CHAPTER 1

Introduction
1. INTRODUCTION

A thorough balance of protein synthesis and degradation is prerequisite for survival of any cell. Degradation of misfolded, damaged, or otherwise malfunctioning proteins is certainly of utmost significance for maintaining cellular homeostasis. Degradation of cellular proteins is a highly complex, controlled and tightly regulated process that plays major roles in a variety of basic cellular pathways during both cell life and cell death (Balch et al., 2008).

Ubiquitin is used by cells as a tag covalently attached to proteins in a polymeric form and marks them for degradation (Ciechanover et al., 1980). Ubiquitin along with the enzymes catalyzing ubiquitination, namely ubiquitin activating enzyme E1, ubiquitin conjugating enzyme E2, ubiquitin ligase E3 and the multi-subunit protease complex proteasome, which act in succession constitute the ubiquitin proteasome system (Varshavsky, 1997; Weissmann, 1997). The ubiquitin proteasome system regulates removal of many intracellular proteins, including those which control cell cycle progression, apoptosis, signal transduction and induction of the inflammatory response.

Subsequently many functions of ubiquitin other than its role in proteasomal degradation have been discovered, which include DNA repair, regulation of transcription and translation, cell signaling, endocytosis and autophagy. Hence, the role of ubiquitin in posttranslational modification in addition to the one in protein degrading system is gaining importance. Ubiquitination can therefore be compared to other post-translational modifications like phosphorylation, methylation and acetylation.

The essential components of ubiquitin structure, ubiquitination system along with its functions and UPS are described in the initial part of introduction. Failure of
UPS, different diseases associated with it and targeting UPS for drug design are highlighted in the later part of introduction.

1.1. The structure of ubiquitin

Ubiquitin is a small, globular and single domain protein. It is made up of 76 amino acid residues and has a molecular weight of 8565 Da. It also does not contain any cofactors. Because of these properties, it has been a good model system for the study of protein folding and stability. Its 3D structure was first determined using X-ray crystallography (Vijay-Kumar et al., 1987; Vijay-Kumar et al., 1985). It contains five strands of β-sheet harbouring two parallel and three antiparallel strands, 3.5 turns of α helix and 3_{10} helix characterized by a particular fold denoted as SSHSSS (Fig. 1.1). The compact structure of ubiquitin has nine reverse turns and two β-bulges residing at N and C terminus of the protein (Vijay-Kumar et al., 1987; Vijay-Kumar et al., 1985).

![Fig. 1.1. Secondary and tertiary structure of ubiquitin from the 1.8 Å resolution. Shown in yellow are the five β-strands that make up the mixed parallel-anti-parallel β-sheet, shown in red are the major a-helix and the 3_{10}helix (Jackson, 2006 ).](image)

The structure was found to be tightly packed and globular with β-grasp fold, in which mixed β-sheets pack against an α-helix to form highly stable hydrophobic core.
Due to this, ubiquitin is found to be stable from pH 1.2 to 8.5 and temperature from 23° to 80°C (Lenkinski et al., 1977). Its C-terminal end comprising LRLRGG amino acid residues protrudes from the core to interact with ubiquitin itself and other proteins.

### 1.1.1. β-bulges in ubiquitin

β-bulges are regions in a β sheet where two residues in one β strand form branched hydrogen bonds with a single residue of a second strand (Milner-White, 1987). The extra residues bulge out from the β strand, giving it the name β-bulge (Richardson et al., 1978). The two residues bulging out on the first strand are labelled “1” and “2”, and the residue on the other strand is labelled “X” (Fig. 1.2).

![Image of β-bulge arrangement](image)

**Fig. 1.2. Arrangement of β-bulge where one β-strand bulges out from the register (Axe et al., 1999)**

Bulges are divided on the basis of conformation of the residues involved in the structure and the type of amino acid they have inside the bulge. The β-bulges are classified into five types, namely classic, G1, wide, bent and special types. Classic and wide include both parallel and antiparallel β bulges and G1 bulges are mostly antiparallel. G1 bulges are further classified into: (1) G1G type with glycine at position first of the bulge, (2) G1T type with glycine at first position and a Type I’ or Type II β turn between position 2 and any other residue, with glycine at position i+2 of turn (Verma et al., 2002). G1A type has any amino acid (excluding glycine) at first position (Chan et al., 1993).
G1 type β-bulge displays glycine in position 1 and it has \( \varphi, \psi \) values of 85°, 0° which is favourable only for glycine (Vijay-Kumar et al., 1987). Position 2 of G1 bulge is within the usual β region, but centred on \( \varphi = -90°, \psi = 150° \). G1 type bulges are found within an interlocking structure in which the glycine in position 1 of the G1 bulge is also the required glycine in position 3 of a type-II tight turn (Vijay-Kumar et al., 1987).

