Chapter 1

Introduction and Review of Literature

Proteasome is unusually large, multi-subunit, ATP dependent compartmentalized protease. By regulating the functional concentration of cellular proteins, it regulates almost all the cellular events. It also takes care of misfolded and damaged proteins and in this way maintains homeostasis. Unlike other proteases, proteasome most of the time completely degrade globular protein into small peptides. Proteasome also play an important role in protein processing and antigen presentation. The most abundant proteasome in eukaryotic cells are known as 26S proteasome consisting of protease (20S) and regulatory component (19S). Globular proteins are targeted to proteasome mainly by the covalent post translation modification, poly ubiquitination or by adaptor proteins. Exposed unordered region in the protein is essential for efficient degradation. Proteasome bound substrates are then presumably unfolded by proteasomal ATPases which is further traslocated to the catalytic core. Substrate binding, release, chain unfolding and translocation are all ultimately linked to degradation. Protein modifications due to mutations, amino acid repeats and aberrant levels of substrate or proteasome regulatory protein (E3 ligases) can adversely affect the efficiency at which substrates are cleared and in turn could lead to several pathological consequences. In recent years with the help of very few model substrates, some fundamental questions are being addressed such as: a) what is the minimal length of Ub required, b) which processes are ATP dependent, c) whether regulatory component communicate with the proteolytic component, d) what are the requirement for efficient degradation etc. But, still we are far from understanding the mechanistic details of protein degradation by proteasome. There are two essential elements in the proteasomal degradation, proteasome localization and presence of unordered or floppy region. Nevertheless, some protein localized to proteasome but escape degradation. Some of the other unanswered question are- a) how does unstructured region originate in a substrate, b) what is the

fundamental mechanism underlying the various steps involved in the degradation of a substrate, its thermodynamic and kinetic parameters. Moreover, direct role of 26S ATPases in protein unfolding has not been demonstrated till date.

1.1 Discovery of proteasomal machinery: Proteins are the essential macromolecules that participate and regulate virtually every cellular process. The first breakthrough in the field of protein degradation came in 1942, when Rudolph Schönheimer postulated in his book 'The Dynamic State of Body Constituents' that proteins are being constantly build up and broken down. Initially, lysosome was believed to be solely responsible for cellular protein quality control machinery. In 1977, while working on reticulocyte that lack lysosomes on maturation, Etlinger and Goldberg for the first time established a cell free ATP dependent proteolytic system (Etlinger and Goldberg, 1977). Ciechanover and Hershko fractionated the crude reticulocyte cell extract on an anion-exchange resin as (APF)-I and II and found that combination of both the fractions reconstituted the energy-dependent proteolytic activity (Ciehanover et al., 1978; Hough et al., 1987). Later, APF-II was further sub-fractionated into APF-IIa and APF-IIb. APF-IIb contained the E1-E3 ubiquitin conjugating enzymes and APF-IIa was shown to contain proteasomes (Hershko et al., 1979). APF-I was identified as ubiquitin (Wilkinson et al., 1980). Eventually, in 2004 Avram Hershko, Aaron Ciechanover and Irwin rose were awarded for Nobel Prize in Chemistry for the discovery of 'Ub dependent degradation of protein'.

1.2 Proteasome regulates various cellular processes: For the healthy cell survival, tight spatiotemporal regulation of cellular processes is an absolute mandate. There are several ways by which cellular processes are regulated tightly. One of the most important is degradation, which is by nature an irreversible process. In higher eukaryotes, only membrane-associated proteins and alien proteins such as those of

bacterial and viral origin are destroyed by hydrolytic enzymes in lysosomes. Degradation of the majority (80-90%) of intracellular proteins is regulated by the 26S proteasome (Craiu et al., 1997; Rock et al., 1994). Depending on their function, the half-life of cellular protein varies from minutes to years. Structural proteins are long lived, while regulatory proteins are degraded in few minutes. For example, for cell cycle progression specific cyclin must be quickly degraded so that another cyclin could play its part. The proteasomal system is so specific and perfect that it only degrades cyclin A or B from Cyclin-CDK complex and the CDKs are untouched (Nishiyama et al., 2000). Proteasome regulates almost all the cellular processes such as cell cycle progression, differentiation, apoptosis, DNA repair, cellular quality control, autophagy, regulation of transcription and generation of peptides for antigen presentation (Glickman and Ciechanover, 2002) (**Table. 1.1**). Proteasome maintain cellular homeostasis by degrading globular proteins with high specificity. It not only regulates the functional concentration of proteins in cytoplasm and nucleus but also takes care of misfolded, unfolded, oxidized and unwanted proteins, thereby acting as 'molecular sweeper' (Glickman and Ciechanover, 2002) (**Fig 1.1**).

Table 1.1 Some of the cellular processes regulated by proteasome

Cellular process	**Regulatory proteins/proteasome substrate**
Cell cycle	Cyclins A, D, and E, p53, p21, p27, mdm2, HIF-1a (major E3 ligases are SCF and APC complex)
DNA transcription	I kappa B alpha, c-myc, c-Jun, c-fos, AP-1, STAT-1
DNA repair	DNA-PKcs, rad 23
Apoptosis	p53, p21, mdm2, bcl2, bax, caspase-3
Inflammation and immunity	I kappa B, tumor necrosis factor-R1, processing of p105
Cell growth/ signal transduction	Epidermal growth factor receptor, insulin-like growth factor receptor, and platelet-derived growth factor receptor
ER associated degradation	CFTR, misfolded protein

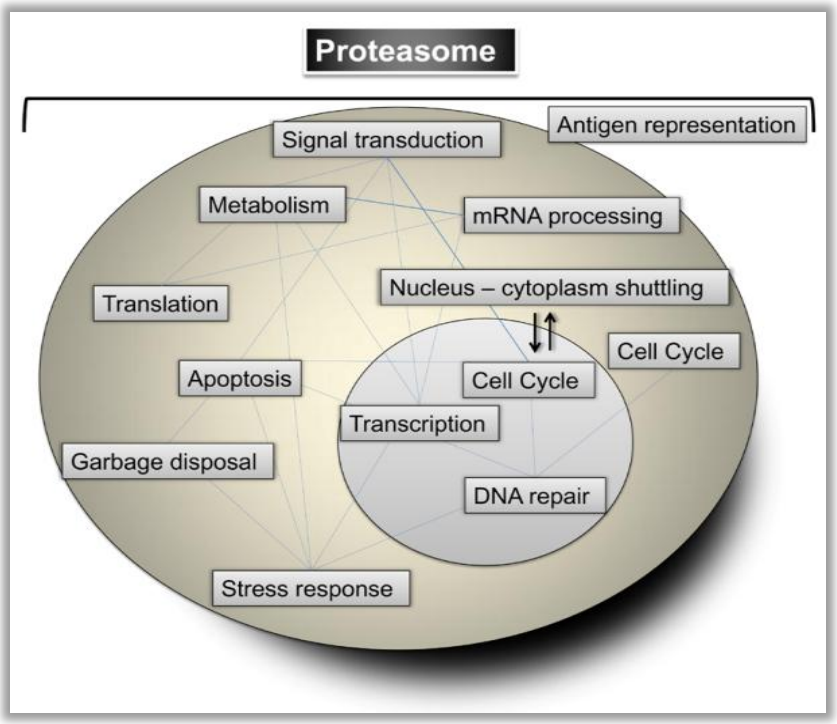

Figure 1.1 Proteasome regulates all most all the cellular process in the cell

The cellular machinery crosstalk with each other and are interdependent, any defect in the master regulator (Ub-proteasome system) may lead to an imbalance thereby resulting in several diseases and disorders (Dahlmann, 2007). The first human disorder identified due to defect in the proteasome system was Angelman Syndrome. Mutations in the E3 ligase, E6-AP has been shown to be the cause of this syndrome that is characterized by mental retardation, seizures and abnormal gait (Kishino et al., 1997). The G201V mutation in β4i (PSMB8) has been shown to be associated with Nakajo-Nishimura syndrome (NNS); symptoms include periodic fever, skin rash, partial lipomuscular atrophy and joint contracture (Arima et al., 2011). Also, uncontrolled proteasomal degradation of tumor suppressors and stabilization of oncogenes have been correlated with various cancers. One of the proteasome inhibitor, PS 134 (Velcade- a boronic dipeptide) is being used for the treatment of leukemia. Apart from this, several proteasome inhibitors are in different phases of clinical trials for several solid tumors and diseases.

1.3 Architecture of 26S proteasome: 26S proteasome belongs to the family of compartmentalized proteases where regulatory component and proteolytic components may exist in free form (Baumeister et al., 1998). Substrate specificity is maintained by regulatory component and their access to the free proteolytic core is generally forbidden by close gate (Baumeister et al., 1998).

26S (S=Svedberg sedimentation coefficient) proteasome is 2.4 MDa multi-specific, multi-subunit ATP dependent protease. It is composed of ~700 kDa 20S core particle (CP) harboring proteolytic activity, sandwiched between two 19S regulatory particles (Glickman and Ciechanover, 2002). The most abundant form of proteasome found in the eukaryotic cell is the 26S proteasome. The 19S regulatory particle (RP) is responsible for substrate recognition and unfolding of globular protein so that it can be translocated to small gate at 20S core particle for degradation in to small peptides (Glickman and Ciechanover, 2002). 20S CP can also associate with other regulatory complexes like 11S and PA200 (Glickman and Ciechanover, 2002).

