Chapter 1

Introduction
1 Introduction

1.1 Proteins

The smallest organized form of life is a cell. Cell is made up of macromolecules such as carbohydrates, proteins and lipids. Among all these, proteins gain a lot of importance because of their functional diversity, astounding structural variety and physico-chemical properties. By enacting the message hidden in DNA, proteins make a living cell functional. Intricate structures of proteins give proteins their function. Since they are involved in almost all cellular processes directly or indirectly, a highly sophisticated network exists in a cell to maintain homeostasis between their synthesis and degradation.

1.2 Protein Degradation

In cell, protein degradation takes place for elimination of abnormal, nonfunctional, denatured and potentially toxic proteins. This may arise from nonsense or misense mutations, errors in gene expression or post translational damage. Protein degradation is also important in the regulation of cellular metabolism (Wickner et al., 1999), recycling of essential amino acids and provision of nutrients during nutritional deprivation. The process of protein degradation is executed by several proteolytic enzymes i.e. proteases found in both prokaryotes and eukaryotes. There are several proteases and proteolytic signals found in both prokaryotes and eukaryotes as discussed below. The purpose of entire degradation machinery is to maintain the cellular quality control.

1.2.1 Protein degradation in prokaryotes

*E. coli* has at least five ATP dependent proteases as Lon, FtsH, ClpAP, ClpXP and HslUV. These enzymes appear to have distinct substrate preferences (Gottesman et al., 1998). In prokaryotic systems, HslUV and ClpP are proteases analogous to the 20S proteasome in eukaryotes. HslUV is composed of the HslV protease and the HslU

---

*Studies on the degradation determinant signals of Ornithine Decarboxylase (ODC)*

Page 1
ATPase. Hsl UV protease is ATP dependent and degrades proteins with the help of chaperone. HslV is made up of two hexameric rings stacked one above the other, with their catalytic sites facing the internal cavity. Association with hexameric rings of HslU on either side, enhances the weak peptidase activity found with HslV several fold. HslU is a chaperon of the Clp/Hsp100 family. HslU has also been shown to recognize a carboxy terminal tag sequence for a substrate called SulA, showing functional similarity with ClpX and ClpA, although the ClpP and HslV proteases are dissimilar in many respects (Sousa et al., 2000).

Carboxy terminal ssrA tagged proteins are actively degraded by ClpA or ClpX and then translocated to the ClpP proteasome for degradation (Keiler et al., 1996; Gottesman et al., 1998) like ubiquitin proteasomal system of eukaryotes. SsrA or transfer messenger RNA (tmRNA) is a bacterial ribonucleic acid that adds a C terminal degradation peptide to nascent polypeptide chains on stalled ribosomes. These peptides are degraded by RNase R (Reynald Gillet and Brice Felden, 2001). tmRNA which is a unique hybrid between a tRNA and a mRNA is exclusively present in eubacteria. Tagging by tmRNA is seen to be operative in three cases, namely, when a translating ribosome encounters a cluster of rare codons while translating an intact mRNA, when a ribosome reaches a termination codon or when there is translational ‘read through’ at the termination codon due a suppressor tRNA. Sometimes, a stable secondary RNA structure in an intact mRNA may also lead to tagging. tmRNA initially acts as a transfer RNA, whose 3’ end is aminoacylated by an alanyl tRNA synthetase. An associated protein SmpB is required for the tmRNA to bind to a stalled ribosome at the ‘A’ site. The nascent polypeptide is then transferred to the tRNA portion of the alanylated tmRNA and the defective message is replaced by tmRNA ORF. Translation resumes on tmRNA internal ORF having a termination codon. The defective polypeptide I releases and might be degraded subsequently by RNase R. Tagged proteins are released by termination factors. Thus, tmRNA participates in intracellular protein quality control.

In bacteria, degradation signals are generally found to be at the N or C-terminal, which are recognized by the target protease itself. Studies have shown the presence of Tail
Specific Proteases (Tsp) in the periplasm of *E.coli* (Keiler et al., 1996). Tsp is an endoprotease that cleaves proteins at internally discrete sites, specifically degrading proteins of C-terminal tails having WVAAA residues. C terminal tails having small uncharged residues (Ala, Cys, Thr, Ser, Val) are preferred for cleavage by Tsp. Non-polar residues are also preferred in the second and third positions, but larger and more hydrophobic side chains are also cleaved. Tsp directly recognizes folded proteins by using a binding site a free α carboxyl group, at the last three positions of the substrate. Since most damaged or misfolded proteins do not have an intact C terminal sequence, the Tsp protease probably does not remove such abnormal proteins from the cell. Recent studies have shown the presence of a peptide tagging system, where proteins translated from damaged mRNA are modified by C terminal addition of a peptide with the sequence AANDENYALAA and then presented to the Tsp protease for degradation.

1.2.2 Protein degradation in eukaryotes

Eukaryotic protein degradation can be subdivided into specific and nonspecific protein degradation. Specific protein degradation is carried out by a multimeric 26S proteasome (Hershko and Ciechanover., 1998) while non specific protein degradation is carried out by lysosomes. Proteasomal substrates include short lived or regulatory proteins such as proteins involved in cell cycle, gene transcription and signaling pathways (Pickart., 2001). Proteasomes can also degrade damaged or misfolded polypeptides (Pickart et al., 2004). Lysosomal substrates include long lived proteins as well as major cytoplasmic organelles. Most of the cytoplasmic components are degraded by autophagic lysosomal pathway (Seglen et al., 1992). Present study is focused on selective protein degradation.

1.3 Proteasomal protein degradation

Proteasomes are multicatalytic proteinase complexes or macro machines of the degradation system of cell (Cuervo et al., 1995). Proteasomes are present in nucleus as well as cytoplasm of the cells (Rivette and Knecht., 1993). In literature many studies...
have focused on structure of proteasome (Groll et al., 1997; Adams et al., 2003; Unno et al., 2002).

