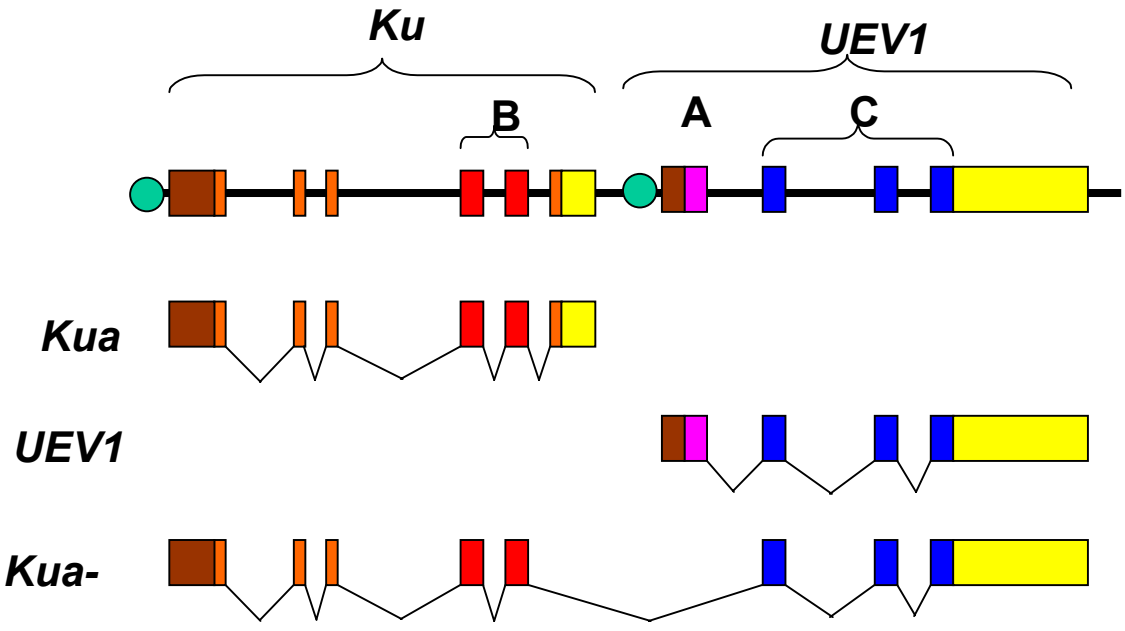
Universitat Autònoma de Barcelona

Studies on UEV, a new regulator of polyubiquitination: Functional aspects and genomic analysis



Noureddine Loukili

Universitat Autònoma de Barcelona

Instituto de Biología Molecular de Barcelona
Consejo Superior de investigaciones Científicas

Departament de Bioquímica y Biologia Molecular
Universitat Autònoma de Barcelona

Studies on UEV, a new regulator of polyubiquitination: Functional aspects and genomic analysis.

Memoria para optar al grado de Doctor en Biología presentada por

Noureddine Loukili

Programa de doctorado de Bioquímica y Biología Molecular

Barcelona, Febrero 2002

Universitat Autònoma de Barcelona

El Doctor Timothy Thomson Okatsu, Doctor en Medicina y Científico Titular del Consejo Superior de Investigaciones Científicas certifica que: la tesis *Studies on UEV, a new regulator of polyubiquitination: Functional aspects and genomic analysis*, de la que es autor el Licenciado en Biología, Noureddine Loukili, para optar al grado de Doctor en Biología, ha sido realizada bajo su dirección y se encuentra en condiciones de ser defendida ante el tribunal correspondiente.

Barcelona, 26 de Febrero de 2002

Dr. Timothy Thomson Okatsu

El interesado

Noureddine Loukili

Agradecimientos

Cuando hago balance de los años que he vivido y de todo lo que he aprendido tanto personalmente como profesionalmente en este maravilloso país, me alegro de haber venido aquí y me resulta muy difícil expresar con palabras mi agradecimiento a todos aquellos que me han abierto los brazos y me han dado mucho más de lo que necesitaba o lo que pedía. Han sido muchos años y en mi camino me he encontrado con mucha gente inolvidable.

En primer lugar quiero dar las gracias Dr. Timothy Thomson por haberme aceptado en su grupo, por su paciencia conmigo y por todo lo que me ha enseñad. Han sido muchos años, muchos proyectos y mucho trabajo, espero que este proyecto continúe adelante y tenga mucho éxito. Espero continuar aprendiendo de ti.

Gracias a Victor y a Rosana por su colaboración continua en este proyecto.

Gracias a Jaume, Francina y Patricia por su apoyo, por sus consejos y por su preocupación por mi vida tanto profesional como personal.

Gracias a todos mis grandes amigos: Miguel Angel, Víctor Barberá, Isabel, Hajji , Thierry, Malika, Guraich Mohammed, Meskini , Jose Piedra, Fausto, Santi, Silvia Gómez, Silvia Pitarch , a Elena y Edu. ***Lee Iacocca.*** dijo: "*Mi padre siempre solía decir que al morir, si contabas con cinco buenos amigos entonces tuviste una vida grandiosa*." Yo que tengo más de cinco buenos amigos considero que tengo una vida grandiosa.

Gracias a toda mi familia de l'IMIM, mi estancia entre vosotros ha marcado mi vida, nunca os olvidaré:

Gracias al Dr Paco (todavía me acuerdo de Antonio reírse cuando contesté una día al teléfono y decía el Doctor Paco no está!). Nunca olvidaré tus llamadas al gobierno civil cuando no podía volver a Marruecos, ha sido un gesto natural para tí pero para mí ha sido un apoyo a una persona desesperada y indignada. Gracias por tu interés por mi vida personal y profesional.

Gracias a Antonio G.H: por tu apoyo y por todo el tiempo que has perdido conmigo.

Gracias a la Dra. Victoria Nogués por su comprensión y su ayuda en el principio y al final de mi tesis.

Gracias a (Carmen de Bolos, Carmen Torns y Marta Garrido: nunca olvidaré vuestra preocupación por mi salud cada vez que os necesitaba y las movidas que hacíais para solucionar mis problemas. Os quiero mucho).

Gracias a Mercè, Teresa, Maria del Mont, Olga, Montse Torà, Clara, Maya y a Leo: siempre he notado vuestro cariño, aunque esté lejos siempre podréis contar conmigo.

Gracias a mi grupo de teatro de l'IMIM: me habéis dado la oportunidad de conocer a grandes artistas. Espero que las obras continúen dando alegría a todo el Hospital y espero veros más en el escenario.
Gracias a mis compañeros de la Vall d'Hebron (a mi harén: Marta, Ester, Vanessa, Johanna y Anna. Por compartir mis alegrías y mis problemas en todo momento. Espero que no me olvidéis aunque esté lejos.)

Gracias(al equipo de campeones David, Oscar, Juan, Luis, Jesús, Josep y el Bifurquense: con vosotros han sido unos años gloriosos en las carreras inter-hospitalarias).

Gracias a Silvia, Anna, Pilar y a las enfermeras Olga, Maruja y Isabel por los Tés con menta y por vuestra simpatía).

A la nueva generación (Alba, Tomás, Maite y a todos los nuevos, os deseo lo mejor en esta carrera que estáis empezando).