1.3.1 The 20S Core Particle: The 20S CP can be found either isolated or associated with the 19S RP. Purified 20S is relatively stable than that of 26S. High resolution structures of 20S proteasome are similar in archaea, mycobacteria, yeast and mammal (Groll et al., 1997; Unno et al., 2002). 20S CP is structurally hollow barrel-shaped composed of 28 subunits, arranged into four hetero-heptameric rings of ~150 Å X ~120 Å dimension (Groll et al., 1997). The complex is assembled from 14 gene products and exhibits twofold (C2) symmetry with four stacked rings - two inner β rings and two outer α rings, achieving overall stoichiometry of $\alpha_{1\text{-}7}\beta_{1\text{-}7}\beta_{1\text{–}7}\alpha_{1\text{–}7}$. The interior surface of 20S CP forms a central channel along with three large cavities separated by narrow constrictions. The two cavities between the α and β-subunit rings of ~40 Å x ~50 Å dimension is known as antechamber and apparently, store the substrate in

unfolded conformation (Ruschak et al., 2010; Sharon et al., 2006). The central catalytic chamber of ~40 Å x ~55 Å dimension harbors the protease active sites. Proteasomes belong to the N-terminal nucleophile (Ntn) family of hydrolases (Seemuller et al., 1995). The γ-oxygen atom of the N-terminal Thr1, liberated by autolytic removal of pro-peptide acts as the nucleophile, the α-amino group most likely acts as proton acceptor in the hydrolysis of peptide bonds, while a water molecule is thought to shuttle a proton between them (Seemuller et al., 1995). In eukaryotes only three of the seven β-type subunits harbor N-terminal threonine residue required for proteolytic activity making six active sites per 20S molecule. Like several other proteases, some of the β-subunits are translated in premature form and later processed to the mature form. Each active site can cleave a broad range of peptide sequences, β1 cleaves after acidic residues (caspase like), β2 after basic residues (trypsin like), and β5 after hydrophobic residues (chymotrypsin like) (Borissenko and Groll, 2007). Upon IFN-γ induction, β1, β2, β5 are replaced by β1i, β2i, β5i and are referred to as immunoproteasome (Yewdell, 2005). Immunoproteasome is responsible for generation of peptides for antigen presentation by MHC class I molecules (Yewdell, 2005). Combination of different active sites of broad specificity ensures that the encounter of an unfolded substrate with them will result in digestion of the substrate in small pieces. Repetitive sequences like poly Q or Gly-Ala repeats have been shown to be resistance for proteolytic cleavage by 26S proteasome (Hoyt et al., 2006; Venkatraman et al., 2004). The size distribution of released peptides is broad, ranging from 4 to 25 residues, with an average length of 7 to 9 residues.

Because the active sites face the interior surface of the CP, the substrates must gain access to this space. The entry to 20S CP chamber is regulated by the N-terminal tails of the α-subunits which form a 13Å gate (Groll et al., 1997) (**Fig 1.2**). In free 20S,

this gate remains close blocking nonspecific protein degradation. In 26S proteasome, opening of the gate is regulated by the ATPases in the base subcomplex of the 19S regulatory particle. At α-α interfaces of 20S there are seven "α-pockets" which provide binding sites for HbYX motif (C-terminal hydrophobic-tyrosine) of three of the 19S ATPase (Smith et al., 2007). This binding to 20S CP pockets induces a rotation, resulting in displacement of a reverse-turn loop. This leads to the stabilization of open-gate conformation (Rabl et al., 2008). In addition to the 19 RP, several proteins, protein complexes, small peptides and other factors can open the 20S CP gate.

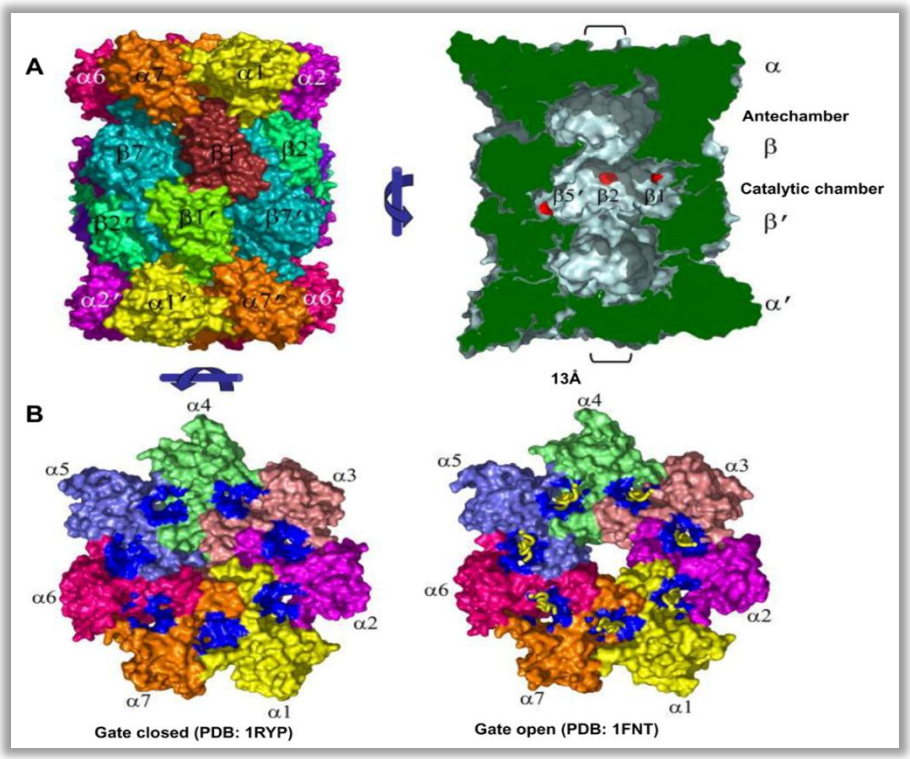

Figure 1.2 20S CP structure: *The heptameric ring arrangement of 20S CP subunits, interior surface of 20S CP contain antechamber and central catalytic chamber (catalytic subunit-red) (PDB: 1RYP) (A). Entry in to CP is guarded by a narrow pore that remains closed (B), unless it will interact with RP subcomplex. The top view of 20S CP closed gate (PDB: 1RYP) and open gate conformation (PDB: 1FNT) (B).*

1.3.2 19S regulatory particle: The proteolytic action of 26S proteasomes is confined to 20S CP, but the substrate selection and the mechanism of degradation are mainly governed by regulatory particles. The 19S RP appears to serve multiple roles in

regulating proteasomal activity e.g.: a) substrates recognition, b) presumably unfolding and c) translocation to the 20S catalytic particle. It may even influence the nature of products generated by proteolysis (Glickman and Ciechanover, 2002). The RP contains about 19 subunits, which has been subdivided into the lid and base sub-complexes. A 10-protein 'base' sub-complex binds directly to the 20S α-ring activating gate opening, and a 9-protein 'lid' sub-complex is involved in substrate recognition.

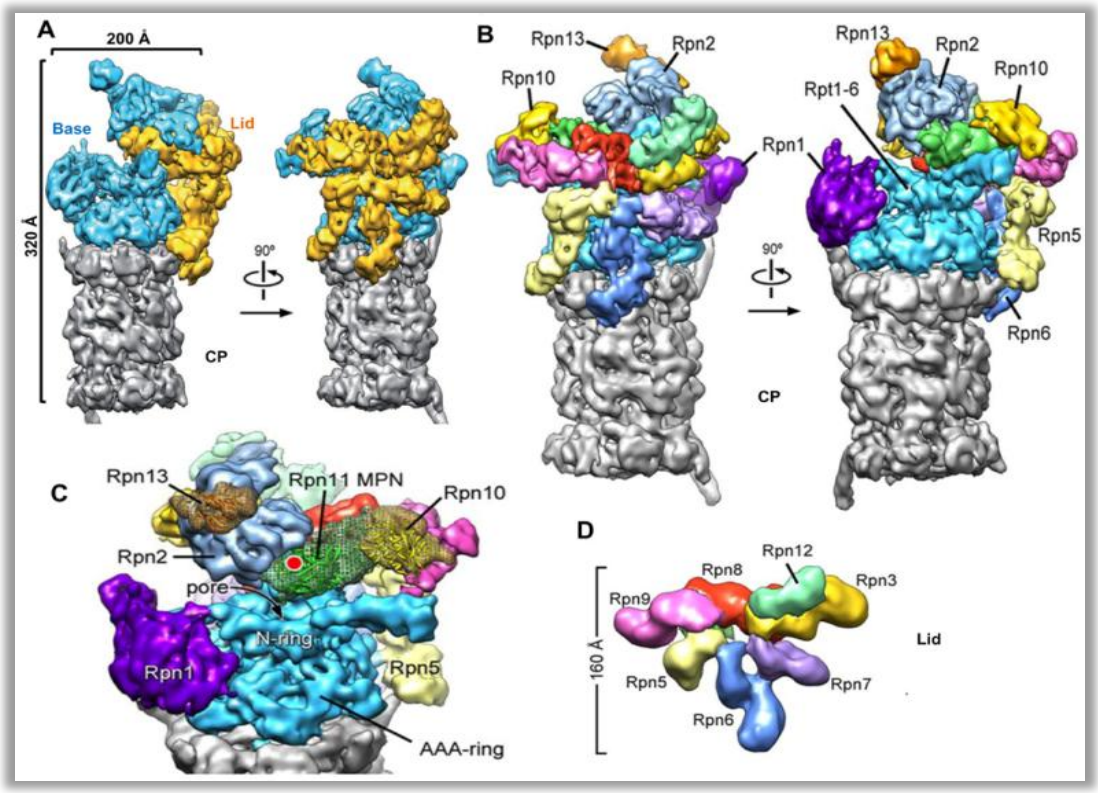

Figure 1.3 Architecture of 26S proteasome: *The high resolution EM map of 26S proteasome (A and B), lid subcomplex makes extensive interaction with base and the CP (A). ATPases of base subcomplex for N-ring and AAA ring with a central pore for substrate unfolding and translocation (C). Lid subcomplex form hand like structure and play major role in placing Ub receptors Rpn 10 and 13 as well as DUb RNP11 (B and D); on top of central channel formed by ATPases (C). [Adopted from- (Lander et al., 2012)].*