1.3.1 Structure of proteasome

Proteasome is giant macromolecular degradation machinery having size of 2.5 mega Daltons (Coux et al., 1996). 26S proteasomal complex is made up of two different subunits i.e. 20S core particle and 19S regulatory particle (Adams et al., 2003).

![Pictorial representation of 20S proteasome](image)

Figure 1.1 Pictorial representation of 20S proteasome (Groll et al., 1997).

1.3.1.1 Core particle

Core particle (CP) is made up of α and β subunits. The barrel shaped structure of 20S core particle is formed by four rings formed by α and β subunits. Two outer rings are formed by seven α subunits in each ring and two inner rings are formed by seven β subunits in each ring (Unno et al., 2002) (Figure 1.1). Out of seven, three β subunits of inner ring i.e. β1, β2 and β5 contain proteolytic active sites. Each active site functions specifically. β1 cleaves at the C terminus of acidic amino acid residue glutamate, β2 cleaves after tryptic residue, while β5 prefers to cleave after hydrophobic residue (Borisssenko et al., 2007). So subunits β1, β2 and β5 possess post glutamyl peptidyl hydrolase (PGH) like, trypsine like and chymotrypsin like activities respectively.
Length of 20S proteasome is 15nm and diameter is 11nm (Groll et al., 1997). Width of degradation channel is 10 to 15 Å. This narrow channel ensures entry of only unfolded polypeptide inside the 20S proteasome thereby increasing the specificity of protein degradation (Larsen et al., 1997).

1.3.1.2 Regulatory particle

Regulatory particle (RP) contains 19 subunits. The subunits are again subdivided into distal lid and proximal base particles (Glickman et al., 1998). Base is composed of 10 components, 6 of which are ATPases (Rpt), Scaffolding proteins Rpn 1, Rpn 2 and ubiquitin receptors Rpn 10 and Rpn 13. Base subcomplex is involved in substrate unfolding before proteasomal entry. Lid subunit is composed of 9 components including a well studied Rpn11 which has deubiquitinating activity (DUB). Lid subcomplex is involved in recognition of substrate and processing of ubiquitin from substrate before proteasomal entry.

1.4 Ways of presenting substrates to proteasome

There are mainly three ways by which substrate protein is presented to proteasome for degradation. Tagging by ubiquitin is the most preferred mechanism used by eukaryotic cells for selective degradation of substrate proteins (Thrower et al., 2000). This is referred as ubiquitin dependent proteasomal degradation. In another way adaptor protein acts as a bridge between substrate and protease and substrate is brought in the vicinity of proteasome (Hartmann-Petersen et al., 2003). Apart from mediator such as ubiquitin tag or adaptor protein, few protein substrates are directly recognized by proteasome for degradation (Orlowski et al., 2003). This is possible because these proteins are equipped with some specific amino acid sequences in their primary structure which can act as degradation signals. Second and third ways of degradation are referred as ubiquitin independent proteasomal degradation. Proteasomes do not degrade protein into amino acids but into smaller peptides.
1.4.1 Ubiquitin dependent protein degradation

Ubiquitin dependent protein degradation is a complex and tightly regulated process (Wolf et al., 2004). Sequence of events such as initiation of ubiquitination, targeting towards proteasome, degradation and recycling of ubiquitin are discussed below.

1.4.1.1 Initiation of ubiquitination

Many proteins are tagged by ubiquitin for degradation after phosphorylation. Many rapidly degraded proteins contain PEST sequence which serve as phosphorylation sites important for degradation. In some cases phosphorylation at multiple sites is required for ubiquitination (e.g. yeast G1 cyclin Cln3 and Cln2) while for ubiquitination of mammalian G1 regulators cyclin E and cyclin D1 phosphorylation at single specific site is required. IκBα which is an inhibitor of NF-κB transcriptional regulator requires phosphorylation at two specific sites. Different ubiquitin ligases recognize different phosphorylation patterns (Hershko et al., 1998).

N end rule is an important finding which explains a crucial role of N terminal amino acid as a stabilizing or destabilizing residue. Methionine is considered as most stabilizing amino acid residue while proline as most destabilizing residue (Varshavsky., 1996).

Mitotic cyclins and certain cell cycle regulators possess destruction box for the purpose of degradation. Destruction box is 9 amino acid residues long sequence which is located approximately 40 to 50 amino acid residues from N terminus of protein and required for ubiquitination and degradation of protein (Hershko et al., 1998).
1.4.1.2 Ubiquitination and proteasomal degradation

The process of tagging the substrate by ubiquitin is executed by three different types of enzymes named as E1, E2 and E3 (Figure 1.2). E1 is ubiquitin activating enzyme, E2s are ubiquitin conjugating enzymes while E3s are ubiquitin protein ligases. A single E1 (Hershko and Ciechanover, 1998; Pickart, 2001) transfers Ubiquitin to approximately 20-30 distinct E2 enzymes (Schuffner et al., 1998), which in turn interact with several E3’s (Wilkinson, 2000; Pickart, 2001). The specificity of the system is maintained by approximately 1000 different E3’s which interact with specific substrates which may undergo a post translational modification (phosphorylation or oxidation) or an alteration in its structure (denaturation or misfolding). Carboxy terminus of 76th glycine residue of ubiquitin is attached to ε-amino group of lysine residue or amino terminus of target protein to be degraded by isopeptide or peptide bond (Varshavsky., 1997).

Further at least three ubiquitin moieties are added on lysine residue in 48th position of first ubiquitin ligated to the substrate protein. This substrate protein tagged with polyubiquitin chain is presented to the proteasome for degradation. Lid subunit of regulatory particle recognizes the polyubiquitin chain on substrate protein. Rpn 10 and

Figure 1.2 Ubiquitin mediated proteasomal degradation of protein.