Gracias a los colegas del Lab 2: Montse, Cristina A. Cristina C, Olga y Anna Meseguer por vuestras ayudas siempre cuando os necesitaba.

Gracias a Sagrario, Ana Cris, Naul y a Marcel por su amistad y su ayuda desinteresada.

Gracias a los amigos del hospital de Elche por estar al tanto siempre de mis noticias y de apreciarme como yo también os aprecio.

Gracias a Mostafa y a Hafid por todo lo que estáis compartiendo conmigo y por la amistad.

Gracias a Marie Hélène y a Yasser por los buenos momentos que pasamos juntos.

Gracias a Marta por creer y confiar en mí independientemente de mi religión, raza o cultura. Ojalá todo el mundo fuera como tú.

Gracias a Paquita y a Pere por tener una hija como Marta.

Algunos se preguntan porqué no me he integrado en España, porqué hago el Ramadán, no bebo alcohol y no como cerdo: Sólo les quiero decir que la prueba de mi integración son estos amigos que tengo. Alguien que no se integra no puede tener tantos amigos. No necesito cambiar mi cultura para integrarme porque todos somos diferentes y todos somos iguales: Una neurona y una célula epitelial son muy diferentes, pero si nos fijamos en los cromosomas y los genes son ambas iguales. Las dos células hacen una función vital para el cuerpo y ninguna puede hacer la función de la otra, pero también las dos pueden ser tumorales si sufren un daño.

Gracias a todos: me habéis aceptado como si fuera uno de vosotros: Habéis vivido mis buenos y malos momentos como amigos, como hermanos, y como padres, os quiero confesar que nunca me sentí extranjero entre vosotros y os lo agradezco mucho.

A mi querida familia de Marruecos, que me perdonen por no estar con ellos.

A la memoria de mi madre, que hubiera estado orgullosa de mí por más pequeño y humilde que sea mi trabajo.
A todas las madres.

Abbreviations

cdk1:	Cyclin dependent kinase 1
E1:	Ubiquitin activating enzyme
E2:	Ubiquitin conjugating enzyme
E3:	Ubiquitin ligase
E6-AP:	E6 associated protein
EST:	Expressed sequence tag
GFP:	Green fluorescent protein
HA:	Hemaglutining
HECT:	Homologue to E6-Ap carboxy terminal
HGT:	Horizontal gene transfer
IKK:	I kappa B kinase
K29:	Lysine at position 29
K48:	Lysine at position 48
K63:	Lysine at position 63
LGT:	Lateral gene transfer
MMS:	Methyl methane sulfonate
RING:	Really interesting gene
Ub:	Ubiquitin
UEV:	Ubiquitin conjugating E2 enzyme variant
UV:	Ultraviolet

I./ Introduction

I.1. Discovery of ubiquitin

About 27 years ago, Goldstein isolated a protein which he identified as a lymphocyte differentiation promoting factor, and which he called ubiquitin (Goldstein and Dayhoff, 1975) as he thought it was probably ubiquitous to living cells. A few years later, the non-histone protein component of the nuclear protein A24 was identified as ubiquitin (Hunt *et al*., 1977) and A24 was found to be a covalent adduct of ubiquitin and histone H2A. The ubiquitin is conjugated to H2A by formation of an isopeptide bond between a lysine side-chain ε-amino group and the carboxyl terminus of ubiquitin (Goldknopf *et al*., 1977). So this appeared to be a new role for ubiquitin, and established that ubiquitin could be conjugated to other proteins.

While protein synthesis was well understood, the breakdown of proteins back to amino acids in the cell was only poorly understood. However it was realised that proteins were continually synthesised and degraded in cells, and that some proteins turned over more rapidly than others. The lysosome system of mammalian cells was known to degrade intracellular proteins but this didn't give the clue about how could protein degradation occur in rabbit reticulocytes (the precursors of erythrocytes) since this cells lack lysosomes. Therefore, rabbit reticulocytes became a useful cell model to study non-lysosomal intracellular proteolysis. A soluble (cytosolic) proteolytic system in reticulocytes was found, surprisingly, to be ATP-dependent (Etlinger and Goldberg, 1977) Fractionation of reticulocyte "lysate" (cytosol) generated two fractions (I and II) which were required for at least some of the ATP-dependent proteolysis of some test

proteins. The active factor in Fraction I called APF-I was purified, and was found to be covalently conjugated to proteins in the presence of ATP and Fraction II (Ciechanover *et al*., 1980). APF-I was then shown to be identical to ubiquitin (Wilkinson *et al*., 1980).

I.2. What is the ubiquitin (Ub)?

The ubiquitin is a heat-stable small molecule (8.6 kDa) with 76 amino acids, that adopts a stable compact globular conformation with four strands of β-sheet and a single α-helix. Its name comes from the fact that it is abundant in all eukaryotes, and extraordinarily well conserved in creatures as divers as yeast and human. The three carboxy-terminal residues, -Arg-Gly-Gly, are flexible and extend into the solvent, which makes this molecule very soluble. The lysine amino acids are very important in the ubiquitin since they allow the conjugation of itself to substrates proteins or to other molecules of ubiquitin to form polyubiquitin chains. There are seven lysines in the ubiquitin and the consequences for the use of one lysine or another are dramaticaly different (**Figure I.1**).

MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLH

Figure I.1. Primary structure of the ubiquitin molecule, showing the the seven lysines for conjugation to the substrate molecules or to other ubiquitin molecules

The ubiquitin is the most conserved protein known in eukaryotes (Özkaynak *et al*., 1984). Although it is apparently absent in most prokaryotes, including *E. coli*, ubiquitin is present in the cyanobacterium *Anabaena variabilis* (Durner and Boger, 1995), and may be present in archaebacteria (Wolf *et al*., 1993). In *Anabaena*, ubiquitin

can conjugate to endogenous proteins including dinitrogenase reductase in extracts prepared from differentiated nitrogen-fixing heterocysts (Durner and Boger., 1995).

I.3. Protein modification by ubiquitin

Protein ubiquitination is a postranslational modification which plays a major role in the degradation and regulation of activity of cellular proteins. The ubiquitination system is involved in diverse biological processes, including cell cycle progression (Koepp *et al*., 1999), oncogenesis (Joazeiro and Weissman., 2000; Joazeiro *et al*., 1999), and antigen presentation (Rock and Goldberg., 1999). The majority of cellular ubiquitin conjugates appears to be targeted to the 26 S proteasome, which degrades substrates to small peptides, whereas ubiquitin is recycled. The ability of ubiquitin to signal substrate proteolysis by the 26S proteasome underlies many of ubiquitin's cellular functions (Hershko and Ciechanover, 1998). This happens for example for the substrates relevant to the cell cycle, and that fall in two broad categories: those whose destruction is required for cell cycle progression (e.g., Sic1, Pds1, and B-type cyclins), and those whose destruction is not essential but is important for cellular homeostasis (e.g., Cdc6, Cdc20, and G1 cyclins) (Koepp *et al*., 1999). However, certain cell surface proteins, when they are modified by ubiquitination, appear to be targeted for lysosomal degradation via the endocytic route (Hicke, 1999; Strous *et al*., 1996): ubiquitination of the growth hormone receptor (GHR) cytoplasmic tail is likely the molecular event to which triggers receptor endocytosis and the degradation of both exoplasmic and cytoplasmic portions within endosomal/lysosomal compartments (Strous *et al*., 1996).