Ubiquitin contains two β-bulges. The first G1 β-bulge is present at the N-terminal end of the protein on two antiparallel β-strands and involves Gly10 (1), Lys11 (2) and Thr 7 (N). It is present along with Type I turn. Two peptides covering the first β-bulge of ubiquitin were found to fold into β-hairpin conformation autonomously in aqueous solvent and also in methanol (Cox et al., 1993; Zerella et al., 1999) establishing the potential of first β-bulge of ubiquitin to act as an initiation site for its folding. Elimination of Gly from this turn results in loss of structure (Chen et al., 2001). (Cox et al., 1993)

Importance of strand-strand interaction in hairpin formation has been investigated by mutating Thr at 9th position to Asp. The results underscored the role of charge-charge interaction in the stability of ubiquitin (Zerella et al., 2000).

The second G1 β-bulge is present between the C-Terminal end and the N-terminal region of protein on two parallel β-strands and involves Glu64 (1), Ser65 (2) and Glu2 (N). It is present along with a type II turn. Glu64 is the third and first residue in a type II turn and in parallel β-bulge respectively (Vijay-Kumar et al., 1987). Presence of Glu in 64th position is unusual with \( \Phi \) and \( \Psi \) angles which are not under the allowed region of Glu in Ramachandran plot. Generally the first residue in a parallel β bulge is a Gly (Chan et al., 1993). The reason for which nature selected Glu instead of Gly and conserved through evolution has been answered by replacing Glu at 64th position by more preferred Gly. Replacement caused less helicity and increased
stability in the structure of ubiquitin (Mishra et al., 2009). Two more residue of second β bulge were also mutated to more preferred amino acid residues, which exhibited minor alteration in the structure and significantly affected the functionality in vivo (Mishra et al., 2011). Ile61 has been reported to show protection early on in kinetic refolding experiments (Briggs and Roder, 1992) and Lys63 has been reported to play an important role in DNA repair mechanism (Spence et al., 1995).

1.2. Stability of ubiquitin

Ubiquitin’s valine, leucine, isoleucine, and methionine residues are buried in the hydrophobic core of the molecule, contributing to the extreme stability of ubiquitin (Thomas and Makhatadze, 2000). Site directed mutagenesis is successful in elucidating functions of many residues contributing to the stability of ubiquitin (Loladze et al., 2001). Makhatadze et al., in 2003 showed different salts increase the thermostability of ubiquitin through anion binding.

Amino acid replacement studies involving lysine to arginine substitutions have established that surface charges do not play a role in making ubiquitin stable (Makhatadze et al., 2003).

1.3. Folding of ubiquitin

A protein always folds so that it achieves the lowest possible energy (Anfinsen, 1973; Hagen, 2007). Theoretically a chain of 100 amino acids will take $10^{40}$ years to attain a thermodynamically stable native state if it has to run through all possible stable conformations one after another (Levinthal, 1968). Protein folding is often described by folding funnels and energy landscapes (Bryngelson and Wolynes, 1987; Dill, 1985). In contrast, protein folding in vivo takes several seconds to a few minutes. The rate and efficiency of folding of a protein in vivo can be attributed to assistance
of chaperones, post-translational modifications and various inter and intramolecular interactions.

![Folding funnel representing various state of energy level](image)

**Fig. 1.3. Folding funnel representing various state of energy level (Se’rgio T. Ferreira, and Felice., 2001)**

Several models have been proposed to explain the process of protein folding, such as the Framework model (Karplus and Weaver, 1976) the diffusion collision model (Kim and Baldwin, 1982) and the hydrophobic collapse model (Chan and Dill, 1990).

Folding studies of ubiquitin have indicated that the N-terminal β-sheet and α-helix are formed early. Even though the C-terminal half of the protein shows slow folding kinetics, residues 59, 61, and the beginning of a $3_{10}$-helix also exhibit protection in the early phase (Briggs and Roder, 1992). Studies on partially folded state of ubiquitin revealed formation of native secondary structural elements in the N-terminal half (Ub 1-21 and Ub 1-35). The C-terminal half (Ub 36-76) which has β strand character undergoes a transition to helical state (Brutscher et al., 1997; Harding et al., 1991; Stockman et al., 1993)

Current model for folding of ubiquitin proposes that the N-terminal portion of the protein (spanning residues 1-35) serves as an autonomously folding chassis,
1.4. Ubiquitin – The highly conserved protein (Sequence Homology)

Among eukaryotes, the sequence ubiquitin protein is highly conserved (Gavilanes et al., 1982; Schlesinger and Goldstein, 1975; Watson et al., 1978). This can be due to its central role in the cellular metabolism. In the sequences of ubiquitin of yeast, human and oat (Vierstra, 1986) replacement of amino acids is seen in only three places. The sequences of yeast, plant and human ubiquitins are shown Fig. 1.4, highlighting the positions of the residues where there are differences.