Studies on the degradation determinant signals of Ornithine Decarboxylase (ODC)
Rpn 13, subunits of regulatory complex, are involved in recognition of ubiquitin tag (Deveraux et al., 1994). Subunits Rpn 10 in yeast and S5a in mammals are localized at base lid interface of 19S particle. The interaction of these recognition subunits occurs through ubiquitin interacting motif (UIM) domain (Hofmann et al., 2001). Other subunits such as Rad 23 and Dsk 2 were also found to be involved in recognition of multiubiquitin chain and presenting the substrate to proteasome (Rao et al., 2002). Once polyubiquitin chain is recognized by proteasome, an unstructured region is required by proteasome to thread the polypeptide chain into proteolytic core (Prakash et al., 2004). The process of unfolding of substrate protein so as to enter in proteolytic core is catalyzed by regulatory particle of proteasome. At the same time trimming of polyubiquitin moiety from substrate protein by de-ubiquitinating enzymes is required for recycling of Ubiquitin. Deubiquitinating enzymes do not allow polyubiquitin chain to enter in proteolytic core (Yao et al., 2002). The process of unfolding is energy dependent (Lee et al., 2001; Rotanova et al., 2008). Stages of unfolding and translocation of protein targets are coupled to ATP hydrolysis. Once a small unstructured segment of polypeptide chain enters in core particle, remaining portion is taken up smoothly and subjected to degradation (Wolf et al., 2004). In this way, multitasking machine of the cell works for protein degradation as shown below in figure 1.3.

Figure 1.3 Execution of protein degradation process by proteasome in various steps.
Membrane and secretory proteins are folded in endoplasmic reticulum (ER). Nascent proteins which fail to fold in ER are degraded by ER associated degradation (ERAD) (Brodsky et al., 1999). Substrates of ERAD are translocated across the membrane of ER and degraded by cytoplasmic proteasomes (Kopito et al., 1997).

An important step in ubiquitin mediated proteasomal degradation of protein is recycling of ubiquitin moieties. Homeostasis of ubiquitin itself is described in detail in chapter 5.

1.4.1.3 Role of adaptor proteins

As discussed above Rpn 10 in yeast and S5a in mammals (subunits of regulatory complex) are involved in recognition of ubiquitin tag. Events of recognition of polyubiquitin chain, dragging of substrate protein inside the core particle and deubiquitination are carried out by proteasomal assembly. So the first event i.e. recognition of polyubiquitin chain is important to start with degradation process. This indicates Rpn 10 or S5a are indispensible for degradation. But yeast strains with deletion mutants of genes coding for these recognition subunits were viable and showed impaired protein degradation. This has given a new notion that apart from proteasomal subunits, another set of proteins exist which are involved in recognition and presenting the ubiquitinated proteins to proteasome (Rao et al., 2002). Two such extra proteasomal recognition proteins i.e. Rad23 and Dsk2 were studied. Rad23 and Dsk2 can bind ubiquitinated substrate and also to proteasome and act as bridge between proteasome and its substrate (van Laar et al., 2002). Rad 23 and Dsk 2 utilize UBA (Ubiquitin pathway Associated) domain to interact with ubiquitin chain. Ubiquitin interaction site is situated in a hydrophobic patch of UBA domain (Mueller et al., 2002). This UBA domain has weaker affinity for free monoubiquitin moiety indicating specificity of these adaptor proteins for polyubiquitinated substrates. Elsasser et al has explained an important finding that Rad23 and Dsk2 interact with Rpn1 subunit of 19S particle (Elsasser et al., 2002). Another important finding explains that Rpn10, a proteasomal subunit of yeast exists in two forms i.e. as a proteasomal subunit and also as a free form (Wilkinson et al., 2000).
1.4.1.4 Role of unstructured region for initiation of degradation

A large number of proteins involved in transferring the substrate protein to the 26S proteasome were studied in detail. Along with ubiquitin, subunits of proteasome and proteasome associated proteins another requirement to initiate the actual degradation process is presence of an unstructured region. Only recognition of ubiquitinated protein by proteasomal assembly is not sufficient but second degradation signal in the form of an unstructured region is also required (Prakash et al., 2004). Unstructured region acts as initiation site for degradation. It was proposed that unstructured region engages the unfolding machinery of proteasome. As discussed above in ‘Role of adaptor proteins’ Rad 23 is an adaptor protein which can bind to ubiquitinated substrates as well as proteasome. It associates with proteasome but escapes degradation (Schauber et al., 1998). This is because it lacks an effective initiation region. So proteasome can not engage the protein and unfold it. Several internal unstructured regions are present in Rad 23 but these are too short to act as initiation regions for unfolding. The short initiation regions cannot solve the purpose of docking in proteasome. Lack of effective initiation region renders it stability against degradation. A fusion protein of Rad23 along with DHFR is also stable. When a longer unstructured tail was attached to the C terminus of Rad23 the fusion protein is degraded efficiently (Fishbain et al., 2011). This could be explained by availability of sufficient length initiation region for proteasomal docking.

Efficiency of proteasomal degradation also depends on the structure of polypeptide adjacent to unstructured initiation region. If initiation region is followed by α helical region, polypeptide is degraded efficiently and if initiation region is situated after β sheet structure, degradation is impaired (Lee et al., 2001; Godderz et al., 2011). Experiments by Inobe et al revealed the details of length of unstructured initiation region required for initiation of degradation. Minimum length required for efficient degradation of reporter protein DHFR was checked by tagging 4 ubiquitin moieties and single Ubl domain at N terminus of DHFR by in frame fusion. In case of Ub4 tagged DHFR, protein with initiation region having length of 34 or more amino acid residues was degraded while protein with initiation region having length of 29 or less amino acid residues remained stable.
residues was stabilized. In case of UbL tagged DHFR, protein with initiation region having length of 44 or more amino acid residues was degraded while protein with initiation region having length of 34 or less amino acid residues was stabilized (Innobe et al., 2011). These results indicate mutual contribution of tag and initiation region for efficient degradation of any protein.