Many studies have demonstrated that ubiquitin also serves as a nonproteolytic signal in DNA repair (Jentsch *et al*., 1987; Spence *et al*.,1995) and IKB kinase (IKK) activation (Chen *et al*., 1996).

Some substrates are conjugated to just one ubiquitin, others are conjugated to multiple ubiquitin molecules in the form of a polyubiquitin chain. The biologically active ubiquitin signal frequently consists of a polyubiquitin chain in which successive ubiquitins are joined through isopeptide bonds involving specific lysine residues of ubiquitin. In some cases, the fate of a ubiquitin-conjugated protein can be correlated with the chemical structure of the polyubiquitin chain (Chau *et al*., 1989; Deng *et al*., 2000; Finley *et al*., 1994; Hofmann and Pickart., 1999; Koegl *et al*., 1999; Pickart., 2000; Spence *et al*., 1995). For example, polyubiquitin chains linked through Lys48 (K48-chains) are the principal signal for proteolysis by 26S Proteasomes (Chau *et al*., 1989 ; Finley *et al*.,1994), whereas Lys63-linked chains (K63-chains) are required for postreplicative DNA repair (Hofmann and Pickart., 1999; Spence *et al*., 1995), IKK activation (Deng *et al*., 2000), translational regulation (Spence *et al*., 2000), and certain cases of ubiquitin-dependent endocytosis (Galan and Haguenauer-Tsapis., 1997).

I.3.1. Monoubiquitination

Monoubiquitin serves as a signal for the endocytosis of plasma membrane proteins. In mammalian cells it promote the internalization of plasma membrane proteins when it is fused to their N-terminal cytoplasmic domain (Nakatsu *et al*., 2000). The single fused ubiquitin carries within its three-dimensional structure all of the information necessary for regulated endocytosis (Shih *et al*., 2000). In contrast, monoubiquitin is an inefficient signal for proteasome recognition (Thrower *et al*.,

2000). Another important role for the ubiquitin is the monoubiquitination of histones, which plays a role in meiosis in yeast and the development of Drosophila embryos (Pham and Sauer., 2000; Robzyk *et al*., 2000), and monoubiquitination of a Fanconi anemia protein is linked to DNA repair and localization to nuclear foci (Garcia-Higuera *et al*., 2001). Monoubiquitination of the retroviral Gag protein is required for a late step in virus budding (Patnaik *et al*., 2000; Strack *et al*., 2000).

I.3.2. Chain assembly of ubiquitin: Polyubiquitination

Chain connected through three of ubiquitin's seven Lysine residues (Lys 29, Lys 48 and Lys 63) have been identified *in vivo* in Saccharomyces cerevisise (Chau *et al* 1989; Spence *et al*., 1995). Others have been synthesized in vitro and may exist in higher eukaryotic cells (Baboshina, and Haas., 1996). Not all the chains of ubiquitin function in the equivalent manner: defined structures of Lys 48-linked tetraubiquitin is predominant in proteasome proteolysis, however specific DNA repair defect is observed in all cells expressing K63R mutant ubiquitin (Hofmann and Pickart., 1999).

It is attractive to hypothesize that alternatively linked chains are selectively recognized, leading to distinct functions for different chains. This would be possible if different chains have different structures, and thus present distinct recognition elements on the chain surface but this model remains speculative.

Preassembled chains through Lys 48 have been observed in cells, this should thus make a significant contribution to substrate conjugation whenever the level chains is significant relative to the level of monoubiquitin. This condition does not appear to be met in differentiating erythroid cells (Haldeman *et al*.,1995), but may will be achieved in other tissues (Van Nocker and Vierstra., 1993).

On the other hand, too massive an accumulation of unanchored chains is deleterious: *In vivo* and *in vitro*, unanchored chains inhibit degradation by competing with polyubiquitinated substrates for binding to the 26 S proteasome (Papa and Hochstrasser, 1993; Beal *et al*.,1996).

I.3.3. Canonical polyubiquitination (chain assembly through Lys 48)

Efficient recognition of ubiquitinated substrates by the 26S proteasome requires a minimum targeting signal consisting of four ubiquitin moieties linked to each other through isopeptide bonds between Gly 76 and Lys 48 (Chau *et al*., 1989; Thrower *et al*., 2000). The Leu 8, Ile 44, and Val 70 amino acids in the ubiquitin polypeptide, known collectively as the hydrophobic patch, are critical for proteasomal degradation, although mutations in these residues have little effect on the formation of ubiquitin conjugates (Sloper-Mould *et al*., 1996).

There are two limiting mechanisms for ligation of Lys 48 linked chains to substrates, and both of them appear to operate. First, unanchored chains may be assembled, using monoubiquitin as the chain initiator, by ubiquitin-specific conjugating enzymes, and then used in this preassembled form by substrate-specific conjugating enzymes. Unanchored chains exist within cells: ubiquitin immunoblots of extracts from plant and animal tissues revel significant levels of unanchored chains (Van Nocker and Vierstra, 1993; Haldeman *et al*., 1995; Spence *et al*., 1995), and only the K48-G76 linkage is detected when plant tissue-derived diubiquitin is analyzed by triptic peptide mapping (Van Nocker and Vierstra, 1993). In yeast, mutations of Lys residues other than Lys 48 does not decrease the level of unanchored chains, providing additional evidence that these species harbor primarily K48-G76 linkages (Spence *et al*., 1995).

Alternatively; the chain initiator may be the first ubiquitin ligated to the target by substrate –specific conjugating enzymes. In this second mechanism, chain elongation could be elongated by the same substrate-specific conjugating enzyme, by distinct ubiquitin-specific conjugating enzymes, or by both.

In conclusion, the conjugation through Lys 48 of at least for molecules of ubiquitin to any substrate molecule leads to its recognition and degradation via the proteasome, and this polyubiquitination form is called "canonical" polyubiquitination and is the most frequently used and best known form of ubiquitination.

I.3.4. Variant polyubiquitination (chain assembly through Lys 63)

The polyubiquitination that uses the lysine at the position 63 of the ubiquitin molecule for conjugation to its substrate molecule is called variant polyubiquitination (Hoffman and Pickart, 1999). A very important difference in the outcome of the substrate proteins modified by this form of polyubiquitination is that the K63 polyubiquitin chains are not recognized by the proteasome, and therefore the substrates are not tagged for degradation (Spence *et al*., 2000; Deng *et al*., 2000). However this form of polyubiquitination has been demonstrated to be involved in many other pathways, such as (*i*) internalization of some surface proteins and vacuolar degradation, as in the cases of the general amino acid permease (Gap1) (Springael *et al*., 1999), or the plasma membrane protein Fur4 (Galan *et al*., 1996); (*ii*) the regulation of the activity of other proteins, as in the case of the formation of polyubiquitin chains linked through Lys 63 on the component of the large ribosomal subunit L28 protein, a modification that regulates ribosome activity and allows proper translation of new mRNAs (Spence *et al*., 2000). When this modification is absent, a hyper-sensitivity to translational inhibitors

and polysome instability is observed (Spence *et al*., 2000); (iii) the variant polyubiquitination of the adapter proteins TRAF6 and TRAF2, mediated by the UBC13-UEV heteroduplex (see I.4.4.), activates the upstream kinases for the cascade of phosphorylation of the NFκB negative regulator IκB by activation of the IκB kinase (Deng *et al*.,2000).