**YEAST UBIQUITIN PROTEIN SEQUENCE**

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MQIFVKLTGKTITLEVESSDTIDNVKSKIQDEGIPPDQQRLIFAGKQL</td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

**PLANT UBIQUITIN PROTEIN SEQUENCE**

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MQIFVKLTGKTITLEVESSDTIDNVKAKIQDEGIPPDQQRLIFAGKQL</td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

**HUMAN UBIQUITIN PROTEIN SEQUENCE**

<table>
<thead>
<tr>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MQIFVKLTGKTITLEVEPSDTIENVKAKIQDEGIPPDQQRLIFAGKQL</td>
</tr>
<tr>
<td>20</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1.4. The amino acid sequences of yeast, plant and human ubiquitins. The residues which were changed have been highlighted
1.5. Ubiquitin gene family

In *Saccharomyces cerevisiae*, ubiquitin is expressed from four different genes namely *UBI1*, *UBI2*, *UBI3* and *UBI4*, as a natural translational fusion to either ubiquitin or other proteins. Ubiquitins encoded by *UBI1*, *UBI2* and *UBI3* are C-terminally fused to ribosomal subunit L40 and S31 respectively (Finley et al., 1987). *UBI4* is expressed as polymeric head to tail fusion containing five ubiquitin repeats (polyubiquitin). Ubiquitin C-terminal hydrolase cleaves the fusion proteins at the C-terminal end of ubiquitin, liberating free ribosomal proteins and individual ubiquitin monomers. During exponential phase all four ubiquitin genes are expressed while during stationary phase the expression of *UBI1* and *UBI2* is suppressed and expression of *UBI4* gene increases (Fig. 1.5). Polyubiquitin gene *UBI4* is an important component of the stress response system and it helps the cell to cope up with stress. The ‘heat shock box’ element upstream of the coding region senses adverse conditions like various stresses including heat stress, nutritional stress (Ozkaynak et al., 1987), UV stress and antibiotic stress (Finley et al., 1994). The essential function of *UBI4* is to provide more ubiquitin during stress conditions. Null phenotype of ubiquitin is lethal but *S. cerevisiae* can survive if the polyubiquitin gene *UBI4* is deleted from the system. The *ubi4* deletion mutant of *S. cerevisiae* becomes hypersensitive to stress conditions (Finley et al., 1987).

![Fig. 1.5. Three ubiquitin precursors in *S. cerevisiae* ubiquitin fused to L40, S31 and itself (Finley et al., 1987)](image_url)
1.6. Ubiquitin lysine linkage – The Molecular Signal

There are seven lysine residues in ubiquitin, namely K6, K11, K27, K29, K33, K48 and K63, all of which participate in ubiquitination (Johnson et al., 1995; Kim et al., 2007; Peng et al., 2003). Ub chains that form through Lys 48 or K48 serve as a tag for degradation of substrate protein (Hershko and Ciechanover, 1998). Different Ub linkages can be formed through Lys 6, Lys 11, Lys 27, Lys 29, Lys 33 and Lys 63 leading to a variety of conformations of Ub chain, which function as different molecular signals in the cell. Depending on the type of linkages Ub chains have been divided into several classes (Fig. 1.6). Ubiquitin with mixed linkages were also identified in a low proportion compared with conjugated mono Ub or different types of homotypic chains (Peng et al., 2003).

![Ubiquitin Lysine Linkages](image)

(I)

![Polyubiquitin chains displaying different kinds of linkages](image)

(Francisca E. Reyes Turcu. et al.)

Fig. 1.6. Polyubiquitin chains displaying different kinds of linkages

(I) shows presence of different lysine residues in ubiquitin, (Francisca E. Reyes Turcu. et al.) shows different classes of uniquitination (A) homotypic, showing only one kind of linkage (B) Mixed, showing multiple types of linkages (C) Heterologous,
involving ubiquitin and ubiquitin like proteins and (D) showing multiple monoubiquitinations (Fumiyrio Ikeda and Dikic., 2008).

1.7. Role of different lysine linkages inside the cell

Lys48 linked chains are the mainly responsible for delivering proteins to proteasome and K48R mutation is lethal in yeast (Chau et al., 1989; Finley et al., 1994; Hochstrasser et al., 1991). Lys63 linked chains participates in DNA repair (Spence et al., 1995; Ulrich, 2002), the inflammatory response (Sun and Chen, 2004), protein trafficking (Hicke and Dunn, 2003) and ribosomal protein synthesis (Spence et al., 1995). Lys11 linked chains can signal proteasome degradation in vitro (Baboshina and Haas, 1996). Lys6 linked chains shows conjugation to tumor suppressor protein BRCA1 implicated in the pathogenesis of breast and ovarian cancer (Nishikawa et al., 2004). The functions of polyubiquitin chains are not fully understood (Kim et al., 2007).

![Diagram of polyubiquitin chains through different ubiquitin lysine residues (Pickart and Fushman, 2004)](image_url)

Fig. 1.7. Polyubiquitin chains through different ubiquitin lysine residues (Pickart and Fushman, 2004)
1.8. Ubiquitin Surface residues

Surface residues of ubiquitin that have been found to be essential for vegetative growth of yeast are concentrated in a hydrophobic patch containing L8, I44 and V70, a patch containing F4 and surrounding residues, and the C terminal tail region of ubiquitin (Fig. 1.8). While the I44 patch has been implicated in proteasomal degradation and endocytosis (Beal et al., 1996; Shih et al., 2000), F4 patch has been implicated in endocytosis but not in proteasomal degradation, and C terminal tail is involved in most functions of ubiquitin, because it covalently links ubiquitin to substrate proteins (Katherine E. et al., 2001).