Although, 19S subunits have been identified and specific functions have been assigned to some of them, the organization of the 19S complex remains unclear. However, recent cryo-EM map by single particle analysis has provided glimpse of the

architecture of 26S holo-complex (9Å to 7.4Å) (Beck et al., 2012; Lander et al., 2012; Lasker et al., 2012) (**Fig. 1.3 and table 1.3)**. With the help of these EM maps, the position of ATPases and Ub binding domain containing subunits could be assigned in 26S proteasome. One of the other key finding was the interaction of some of the lid subunits directly to 20 CP. (**Fig 1.3 A and B**).

Table 1.2 19S regulatory particle, known domains and effect of deletion of on survival

Budding yeast	Human	Activity or domain type	Lethality	Comments (other names)
Rpt1	PSMC2	AAA	L	ATPase
Rpt2	PSMC1	AAA, HbYX	L (NL)	ATPase and gate opening
Rpt3	PSMC4	AAA, HbYX	L (L)	ATPase and gate opening
Rpt4	PSMC6	AAA	L	
Rpt5	PSMC3	AAA, HbYX	L (L)	ATPase and gate opening
Rpt6	PSMC5	AAA	L	
Rpn1	PSMD2	PC	L	PIPs scaffold
Rpn2	PSMD1	PC, NLS	L	PIPs scaffold
Rpn3	PSMD3	PCI, PAM	L	
Rpn5	PSMD12	PCI	L	
Rpn6	PSMD11	PCI, PAM	L	
Rpn7	PSMD6	PCI	L	
Rpn8	PSMD7	MPN	L	
Rpn9	PSMD13	PCI	NL	
Rpn10	PSMD4	VWA, UIM (2)	NL (L)	Ub receptor
Rpn11	PSMD14	MPN, JAMM	L	DUb
Rpn12	PSMD8	PCI	L	
Rpn13	ADRM1	Pru	NL	Ub receptor, Uch37 recruit
Shuttling proteasomal subunits				
	PSMD10	Ankyrin		
	PSMD9	PDZ	NL	PIP

AAA: ATPase associated with diverse cellular activities, ANK: ankyrin repeats, DUb: Deubiquitylating enzyme, L: Deletion result in lethal phenotype in yeast, (L): Deletion result in embryonic lethality in mouse, NL: non-lethal in yeast, (NL): non-lethal in mouse, MPN: Mpr1, Pad1 N-terminal, NLS: Nuclear localization signal, NL: Deletion non-lethal in yeast,(NL): Deletion embryonically Non-lethal in mouse, Ntn: N-terminal nucleophile hydrolase, PAC: Proteasome assembling chaperone, PAM: PCI associated module, PC: proteasome/cyclosome repeat, HbYX: hydrophobic-tyrosine-X, PCI: proteasome, COP9,eIF3, PDZ:DLG/ZO-1, PIPs: Proteasome interacting proteins, Pru: Pleckstrin-like receptor for ubiquitin, VWA: Von Willebrand factor type A.

1.3.2.1 The Proteasome Base sub-complex: The ten components of the base include six paralogous AAA^+ ATPases, referred to as Rpt (**R**egulatory **p**article **t**riple A-ATPase) proteins in yeast and PSMCs in mammals. The Rpts are critical for RP-CP complex formation, as the C-termini of the Rpts insert into the above-described α-subunit cavities (Smith et al., 2007). Degradation of typical physiological proteasome substrates is an ATP dependent process, indicating a central role for Rpts in proteasomal degradation. Apart from ATPases, the base subcomplex also contain four non-ATPase subunits (Rpn-**R**egulatory **p**article **n**on-ATPase) Rpn1, Rpn2 and the ubiquitin receptors Rpn10 and Rpn13 (Glickman et al., 1998) (**Table. 1.2**).

1.3.2.1A ATPases in Base: Proteasomal ATPases are the members of AAA^+ family (**A**TPases **a**ssociated with a variety of cellular **a**ctivities). The six ATPases are the product of distinct genes and share substantial sequence similarity. At the N-terminus, Rpt subunits contain oligosaccharide binding domain (OB) followed by coiled coil domains. The conserved Pro residue between the OB and the coiled coil domains determines the relative orientation of the two Rpt subunits (Park et al., 2010). Preceding the coiled coil domains are the AAA^+ domain. It is composed of conserved Walker A and Walker B motifs (Rubin et al., 1998). The Rpt heterohexamer forms a trimer of dimers with the pairs Rpt1/Rpt2, Rpt6/Rpt3, and Rpt4/Rpt5 each held together by coiled coil domain (Tomko et al., 2010) (**Fig. 1.4A**). Each dimer is stabilized by salt bridges between Asp of one subunit and Arg from other subunit within OB domain (Tomko et al., 2010). The N-terminal domains of the ATPases form a separate hexametric ring (N-ring) that consists of OB domains and the three coiled coils protrude from the ring (Lander et al., 2012; Zhang et al., 2009a) (**Fig 1.4 A and B**). The AAA^+ of Rpts 1-5 are oriented in a spiral staircase around the hexameric ring with Rpt3 at the top and Rpt2 at the bottom of the staircase (Lander et al., 2012). The AAA^+ domain of Rpt6 adopts a

tilted orientation, bridging Rpt2 and Rpt3 (**Fig. 1.4 C**). The coiled coils formed by Rpt1/2 and Rpt4/5 are flexible and can undergo 40º swinging motions. It is hypothesized that this will allow them to grab substrates bound to Rpn10 and Rpn13 via ubiquitin chains (Lasker et al., 2012) (**Fig. 1.4A**).

Table 1.3 Summary of 26S proteasome EM Map

Reports	(Lander et al., 2012)	(Lasker et al., 2012)	(Beck et al., 2012)
Organism	*S. cerevisiae*	*S. pombe*	*S. cerevisiae*
Resolution	9 Å	8.5 Å	7.4 Å
Source	Recombinant lid and 26S was purified for cell	Purified	Purified
Assumption	C2 symmetry	C2 symmetry	No C2 symmetry
Analysis	Reconstituted lid, helped in localization of lid subunits. Subunit position was determined by means of fusion constructs and automated segmentation methods. Based on PAN structure, Rpts were docked on EM map of the 19S.	Yeast two hybrid, cross-linking data, known subunit structure were compiled, extensive computational methods were used to get and score possible configurations. These data were further used for subunit position docking of comparative models of the subunits in EM map and refined by flexible fitting.	C2 symmetry is confined to 20S due to the presence of PIPs. both 19S may not be identical. EM map in conjunction with molecular dynamics-based flexible fitting to build holocomplex model.
Lid	Lid subunits from Hand-like structure where five PCI subunits (Rpn3,7,6,5 and 9) form the fingers. Rpn11 lies in the palm of the hand and interact with Rpn8 which connects Rpn3 and Rpn9. As compared to free lid, large conformational changes were observed in Rpn8 and 11 in holocomplex. This may regulate DUb activity of Rpn11.	The six PCI subunits (Rpn3,5,6,7,9 and 12) form horseshoe like structure covering a large part of the ATPase (Rpt3,6, and 4) as roof. The subunit order in the horseshoe is Rpn9/5/Rpn6/7/Rpn3/2.	The PCI subunits form a scaffold that positions the Rpn8/11 heterodimer in close vicinity to the mouth of ATPase and anchor the RP to the CP.
ATPase	The C-terminal 'small AAA^{+}' subdomains (Rpt 1-5) arrange in one plane above the 20S core, 'large AAA^{+}' subdomains (Rpt1–5) are oriented in a spiral staircase around the hexameric ring, with Rpt3 at top and Rpt2 at the bottom. The AAA of Rpt6 adopts a tilted orientation and bridge Rpt2 and Rpt3. Several lid subunits (Rpn6, 5 and 7) interact directly with AAA^{+} domains of the Rpts.	Rpn10 and Rpn13 are positioned above the coiled coils of the Rpt4/5 and Rpt1/2 dimers, respectively. Rpt1/2 and Rpt4/5 coils can undergo 40º swinging motions to grab substrates bound to Rpn10 and Rpn13 via ubiquitin chains.	The coiled coil of Rpt6/3 is less flexible as it interacts with PC-domain of Rpn2 and a large coiled-coil bundle formed by the C-termini of the lid. Rpt1/2 coiled coils appear to interact with Rpn1 while the coiled coil of Rpt4/5 doesn't interact with any Rpn subunit.