I.4. Mechanisms of ubiquitination

The ubiquitination of a protein substrate involves the formation of an isopeptide bond between a substrate lysine residue and the C-terminal carboxyl group of the Gly 76 of the ubiquitin. This reaction is accomplished through the sequential actions of three enzymes: ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2 or Ubc), and ubiquitin-protein ligase (E3) (Hershko and Ciechanover, 1998).

I.4.1. Ubiquitin activating enzyme (E1)

Ubiquitin activating enzymes are abundant proteins of the cytosol and the nucleus. Genes encoding E1 enzymes have been cloned from various organisms including humans and yeast. In most organisms, including humans and the yeast Saccharomyces cerevisiae, a single E1 enzyme activates for the entire array of downstream conjugating enzymes (Zacksenhaus and Sheinin, 1990; McGrath *et al*., 1991). Diubiquitin and higher chains are activated by E1 (ubiquitin-activating enzyme), and transferred to E2 (ubiquitin-conjugating enzyme or ubc), with the same kinetics as monoubiquitin (Chen and Pickart, 1990).

E1 forms an initial thioester bond with the carboxyl terminus of ubiquitin in an ATP-dependent reaction. The carboxy-terminal Gly residue of ubiquitin is activated in an ATP-requiring step by an E1 enzyme. This step consists of an intermediate formation of ubiquitin adenylate, with the release of PPi, followed by the binding of ubiquitin to a Cys residue of E1 in a thiolester linkage, with the release of AMP. Those thiolester linkages can be broken by reductive agents like: mercaptoethanol, dithiothreitol and hydroxylamine.

I.4.2. Ubiquitin conjugating enzymes (E2s)

Ubiquitin conjugating (UBC) enzymes are encoded by a gene family and differ in their properties and intracellular localization (Jentsh *et al*., 1990; Jentsh, 1992 a,b). Structurally, E2 are related proteins bearing a (~160 amino acid) highly conserved (35-40% identity) catalytic domain called the UBC domain. Within this domain, E2 enzymes possess a specific (active site) cysteine residue required for ubiquitin-E2 thiolester formation.

Many E2 enzymes are small proteins of roughly 16 kDa and consist of the UBC domain only (Class I E2s). Some E2 enzymes have additional carboxy-terminal extensions (Class II E2s). Such sequences can be important for UBC function and may mediate substrate specificity or intracellular localization (Jentsch *et al*., 1987; Sung *et al*., 1988; Goebl *et al*., 1988; Kolman *et al*., 1992; Silver *et al*., 1992; Sommer and Jentsch, 1993). Other E2 enzymes lack carboxy-terminal extensions but contain additional amino-terminal sequences (class III E2s) (Matuschewski *et al*.,1996; Nuber *et al*., 1996). Lastly, there are some E2s that have both amino-and carboxy-terminal extensions (class IV E2s) (Jentsch *et al*., 1991)

There are 13 E2s in *S. cerevisiae*, and even more in higher organisms, reflecting multiple isoforms of some E2s (Jensen *et al*., 1995; Rajapurohitam *et al*., 1999), as well as the evolution of new E2s. The E2 protein family is structurally well characterized. The core domain consists of four standard helices α and four-stranded antiparallel β-sheet. The β-sheet and α2 form a central region that is bordered by α1 at one end and α3/α4 at the other (**Figure I.2**).

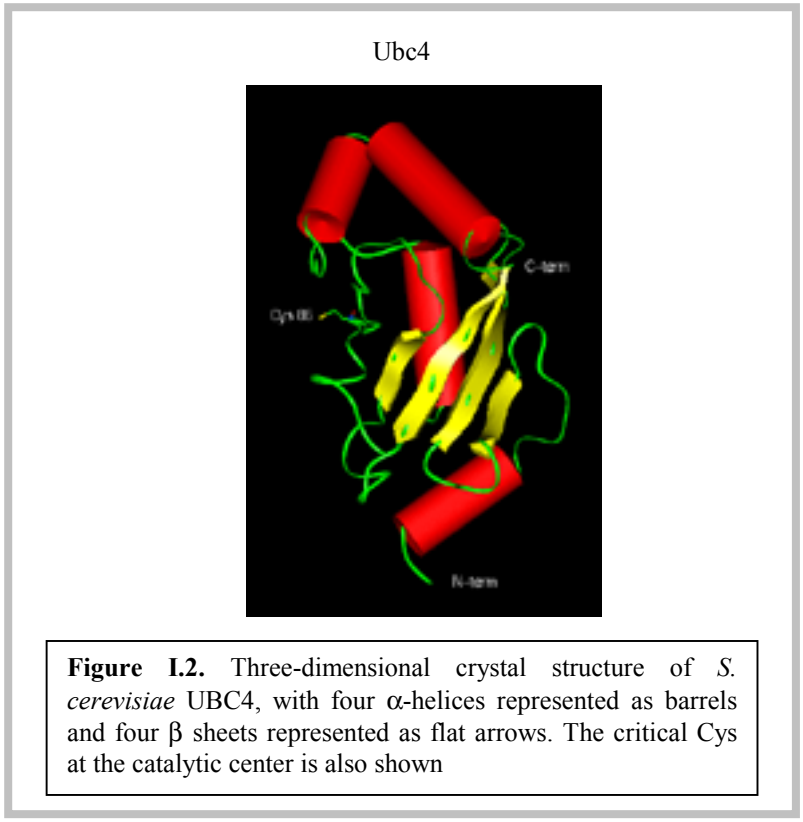


Figure I.2. Three-dimensional crystal structure of *S. cerevisiae* UBC4, with four α-helices represented as barrels and four β sheets represented as flat arrows. The critical Cys at the catalytic center is also shown

The cysteine residue of the active site (Cys 86 in UBC4) lies in a long loop that connects S4 to α2. Many of the most highly conserved E2 residues surround the active site cystein (Cook *et al*., 1993; Jiang and Basavappa, 1999). Some of these residues interact with ubiquitin, and others presumably interact with E1. Most of the poorly conserved E2 residues cluster on the face of the molecule that is opposite the active site.

Some of these residues may have diverged under low selective pressure, but others are likely to mediate interactions relevant to the specific functions of individual E2s.

Ubiquitin activated by E1 enzyme is transferred to the active site cysteine of the E2 in a thiolester linkage E2-S-Ub. Isopeptide bound formation results from attack on the E2 bound ubiquitin by a lysine residue of the substrate. This final reaction usually requires the participation of an E3 that binds both the E2 ubiquitin thiolester and the protein targeted for ubiquitination (Hershko and Ciechanover, 1998). Lysine residues within ubiquitin may also serve as substrates, leading to the formation of diubiquitin and, eventually, polyubiquitin chains (**Figure I.3**).