Fig. 1.8. Ubiquitin surface residues

(A) ubiquitin residues required for life in yeast. Essential ubiquitin amino acids are shown in blue. All essential residues are visible; none lie on the hidden surface. (B) Internalization information carried by the ubiquitin molecule. Primary residues required for internalization, Phe4 and Ile44, are shown in magenta; residues that play a minor role in endocytosis are shown in pink (Katherine E. et al., 2001).

1.9. Ubiquitin Binding Domains

Ubiquitin-binding domains (UBDs) recognise different parts of ubiquitin, and hence form an important part of proteins that interact with ubiquitin, and may be
responsible for high sequence conservation in ubiquitin (Hicke et al., 2005). UBDs are structurally diverse and have different biological functions.

The first ubiquitin-binding protein to be characterized was S5A/RPN10 subunit of the proteasome (Young et al., 1998). Nine UBDs have been characterised up to now (Wang et al., 2003). Mechanisms of action of UBDs are not clearly known, but their properties hint at their way of functioning in vivo.

1.10. Ubiquitin Proteasome System – Key Players of UPS

1.10.1. Ubiquitin

Ubiquitin is found ubiquitously among eukaryotes across species. The ubiquitination machinery uses ubiquitin as a post-translational modifier by covalently linking a lysine on target protein with the C terminal glycine of ubiquitin. C-terminal end of ubiquitin is activated by E1 enzyme as a preparation towards ubiquitination process.

1.10.1.1. Enzymes of Ubiquitination process

Ubiquitination serves as a molecular signal on target proteins for the execution of specified role. If single ubiquitin is attached to target protein, it is called monoubiquitination (Di Fiore et al., 2003; Haglund and Dikic, 2005; Hicke, 2001). If more than one ubiquitin is attached to the same protein at different places, it is called multiubiquitination, and if ubiquitin linked to substrate protein is further ubiquitinated, leading to the formation of polyubiquitin chains, it is called polyubiquitination (Haglund and Dikic, 2005; Welchman et al., 2005).
1.10.2. E1 (Ubiquitin activating enzyme)

E1 is a monomer of 110 kDa and its sequence has a weakly conserved 2 fold repeat (McGrath et al., 1991). It contains six domains which include adenylation domain, catalytic cysteine domain, four helix bundle and ubiquitin fold domain. All of them pack together to generate large canyon which at one end recruits a ubiquitin molecule.

E1 catalyzes the first step in the ubiquitination process of a target protein by activating C-terminal glycine of ubiquitin. This step requires Mg\(^{+2}\), ATP and ubiquitin. Ubiquitin is converted to adenylate intially, and subsequently transferred to the active site cysteine in the E1 (Haas and Rose, 1982). The thiol-linked ubiquitin is transferred to the next enzyme in the cascade, the E2 (Fig. 1.9).

![Diagram of E1 enzyme catalyzing ubiquitination process](image)

**Fig. 1.9.** The sequence of how the ubiquitin activating enzyme attaches to ATP and the ubiquitin substrate. It also shows how two ubiquitin substrates can be bound at one time (Haas and Rose, 1982).
1.10.3. E2 (ubiquitin conjugating enzyme)

E1 transfers ubiquitin to E2 by covalently linking C terminal glycine of the ubiquitin to a cysteine in E2. The diversity of E2s is much more compared to one or two E1s present in a cell.

The E2 enzymes have a 150 amino acids long core domain that is composed of four standard α helices, a four stranded antiparallel β sheet, and one 3_10 helix, and is conserved across all E2s (Tyers and Jorgensen, 2000). Some of the residues surrounding the active site cysteine are very highly conserved (Cook et al., 1993; Jiang and Basavappa, 1999; Worthylake et al., 1998). In addition to the core domain, N or C terminal extensions are found in some, though not all E2s. It is believed that these extensions are responsible for interactions of these E2s with various E3s.

![Diagram of ubiquitination process](image)

**Fig. 1.10. Variety and substrate specificity of enzymes involved in ubiquitination process**

1.10.4. E3 ubiquitin Ligase

Next enzyme in the ubiquitin proteasome cascade is E3 also known as ubiquitin ligase. E3s carry out the final step in ubiquitination, i.e. covalently linking C terminal glycine of ubiquitin taken from an E2, to a lysine on the substrate protein to
be ubiquitinated. E3 ligases are the most abundant and diverse proteins in a cell and are encoded by several hundred genes (Deshaies RJ, 2009). This diversity facilitates selective ubiquitination of proteins, consequently regulating cellular activities.

Fig. 1.11. Substrate recognition and transfer of ubiquitin to the substrate (Nalepa et al., 2006)

There are two main classes of ubiquitin E3s: the RING (really interesting new gene) (Borden, 2000) and HECT (homologous to E6-associated protein C terminus) domain families of E3s. The difference between HECT and RING domain family is that HECT E3s contain a conserved cysteine residue within the domain and accept ubiquitin from an E2 subsequently donating it to substrate protein. In contrast to HECT, RING domain serves as an adaptor between substrate and ubiquitin-loaded E2, binding both the proteins. Ubiquitin bound to E2 is directly transferred to substrate in this case.