It has been shown that ATP binding and hydrolysis play different roles in 26S proteasome. In presence of non-hydrolysable analogue of ATP (ATPγS), 26S proteasome was able to translocate and degrade unfolded and denatured proteins, while there was no degradation of globular proteins (Benaroudj et al., 2003). This observation suggests that unfolding requires energy from ATP hydrolysis, whereas ATP binding alone may be sufficient for 19S-20S association, gate opening, and probably translocation of unfolded substrates (Liu et al., 2006).

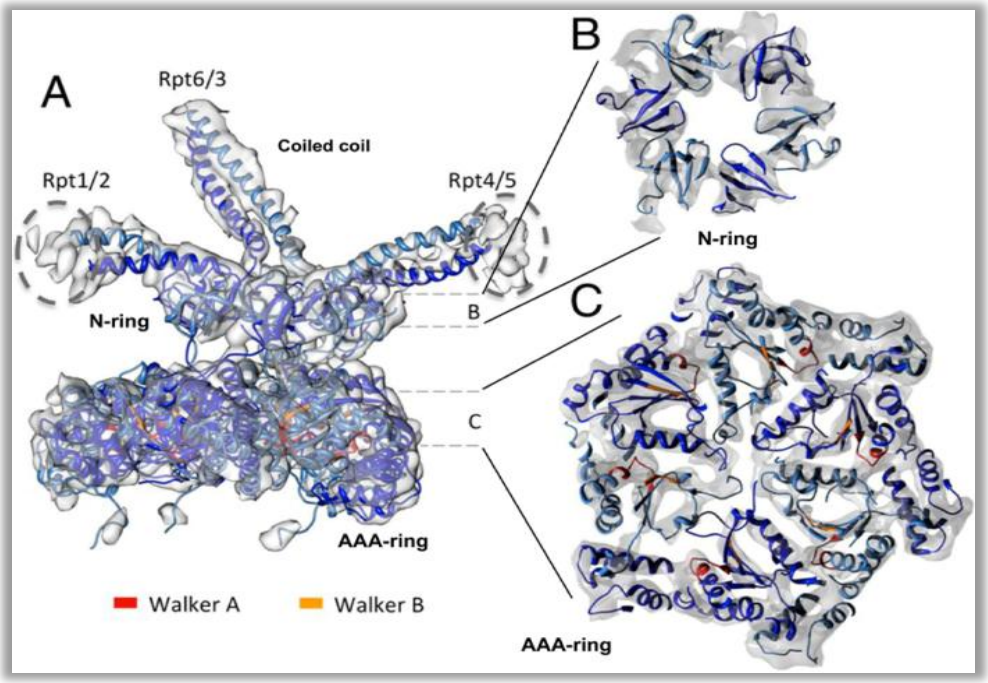

Figure 1.4 Architecture of proteasomal ATPases*: Proteasomal ATPases form dimer of trimer (Rpt1/2, Rpt 3/6 and Rpt 4/5) (A). The atomic model of the AAA-ATPase, based on EM map showed that ATPase complex form N-ring of larger pore size (B) while AAA-ring which is arranged as spiral staircase like structure (C) Rpt1/2 and Rpt4/5 coils can undergo 40º swinging motions to grab substrates bound to Rpn10 and Rpn13 via ubiquitin chains (A). (Dark blue: Rpt1/6/4; light blue: Rpt2/3/5, Walker A-red and Walker B-orange). [Adapted from- (Beck et al., 2012)].*

1.3.2.1B Non-ATPase in base: The non ATPases of base subcomplex include the scaffolding proteins Rpn1 and Rpn2 and the ubiquitin receptors Rpn10 and Rpn13 (Glickman et al., 1998). Rpn1 and 2 are the largest proteasomal subunits and contain proteasome cyclosome (PC) repeats (Kajava, 2002) (**Table. 1.2**). In recent EM

structure, the PC-domain of Rpn2 has been shown to interact with N-terminal end of the coiled-coil pair of Rpt6/Rpt3 while Rpt1/Rpt2 coiled coil appear to interact with Rpn1 (Beck et al., 2012) (**Fig 1.3B**). The Ub binding subunits Rpn 11 and 13 have been assigned a position above the coiled coil of the Rpt4/5 and Rpt1/2 dimers, respectively (**Fig 1.3B and C**). The distance between Rpn11 and 13 subunits (in 3D) is approximately 90 Å, which could be bridged by a tetra Ub moiety (Lander et al., 2012). This relative assignment of Ub receptors offers an explanation as to why polyubiquitin chains needs to be comprised of at least four Ub to function efficiently as a degradation signal. Several Ub adaptor proteins like Rad23, Ddi1 and Dsk2 as well as non-essential DUb protein like Ubp6, interact with Rpn1. Rad23, Dsk2, and Ddi1 share a common domain at their respective N-termini, known as ubiquitin-like domain (Ubl), and this domain mediates recognition by Rpn1 (Gomez et al., 2011). The function of Rpn2 is still not clear but it was found to interact with Hul5, a HECT-domain containing ubiquitin ligase (Kohlmann et al., 2008).

1.3.2.2 Lid subcomplex: The lid subcomplex is composed of at least 9 non ATPase subunits. Based on amino acid sequence similarity, these can be divided into two categories: a) the MPN (**M**pr1 and **P**ad1 in the **N** terminus) domain containing subunits Rpn8 and Rpn11 (Verma et al., 2002),

b) The PCI (**P**roteasome-**C**OP9-e**I**f3) domain containing subunits Rpn3/5/6/7/9/12 (Finley, 2009).

Of the nine lid subunits, the function of only Rpn11 is known. Rpn11 has a metalloprotease-like deubiquitinating (DUb) activity which removes proximal ubiquitin from substrates (Verma et al., 2002). While the MPN domain of Rpn8 is very similar to that of Rpn11, it lacks crucial catalytic residues. A free lid subcomplex has the ability to bind ubiquitinated substrate while free Rnp11 or purified lid does not have DUb

property (Verma et al., 2002). In light of recent high resolution EM map, large conformational change was observed in Rpn8 and 11 between free lid and in holocomplex (Lander et al., 2012). This change has been proposed to regulate DUb activity of RNP11. In 26S holocomplex, Rpn8 interacts with Rpt3/6 pair while Rpn11 interacts with Rpn1 of base subcomplex. Some of the lid subunit have been shown to directly interact with AAA^+, e.g. Rpn 7 interact with AAA^+ domain Rpt2 and 6 while Rpn6 and 5 with Rpt 3 (Beck et al., 2012). Lid subunits form hand-like structure where five PCI subunits (Rpn3, Rpn7, Rpn6, Rpn5 and Rpn9) form the fingers and Rnp11 palm (Lander et al., 2012). Rpn8 not only connects Rpn3 and Rpn9 but also the palm of the hand Rpn11 (Lander et al., 2012). The six PCI subunits form horseshoe-like structure covering a large part of the ATPase (Rpt3, Rpt6, and Rpt4) in form of a roof (Lasker et al., 2012). The subunit order in the horseshoe heterohexamer is Rpn9/Rpn5/ Rpn6/Rpn7/Rpn3/Rpn12 (Lasker et al., 2012). PCI subunits have scaffolding function and also play important role in maintaining inter subunit contacts, thereby stabilizing base, lid and 19S with 20S. The scaffold formed by PCI subunits also positions the Rpn8/Rpn11 heterodimer in close vicinity to the mouth of ATPase ring, so that the engaged substrate can be deubiquitinated before entering translocation channel. The Rpn2 from base is thought to help in stabilization of lid conformation with the help of Rpn3, Rpn8 and the Rpn11 which extend towards the base. Unlike conventional model recent EM Map also showed that lid subunits interact directly with 20S CP (Lander et al., 2012). It seems that the function of PCI subunit is to bring the essential machinery in 3D space, necessary for efficient degradation (**Fig. 1.3A and B**).

1.4 Proteasome degradation pathways: In eukaryotic cells proteins are degraded by proteasome by various ways but broadly it could be divided into-

a) Classical pathway (Ub dependent) and

b) Non classical pathway (Ub independent and processing). While Ub dependent pathway is well understood, recent reports suggest proteasome can also degrade substrate without Ub tag.

1.4.1 Ubiquitin dependent proteasomal degradation: Ubiquitin (Ub) is an evolutionarily conserved 76-residue protein. Protein ubiquitination is one of the very important post translation modifications. There are 7 Lys residues in Ub, while chains of four or more Ub linked mostly through K48, tag the protein for degradation, other linkages and mono Ub, are associated with other cellular functions (Hicke, 2001; Weissman, 2001). Ub modification is multistep, energy consuming process. In the first step, Ub gets activated by E1 enzyme (activating enzyme) in the presence of ATP forming Ub-E1 thiolester. This can be recognized by several E2 enzymes (conjugating enzyme) and Ub is transferred by another thiol ester linkage. E2 enzymes help in bringing activated Ub to the substrate. Finally, E2 enzymes associate with E3 enzymes (Ub ligase) which are responsible for the final target selection and specificity. With some exception, E3 enzymes can be broadly classified into two distinct families: a) HECT domain (**H**omologous to **E**6-AP **c**arboxyl **t**erminus) family, which forms a covalent (thiolester) bond with the ubiquitin before transferring it to the substrate. b) RING domain (**r**eally **i**nteresting **n**ew **g**ene) contain Cys and His residues that co-ordinate two metal atoms (Pickart, 2001; Weissman, 2001). As compared to HECT domain ubiquitination by RING finger E3 family involves direct transfer of Ub from Ub-E2 complex to the targeted protein without the addition of thiolester.