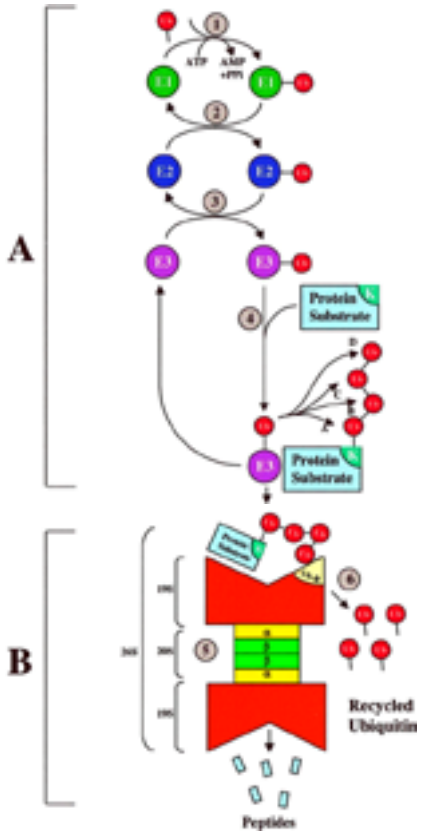


Figure I.3. The ubiquitin proteasome pathway. (A) Conjugation of ubiquitin to target molecule. (B) Degradation of the tagged substrate by 26 S proteasome. (1) Activation of ubiquitin by E1. (2) Transfer of activated ubiquitin from E1 to a member of the E2 family. (3) Transfer of activated ubiquitin from E2 to substrate- specific E3. (4) Formation of substrate-E3 complex and biosynthesis of a substrate-anchored polyubiquitin chain. (5) Binding of the polyubiquitinated substrate to the ubiquitin receptor subunit in the 19 S complex of the 26 S complex. (6) Recycling of ubiquitin via the action of isopeptidases. (from Ciechanover, 1998)

I.4.3. Ubiquitin protein ligases (E3s)

In any given organism, there is usually a single class of E1 enzymes, but there are many classes of E2s and multiple families of E3 or E3 multiprotein complexes. Specific E3s appear to be responsible mainly for the substrate selectivity of ubiquitin-protein ligation. They do so by binding specific protein substrates that contain specific recognition signals. In some cases, binding of the substrate protein to an E3 is direct or it can also happen via an adapter protein. The E3 enzymes do not discriminate between monoubiquitin or polyubiquitin chains conjugated to the E2 proteins.

Different types of E3s may carry out the transfer of the ubiquitin to the substrate protein by two different mechanisms. In some cases, such as with the HECT-domain family of E3 enzymes, ubiquitin is first transferred from an appropriate E2 to an active site Cys residue of the E3 enzyme. This E3-ubiquitin thiolester is the donor for amide bond formation with the protein substrate (**Figure I.4**).

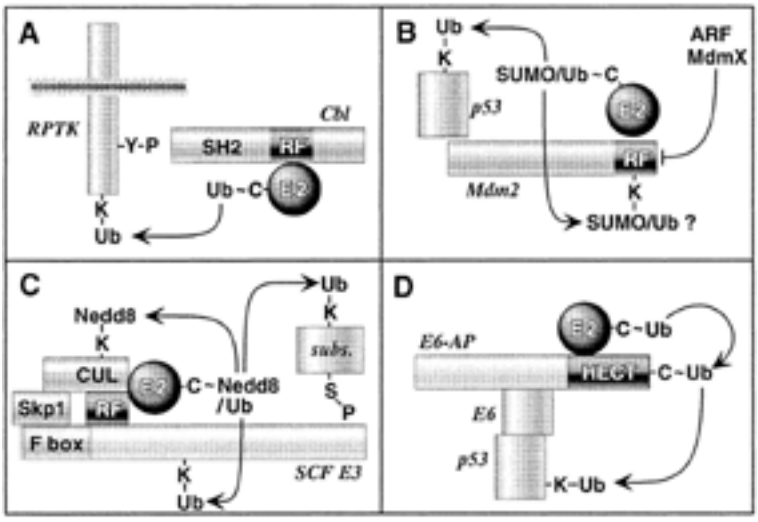


Figure I.4. Variety of action of E3, proteins that recognize specific substrates and mediate ligation of the latter to ub- or Ub-like proteins. (**A**) Ubiquitination of active RPTK by Cbl. (**B**) Ubiquitination of p53 by Mdm2 and modification of Mdm2 itself with SUMO-1 or Ub. (**C**) Multisubunit SCF-type E3: ubiquitination substrate and F box protein subunit; modification of Cullin subunit with Nedd 8. (**D**) Ubiquitination of p53 by HECT-type E3 ligase E6-AP: dependence on human papilloma virus E6 protein and formation of intermediate Ub thioester with E3. Abbreviations: Y, tyrosine; K, Lysine; S,serine; C, cysteine; P, phosphate group; ~, thiol-ester bond (From Jozaeiro *et al*., 2000).

In other families of E3 enzymes, E3 ubiquitin thiolester formation can not be demonstrated. Since E3 enzymes bind cognate E2s tightly and they also bind their appropriate protein substrate, ubiquitin can be directly transferred from E2 to the protein substrate (**Figure I.4**).

Therefore, the two known E3 enzymes are members of two families, HECT domain E3s and RING E3s (**Table I.1**). The catalytic modules of the two families are unrelated in sequence or structure (Borden, 2000; Pickart *et al*., 1994). Nonetheless, certain E2s interact well with E3s from both families. UbcH7, for example, efficiently binds to the HECT domain enzyme E6-AP (Huang *et al*., 1999; Nuber *et al*., 1996; Kumar *et al*., 1997) and to several known or presumptive RING finger containing E3 (Zheng *et al* 2000; Moynihan *et al*., 1999; Yokouchi *et al*., 1999).

I.4.4. UEV Ubiquitin conjugating enzymes E2 Variants a new class of protein which participate in K63 polyubiquitin chain formation

UEV proteins (ubiquitin conjugating enzymes E2 variant) form a new class of proteins related to the ubiquitin conjugating enzymes (E2), very similar in sequence and tridimensional structure to E2 enzymes. These proteins were first named CROC-1 proteins and shown to participate in intracellular signaling pathways involved in induction of transactivation of the human c-fos proto-oncogene promoter (Rothofsky and Lin, 1997). Subsequently, UEV proteins have been demonstrated to participate in different cellular processes, including cell cycle regulation (Sancho *et al*., 1998), DNA repair (Thomson et al., 1998; Broomfield et al., 1998) and signal transduction (Deng et al., 2000).