Apart from significant contribution in the initiation of degradation process further details of unstructured region are described in detail in section 1.7 under intrinsically disordered proteins or regions.

Hsp 70 chaperon was also found to associates with proteasome via Bag1 protein for the delivery of aggregation prone substrates for degradation (Bimston, et al., 1998). Some other proteasome associated proteins having deubiquitinating activity are also discussed including Uch37, Ubp6, Hul5 (Finley, 2009).

1.4.2 Ubiquitin independent protein degradation

There are some proteins which are degraded by proteasome but do not require ubiquitin modification (Hoyt et al., 2004; Orlowski et al., 2003,). These include ornithine decarboxylase (ODC) (Murakami et al., 1992), thymidylate synthase, calmodulin, c-Jun, TCR α subunit and troponin C (Jariel-Encontre et al., 1995; Tarcsa et al., 2000; Benaroudj et al., 2001; Sheaff et al., 2000; Asher et al., 2002).

Present study is focused on ornithine decarboxylase (ODC). ODC is well known example of ubiquitin independent proteasomal substrate. Degradation signals which lie inside ODC itself along with another protein termed as antizyme are sufficient to carry out degradation of ODC without the help of ubiquitin. Thus in case of ODC degradation, targeting role of tag is performed by antizyme. Unstructured region plays an important role in targeting protein to proteasome.

Proteins are degraded by 26S as well as 20S proteasomes in ubiquitin independent manner. Protein degradation by proteasome is also observed in eubacteria and archaeabacteria (Zewick et al., 2000). Examples of proteasomal substrates include natural
Chapter 1: Introduction

unfolded proteins, some short-lived regulatory proteins, long-lived proteins, and oxygen-damaged, misfolded, mutated, or damaged proteins. Targeting of substrate towards 20S proteasome is brought by sequence within substrate itself or accessory molecule (Orlowski et al., 2003).

Janse et al proposed a different notion that once a substrate is localized to the proteasome additional signals are not required for degradation to occur (Janse et al., 2004). Proteasomal degradation of proteins without the help of ubiquitin is referred as one of the evolutionarily conserved mechanism (Hoyt et al., 2004).

1.4.2.1 An event of unfolding

During ATPase cycle the ring structures formed by AAA⁺ modules undergo significant conformational changes resulted in mutual rotation of these rings changes in size of their axial opening. So AAA⁺ proteins transform energy of ATP binding and hydrolysis into own mechanical movements. This develops tensions in associated protein molecules which results in unfolding of polypeptide chains (Prakash et al., 2004).

As mentioned in above paragraphs, specific as well as non specific protein degradation machinery of cell works for protein degradation.

1.5 The need for Targeted Protein Degradation

The purpose of such complicated degradation machinery is to maintain quality control in the cell. Studies on molecular mechanism of many diseases established protein degradation to have a critical role. Either mutation or accumulation of proteins in cell may lead to defective cellular processes. Failure of degradation of such proteins by UPS system results into diseases like the Liddle syndrome, certain kinds of cervical cancers, Parkinson’s, Alzheimer’s, Huntington’s, cystic fibrosis, and multiple myeloma (Lowe et al., 1988; Ciechanover et al., 2003). To combat such diseases where protein degradation is defective, attempts are being made to develop targeted or specific protein degradation. Manipulations at gene level such as gene knockout or antisense oligonucleotides and small interfering RNA (si RNA) were found to have various drawbacks.
This prompted attempts for manipulation of protein degradation pathway of proteasome. Probing the pathways of targeted protein degradation has been subject of much attention in biomedical, pharmaceutical and microbial disciplines.

1.5.1 Drawbacks of Ubiquitin-Proteasome system

Efforts to develop targeted protein degradation with ubiquitin-proteasome system have not met with much success. The process of ubiquitination requires the action of 3 different enzymes, namely E1, E2 and E3. E3s are highly substrate specific to carry out ubiquitination. The 3D structure of E3 ligase complexes with their substrate protein reveals a precise geometric relationship between substrate & ligase. Further, E3s are specific to certain E2s for catalytic addition of ubiquitin onto susceptible lysine residues in target protein. These structural constraints suggest that attempts to induce ubiquitination of target proteins may prove difficult because of incompatibility between substrate & ligase (Matsuzawa et al., 2005).

Direct method for targeted protein degradation without ubiquitination could overcome this limitation. A small number of proteins have been identified which act as substrates of proteasome though they do not require ubiquitin for their degradation (Marian et al, 2003, Coffino et al, 2004). Ornithine Decarboxylase (ODC) being one of them undergoes ubiquitin independent proteasomal degradation pathway (Murakami et al., 1992). ODC is an unusual proteasomal substrate as it is degraded by ubiquitin independent mechanism. Some degradation signals lie within the ODC, which along with another protein termed as antizyme, target it for degradation (Coffino, 2001). Antizyme recognizes a stretch of 24 amino acids known as Antizyme binding element or AzBE in the $\alpha/\beta$ barrel domain of monomer and destabilizes ODC. The last 37 amino acids in the C terminus of mammalian ODC help in docking the protein on proteasome. Cys 441 in the C terminus, if mutated, stabilizes ODC (Coffino et al., 2007). The yeast ODC has two intrinsic degradation determinant signals, namely the N-terminal region of 1-50 residues and Antizyme Binding Element (AzBE) present on the surface helices of the $\alpha/\beta$ barrel.
domain, from 164 to 187 residues. The stretch of last 37 amino acids of mammalian ODC which helps in docking the protein on proteasome is not present in yeast ODC. There are evidences of N terminal being important for degradation via Antizyme (Gandre et al., 2002). The N terminal region that is unique to yeast ODC acts similar to the C-terminal region of the mammalian enzyme in mediating its recognition by the proteolytic machinery. All these signals make ODC a favourite system for present study.