E3 ligase sequentially attaches other Ub molecules on the Lys residue of the first Ub molecule. Poly-ubiquitination of some of the substrates is assisted by U-box-domain containing E3 ligases (also called E4 ligases) (Hoppe, 2005). After the attachment of multi Ub degradation tag, substrates are recruited to proteasome. There are five Ub

receptors currently known to be associated with the proteasome, two of them are proteasome subunits (Rpn10 and Rpn13). Rad23, Dsk2, and Ddi1 contain Ub binding domain as well as Ubl domain with which these 'shuttling' or adaptor proteins interact with 19S RP (Gomez et al., 2011). The Ub tags are removed before substrate unfolding and degradation by Cys protease family member called DUb enzyme (Rnp11) (**Fig. 1.5**) (Verma et al., 2002).

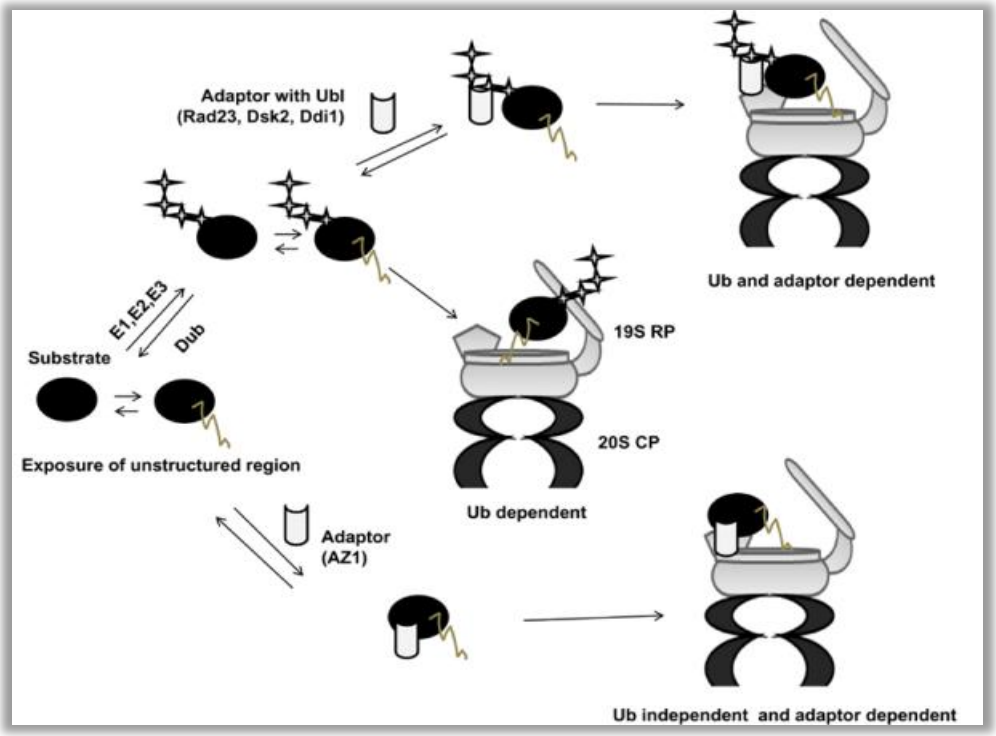

Figure 1.5 Ub dependent and independent proteasomal degradation: *When the substrates are ready for degradation, they are either targeted to proteasome through Ub dependent or independent manner. Several adaptor proteins may help in recruiting substrate to the proteasome. The essential loosely structured or unstructured (Cis-acting) element may originate from substrate during or before targeting it to the proteasome system.*

1.4.2A Ub independent proteasomal degradation: Most of the well-studied proteasomal substrates are degraded by Ub-dependent pathway. However, recently several substrates have been identified which do not follow the classical pathway. The non-classical proteasomal degradation can be broadly grouped into following: a) Ub-

independent degradation (also include degradation directly by 20S CP) b) Protein processing. Generally, unfolded or oxidized proteins do not need the help of 19S RP. The role of Ub is presumably done by exposed hydrophobic regions in the unfolded proteins. In case of Ub independent degradation of globular proteins, these are recruited to proteasome by 'adaptor' or 'shuttling proteins' (Chattopadhyay et al., 2001) (**Table 1.5**). Apart from adaptor protein, substrate must contain an 'unstructured region' for efficient degradation (Prakash et al., 2004) (**Fig. 1.5**). Recently, several non-classical proteasome substrates have been identified, for example-ornithine decarboxylase (ODC), Thymidylate synthase, α-synuclein, $p21^{Cip1}$, tau, proteasome substrate RPN4, p53, p73, HIF-1α, Rb, p105 subunit of NF-kB, pertussis toxin, NFAT5, Aurora-Akinase, pp89, KLF5, hepatitis C virus (HCV) F protein, c-Jun, calmodulin (CaM) and troponin C (Finley, 2009; Sorokin et al., 2009). Many of these substrates are degraded by both Ub dependent as well as independent pathways. In these cases, it is difficult to predict which pathway is default; probably it is mainly dependent on the signal and cell type. Several viral proteins also have been found to play key role in proteasomal degradation by regulating the dependence of Ub for proteasome degradation. Few important substrates that follow Ub independent pathways is discussed below (and **Table 1.4**):

- **ODC:** Ornithine decarboxylase (ODC) was the first example of Ub-independent degradation of a globular protein (Bercovich et al., 1989; Rosenberg-Hasson et al., 1989). ODC degradation was ATP-dependent and regulated by a protein called antizyme 1 (AZ1) (Pegg, 2006). ODC is the initial enzyme in the polyamines biosynthesis. In response to excess cellular polyamines, level of AZ1 increases and binds to ODC, resulting in destruction of enzymatically active homodimer of the same. This eventually leads to ODC degradation. It was shown that the formation of the

ODC/AZ1 complex induces the exposure of the 37-residues from ODC C-terminus, which increases the efficiency of proteasomal recognition (Ghoda et al., 1989). C-terminus of ODC has been shown to harbor two recognition elements - the sequence ARINV along with C441 forms one recognition element while AZ1 in conjunction with other residues constitute a second recognition element. ODC/AZ1 complex compete with Ub for proteasome interaction (Takeuchi et al., 2008). AZ1 was later shown to regulate the intracellular half-life of several substrates (Sorokin et al., 2009). Fusion of ODC C-terminus to the GFP has been shown to shorten its intracellular half-life (Corish and Tyler-Smith, 1999).

- **p53:** Wt p53 is the most studied tumor suppressor that accumulates in response to DNA damage as well as other types of stress and induce either growth arrest or apoptosis. Mutation in p53 and its regulator has been associated with at least 50% of the cancers. p53 has been shown to be regulated both by Ub dependent and independent processes. Several E3 ligases have been shown to regulate half-life of p53. Some of these include Mdm2, E6-AP/E6, ARF-BP1, COP1, CHIP, Pirh2 and Topors (Sorokin et al., 2009). The role of these ligases is mostly cell type dependent and their elevated levels have been directly correlated with different types of cancers. Mdm2, one of most popular E3 ligases, belongs to RING domain E3 ligase. It directly binds to N-terminus of p53 and targets it for Ub-dependent degradation; or associate with gankyrin (PSMD10) and PSMC4 to promote p53 ubiquitination and degradation (Brooks and Gu, 2006; Marine and Lozano, 2010). In cervical cancer, E6 protein of HPV (**H**uman **p**apilloma **v**irus), p53 and E6-AP (**E6-a**ssociated **p**rotein or Ubiquitin-protein ligase E3A) form a ternary complex and E6-AP ubiquitinates p53 for degradation (Camus et al., 2007). In p53 there are two binding sites for E6 protein; one is in the DNA-binding domain and other at the C-terminus. The existence of alternative pathway was observed

when in Ub-defective system (thermo sensitive E1 or ectopic expression of lysine-less Ub), p53 was not completely stabilized. E6 has also been shown to degrade p53 in Ub independent manner (Camus et al., 2007). Another p53 regulator is NQO1 (**N**ADH **q**uinone **o**xidoreductase 1), which binds to p53 in NADH-dependent manner. Interestingly, the presence of competitive inhibitor of NQO1 (dicoumarol) results in dissociation of the NQO1-p53 complex and p53 becomes highly unstable and degraded by 20S CP (Anwar et al., 2003; Asher et al., 2002; Asher et al., 2003).

NQO1 has been found to co-elute with 20S but not with 26S proteasome. This association has been shown to prevent the degradation of several proteins, such as p53, p73, and ODC. NQO1 plays the role of 'gatekeeper' of the 20S proteasome and degradation is dependent on NADH concentration, at high levels of NADH substrates are protected and does not enter the 20S CP (Asher et al., 2005).