E2	Cognate E3	Function(s)/Substrates[b]
S. cerevisiae		
Ubc1	Unknown	Sporulation; essential in ubc4Δubc5Δcells
Ubc2/Rad6	Ubr1(RING) Ubr1 Rad 18 (RING) Unknown	N-end proteolysis; substrate: cohesin Regulation of peptide import; substrate: Cup9 Postreplicational DNA repair; DNA damage response [substrate: HO endonuclease?] Meiosis/meiotic cell growth, substrate: histone H2B
Ubc3/Cdc34	SCF^{Cdc4} SCF^{Grr1} SCF^{Met30}	Regulates GI/S transition; substrates: Sic1, others Cell cycle; substrate: G1 cyclins Transcriptional regulation; substrate: Met4
Ubc4,Ubc5	Unknown Ufd4(HECT) Rsp5(HECT) Rsp5	Turnover of short-lived and abnormal proteins Degradation of ubiquitin-fused substrates DNA damage response; substrate: pol II LS Endocytosis; substrates: Gap1, Fur4, Ste2
Ubc6	Unknown	Integral protein of endoplasmic reticulum (ER) membrane; substrate: mutant Sec61; associates with Ubc7
Ubc7	Der1 (RING) Unknown	ER degradation; localized to ER membrane via Cuel; associates with itself and Ubc6 Stress/cadmium resistance
Ubc8	Unknown	Glucose-induced proteolysis, substrate: fructose-1,6-bisphosphatase
Ubc10/ pas2	Unknown	Peroxisome biogenesis
Ubc11	Unknown	Unknown; similar to E2-C/UbcH10
Ubc13	Rad5? (RING)	Postreplic. DNA repair; complexed with Mms2
Humans		
$E2_{14K}$/UbcH1	$E3_{\alpha}$Ubr1 Unknown	Related to ScUbc2;N-end rule proteolysis; up-regulated in wasting muscle Implicated in spermatogenesis
UbcH5A	Hect/RING E3s	Interacts with HECT and RING E3s; murine ortholog essential
UbcH5B	Hect/RING E3s	Interacts with HECT and RING E3s; induced in spermatocytes
UbcH5C	RING E3s	Interacts with SCF E3; induced in spermatocytes
$Ubc4_{testis}$	Unknown	Expressed only in testis
UbcH6	Unknown	Interacts with some HECT E3s
UbcH7/E2-F1	Hect/RING $E3_S$	Interacts with HECT and RING $E3_S$
UbcH8	Hect/RING $E3_S$	Interacts with HECT and RING $E3_S$
Ubc13	TRAF6 (RING)	IKK activation; complexed with Uev 1A
NCUBE1	Unknown	Related to ScUbc6
$E2_{17K}$	Unknown	Related to ScUbc7; ER degradation
E2-EPF	Unknown	Highly expressed in keratinocytes
E2-C/UbcH10	APC	Mitotic cyclin degradation; similar to ScUbc11
$E2_{20K}$/UbcH2	Unknown	Related to ScUbc8; up-regulated in erythroid differentiation
$E2_{25K}$	SCFE3s	Assembles K-48-linked polyubiquitin chains in vitro
$E2_{35K}$		Related to ScUbc3; numerous substrates identified
$E2_{230K}$	None?	Up-regulated in erythroid differentiation; fused E2-E3?
BRUCE	Unknown	528 kilodaltons; Golgi-localized; BIR domain

Table I.1. Selected E2 enzymes with their corresponding E3s in yeast and human and their possible function.

Hoffman and Pickart (1999) showed that the yeast UEV protein (that in yeast has been called Mms2; Broomfield *et al*., 1998) associates physically to the canonical E2, Ubc13, and forms an heterodimer with a central role in the assembly of K63-linked

polyubiquitin chains onto a target protein both in yeast and in human cells (Hoffman and Pickart, 1999; Deng *et al*., 2000). Both the yeast and the human UEV-Ubc13 heterodimers interact with a RING domain protein that may be either a cognate E3 or a substrate of the respective E2-UEV heterodimers (Deng *et al*., 2000; Ulrich and Jentsch, 2000). Although UEV proteins were initially thought to be catalytically inert (Broomfield *et al*., 1998; Koonin and Abagyan, 1997; Sancho *et al*., 1998), the Mms2-Ubc13 and UEV1A-Ubc13 heterodimers catalyse the assembly of K63-chains in the absence of any other factor besides E1 (Deng *et al*., 2000; Hofmann and Pickart, 1999). The crystal structure of this complex (**Figure I.5**) shows that UBC13 adopts a canonical E2 fold, consisting of four-stranded antiparallel β-sheet flanked by four α-helices (VanDemark *et al*., 2001; Moraes *et al*., 2001).

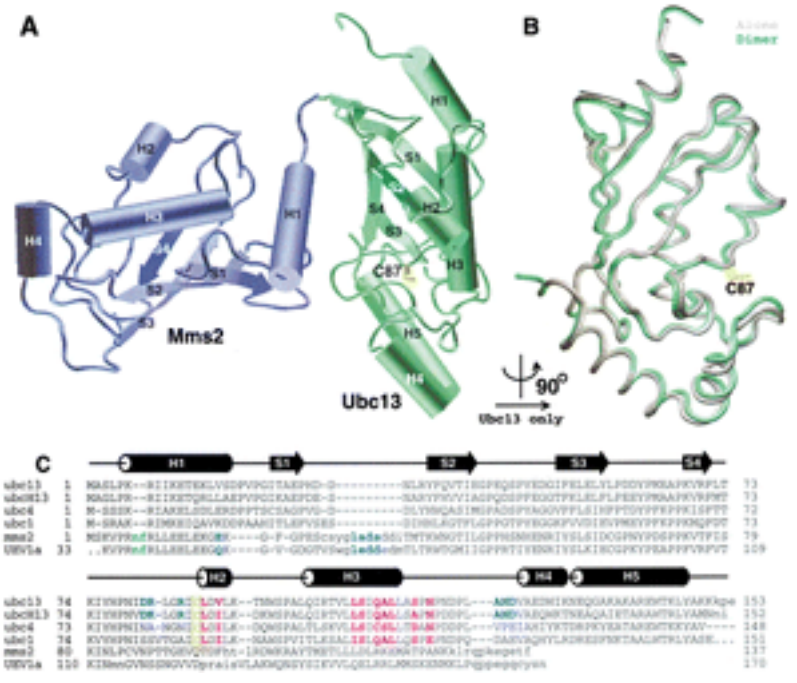


Figure I.5. Structure of the Mms2/Ubc13 Complex
(A) Cartoon representation of Mms2/Ubc13 complex with Mms2 in blue and Ubc13 in green. Helices and strands are labeled. The side chain of the active site residue Ubc13-Cys87 is shown in yellow.
(B) Alignment of Ubc13 from complex, green, with Ubc13 alone structure, gray. Only Cα are shown, plus Ubc13-Cys87 for orientation.
(C) Primary sequence alignment of selected E2s and UEVs. The E2 active site cysteine position is outlined by a tan box; the Mms2 insertion containing F8 is shown in green. Ubc13 surface residues that are part of channel 1 are shown in magenta. Residues of UEVs and Ubc13s that make up channel 2 are shown in blue. (From VanDemark.*et al*., 2001)

The UEV protein adopts a fold similar to UBC13 and other E2 enzymes, but with several notable differences. The seven amino-terminal residues of UEV, which participate in complex formation with UBC13, adopt a conformation that is atypical of E2s. This conformation allows the N terminus to participate in the dimer interface (**Figure I.5**).