HECT domain proteins contain a 350-amino acid residue sequence homologous to the C-terminal domain of the family E6-AP (E6-associated protein) (Huibregtse et al., 1995). This domain contains a conserved Cys residue, which interacts with the activated ubiquitin moiety transferred from E2 (Scheffner et al., 1995). The N-terminal domain, which varies among the different HECT domain proteins, is probably involved in specific substrate recognition.
The RING domain family of ubiquitin E3s is the largest class having near about 600 members. Within this family are multi-subunit cullin-RING ubiquitin E3s. In the characteristic RING sequence, four cysteines coordinate one Zn$^{2+}$ ion and another Zn$^{2+}$ ion is coordinated by three cysteines and a histidine. Hence the RING domain is structurally related to zinc finger domain. A well-known examples of RING E3s are SCF (Skp1/Cullin/Fbox) complex APC/C (anaphase-promoting complex/cyclosome).

Fig. 1.12. Components of the ubiquitin E3 ligases - Ubiquitin E3s are divided into the RING (really interesting new gene) and HECT (homologous to E6-associated protein C terminus) domain families. RING domain includes the Skp1/Cullin/F-box (SCF) and anaphase-promoting complex/cyclosome (APC/C) complexes. U-Box proteins are a distinct set of RING-like ubiquitin E3s (Rotin and Kumar, 2009).

1.10.5. Proteasome – The Nano Machine

Proteasome is a cytoplasmic protease, which specifically degrades certain proteins instead of blindly targeting every protein encountered. It does this by specifically recognizing ubiquitinated proteins and degrading them. Besides the degradation of intact proteins, proteasome also degrades misfolded and denatured proteins that are formed during conditions of stress like high temperature, UV
exposure or starvation. Certain non-ubiquitinated proteins are also degraded by proteasomes with the help of other mechanisms.

Proteasome is composed of a 20S protein cylinder and two 19S lids (Peters, 1994). The active sites are situated inside the cylinder. Hence the lids are responsible for imparting specificity to proteasome as they mostly recognize ubiquitinated proteins.

**Fig. 1.13. Structure of the 20S core complex in α7β7 β7 α7 architecture (Dieter H. Wolf and Hilt, 2004)**

The cylindrical part, which is called the 20S complex, is 15 nm long and 11 nm wide, and is made of two α rings and two β rings. The two α rings are joined together and make up the central portion of proteasome, while one β ring is situated on either end of the proteasome, attached to one α ring (Groll et al., 1997; Lowe et al., 1995). Both α and β rings consist of seven subunits each. Only three β subunits in each β ring have a proteolytic site. The three sites in each ring are different from each other, and are called β1, β2 and β5. β1 site cleaves near acidic residues, β2 site has trypsin like activity and β5 site has chymotrypsin like activity (Arendt and Hochstrasser, 1997; Heinemeyer et al., 1994). Proteins are degraded by these
proteolytic sites in the central chamber formed by β rings. Hence this chamber is called catalytic chamber.

![Diagram of proteasome mechanisms](image)

**Fig. 1.14. Mechanisms executed by the proteasome (Dieter H. Wolf and Hilt., 2004)**

Two 19S complexes are found in proteasome, one attached to either end of the 20S complex. These 19S complexes regulate the entry of proteins into proteasome. It is composed of two parts, namely the base, and the lid, and is made up of seventeen subunits (Peters, 1994). The base has eight subunits and is responsible for unfolding substrate proteins and inserting them into 20S complex. Six of the eight subunits in the base have ATPase activity, and provide the energy required for unfolding of the substrate proteins. These subunits are Rpt1 to Rpt6. Rpn1 and Rpn2 are the two non-ATPase subunits. There are eight subunits in lid as well, which are Rpn3 and Rpn5 to Rpn11. Rpn10 is responsible for attaching the lid to the base (Glickman et al., 1998). The polyubiquitin tail must have at least 4-5 ubiquitins for it to be recognised by the 19S complex. There are only a few exceptional cases in which non-ubiquitinated proteins are recognised by the 19S complex, the best example of which is ornithine decarboxylase (Coffino, 2001; Hoyt and Coffino, 2004). Rpn10  (Deveraux et al.,
1995; Mayor et al., 2007) and Rpn13 (Husnjak et al., 2008) are responsible for binding to polyubiquitinated proteins.

1.10.5.1. Proteasome mediated degradation

For many short lived eukaryotic proteins, conjugation to ubiquitin is an obligatory step in their degradation. Ubiquitin is joined reversibly to proteins by covalent (isopeptide) linkage between the carboxyl-terminus of ubiquitin and lysine e-amino groups of the acceptor proteins (Hochstrasser et al., 1999). For proteolytic substrates, assembly of ubiquitin chains on the protein is generally observed. Ubiquitinated proteins are in a dynamic state, subject to either further rounds of ubiquitin addition, ubiquitin removal by deubiquitinating enzymes, or degradation by a complex multicatalytic proteinase called the 26S proteasome (Fig. 1.16). The proteasome breaks down targeted substrates to short peptides but recycles the ubiquitin molecules.