Table 1.4 Examples of Ub-independent degradation and their regulators

Substrate	**Proteasome**	**Regulators**	**Remarks**
ODC	20S and 26S	NQO1(-) and AZ1(+)	C-terminal 37 amino acids are essential
Cyclin D1	26S	AZ1 (+)	*in vitro* experiment
Aurora-A kinase		AZ1 (+) and Aurora-A kinase-interacting protein 1 (+)	
$p21^{CIP1}$	PA28γ-20S, 20S and 26S	Cyclin D1, HSP90/WISp39 (-) and MDM2 (+)	Binds α7 subunit of 20S directly, co-eluted with both 20S and 26S
c-Fos		Phosphorylation of C-terminal S364 and 374 by Rsk1/2 and Erk1/2 (-)	C-terminal destabilization region (Ub-independent degradation of nuclear c-Fos)
Fra 1		C-terminal 30-40 residues are essential	C-terminal destabilization region
p53	20S and 26S	NOQ 1(-), E6/E6AP (+)	Exposed C terminus result in E6 and NOQ1 independent degradation.
HIF-1α	20S and 26S	Acetylated Hsp90 and Hsp70 (+) leads to improper folding of HIF protein.	Depending on oxygen concentration follow deferent degradation path
Rb	20S	MDM2(+)	Rb MDM2 bind to α7 of 20S
Thymidylate synthase		No known regulators	N-terminal 30 residue are important
SRC-3/AIB1	PA28γ-20S		HAT domain of SRC-3 binds to PA28γ

- **Rb:** Both Ub-dependent and Ub-independent degradation mechanisms have been described for Rb. Several viral proteins like E7 (HPV), EBNA3C (Epstein Barr virus) and NS5B (hepatitis C virus) have been shown to enhance the degradation of Rb. Human cytomegalovirus (HCMV) protein pp71 has been shown to help in Ub independent degradation of Rb (Kalejta and Shenk, 2003). There are two contradictory reports regarding the role of Mdm2 in Rb degradation. It was demonstrated that Mdm2 binds to the hypophosphorylated Rb and enhances its ubiquitination (Sdek et al., 2004; Uchida et al., 2005). On the other hand, Rb was found to be part of 20S CP/ Mdm2 complex. Both the RING-domain of Mdm2 and C-pocket of Rb were found to be necessary for binding of α7 of 20S CP, leading to degradation of Rb (Sdek et al., 2005). There are two possibilities, either Mdm2 regulates Rb degradation by both the pathway simultaneously or its Ub dependent or independent action is cell type dependent.
- **$p21^{Cip1}$:** $p21^{Cip1}$ is the inhibitor of cyclin-dependent kinase that was shown to be degraded in Ub-dependent manner. But, Lys less $p21^{Cip1}$ (non-ubiqutinable) has also been shown to be degraded by proteasome (Sheaff et al., 2000). In some cell lines, different RP like PA28γ has been found to be associated with Ub-independent degradation of $p21^{Cip1}$ (Chen et al., 2007). In some cases $p21^{Cip1}$ was degraded directly by 20 CP, by interaction of its C-terminus with α7 of the 20S CP (Touitou et al., 2001).
- **M140:** Murine cytomegalovirus (MCMV) US22 family genes M36, M139, M140, and M141 promote efficient replication of the virus in macrophages. A PEST sequence was identified in the C-terminal region of pM141 (360-378) which was responsible for its proteasomal degradation. The residues 306-380 of pM141 contain pM140 binding region (Bolin et al., 2010). Binding of pM140 protects its binding partner, pM141, from degradation by the proteasome in ubiquitin-independent manner (Bolin et al., 2010).

- **Thymidylate synthase (TS):** It catalyzes the conversion of dUMP to dTMP, essential for DNA replication. The intracellular half-life of TS is 10-12 hour (Kitchens et al., 1999). The proteasomal degradation of Lys less mutant of TS showed that TS degradation is independent of Ub. Mutagenesis studies have demonstrated that the N-terminal 30 residue contains the primary half-life determinant of TS and that penultimate Pro2 residue is necessary for degradation (Pena et al., 2006).

1.4.2B Protein Processing: The classical proteasomal degradation mechanism involves substrate localization to proteasome, unfolding and degradation into small peptides. The degradation progresses through N to C or C to N terminus and in some case by endoproteolysis. Sometimes proteasomal degradations are non-processive and a stable domain may escape degradation. It was shown that three polypeptide chains may enter the catalytic chamber simultaneously. The endoproteolytic activity and non-processive proteasomal degradation was demonstrated with the help of synthetic substrate where proteins α-synuclein and/or p21Cip1 were fused to the C and/or N-termini of GFP protein (GFP-α-synuclein, GFP-p21Cip1 or α-synuclein-GFP, p21Cip1-GFP) (Liu et al., 2003). The α-Synuclein and p21Cip1 are unstructured proteins and have been shown to be degraded by 20S or 26S proteasome in ub-independent manner. GFP on the other hand is a structured protein. The 20S and 26S proteasomes were able to degrade only the unstructured part (α-synuclein or p21Cip1) leaving GFP intact. This demonstrated that the substrate can be degraded by the proteasome from either end of the polypeptide chain and during this process encounter with stable domain may lead to incomplete degradation leaving stable part intact. When GFP was fused to both the termini (GFP-α-synuclein-GFP and GFP-p21Cip1-GFP) only unstructured protein was degraded leaving GFP intact (Liu et al., 2003). This clearly indicates that proteasome can degrade substrate even from internal region (endoproteolysis).

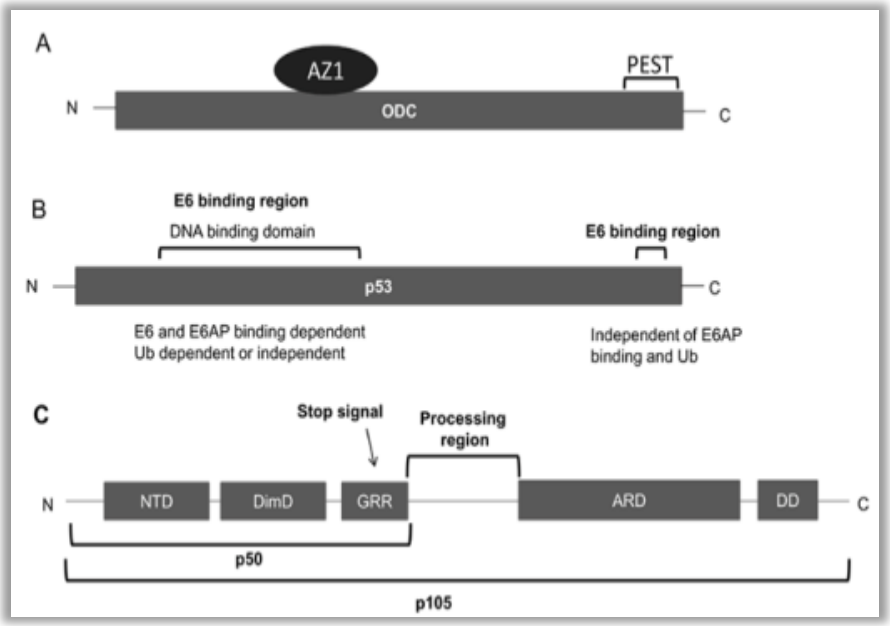

Figure 1.6 Classical and non-classical proteasomal degradation: *Due to elevated intracellular level of product formed by ODC, AZ1 level increases, it binds to ODC. Binding of AZ1 to ODC exposes C-terminal PEST sequences essential for degradation in Ub independent manner (A). p53 is targeted for degradation by various E3 ligases (Mdm2), it was also degraded with the help of viral protein E6 and cellular protein E6AP by both Ub dependent or independent manner (B). Sometimes proteasome help in generation of active form of the protein. Processing of p150 starts from an internal region in the protein and N-terminus is protected by stop signal GRR, while C-terminus is degraded by proteasome (C).*

Proteasome regulates the function of several proteins by processing and/or providing mature polypeptide; this action can also be dependent or independent of Ub. Some of the examples are given below-

- **NF-kB p105:** NF-kB is a dimeric transcription factor. It controls the expression of a wide range of genes. Besides, depending on cell type and condition they regulate cell proliferation or apoptosis. The NF-kB family consists of five members: p50, p52, p65/RelA, c-rel, and RelB. Of these p50 and p52 are formed as a result of processing from precursors p105 and p100 respectively. The p105 processing can be performed both by the 26S proteasome and 20S CP (Coux and Goldberg, 1998). p105 is composed of 4 domains: the N-terminal domain (NTD), the dimerization domain (DimD), ankyrin-

repeat domain (ARD) and a death domain (DD). The DimD and ARD domains are connected through a 180-residue long flexible region and acts as site of processing and the N-terminus of this region contains a glycine-rich region (GRR) that acts as stop signal (Orian et al., 1999). External stimuli activates IkB kinase β, which phosphorylates p105 at residues Ser927 and Ser932, thereby triggering its proteasomal degradation. 20S CP was also shown to generate p50 from p105 *in vitro* (Moorthy et al., 2006). In this case, the processing begins from the unstructured region of p105 by endoproteolysis. However, due to the presence of GRR the degradation of N-terminal region is prevented while C-terminal region was degraded.