The discovery of the role of UEV-Ubc13 in post-replicative DNA repair heterodimer was a big help in understanding its mode of action (Ulrich and Jentsch, 2000; Broomfield *et al*., 1998, Thomson *et al*., 1998; Hoffman and Pickart, 1999). The yeast UEV protein or Mms2, is required together with Ubc13 for RAD6/RAD18-dependent postreplicative DNA repair in yeast (Broomfield *et al*., 1998; Hofmann and Pickart, 1999). In this mechanism, the two chromatin-associated RING finger proteins, RAD18 and RAD5, play a central role in mediating physical contacts between Ubc13-UEV complex to DNA by means of its RING finger domains. Interaction between the two RING finger proteins thus promotes the formation of heterodimeric complex with the two distict ubiquitin conjugating activities of the E2, RAD6 and the UBC13-UEV can be closely coordinated (Ulrich and Jentsch, 2000).

A second major breakthrough in the understanding of the function of these two proteins came from studies on the mechanisms of activation of the kinase cascades after ligand-dependent signaling through the TNFα and IL-1 receptors (Deng *et al*., 2000). The quiescent forms of the transcriptional regulator NFκB can be rapidly activated by a remarkable large number of extracellular signals, including a number of cytokines such as tumor necrosis factor-α (TNF-α) and interleukin 1 (IL-1), processes such as antigen-dependent T-cell activation, viral or bacterial infection, and stress such as UV irradiation and reactive oxygen. As a consequence of an extracellular signal the receptors for TNFa or IL-1 interact with the adapter protein TRAF6, which in turn

activates the protein kinase IκB kinase (IKK) (Cao *et al*., 1996; Ishida *et al*., 1996; Lomaga *et al*., 1999; Naito *et al*., 1999). Very recently, it was shown that the synthesizes of K63-Ub chains mediated by the human UEV1A/Ubc13 heterodimer is required for the IKK activation (Deng *et al*., 2000). The binding of the interleukin-1 to the extracellular domain of the interleukin-1 receptor leads to the recruitment of the IRAK protein. IRAK is an adapter protein that binds TRAF6, which forms an homodimer and ubiquitinates itself. TRAF6 interacts with the UEV1A/Ubc13 heterodimer through its RING finger protein domain, and this heterodimer polyubiquitinates TRAF6 with a K63 polyUb chain. (Finley, 2001). The formation of K63 polyUb chains on the TRAF6 activates in an unknown manner the kinase TAK1 (TGF-beta-activated kinase 1) which in turn phosphorylates and activatesIKK. (Wang et al., 2001; Finley, 2001). The intervention of the heterodimer UEV-UBC13 in NFκB pathway is summarized in **Figure IV.3**.

I.5. Proteasome degradation

I.5.1. The 26S proteasome

In eukaryotic cells, most proteins in the cytosol and nucleus are degraded via the ubiquitin-26 S proteasome pathway **(Figure I.3*)***. The 26 S proteasome is a 2.5 MDa molecule machine built from approximately 31 different subunits, which catalyzes protein degradation. It contains a proteolytic core complex (the 20S proteasome), capped at one or both ends by 19S regulatory complexes, which recognize ubiquitinated proteins. The 19S regulatory complexes are also implicated in unfolding and

translocation of ubiquitinated targets into the interior of the 20S complex, where they are degraded to oligopeptides. The structure, assembly and enzymatic mechanisms of the 20S complex have been elucidated, but the functional organization of the 19S complex is less well understood. Most subunits of the 19S complex have been identified; however, specific functions have been assigned to only a few. A low-resolution structure of the 26S proteasome has been obtained by electron microscopy, but the precise arrangement of subunits in the 19S complex is unclear (Voges et al., 1999; Zwickl et al., 1999).

I.5.2. The 20S proteasome

The 20 S proteasome is ATP-independent and only degrades unfolded polypeptides, whereas the 26S proteasome degrades folded proteins in an ATP-dependent manner. Since its first sighting in eukaryote extracts in 1968 (Harris, 1968), the 20 S proteasome has accumulated more than 20 different names - including cylindrin (Harris, 1988), alkaline protease (Hase *et al*., 1980), multicatalytic proteinase MCP (Orlowski and Wilk, 1988), ingensin (Ishiura *et al*., 1985), prosome (Schmid *et al*., 1984), low-molecular-weight protein LMP (Monaco and McDevitt, 1984), macropain (McGuire and DeMartino, 1986) - and many functions illustrating clearly its important role in many cellular processes (Confalonieri, et al., 1995; Zwickl et al., 1999).

I.6. Ubiquitin-like proteins

The number of genes known to encode ubiquitin related proteins increases day after day (**Figure I.6**). The ubiquitin homology (UbH) family is structurally and

functionally diverse. The high evolutionary conservation of ubiquitin enabled the discovery of many ubiquitin-related proteins.

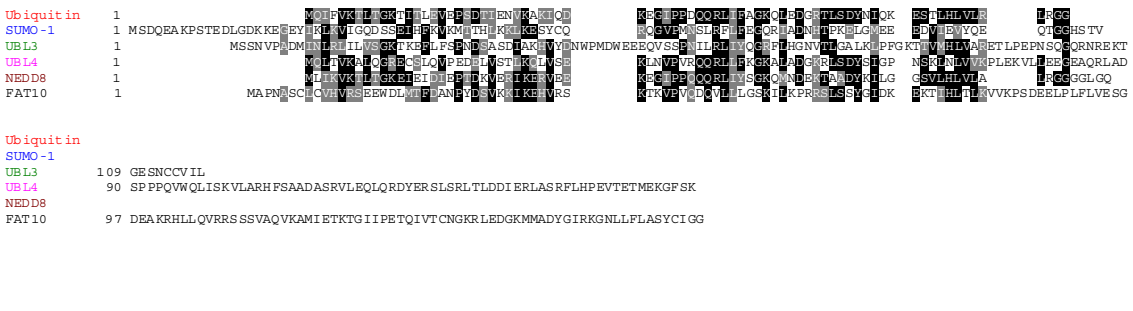


Figure I.6. Sequence alignment of Ubiquitin with some Ubiquitin-like proteins.

Some, such as parkin, which is implicated in the pathogenesis of certain forms of Parkinson's disease (Kitada *et al*., 1998), are larger than ubiquitin and possess ubiquitin-like domains that display only slight homology to ubiquitin, they lack the carboxy-terminal Gly and cannot be conjugated. Their physiological significance still obscure.

A second group contains small proteins with higher degree of homology to ubiquitin which are involved in post-translational and single or multiple modification of target proteins with serves non-proteolytic purposes (Hochstrasser, 1998). UCRP is an interferon-inducible 15 kDa protein that resembles two tandem repeats of ubiquitin and may be involved in targeting proteins to the cytoskeleton (Loeb and Haas, 1994).

Small ubiquitin-related modifier.1 (SUMO-1) is an 11.5 kDa polypeptide involved in targeting RanGAP1 to nuclear pore complex (NPC) protein RanBP2 (Mahajan *et al*., 1997). RanBP2 is a GTPase required for the transport of proteins and ribonucleoproteins across the NPC. Its guanosine 5'-triphosphate/diphosphate

(GTP/GDP) cycle regulated by RanGAP1. Localization of Ran GAP1 to the NPC is dependent on its single, stable covalent modification by SUMO-1. The SUMO-1-RanGAP1 conjugate generates a complex with RanBP2 that is essential for the function of RanBP2. SUMO-1 modification of IκBα stabilizes the protein and inhibits NF-κB activation (Desterro *et al*., 1998). Here SUMO-1 acts antagonistically to ubiquitin by generating a degradation resistant protein. SUMO-1 is identical to Sentrin involved in protecting cells against anti-Fas/TNFα-induced apoptosis and, like ubiquitin, can generate multiply modified conjugates with cellular proteins (Kamitani *et al*., 1997a).