![Proteasome Pathway Diagram](image)

**Fig. 1.15.** The ubiquitin proteasome pathway (Proctor et al., 2007)
1.10.6. Deubiquitinating Enzymes (DUB’S) and their classes

Deubiquitinating enzymes are proteases that recognize the glycine residue of C-terminal end of ubiquitin and cleave it either from ubiquitin itself or from other target proteins. DUBs negatively regulate ubiquitination by cleaving ubiquitin from target proteins, maintain a pool of free ubiquitins by breaking down polyubiquitin chains (Amerik et al., 1997; Dayal et al., 2009). Besides this, polyubiquitins are also spontaneously produced by the ubiquitin system without being attached to a substrate protein. These free polyubiquitin chains are also broken down by DUBs. DUBs are also responsible for cleaving ubiquitin fused to ribosomal proteins. Five different families of DUBs, together containing around 100 DUBs are so far known in humans (Nijman et al., 2005).

![Fig. 1.16. Proteasome bound deubiquitinating enzymes.](image)

Deubiquitinating enzymes are indicated in red (POH1 and USP14), POH1 catalyzes the release of a polyubiquitin. RPN10 (yellow) binds the polyubiquitin chain and the distal end of the chain can be removed by the action of UCH37 bound to ADRM1 (orange). USP14 is bound to the proteasome via interactions with RPN1 (purple) and probably removes mono-ubiquitin attached to the substrate (Francisca E. Reyes Turcu. et al., 2009).
1.10.7. Ubiquitin like modifiers (ULM’S)

Molecules other than ubiquitin were identified which mimic the structure and functions of ubiquitin to some extent hence called ubiquitin like proteins, molecules or modifiers (ULMs) (Welchman et al., 2005).

The family structural homology of ubiquitin-like protein modifiers (ULMs) is much more because of common 3D structure, the ubiquitin fold and a C-terminal protruding tail rather than the sequence homology. Hence, they are conjugated to proteins and function in “ubiquitin-like” manner. The attachment of ULMs to substrate protein changes its topology in such a way that it may facilitate or inhibit the binding of the protein to another molecule, affect enzymatic activity or sub-cellular localization and ultimately determine the half-life of the protein.

Ten different ULMs have been identified so far in mammals (Yamanaka et al., 2000). They are small Ub-like modifier, interferon stimulated gene 15 (ISG15), autophagy 8 (ATG 8) and neural precursor cell expressed, developmentally down regulated 8 (NEDD8). Ubls have a similar structural-fold to Ub, but use specific conjugation machineries and are recognized by distinct Ubl-binding domains. So far, SUMO2/3 and NEDD8 are the only Ubls known to participate in chain formation (Tatham et al., 2001; Xiromimas et al., 2004). Ubls are implicated in the regulation of many cellular processes such as gene transcription, signal transduction, autophagy and cell-cycle control (Kirsch et al., 2006).

1.10.8. Non-proteosomal functions of ubiquitin

Besides the role played by ubiquitination in proteolysis, it performs several other functions such as protein kinase activation, DNA repair, regulation of chromatin dynamics through histone modification, vesicle trafficking and endocytic machinery.
For example, K29 or K33 linked polyubiquitin chains inhibit AMP-activated protein kinase and (AMPK)-related kinases (A.-Hakim et al., 2008).

1.10.8.1. Ubiquitination and replication

Proliferating cell nuclear antigen (PCNA) influences DNA polymerase recruitment to DNA during replication. Monoubiquitination of PCNA seems to shift recruitment preference from replicative to trans-lesion polymerases (Bienko et al., 2005; Kannouche et al., 2004), whereas UbK63 linked polyubiquitination signals error-free DNA repair (Friedberg et al., 2005).

1.10.8.2. Ubiquitination as an endocytic signal

Mono-ubiquitination signals for the internalization of the membrane proteins in yeast and growth hormone receptor has been demonstrated in animal cells (Strous et al., 2004). Therefore, modulation of the endocytic internalization route by Ub could be important for signal transduction and receptor down-regulation. Ubiquitination also functions as a sorting signal at endosomes in the case of certain cargo proteins (Komada and Kitamura, 2005).

1.10.8.3. Ubiquitination involved in Transcription regulation

Ub proteasome pathway is involved in transcription regulation as many transcription factors are unstable proteins that are degraded by UPS (Elsasser and Finley, 2005). In Most transcription factors, the degradation signal which is recognised by E3s overlaps closely with a transcriptional activation domain (TAD) (Dupre and Haguaenauer-Tsapis, 2001).

1.10.8.4. Ubiquitin in Autophagic Degradation of Protein and organelle

Proteins and several other organelles have been found to be degraded following ubiquitin signalling. They are Mitophagy, Xenophagy, Ribophagy and others.
1.10.8.5. Ubiquitination involved in nonproteasomal removal of proteins

Protein misfolding results in the exposure of hydrophobic residues which are recognized by molecular chaperones of the heat shock protein (HSP) family, which bind to and shield exposed hydrophobic surfaces from the cytosol while promoting protein refolding (Goldberg, 2003). In addition, HSPs interact with Ub E3 ligases, such as CHIP and Parkin (Imai et al., 2002), which promote substrate polyubiquitination. These E3 ligases may play protective role by preventing accumulation of misfolded protein aggregates (Bjorkoy et al., 2005). Otherwise, the aggregates of proteins may inactivate the proteasome and mediate cytotoxicity. Thus, inhibition of the proteasome potently induces autophagy, which serves as a compensatory mechanism for degradation of accumulating polyubiquitinated misfolded proteins.

