan et al., 1999; Pickart, 2001) and the SCF (Skp1-Cdc53/Cullin-F-box protein family). These complexes contain a catalytic core and substrate-specific adapter proteins (Tyers and Willems, 1999). For example, in SCF E3s, the catalytic core is formed by three subunits: Cullin1, the RING finger subunit Rbx1 and an E2. The adapter protein Skp1 recruits (via the F-box motif) F-box proteins, which themselves recruit specific protein substrates through protein-protein interaction domains such as leucine-rich repeats or WD-40 domains.

**Cullins**

The name cullin derived from the fact that each member of the family ‘culls or sorts’ the substrates for ubiquitinylatation. Budding yeast have 3 cullin proteins CulA or Cul1, CulB or Cul3 and CulC or Cul8 and Metazoans have at least five cullins (Cul1-5). Sequence homology extends across the whole length but greatest homology lies in the C terminus that contains a cullin domain. Cullins derived from duplication events of ancestral cullin gave rise to two main branches, Cul 1,2 and 5 and Cul3/4. *C.elegance* Cul6 (absent in vertebrates) appears to be derived by a duplication event of ancestral Cul1 gene, while a later duplication separated mammalian Cul4A and Cul4B gene. An additional cullin member Cul 7 has been identified in humans. At least two additional cullin domain-containing proteins exist in mammals. Parc, which interacts with p53 controlling its subcellular localization (Nikolaev et. al., 2003) and APC2 a subunit of another E3 ubiquitin ligase- Anaphase promoting complex/ cyclosome (APC/C complex) (Yu et al. 1998). Most cullin members have been shown to interact with RING finger proteins. Cul7

73
interacts with Roc1 or its related protein Roc2 (Ohta et al. 1999), Apc2 interacts with Apc11 another homologue of Roc1. The cullin RING finger protein complex shows intrinsic ubiquitin ligase activity \textit{in vitro} since purified recombinant Cullin-Roc complex are able to catalyse substrate independent ubiquitin ligation by E2s forming free ubiquitin chains. Similarly Parc contains two RING domains that are likely responsible for its intrinsic ubiquitin ligase activity (Nikolaev et. al., 2003). CDLs are positively regulated by covalent conjugation of the ubiquitin like molecule Nedd8 to a specific lysine residue present in the cullins (Hori et al 1999). Conversely deconjugation of Nedd8 from cullins by the isopeptidase activity of COP9 signalosome inhibits their ubiquitin ligase activity (Schwechheimer and Deng, 2001). The best characterized mammalian cullin family member is Cul1, which is a component of a multi protein ubiquitin ligase complex SCF (Skp1, cul1, Fbox protein) or Cdl 1 (Feldman et al. 1997 and Skowyra et al 1997). In this complex, Cul1 acts as a scaffold protein. The N-terminus of Cul1 interacts with the Skp1 and the C-terminus cullin domain interacts with the Roc1 or 2 and a specific ubiquitin conjugating enzyme (Ubc3, 4, or 5). Skp1 in turn binds with the F box protein that is known so because it contains a 40 amino acid motif known as F box, through which it interacts with Skp1 (Cenniciarelli et al.1999 and Winston et al 1999). Each F box protein can interact with specific substrate through specific protein-protein interaction domain (WD40 domain, leucine rich repeats etc.). FBP family has approximately 70 members, and therefore about 70 SCF complexes are expected to be present in the mammalian cells, but only three of them are extensively studied and corresponding substrates are identified. \textbf{Fig.6} describes all the identified SCF and their substrates. Thus, the F box proteins provide substrate specificity to the SCF complex. Complexes similar to SCF are formed by analogous modular proteins, which gather to form other CdlS. Cul7, similar to cul 1, assembles an SCF like complex containing Roc1, Skp1 and FBP Fbw6. In contrast to Cul1
Fig. 6 Three SCF complexes contain a common core, composed of Cul1, Skp1 and ROC1, that associates with different FBPs. Each FBP acts as substrate receptor subunit which is coupled to the discrete number of specific phosphorylated substrates through protein-protein interaction domains. The most extensively characterized substrates for these SCF ligase are shown.

(Daniele et al. 2004, Oncogene reviews, Ubiquitination in cancer.)
which can interact with Skp1 by itself, cul7 can not interact with Skp1 till it binds with Fbw6. *C. elegans* Cul6 may also form a SCF like complex since it interacts with Skr3, a Skp1 like protein (Nayak et. al., 2002).