1.6 Ornithine Decarboxylase

Ornithine Decarboxylase (ODC) is the first enzyme in polyamine biosynthesis pathway (Tabor et al., 1984). The rate limiting step in polyamine biosynthesis is conversion of ornithine to putrescine which is catalyzed by ODC. Putrescine is then converted into spermidine and spermine. Spermine is most effective in supporting biological processes. ODC is approximately 450 amino acids long protein having molecular weight of 52kD and requires cofactor Pyridoxal 5'-Phosphate (PLP) for activity. Structure of mammalian (human and mouse) Ornithine Decarboxylase (ODC) is as shown below (Figure 1.4A and 1.4B)

![3D ribbon model of human Ornithine Decarboxylase (ODC) at 2.1 Å](image)

**Figure 1.4 Structure of mammalian ODC**

Figure A shows 3D ribbon model of human Ornithine Decarboxylase (ODC) at 2.1 Å (Almrud et al., 2000). Figure B shows features of secondary structure of mouse ODC (Kern et al., 1999).
Human ODC is active as homodimer having two fold symmetry. Homodimer is formed by head to tail interaction of two monomers. N terminal domain of one monomer interacts with C terminal domain of another monomer and vice versa. Two active sites are formed at the interface of two monomers in a dimer. Monomer of ODC consists of two domains i.e. N-terminal α/β barrel domain and C-terminal β sheet domain. The N terminus has 9 α helices and 8 β-strands. Barrel domain begins with helix H2 and ends with helix H10. β Sheet domain is composed of two sheets i.e. S1 and S2 and two helices i.e. H11 and H12. Perpendicular strands of sheet S1 and S2 forms a central hydrophobic core. Sheet S1 is composed of 6 strands which are named as B2, B18, B3, B12, B17 and B14. Sheet S2 is composed of four strands which are named as B12, B13, B15 and B16. Strand B12 connects sheet S1 to sheet S2. The loops connecting helices and strands are described as disordered loops. Protease sensitive loop is present over monomer-monomer interface towards the interior of the dimer. Lysine residue at 169th position is involved in salt bridge with aspartic acid residue at 364th position. Crystal structures of truncated mouse ODC and *Trypanosoma brucei* ODC were also studied (Kerl et al., 1999; Grishin et al., 1999). Primary sequence alignment and structural comparison of human ODC, truncated mouse ODC and *Trypanosoma brucei* ODC show extensive sequence and structural identity between these 3 eukaryotic ODCs. Sequence of active sites is also conserved between these 3 ODCs. 16 residues involved in active site of ODC are K69, D88, R154, K169, H197, G235-237, R274, R277, Y323, D332, Y389, C360, D361 and N398. The region from 117th to 140th amino acid residues is Antizyme binding element (AzBE) the region where antizyme binds and destabilizes ODC.
Figure 1.5 ODC homodimer and ODC-AZ heterodimer

ODC homodimer is formed by head to tail interaction of two monomers. ODC-AZ heterodimer is formed by interaction of antizyme with AzBE which is a basic patch situated in N terminal α/β barrel domain of ODC.

AzBE region is a basic patch which is buried in dimeric ODC, so antizyme does not bind homodimeric ODC. Major electropositive patch of AzBE is composed of K115, K121, and R144 on 5th and 6th helices. Residues 428-461 are described as disordered in crystal structure of human ODC. These last 37 amino acid residues which are unstructured in nature possess basal degradation element for the degradation of ODC. This C-terminal tail of mammalian ODC is buried in structural core of monomeric as well as dimeric states of ODC. This prevents early degradation of ODC. Binding of antizyme at AzBE region induces a conformational change in β sheet domain in monomer ODC. This results in exposure of C-terminal tail which initiates the degradation of ODC. Thus binding of antizyme makes ODC susceptible to proteolysis through conformational changes. As smaller part of surface area is buried in ODC-ODC homodimer formation, ODC dimer is easily dissociable. This is important for the regulation of ODC by antizyme.
Figure 1.6 Comparison of human and yeast ODC.

As discussed earlier, mammalian ODC has C terminal unstructured region. In yeast ODC first 50 amino acid residues form unstructured stretch. In mammalian ODC AzBE region is situated at 117 to 140 amino acid residues while in yeast ODC AzBE region is present from 164 to 187 amino acid residues (Porat et al., 2008) (Figure 1.6).

Figure 1.7 Comparison of secondary structural features of yeast and mouse ODC (Fonzi et al., 1987).
Chapter 1: Introduction

Yeast ODC amino acid sequence:

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSSTQVGNAL</td>
<td>SSSS TTVLDL</td>
<td>SNSVTVQKQ</td>
<td>YIFDGETLN</td>
<td>LLLELENNQG</td>
<td>LELLPHQAH</td>
</tr>
<tr>
<td></td>
<td>PKIFQALKAR</td>
<td>IGRINNRTCD</td>
<td>PGEEENFFIC</td>
<td>DJGEVEKLFN</td>
<td>NWVKEILPRK</td>
<td>FPYAVKCNFD</td>
</tr>
<tr>
<td></td>
<td>TKVLSLAEVL</td>
<td>GVNFDCASKV</td>
<td>EIDRVLSTMNI</td>
<td>SPDRIYYANP</td>
<td>CFSASFIRYVA</td>
<td>ASENVKMSFT</td>
</tr>
<tr>
<td></td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
<td></td>
</tr>
<tr>
<td></td>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>250</td>
<td>260</td>
<td>270</td>
<td>280</td>
<td>290</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>310</td>
<td>320</td>
<td>330</td>
<td>340</td>
<td>350</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>370</td>
<td>380</td>
<td>390</td>
<td>400</td>
<td>410</td>
<td>420</td>
</tr>
<tr>
<td></td>
<td>430</td>
<td>440</td>
<td>450</td>
<td>460</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YYMKHDVIVG</td>
<td>DWFYFPALGA</td>
<td>YTSSAATQFN</td>
<td>GFEQTADIVY</td>
<td>IDSELD</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.8 Amino acid sequence of yeast ODC.