- **Y-box-binding protein 1 (YB-1):** YB-1 is a DNA/RNA-binding nuclear-cytoplasmic shuttling protein. In the nucleus, it has been implicated in a variety of functions, including transcription, replication, DNA repair, DNA recombination and alternative pre mRNA splicing. In the cytoplasm, YB-1 is involved in various aspects of mRNA metabolism. The cleavage of YB-1 by 20S into two fragments was shown *in vitro,* but in cells mostly the N- terminal fragment is predominant. This may be due to the degradation of C-terminal fragment. The cleavage is performed by the caspase-like activity (after Glu219) of the 20S CP (Sorokin et al., 2005). After the proteasome processing, the N-terminal fragment of YB-1 gets accumulated in the nucleus (C terminal region contain cytoplasmic retention signal). Besides this processing, the full length protein may undergo complete degradation by proteasome. It seems, the functional amount of YB-1 in the cell is controlled by the Ub-dependent degradation, whereas in presence of certain stimuli or condition, it is processed.
- **eIF4G and eIF3a:** elF4G and elF3a are translation initiation factors. In rabbit reticulocyte lysates, the fragments of eIF4G (factor eIF4F) and eIF3a (factor eIF3) were found. It was shown *in vitro* that the 20S proteasome was able to cleave eIF4G and

eIF3a in Ub independent manner and inhibited the translation of mRNAs that are dependent on these factors (De Benedetti and Graff, 2004; Dong and Zhang, 2006).

Table 1.5 Adaptors in proteasomal degradation

Ub-dependent degradation	
Name	**Proposed role**
Rad23 Dsk2 Ddi1 NUB1	Contain N-terminal Uba (binds proteasome) and C- terminal Ubl domain (bind Ub). Act by increasing local proteasome concentration in regions where quick degradation is necessary (at the site of DNA damage) or increase the ubiquitinated substrate concentration for quick degradation.
Other Adaptors of Ub-dependent degradation	
Name	**Proposed role**
Cdc48/p97	ATP dependent chaperone may help in substrate unfolding.
E3 ligases	Increase substrate concentration as well as Ub length.
Ub-independent degradation	
Name	**Proteasome substrates**
AZ1	ODC, Arora Kinase A, Cyclin D1
HBZ	
FAT10	Also proteasome substrate
NOQ1	ODC, p53, p73

1.5 DEGRADATION SIGNALS

Degradation signals are a major determinant of a protein's lifespan. The widely used degradation signal in archaea and prokaryotic system is ssRA. In bacterial system, ssRA tagging occurs when a ribosome gets stuck due to truncated mRNA. It cannot detach from the defective mRNA without a termination codon. A special type of RNA known as ssRA ("**s**mall **s**table **R**NA **A**") or tmRNA ("**t**ransfer-**m**essenger RNA") rescues the ribosome by adding an eleven codon degradation tag followed by a stop codon (Keiler et al., 1996; Tu et al., 1995). This allows the ribosome detachment from truncated mRNA and ssRA tagged incomplete protein then gets degraded by the proteases ClpXP or ClpAP. Although the originally identified ssRA/tmRNA tag had the sequence AANDENYALAA, additional designed degradation tags that differ in the last three residues (AAV, ASV, LVA, LAA) have been found to affect protein half-lives.

In eukaryotic system, ssRA like sequence is not identified till date. The N-end rule degron was the first degradation signal identified in eukaryotes. Other degradation

signals that have been studied in detail are the Ub tagging signal (like phosphorylation), the PEST sequence, and the destruction box. According to the N-end rule the nature of the N-terminal amino acid defines the half-life of a protein. Newly synthesized proteins normally contain an N-terminal methionine. To create an N-end rule degradation signal, the stabilizing Met at the N-terminus has to be removed. Depending on their N-terminal amino acid, the half-lives of proteins range from a few minutes to hours (Bachmair et al., 1986).

Some substrates harbor a non-cleavable N-terminal ubiquitin that functions as a degradation signal itself. Ufd4 and Ufd2, which are an E3 and E4 ligase, respectively, recognize these and help in degradation.

The anaphase promoting complex (APC) is a multi-subunit E3 protein. APC is responsible for ubiquitination of cell cycle regulators at the metaphase-anaphase and mitosis-G1 transitions. APC substrates (like cyclin A and B) contain two degradation motifs: destruction box (D-box- RTVLGVIGD) and KEN box (Glotzer et al., 1991; Yamano et al., 1998). Most APC substrates contain D-box and are degraded quickly by proteasome.

Proteins with short intracellular half-life contain a region rich in Pro, Glu, Ser and Thr, these regions are called as PEST sequences (Rechsteiner and Rogers, 1996; Rogers et al., 1986; Wang et al., 2003). PEST sequences are hydrophilic stretches, about 12 amino acids in length and contain at least one Pro, one Glu or Asp and one Ser or Thr. They are flanked by Lys, Arg or His residues, but positively charged residues should not be present within the PEST sequence.

Most of these degradation signals are the signal for ubiquitination. Degradation signal similar to ssRA have not been identified in higher organism.

1.6 MODELS OF SUBSTRATE UNFOLDING AND TRANSLOCATION BY PROTEASOME

Like other compartmentalized protease and due to the fact that folded proteins cannot enter the narrow catalytic core it was thought that ATPases of 19S RP unfolds the substrate. But whether unfolding and translocation are separate events or concurrent and whether these processes are inter-dependent is largely unknown. There are two schools of thoughts about the unfolding and translocation.

- Model I proposes that unfolding occurs on the surface of the proteasome and that translocation is a distinct process that can begin only when unfolding will generate the unstructured region required to enter translocation channel. Studies carried out with the PAN protease are in agreement with this model (Navon and Goldberg, 2001). This model emphasizes on establishment of multiple contacts between the substrate and the surface of the ATPase ring as well as ATP-driven motion within the ring, which mechanically unfolded the substrate.
- An alternative model proposes that unfolding of the substrate is driven by its interactions within the translocation channel. Unfolding occurs due to translocation, which is induced by pulling action of the translocation motor on the substrate. Model II is simplistic because it reduces unfolding and translocation to a single process for which a single strong interaction site is sufficient (Pickart and Cohen, 2004; Prakash et al., 2004).

The mechanism of substrate unfolding has not been clearly elucidated for 26S proteasome. Some clues can be derived from PAN-20S system, where GFP-ssRA unfolding and degradation has been demonstrated and this process was found to be dependent on ATP hydrolysis. PAN alone was able to unfold GFP-ssRA but PAN-20S accelerated this processes without affecting the rate of ATP hydrolysis (Navon and

Goldberg, 2001). Interestingly blocking the proteolytic activity of 20S affected the rate of unfolding, thereby suggesting that all these processes are interdependent and coupled to one another (Navon and Goldberg, 2001; Zhang et al., 2009b). Series of mutational and structure guided biochemical analysis suggest that two PAN ATPase - subcomplex I (formed by OB domain) and II (AAA^+ domain) might play different roles. Substrate binds to the distal face of subcomplex I and degradation tag is then recognized by surface motifs in subcomplex II, cycles of ATP hydrolysis result in a pulling force that is exerted on the folded substrate protein (Zhang et al., 2009b). Translocation is hindered when the folded substrate is pulled against the narrow interface between the two sub-complexes (CC-OB domain) (Zhang et al., 2009b). Continuous pulling forces from subcomplex II generates tension on the leading end of the substrate, resulting in unfolding and resumption of translocation.

In 26S proteasome ATP binding and hydrolysis play different role. For instance, ATP binding is sufficient for 19S and 20S assembly while hydrolysis is required for substrate unfolding and translocation (Liu et al., 2006). Unlike PAN where single ATPase form hexameric complex, eukaryotic ATPases are six different gene products, which form dimer of trimer to form hexameric complex.

1.7 RECENT ADVANCEMENT IN UNDERSTANDING 26S PROTEASOMAL DEGRADATION

Proteasome is a master regulator of virtually all the cellular events and despite its clinical utility; we are only beginning to understand the molecular mechanism of proteasomal degradation. This is mainly due to the fact that there are very few model substrates and system to answer the fundamental questions about proteasomal degradation like: a) what are degradation signals, b) how substrate are recruited to the proteasome, c) how sometimes recruited substrate escape degradation, c) whether 19S

and 20S regulate to each other, if yes than how, d) how substrate are unfolded and traslocated to the catalytic core, e) what is mechanism of unfolding *etc*. Recently, high resolution EM map of 26S proteasome has become available, the salient finding have been discussed earlier. Some of model substrate and system used to understand the proteasomal degradation are listed in table 1.6, most of these systems use reticulocyte lysate as a source of proteasome, substrates are either tagged with tetra Ub, or other localization factor and unstructured regions were derived from other proteins. Some of recent findings, relevant in the present context are summarized below-

a) Pioneer study from Dr. Matouschek's group suggests that short lived proteins usually contain a long unstructured region and degradation initiate from the unstructured region (Prakash et al., 2009; Prakash et al., 2004). Mostly they have used dihydrofolate reductase (DHFR) or barnase as a proteasomal substrate and attached either minimal Ub signal (tetra Ub) or Ub-like domain (Ubl) for proteasome recruitment. Ub tags are recognized mainly by Rpn10 and Rpn13 while Ubl is mainly recognized by Rpn1. Tetra Ub-DHFR was stable for degradation but when long unstructured region was attached, it was degraded by proteasome and unstructured region was the first to get degraded. Later, Dr. Coffino's group also showed that GFP protein when fused with Rpn10 localized to proteasome but it was degraded only when the C-terminal of ODC was fused to it, again emphasizing the requirement of loosely folded structure for efficient proteasomal degradation (Takeuchi et al., 2007).