**Cul1**

Cul1 the ortholog of yeast cdc35 was the first member of cullin family to be identified in metazoans (Kipreos et al 1996). In *C. elegans*, loss of function mutation of cul1 causes hyperplasia of all the tissues with a shortened G1 phase of cell cycle. Cul1 is necessary for proper cell-cycle exit. In mice, loss of Cul1 results in early embryonic lethality (Dealy et al., 1999 and Wang et al 1999). Cul1 deficient embryos implant in the uterine wall but do not develop beyond embryonic day 5.5, before the onset of gastrulation. In addition, apoptosis is increased in the embryonic ectoderm. The early lethality in mouse is mainly due to accumulation of all the substrates degraded by cul1. Different molecules regulated by Cul1 containing SCF complex are depicted in the Fig.6.

**Cul2 and Cul5**

Cul2 and Cul5 assemble in a complex (Cdl3 and Cdl5 respectively) containing a Roc1 protein ElonginB, ElonginC, and one of many SOCS box proteins (SBPs) - where SOCS stands for Suppressor of Cyotkine Signaling (Kamura et al., 1998, 2000 and Schoenfeld 1998). Cdl2 and Cdl5 resemble SCF complex since sequence homologies have been identified between Skp1 and Elongin C, which are both adaptors for the substrate targeting subunit, and SBPs like FBPs are a large family of proteins that contain protein - protein interaction domains useful for binding specific substrates (Hilton et al., 1998). Thus, SCF complexes act as a common core (ElonginB-ElonginC-Cul5/Cul2-Roc), which are coupled to large number of substrates through SBPs to ubiquitinylate a large number of proteins. In
addition, Cdl2 and Cdl5 can also interact with another family of substrate recognizing proteins, which do not have clear SOCS motif but a more loose motif called BC box motif, necessary for binding to Elongin B or Elongin C. It has been shown that the product of the Von Hippel Lindau gene interacts via its BC box to the Cul2 forming a Cdl2 known as VBC complex. Another BCBP known as mediator subunit Med8 has been shown to assemble with ElonginB-ElonginC-cul2-Roc1 via its BC box to form a complex recruiting ubiquitin ligase activity to the RNA polymerase transcriptional machinery (Brower et al. 2002). The SOCS box is contained in more than 40 proteins belonging to different families. The canonical SOCS family comprises Socs1-7, Cytokine inducible SH2 motif containing protein, the ankyrin repeat proteins containing SOCS box, SPRY domain proteins with SOCS box, WD40 motif proteins with SOCS box, Tubby domain containing protein with SOCS box, Rar proteins and neutralizing family of proteins (Hilton et al., 1998). All these polypeptides contain protein–protein interaction domains other than SOCS box. Many of them, such as Socs1, Socs3, Wsb1, Asb2 and Rar proteins, have been shown to interact with Elongin B or C (Kamura et al., 1998). Socs1 and Src2 homology domain (SH2)-containing proteins are negative regulators of cytokine and growth factor signaling and they interact with Cul2 (De Sepulveda et al 2000, Kamizono et al., 2001).

In C. elegans, Cul2 acts as a positive cell cycle regulator. It is expressed mainly in proliferating cells and disruption of Cul2 induces G1 arrest of germ cells (Feng et al. 1999). This block correlates with the accumulation of CDK inhibitor Cki1. Cul2 is also necessary for proper cytoskeleton movement and mitotic chromosome condensation.

It has been reported that besides Cul2, Cul5 is also able to assemble with Roc1, Elongin B, Elongin C and a BCBP to form a complex with ubiquitin ligase activity (Kamura et al. 2001). Several BCBPs and SBPs have been shown to associate with Cul5-ElonginB-ElonginC-Roc1 complex. Cul5-elonginB-ElonginC-Roc1 complex can also bind with two
viral proteins E4orf6 and E1B55k (Querido et al. 2001). It was previously shown that these two proteins interact with p53, inhibiting its transcriptional activity and facilitating its degradation. E4orf6 and E1B55k mediated degradation of p53 is analogous to that induced by Human papilloma virus E6 oncoprotein, which recruits the cellular ubiquitin ligase, E6 associated proteins, to p53 and cause its degradation. Hence in both the cases viral oncoprotein can recruit cellular ubiquitin ligase complex and use it to eliminate tumor suppressor p53.