NEDD8 is a mammalian ubiquitin-like protein that is developmentally down regulated and is expressed in high levels in post-mitotic cells characterized by high protein turnover rate, such as skeletal and heart muscle (Kamitani *et al*., 1997b).

RUB1 is a yeast ubiquitin like protein that was found to modify CDC53/Cullin (Liakopoulos *et al*., 1998), a common subunit of the SCF ubiquitin ligase complex (Skp1-cullin-F-box protein ligase complex). While the modification of the CDC53 does not affect its stability, it may influence the activity of SCF or its specificity towards its different substrates.

Agp12 is another yeast ubiquitin-like protein. Its Agp7 (E1)- and Agp10 (E2)-mediated conjugation to Agp5 is essential for autophagy (Mizushima *et al*., 1998). Conjugation of the ubiquitin-like proteins raises several questions related to the chemical nature of the adduct, the identity of the conjugating enzyme(s) and the specifity of substrate targeting. The C-terminal domain of SUMO-1 is processed proteotically at residue 97 (TGG^{97}▼H^{98}STV) to generate a free-G^{96}G^{97}-COOH that, like the C-terminal Gly^{76} of ubiquitin, is essential for conjugation. Similarly, RUB1, SMT3 and NEDD8 are also processed to yield C-terminal –Gly-Gly. Activation of SMT3 requires at least three proteins: AOS1, UBA2 and UBC9. AOS1 and UBA2 are

homologous to the N-terminal and C-terminal domains of E1, respectively, and are probably heterodimerizing to generate an active E1 (Johnson *et al*., 1997).

UBC9 can serve as the E2 in the modification reaction (Schwarz *et al*., 1998). Conjugation of RUB1 requires ULA1/UBA ULA1/UBA3 that serve as a heterodimeric E1, and UBC12 as an E2 (Liakopoulos *et al*., 1998). While conjugation of the known ubiquitin-related proteins does not require E3, it is not clear that this is the case for all of these modifications. The requirement for E3s probably depends on the breadth of spectrum of substrates, and the functions of each of the modifying proteins.

I.7. Ubiquitin Hydrolases

Release of ubiquitin is necessary to "proofread" mistakenly ubiquitinated proteins and plays an essential role in two processes, degradation and ubiquitin biosynthesis. During degradation it is important to release ubiquitin from Lys residues of end proteolytic products to maintain the free ubiquitin pool in the cytoplasm. Therefore, polyubiquitinated chains have to be also hydrolised trough their Gly-Lys bounds to disassemble polyubiquitin chains. The second process is ubiquitin biosynthesis. Ubiquitin is synthesized in a variety of functionally distinct forms. One of them is a linear, head-to-tail polyubiquitin precursor. Release of the free molecules involves specific enzymatic cleavage between the fused residues. The last ubiquitin moiety in many of these precursors is encoded with an extra C-terminal residue that has been removed in order to expose the active C-terminal Gly. In a different precursor, ubiquitin is synthesized as an N-terminal fused extension of two ribosomal proteins and serve as a covalent "chaperone" that targets them to the ribosome. After their incorporation into the ribosomal complex, ubiquitin is cleaved.

In general, the recycling enzymes are thiol proteases that recognize the C-terminal domain/residue of ubiquitin (Hoschtrasser, 1996; Wilkinson, 1997). All deubiquitination reactions require accurate proteolytic processing at the C-terminal glycine of ubiquitin. The enzymes responsible for these reactions have been called isopeptidases (Matsui *et al*., 1982), ubiquitin carboyl-terminal hydrolases (Pickart and Rose, 1985), ubiquitin thiolesterases (Rose and Warms, 1983), or ubiquitin-specific processing proteases (Tobias and Varshavsky, 1991). The general acronym proposed is DUB enzymes (DUBs) (for deubiquitinating enzymes). They are divided into two classes: ubiquitin C-terminal hydrolases (UCH) and ubiquitin-specific proteases (UBPs; isopeptidases).

UCHs are about 25 kDa enzymes which are involved in co-translational processing of pro-ubiquitin gene products and in the release of ubiquitin from adducts with small molecules, such as amines and thiol groups. UBPs are about 100 kDa enzymes which catalyze release of ubiquitin from conjugates with cellular proteins or from free polyubiquitin chains. Since a large number of UBPs are encoded by the yeast genome and higher eukaryotes, it is thought that some of them may have specific functions such as recognition of distinct tagged substrates. In accordance with the broad spectrum of their functions, they also differ in their characteristics. While some are free, others are subunits or associated with the 19S proteosome complex. Some require ATP for their activity while others act in an energy-independent manner. Their mechanisms of action also differ, since some are sensitive to ubiquitin aldehyde, while others are not. De-ubiquitinating enzymes can either accelerate or inhibit proteolysis. By removing ubiquitin moieties from mistakenly tagged proteins they inhibit proteolysis. Stimulation of proteolysis can be mediated by release of free ubiquitin from biosyntetic precursors and terminal proteolytic products restoring therefore the cellular ubiquitin pool.

Ubiquitin can also be released from polyubiquitin chains that bind to the 26S proteasome and inhibit the later, or by "editing" polyubiquitin chains and "fitting" them better for recognition by 26S proteasome.

Recent experimental evidences indicates that some of these enzymes play an essential role in specific processes and must therefore target specific substrates. The *Drosophila melanogaster* FAT FACETs (FAF) gene affects eye development (Huang *et al*., 1995). Mutant FAF flies have more than eight photoreceptors in each of the compound eye units. The protein is probably involved in generating the inhibitory signal sent by photoreceptor cells to undifferentiated surrounding cells, to stop differentiation and migration to facet unit. Due to the fact that inactivation of FAF can be suppressed by another mutation in a proteasome subunit, it appears that the enzyme stabilizes some unidentified protein(s). A specific serotonin-inducuble UCH has been implicated in activation of cAMP-dependent protein kinase A (PKA) in *Aplysia via* stimulation of the degradation of the inhibitory regulatory subunit of the enzyme (Hegde *et al*., 1997). Degradation is initiated by cAMP that leads to dissociation of the holoenzyme and release of free R subunits. PKA-dependent phosphorylation of a variety of proteins in sensory neurons is responsible for a broad array of morphological changes on the synapse that produce the continuous presynaptic facilitation necessary for long-term behavioral sensitization.

UBP3 has been implicated in gene silencing (Moazed and Johnson, 1996). Actively transcribed genes can be silenced following positioning near heterochromatic regions. SIR4 is one trans-acting factor that is required for the establishment/maintenance of silencing. One identified SIR4-interacting protein is UBP3, an inhibitor of silencing that acts by either stabilizing an inhibitor or by removing a positive regulator.