1.6.1 Ubiquitin independent antizyme mediated proteasomal degradation of ODC

Figure 1.9 Antizyme mediated degradation of ODC (Pegg et al., 2006)
Degradation of ODC is a tightly controlled process because ODC catalyses first step of polyamine biosynthesis. Polyamines are multifunctional organic bases. They play an important role in cell growth, differentiation, apoptosis and malignant development (Pegg et al., 1986). Positively charged polyamines bind DNA and stabilize it. At physiological pH amino group of polyamine is protonated and pairs with negatively charged molecules. Spermine is most and putrescine is least effective in biological processes in which polyamines are involved because functions of polyamines depend on their electrostatic interaction. As polyamines are involved in crucial mechanisms of cell, the level of polyamines is maintained strictly. First step of polyamine biosynthesis catalyzed by ODC is regulatory. ODC has extremely short half life of several minutes to one hour. Subcellular localization of ODC is cell type specific and also depends on physiological conditions such as growth, differentiation, malignant transformation and apoptosis. ODC is mainly found in cytoplasmic staining while antizyme is found in nucleus in immune staining studies (Raymond et al., 2004).

In eukaryotic cells selective degradation of proteins occur via ubiquitin modification. But there are some examples of protein including ODC, which are substrates of proteasome but do not require ubiquitin modification. Being a crucial regulator of polyamine levels a different mechanism is developed by the cell for regulation of ODC.

Degradation of ODC is well studied and characterized. In vivo and in vitro studies have proven that there is no role of ubiquitin in the degradation of ODC (Bercovich et al., 1989, Glass et al., 1987).

Level of ODC is regulated by feedback mechanism (Figure 1.9). ODC catalyzes synthesis of putrescine from ornithine. Increased level of polyamines stimulates the synthesis of antizyme by inducing translational frameshifting. Antizyme binding to ODC is reversible. Polyamine responsive element is present at 5' upstream of first open reading frame of antizyme. When present above threshold level polyamines recognize polyamine responsive element and translation of antizyme is initiated. As discussed under session 1.6.2, there are two open reading frames for antizyme synthesis. First open reading frame
Chapter 1: Introduction

has stop codon. Second reading frame does not have any start codon to start translation. Translational frameshifting should occur for the synthesis of full length active antizyme protein. Before reading the stop codon of first open reading frame, ribosome shifts forward by one base pair and continues translation. This shifting escapes the stop codon and continues with second open reading frame till the end so as to translate entire antizyme protein. In this way increased polyamine level induces synthesis of antizyme. ODC shows more affinity towards antizyme than ODC-ODC homodimer. Thus in presence of antizyme ODC-ODC homodimer is dissociated and ODC-Az heterodimer is formed. AzBE region where antizyme binds with ODC is buried in dimeric state. After dissociation of dimer, AzBE region is exposed in monomeric ODC. Antizyme binds AzBE and induces conformational change in ODC. In monomeric ODC, C terminal tail is buried inside the structural core of the protein. Interaction of antizyme with ODC induces some conformational changes in ODC because of which the buried C terminal tail flips out. This unstructured region acts as docking site for proteasomal entry. After interaction with ODC, antizyme not only destabilizes but directs ODC towards proteasome for degradation. Recognition element situated in regulatory particle of 26S proteasome recognizes antizyme bound ODC. C terminal tail is first taken up by proteasome and then an entire ODC is dragged inside the core particle for degradation. Antizyme is recycled in this process like ubiquitin. Detailed analysis of this degradation of mammalian ODC revealed three important regions inside ODC which play an important role in degradation of ODC itself. These degradation signals within ODC are AzBE (Antizyme Binding Element), last 37 amino acid residues which are unstructured in nature and cysteine at 441\textsuperscript{st} position. These three degradation signals are known to be involved in efficient degradation of mammalian ODC. AzBE is the region where antizyme binds and directs ODC towards proteasome for degradation. C terminal tail of last 37 amino acid residues serves the requirement of unstructured region for initiation of actual degradation process. Role of cysteine at 441\textsuperscript{st} position is not clear exactly but mutation at this position made ODC stable. Cysteine residue when situated at another positions apart from 441\textsuperscript{st} position also made ODC stable (Hoyt et al., 2008).
A study by Matsuzawa et al. has shown that peptides and proteins attached to ODC undergo degradation (Matsuzawa et al., 2005). This portability of ODC as a degradation tag is contributed by degradation signals within ODC.

### 1.6.2 Antizyme

Antizyme is a mediator of ubiquitin independent proteasomal degradation of ODC (Coffino, 2001). C terminal half of antizyme molecule is involved in interaction with antizyme binding element i.e. AzBE situated in N terminal α/β barrel domain of ODC. But only interaction of C terminal part of antizyme with ODC is not sufficient to direct the proteasomal entry. N terminal half of antizyme is involved in directing the ODC-Az complex towards proteasome. Residues 106 to 212 are involved in interaction with ODC while residues 55 to 105 are involved in directing ODC-Az complex towards proteasome. First 70 amino acid residues are not involved in degradation process but are important for transport of Az to various locations (Pegg, 2006).