1.10.8.6. Ubiquitination in Selective Elimination of Ribosomes and mitochondria

Mechanism involving the selective removal of ribosomes is termed as ribophagy (Kraft et al., 2008). GFP labelling of ribosomal proteins have demonstrated that ribosomes are specifically targeted to the vacuole upon starvation. The process is dependent on functional autophagic machinery. Ub-specific protease, Ubp3, and an associated factor, Bre5 were found as proteins responsible for ribophagy.

Conjugation of monoUb/ Ub chains to an exposed mitochondrial protein is hypothesized to lead the mitochondrion towards autophagosome mediated degradation (Kirkin et al., 2009). There is an ongoing debate on the origin of E3 ligases which are responsible for the above ubiquitination as to they are residents of cytosol or mitochondrial membranes (Chu et al., 2009; Li et al., 2008).
1.10.8.7. Ubiquitination as Antimicrobial Mechanism

Mechanism involving removal of intracellular pathogens is referred as xenophagy. Several medically relevant bacteria, including *Mycobacterium tuberculosis*, *Salmonella enterica*, and *Listeria monocytogenes* are degraded by autophagy in vitro (Levine and Deretic, 2007).

Ub signalling found to be linked to xenophagy, Ub was recognized as a vital factor in plant resistance to bacterial pathogens (Nishimura and Somerville, 2002), and association with Ub chains was demonstrated for both gram-negative and gram-positive intracellular bacteria (Perrin et al., 2004).

1.10.9. The Role of Ubiquitin proteasome system in Diseases

Aberrations in the tightly regulated UPS pathway lead to the pathogenesis of many diseases, certain malignancies, and neurodegeneration. Some of the diseases which are linked to ubiquitin proteasomal machinery and the possible cure or drug development is mentioned below.

1.10.9.1. Cancer

Accelerated degradation of tumor suppressor proteins such as p53 and p27 are reported in malignant transformations. Mdm2, is the E3 ligase which binds and ubiquitinates the protein p53 (Michael and Oren, 2003). A reduced level of p53 and overexpression of the Mdm2 or Hdm2 (human counterpart of Mdm2) have been found for many tumors carrying wild-type p53, such as neuroblastoma, acute lymphoblastic leukemia (ALL), melanomas, and colorectal, lung and breast carcinomas (Ciechanover and Schwartz, 2004; Sun, 2006). Many drugs have been formulated to target Mdm2 to inhibit its interaction with tumour suppressor proteins (p53 and p27).
1.10.9.2. Neurodegenerative diseases

Inability of timely removal and accumulation of ubiquitin conjugates and/or inclusion bodies associated with ubiquitin proteasome system have been identified in a broad array of chronic neurodegenerative diseases, such as Alzheimer’s disease (AD), Parkinson’s disease (PD), amyotrophic lateral sclerosis (Hochstrasser et al., 1991) and Huntington’s disease to name a few.

Recent findings have revealed one mutant form of ubiquitin known as [Ub(+1)], which is prevalent in the brain of AD patients (Van Leeuwen et al., 1998). [Ub(+1)] is the result of frame shift mutation in the ubiquitin transcript which leads to extension of the molecule with 20 amino acid residues. [Ub(+1)] is an efficient substrate for polyubiquitination and its degradation is inhibited by 26S proteasomes, which results in the accumulation of the toxic proteins with neuro-pathological consequences. [Ub(+1)] was also described in other similar diseases and in other disorders such as Down’s syndrome (Van Leeuwen et al., 1998) or supra-nuclear palsy (Fergusson et al., 2000), it is evident that it is not entirely specific to AD.

1.10.9.3. Immune and Inflammatory Response

Peptide epitopes presented to T-lymphocytes on class I MHC molecules are generated in the cytosol and processed by the ubiquitin-proteasome pathway (Kloetzel, 2001; Rock and Goldberg, 1999). Presentation of self antigen as non-self may be the potential cause that underlies the pathogenesis of autoimmune diseases.

A broad array of immune and inflammatory disorders is caused by activation of the immune system’s central transcription factor NF-κB. Activation of the factor stimulates transcription of many cytokines, adhesion molecules, inflammatory response and stress proteins and immune system receptors. The factor is activated by the ubiquitin proteasome system’s proteolytic mechanism by limited processing of the
precursor protein to yield the active subunit p50 and signal induced phosphorylation and subsequent degradation of the inhibitor IκB that enables translocation of the factor into the nucleus where it initiates specific transcriptional activity (Karin and Ben-Neriah, 2000)

1.10.10. Drug Development to Target Ubiquitin Proteasome System

Ubiquitin proteasome system plays a central role in a wide array of basic cellular processes, hence it is difficult to target and select cells which show abnormality. Inhibition of the proteasome may affect many processes nonspecifically. Recent experimental evidence strongly suggests that such inhibition may be beneficial in certain pathologies, such as in cancer (Golab et al., 2000), asthma (Elliott et al., 1999), brain infarct (Phillips et al., 2000) and autoimmune encephalomyelitis.