Cul3

In C. elegans, depletion of Cul3 by RNAi causes defects in early embryogenesis resulting in abnormal filament and microtubule organization (Kurz et al., 2002). Cul3 is required for proper development in mice since inactivation of Cul3 locus by homologous recombination results in embryonic lethality prior to 7.5 days of gestation with defects in both embryonic and extra embryonic compartments (Singer et al., 1999). The extra embryonic tissues are completely disorganized and the trophectoderm develops abnormally. Gastrulation is also abnormal in Cul3 deficient embryos. In addition, increased level of free CyclinE (Non-Cdk2 associated) protein was observed in ectoplacental cone and extra-embryonic ectoderm of Cul3-/- embryos (Singer et al., 1999). In an attempt to identify the adaptor protein equivalent to Skp1, which plays a conserved function in Cdl3 complex, C. elegans Cul3 was used as bait in the yeast two-hybrid screen (Xu et al., 2003). In all, 11 Cul3 specific interactions were isolated and all of them have a common domain, previously known as 'Broad complex Tramtrack and Bric-a- Brac' (BTB), which is thought to play a role in protein–protein interaction (Zollmen et al., 1994; Chen et al., 1995). BTB proteins from different species can also contain additional domains which are known to facilitate protein–protein interactions, such as MATH
domain, kelch repeats, WD40 motif and zinc finger domain. Remarkably the structure of Skp1 and BTB domain of human promylocytic leukemia zinc finger protein revealed the presence of conserved structural features. Cul3 along with different BTBs might target many more substrates. One of the BTBPs, Mel 26 binds to the microtubule severing protein Mei-1/Katanin. Mei-1 is an essential component of the C. elegans mitotic spindle. Mei-1 is degraded during meiosis to mitosis transfer and Cul3 is required for the degradation. Cul3 silencing by RNA interference results in abortive cytokinesis and analogous spindle orientation defects with disoriented spindles that become displaced towards one pole (Kurz et al. 2002). These genetic results together with the finding that Mel-26 physically interacts with Mei-1 (Furukawa et al., 2003, Geyer et al 2003, Pintard et al. 2004, Xu etal. 2003) strongly indicate that Cul3 dependent ubiquitin ligase Cdl3^{Mel26} targets Mei-1 for degradation. In developing eye of Drosophila melanogaster, Cul3 plays a major role in degradation of Cubicutus interruptus (Ci), which is a major component of hedgehog signaling pathway (Ou et al. 2002). S. pombe Cul3 interacts with the BTBP, Btb3p the fission yeast ortholog of human Bpoz2 (Geyer et al. 2003), a protein that is thought to be implicated in the growth arrest mediated by the tumor suppressor PTEN (Unoki and Nakamura 2001).

Cul4

The gene encoding Cul4A is amplified in 16% of primary breast cancers, and RNA in situ hybridization analysis indicates that 47% of primary breast cancers over express Cul4A (Chen et al., 1998). Cul4A^{-} mice are inviable and no homozygous mutants are recovered as early as 7.5day dpc. However, Cul4A^{-} blastocysts are viable and capable of hatching and forming both an inner cell mass and trophectoderm. Approximately 20% of the 7.5dpc implantation sites appeared empty, suggesting that cul4 mutant embryos are able to
implant but die between 4.5dpc to 7.5dpc. Despite 82% similarity between Cul4A and Cul4B, the \textit{cul4A}^{-/-} lethal phenotype suggests that Cul4A and Cul4B have one or more distinct functions. In addition, fewer heterozygous pups were observed than expected by Mendelian genetics indicating that many \textit{cul4}^{+/+} embryos also die (Binghui et al. 2002).

It has been shown that there is a modest reduction (about 2.5-fold) in CUL4A expression during granulocytic differentiation and appropriate CUL4A expression is essential for regulating granulocytic and macrophage differentiation and for regulating proliferation during differentiation (Binghui et al., 2003). Since modestly increased CUL4A expression does not alter the cell cycle distribution in cells not induced to differentiate, but dramatically alters the cell cycle distribution in cells induced to differentiate, it is suggested that the cell-cycle regulatory function of CUL4A is inter-connected with differentiation. Groisman et al., 2003 have shown that Cul4A is involved in nuclear excision repair (NER), an essential cellular defense mechanism against the oncogenic consequence of ultraviolet light. They reported that two NER proteins DDB2 and CSA are associated with identical complexes containing Cul4A, Roc1, DDB1 and the COP9 signalosome. These complexes display ubiquitin activity that is regulated in response to UV irradiation.

It has also been shown that Cul4 has a role in maintaining genomic stability by temporally restricting DNA replication licensing in \textit{C. elegans} (Zong et al. 2003). Silencing of Cul4A and Cul4B by RNA interference leads to massive DNA re-replication and S phase accumulation of the replication licensing factor Cdt1 in mammalian and \textit{Drosophila} cells (Leigh et al. 2003). Taken together these results suggest that Cul4 plays an important role in regulating DNA repair, replication as well as differentiation. There is no information available about any specific role for Cul4B as well as how exactly these two protein play a role during differentiation.