![Figure 1.10 Structure of antizyme (Hoffman et al., 2005).](image)

The expression of antizyme mRNA is interesting. There are two open reading frames for antizyme synthesis. First open reading frame has stop codon while the second reading frame does not have any start codon. Event of frame shifting is needed as second ORF
does not have any start codon. Without frame shifting only part of antizyme protein would be synthesized which is coded by first ORF. The efficiency of +1 translational frame shifting is required for synthesis of entire antizyme protein (Figure 1.11). An event of Translational frame shifting is stimulated by increased polyamine levels. Increased ODC activity leads to increased polyamine levels which in turn stimulates antizyme synthesis. Increased antizyme causes ODC degradation. Thus, ODC is regulated by feedback mechanism.

Mechanism of Translational frame shifting- Frame shifting occurs at stop codon of first open reading frame of antizyme.

![ Diagram showing the mechanism of translational frame shifting ]

**Figure 1.11 Frameshifting needed for antizyme synthesis.**

In vertebrates generally the last codon is UCC. This is followed by UGA which is a stop codon. After decoding UCC which is a codon for serine, the ribosome shifts frame by +1 and continues translation. So instead of reading UCCUGA, ribosome reads as UCCGA and next codon’s C. The frame shifting site in *Saccharomyces cerevisiae* antizyme mRNA is GC-GCG-UGA-C (Palanimurugan et al., 2004). Highly conserved nucleotides are involved in stimulation of frame shifting. These nucleotides are CA pair which occurs...
immediately at three prime of UGA stop codon. Number of study has shown that nucleotide immediately 3’ of stop codon is important because it determines the efficiency of translation termination. The CA pair following the stop codon decreases termination efficiency and stimulates +1 frame shifting. In *S. cerevisiae* decoding is carried out by rare arginine as acceptor tRNA and involves a translational pause and facilitates frame shifting (Pande et al., 1995). *S. cerevisiae* does not have a tRNA whose anticodon can form full Watson-Crick or wobble, pairing with GCG. Instead of general pairing anticodon of alanine tRNA uses unusual purine-purine opposition. This poor pairing of codon and anticodon favours frame shifting.

Till now four types of antizyme are described and named as Az1, Az2, Az3 and Az4. Az1 is potent regulator of ODC and has wide tissue distribution. Az1 has molecular weight of 26kD. Mouse Az1 is 227 amino acid residues long while human ODC has 228 amino acid residues (Tewari et al., 1994). Az2 has also equal tissue distribution but is less abundant. Az3 is found only in male germ cells in post meiotic stage of their differentiation to mature sperm. Less is known about Az4. Az1 and Az2 differ in their capacity to direct proteasomal degradation. It was studied that sixteen fold lower concentration of Az1 catalyzed significant degradation as compared to Az2.

Targeted protein degradation using antizyme has shown that proteins attached to antizyme undergo degradation by proteasome (Li et al., 1996).

The crystal structure of mammalian ODC is available. However yeast ODC protein structure has not come in focus by far. Present study tries to reveal the structural aspects as well as unstructured region of degradation signals of yeast ODC.

### 1.7 Intrinsically Disordered Proteins or Regions (IDPs or IDR)

Structure of myoglobin and lysozyme were revealed by using X-ray crystallography before 50 years (Kendrew et al., 1958; Blake et al., 1965). Till then structures of large number of protein were elucidated. Different techniques such as NMR spectroscopy, crystallography, CD (Circular Dichroism) reveal the 3D structure of any protein.
Structure function relationships of proteins are a major research area in biophysical, biomedical as well as pharmacological studies. Though the 3D structure of any protein is important, often some regions remain unfolded or unstructured (Gast et al., 1995; Uversky et al., 1999; Boublik et al., 1970). Earlier it was believed that polypeptide should at least gain some 3D structure to function. But it is not the case that entire polypeptide chain always gains some secondary or tertiary structure before being functional. Many a time some region of polypeptide gains some conformation while few regions are without any conformation. These unfolded regions are called as disordered regions or simply unstructured stretches. These unstructured regions are equally important as structured regions in any protein as far as function is concerned (Huber et al., 1983). Rather the conformational flexibility confers functional ease significantly rather than a rigid structure (Harauz et al., 2009). Most of the proteins have disordered region, in fact half of the eukaryotic proteins have long disordered regions especially the signaling proteins. There are few examples of proteins which are fully disordered yet functionally important such as Myelin Basic Protein (MBP) (Harauz et al., 2004), prothymosine α (Gast K et al., 1995) and microtubule formation promoting Tau protein (Schweers et al., 1994). Unstructured regions are needed in variety of functions such as sequence recognition, molecular interaction (Pontius et al., 1993) and protein degradation (Prakash et al., 2004; Zhang et al., 2004). IDPs are often involved in regulatory/signaling interactions with multiple partners that require high specificity and low affinity (Uversky et al., 2005). Uversky listed some crucial roles of IDPs which include regulation of cell division, transcription and translation, signal transduction, protein phosphorylation, storage of small molecules, chaperone action, and regulation of the self-assembly of large multiprotein complexes such as the ribosome. The functions of IDPs are grouped into four broad classes: 1) molecular recognition 2) molecular assembly 3) protein modification, and 4) entropic chain activities (Uversky, 2010).