Recently, either Ub-DHFR or UBL-DHRF was fused 20 to 150 residues unstructured region derived from cytochrome b2 or F0 ATPase and their degradation was monitored. The minimal length of unstructured fusion for Ub4-DHFR was 29-34 residues, while for Ubl-DHFR it was 33-44 residues (Inobe et al., 2011). Not only the length but their placement from recognition tag also affected degradation. If substrate

was recognized by Ub, unstructured region adjacent to the Ub accelerated degradation. In contrast, in case of recognition by Ubl tag, unstructured region should be long and separated in space from the UbL tag for efficient degradation (Inobe et al., 2011).

The difference in the length of unstructured region and spacing between recognition and unstructured region can be understood in the light of EM map. Ub-recognition subunit (Rpn10 and 13) as well as the Dub subunit, Rpn11 are placed above ATPase ring. UBL recognizing subunit Rnp1 on the other hand is physically separated from translocation channel at ATPase ring, which is `why Ub tagged protein needed shorter unstructured region while Ubl tagged protein not only needed longer unstructured region but also unstructured region was far from recognition region (Ubl).

b) Although, it has been demonstrated long back that proteasome can degrade proteins in N to C, C to N direction as well from the internal region of protein by endoproteolysis (discussed earlier), recent report suggests that the stability of termini provides the direction of degradation (Berko et al., 2012). It was shown that unstructured proteins do not have any preference for direction, while direction of degradation for some substrates are fixed. This property results from interactions of the substrate's termini with the regulatory ATPase and could be predicted based on the calculated relative stabilities of the N and C termini. It was also demonstrated that the direction of degradation affected the peptide generated by proteasome and thereby influence the antigen presentation by MHC class I (Berko et al., 2012).

c) Mono ubiquitination is linked with membrane fusion as well as transcription in a proteolysis independent manner. Few recent reports suggest connection between mono ubiquitination and proteasomal degradation, e.g., Cyclin B1 and phospholipase D. Mono ubiquitinated substrates which are <150 residue in length can be degraded by proteasome but longer protein require poly Ub chain (Shabek et al., 2012). These

observations also suggest that proteasome is versatile in context of substrate recognition.

d) The stability of protein near the degradation signal (Ub and/or long unstructured region) has been shown to affect the efficiency of protein degradation. Coffino's group have shown that rate of ATP hydrolysis (unfolding) does not change no matter whether a stable or unstructured protein is subjected to proteasomal degradation, stable protein takes more time for degradation. Titin I27 protein module or human DHFR as well as their mutants of varying stability were fused to Rpn10 for proteasome localization and C-terminal degron of ODC was attached to it. The rate of ATP consumption and half-life of these proteins were analyzed. As compared to stable protein destabilized protein was degraded faster while ATP turnover did not change (110/min./proteasome) (assuming other processes are not rate limiting and saturating substrate concentration) (Henderson et al., 2011).

e) 19S and 20S are spatially and functionally distinct but they regulate the function of each other. This was experimentally demonstrated by blocking the 20S proteasomal active sites by specific inhibitor which resulted in stabilization of the 26S holo-complex. Moreover, binding of poly Ub proteins to the 19S regulator results in stimulation of the protease activity of 20S proteasome suggesting that binding of polyUb substrates to the 19S regulator transfers the signal of substrate recognition, may be in the form of conformational changes to the 20S proteasome (Bech-Otschir et al., 2009).

Table 1.6 Some model substrate and system used to understand proteasomal degradation

Substrate	Source of substrate	Source of proteasome	ubiquitination	Degradation signal	References
N-degron DHFR-barstar hybrid	*In vitro* translation	Reticulocyte lysate	N terminal Ub, N-end rule pathway	Linker and 18 residue (318-335) of lac repressor	(Prakash et al., 2009)
N-degron-DHFR-barstar and USR barnase heterodimer	*In vitro* translation	Reticulocyte lysate	N terminus DHFR-barstar is ubiquitinated	USR or cytochrome b2	
Ub4-USR2-barstar and USR1 barnase	Purified	Purified proteasome	Ub tag is attached 102 amino acid long linker or 129 amino acid p105	1-95 residues of cytochrome b2	
DHFR, circularly permutated DHFR, barnase, DHFR-barnase	*In vitro* translation	Reticulocyte lysate	Ub is attached to 40 residue linker derived from lac repressor	Lac repressor	(Lee et al., 2001)
N-Usignal-DHFR-barnase-C, N-barnase-DHFR-Usignal-C	*In vitro* translation and partial purified substrate	Reticulocyte lysate	N-end rule pathway	1–95 of cytochrome b2 or residues 1–40 of lac repressor	(Prakash et al., 2004)
N-Usignal-DHFR-unstructured-6Xhis	Purified	Reticulocyte lysate	Ub was attached N to C terminus	1–95 of cytochrome b2 or residues 1–40 of lac repressor	
N-Ubl- DHFR-RBD-DHFR-C, N-Ubl-Uba-RBD-Uba-unstructured region-C	*In vitro* translation and partial purified	Purified They have also done in vivo	Ubl and 2 Uba domain	1–95 of cytochrome b2 or residues 1–40 of lac repressor	(Fishbain et al., 2011)
UbcH10,Ub4 UbcH10, Ub4 PEST UbcH10	Purified	Purified	Tetra Ub	PEST- 37aa of ODC	(Zhao et al., 2010)
Titin I27-cODC, RPN10- Titin I27-cODC, RPN10-DHFR-cODC	Purified	Purified	RPN10 was used for localization	PEST- 37aa of ODC	(Henderson et al., 2011)
p21,α-synuclein	Purified	20S and 26S Purified		Both proteins are natively disordered	(Liu et al., 2003)
GFP-p21, GFP-α-synuclein, GFP-p21-GFP and GFP- α-synuclein-GFP	Purified	20S and 26S Purified	ubi not required	p21,α-synuclein proteins are natively disordered	
Ornithine decarboxylase (ODC), DHFR ODC,	Purified radio labeled	Purified and reticulocyte lysate	Ub-not required. Antizyme 1 (AZ1) is necessary	disordered 37-residue region at the C-terminal end.	(Takeuchi et al., 2007; Takeuchi et al., 2008;

GFP ODC					Zhang et al., 2003)
Fpr1-rapamycin binding domain of Tor1 Rpn10 Fpr1 complex, Tor fused His 3	Purified	Purified	Ub not required. Fpr1 was fused to proteasomal subunit. Tor was fused to His3	Tapamycin hetero dimers Fpr1 and tor1 and His3 translocate to proteasome	(Janse et al., 2004)
Thymidylate synthase	In vivo	Cellular	Ub-independent	Internal N-terminal	(Forsthoefel et al., 2004; Pena et al., 2006)

URS 1- one copy of unstructured region from lac repressor, USR2- 210 amino acid long, 2 copies of unstructured region

Rationale of the study

&

Aim and objectives

RATIONALE OF THE STUDY

All biological processes are tightly regulated by compartmentalized, ATP dependent protease, the 26S proteasome. Substrates are degraded by proteasome either by Ub dependent or independent pathway. The substrates are recruited to the proteasome either by Ub or adaptor proteins. Despite localization to the proteasome, all proteins are not amenable for proteasomal degradation. Many proteins are found be heavily ubiquitinated as well as recruited to the proteasome but escape degradation. Therefore, the question arises whether 'trans acting elements' like Ub or adaptors are sufficient, or the choice lies in the substrate itself ('cis-acting'). One of the possible 'cis-acting' element might be the 'unstructured region' present in the substrates. A clue towards involvement of such regions was obtained from faster degradation of protein, when fused with long unstructured region derived from another protein. This clearly indicated the requirement of an element other than proteasomal binding for efficient degradation. When such elements are buried inside, even if the substrate gets localized to the proteasome it will not be degraded. Now, the question arises, how these elements get exposed or present themselves to the proteasome and whether there is a preference for any specific sequence or conformation. Despite being a necessary regulator of several cellular processes and recent clinical utility, many fundamental aspects of proteasomal degradation are still largely unknown. This is mainly due the limited availability of model systems to understand the process. This is further complicated by the complex architecture of the 26S proteasome and the fact that not all proteins are amenable for degradation *in vitro*.

More specifically, the role of protein sequence, structure, thermodynamic and kinetic aspects of degradation is largely an unexplored arena. To the best of our knowledge the inherent ability to of purified 26S proteasome to degrade a globular

protein without the help of trans-acting elements has not been demonstrated till date. Some of the fundamental question that we are interested in are -

1. What are the cis-acting elements in the substrates?
2. If an unstructured region is necessary, how does it originate within substrate?
3. What are the sequence and structural requirements for protein degradation?
4. What are the rate limiting steps in proteasomal degradation?

We believe that these questions can only be answered if a model system composed of purified components, which could recapitulate the various hierarchical steps involved in the proteasomal degradation, without the help of any trans-acting elements could be created. With this background, the following objectives were proposed.

AIMS and OBJECTIVES

The aim is to develop an *in vitro* model system for the characterization of substrate recognition, global unfolding and degradation by eukaryotic proteasome.

Special focus:

1. To establish an *in vitro* model system that would recapitulate the hierarchical steps in proteasomal degradation.
2. To explore how proteasome senses the presence of substrate and develop an assay with which the affinity of substrate with proteasome could be deciphered.

3. Using structural and biophysical approach for understanding structure function correlation.
4. To identify ‘cis-acting’ element and try to understand so as how it originated in substrate.
5. To understand the effect of local secondary structure on proteasomal degradation.
6. To identify proteasome interacting surfaces in substrate.

Finally, we propose to combine all the information in the form of a model that would help in understanding proteasomal degradation in greater details.