In malignancies, the drugs may act by inhibiting the degradation of different cell cycle inhibitors, whereas in neuroprotection they may act via inhibiting activation of NF-κB, which elicits an inflammatory response. When accelerated degradation of a tumor suppressor results in exposure of cells to malignant transformation, selective inhibition of the recognition machinery can potentially reverse the malignant phenotype. Peptides that bind specifically to HPV-E6 and prevent its association with p53 can interfere with p53 targeting (Michael H. Glickman and Ciechanover, 2001).

In autoimmune diseases, they may act by inhibiting presentation of self-peptides and by interfering with signal transduction cascades. Another approach of targeting may be the development of small molecules that are substrate specific and bind specific substrates or to their ancillary proteins rather than to an E3.
1.11. BRIEF INTRODUCTION TO CHAPTERS

Chapter 1 is an introduction ubiquitin and the enzymes, which are the key players of the ubiquitin proteasome system. Structure of ubiquitin, importance of its surface residues and ubiquitin binding domains in the interacting proteins have been discussed as ubiquitin accomplishes a variety of roles through protein-protein interactions, which explains the reason for its structural conservation. Beside proteasomal functions ubiquitin is also known to mark its presence in other non-proteasomal functions. A brief account on non-proteasomal functions of ubiquitin is presented here. Due to defective functioning of the UPS system, cells are prone to various diseases. Some of the diseases and the strategies for drug development which targeting UPS are also discussed. Central role of ubiquitin in cellular pathways and the need to throw light on high degree of conservation of its structure to its functional relevance are the main reasons to choose this protein in the current study.

Chapter 2 deals with generation of mutant ubiquitins, which are derivatives of dose dependent lethal mutant of ubiquitin UbEP42 and their functional characterization. UbEP42 a mutant form of ubiquitin generated earlier in our laboratory by error-prone PCR reaction harbours four amino acid substitutions (Prabha et al., 2010). In order to understand the importance of the residues which have undergone substitution, mutant forms of ubiquitin carrying single mutations and combinations of mutations were constructed using site directed mutagenesis. All mutant ubiquitins were analyzed for functional integrity in UBI4 deletion strain SUB60 of Saccharomyces cerevisiae, which is hypersensitive to stress conditions. The SUB60 strain which fails to grow under stress, can withstand stress and behave like wild-type if ubiquitin is expressed extra-chromosomally from a plasmid. So if the mutant forms are expressed extra-chromosomally in the same strain in place of wild type ubiquitin and the cells can survive stress conditions, it can be concluded that the mutant forms retain their functionality. Following this line of logic, complementation studies were carried out.
under heat stress and antibiotic stress. Prior to it the effect of over-expression of
mutant ubiquitins on growth and doubling time have been studied to understand the
impact of ubiquitin mutants on the normal functioning of the cells. Our results show
that two of the mutations have compensatory effects over other two mutations which
are drastically detrimental in their nature (Mrinal Sharma and C. Ratna Prabha,
Manuscript under preparation).

Chapter 3 deals with the functional importance of residues in the parallel β-bulge of
ubiquitin. This β bulge is formed by Glu64, Ser65 and Gln2 residues. This structure
of β-bulge complete with residue identities is totally conserved in all ubiquitins.
Interestingly, it displays certain uncommon features. It was hypothesized that if the
individual residues of the β-bulge are replaced with residues which have greater
preference for the same secondary structure, then the structure would not be disturbed
and whatever functional changes are observed will highlight the role of the replaced
residues. Therefore, single mutants of the β-bulge were constructed and characterized
earlier in our laboratory which in spite of minor structural changes showed significant
functional differences (Mishra et al., 2009; Mishra et al., 2011). The above
observations led us to the question of what would be the functional consequences if
more than one residue is changed at once. Will the resultant mutations have effects
which are synergistic in the same line of what was observed earlier or will they be
compensatory? Thus, the chapter presents the generation of mutations in β-bulge in
combinations and their functional characterization in S. cerevisiae by
complementation analysis under various stresses. Our results establish that two of the
mutations UbE64G-S65D and UbQ2N-E64G-S65D are severely debilitating
compared to other two mutations UbQ2N-E64G and UbQ2N-S65D, implying that
E64G and S65D together exert more debilitating effects than Q2N with either E64 or
S65 (Sharma and Prabha, 2011; Mrinal Sharma, Ankita Doshi and C. Ratna Prabha,
Oral presentation in Regional Science Congress held in Vadodara, September 15-16,
2012).
Chapter 4 presents the construction of dose dependant lethal ubiquitin UbEP42 and its derivative Ubl61T in bacterial expression vector. This chapter describes the various steps of purification to achieve purified protein for structural studies. Secondary and tertiary structure characterization of the mutant has been carried out by circular dichroism and fluorescence spectroscopy respectively.

Chapter 5 describes the construction of double and triple mutants of second β-bulge of the ubiquitin in bacterial expression vector and their purification by affinity chromatography. Secondary and tertiary structures of the mutant proteins have been studied using circular dichroism and fluorescence spectroscopy respectively.