Crystallographic study revealed the unstructured nature of C terminal 37 amino acid residues of mammalian ODC (Kern et al., 1999, Almrud et al., 2000). Studies also revealed the crucial role of this unstructured region in the degradation of ODC itself. C
terminal unstructured tail of mammalian ODC acts as a portable degradation signal and
directs degradation of protein to which it is tagged. Reporter proteins such as GFP,
luciferase, DHFR and Ura3 were found to be destabilized when tagged C terminally
with tail of mODC (Li et al, 1998; Zhang et al., 2003; Leclerc et al., 2000; Hoyt et al,
2006).
N terminal stretch of yeast ODC was also found to have a potential to act as degron
when tagged to reporter protein Ura3 (Godderz et al., 2011).
After increasing evidences on significance of unstructured regions or proteins a new
branch named unfoldomics has emerged which covers study on unfoldome (Dunker et
al., 2008).

Unstructured regions do not fold into any specific conformation because of their amino
acid sequence. Hydrophobic amino acid residues contribute for folding of polypeptide
chain by pushing themselves into inner core of structure. Unstructured regions are rich
in polar and charged residues and depleted of hydrophobic residues. Low hydrophobic
force is responsible for less driving force for compact structure formation (Uversky,
2010). Role of unstructured regions or IDPs in neurodegeneration and protein
dysfunction diseases has focused IDPs as potential drug targets.

1.8 Importance of structural characterization

Structural information of any protein has been matter of utmost significance in the field
of structural biology. It is prerequisite in discipline of pharmaceutical research for
establishing the mechanism of action of a drug. It gives a visual idea of physical,
chemical characteristics of protein by providing information on its secondary and tertiary
structure and thus providing base to predict function of a protein. Some proteins are
targets for chemotherapy.

Study of domain organization has been one of the important aspect of structural study.
Domain is a part of polypeptide chain that can fold and function independently. Domain
has a stable tertiary structure with its own hydrophobic core. A single protein may
contain one or several domains. Each domain performs one or many distinct biological
functions. This confers tremendous functional capacity to protein moieties. Domains interact with other molecules in the cell and they are also involved in signal transduction pathways. Formation of new domains and new domain combinations indicates an evolutionary mechanism of protein folding. Similarity in domain architecture of two different proteins indicates possibility of common ancestor. So study on domain helps us to understand the protein interaction, protein function and evolution too.

Sometimes unstructured region of a protein also can be a domain and has an important role in protein assembly and function (Prakash et al., 2004). These stretches are involved in initiation of degradation and recognition of other molecules. Every structure in an entire assembled protein is important with respect to its assigned function. Keeping this in mind, the idea was to focus on the structural aspects of degradation determinant signals of yeast ODC.

When protein is in a native state its structure is maintained. But during denaturation process unfolding of the molecule increases. This increases the protein solvent interface and ultimately stabilizes the structure. Large nonpolar regions and peptide bonds become exposed to the solvent and nature of interaction between protein and co-solvent changes. In the present study CD and Fluorescence spectroscopy was used for above mentioned structural characterization.

1.8.1 Structural characterization by Circular Dichroism (CD)

Circular Dichroism spectroscopy is used to gain information about the secondary and tertiary structure of proteins and polypeptides in solution. The chromophores for the protein study include the peptide bond (absorption below 240 nm), aromatic amino acid side chains (absorption in the range 260 to 320 nm) and disulphide bonds (weak broad absorption bands centred around 260 nm). Different types of secondary structure (i.e. a helix, β sheet, turn) found in proteins give rise to characteristic CD spectra in the far UV region (figure 1.12).
1.8.2 Structural characterization by Fluorescence spectroscopy

Fluorescence has an important role in the structural determination of proteins. Fluorescence is the emission of radiation that occurs when a molecule in an excited electronic state returns to the ground state. For fluorescence study intrinsic (e.g. tryptophan residue) as well as extrinsic (e.g. ANS) fluorophores are used.

Protein Purification by Affinity Chromatography

Figure 1.13 Protein purification by affinity chromatography

Studies on the degradation determinant signals of Ornithine Decarboxylase (ODC)
1.9 Importance of functional characterization

Present study tries to provide new options to carry out targeted protein degradation with the help of peptides derived from yeast and mouse ODC. As the selected peptides in the present study are degradation signals of yeast ODC, it would be interesting to check their portability. Na/β and N50 peptides derived from yeast ODC and 37mODC derived from mouse ODC are chosen for functional characterization. Na/β peptide is an independent domain of yeast ODC which carries 2 degradation signals namely N terminal unstructured region and AzBE region situated in α/β barrel domain. Portability of Na/β peptide was checked by tagging it to reporter protein and then checking the stability of reporter protein. As N50 is an unstructured stretch, it would be interesting to check the potential of this peptide to act as initiator of degradation process. As discussed in chapter 5, ubiquitin is a natural tag for proteins to be degraded but regulation of ubiquitin itself is also equally important. Natural ways for degradation of ubiquitin are described in literature, but attaching a tag of another unstructured region derived from mouse ODC i.e. 37 mODC will provide a new means of targeted ubiquitin degradation (Figure 1.14), bringing together two degradation signals effectively.
The proposed peptides are:

1. Entire N terminal domain of yeast ODC (Na/β peptide)
2. α/β barrel domain of yeast ODC (α/β peptide)
3. N terminal stretch of 50 amino acid residues (N50 peptide)
4. Ubiquitin with C-terminally tagged mODC peptide (UbmODC fusion protein)

**Figure 1.14  Schematic representation of proposed objectives**

Proposed peptides vary in structure when present as a part of ODC. Na/β peptide and α/β peptide are domains while N50 and 37mODC are unstructured in nature. Na/β peptide carries two degradation signals and composed of α/β barrel domain and N terminal unstructured stretch. It would be interesting to know the structure gained by these peptides when expressed independently from the rest of the protein. It would also be important in functional evaluation of Na/β, N 50 and 37mODC peptides.
Chapter 1: Introduction

Pictorial representation of proposed fusion proteins

Figure 1.15 Pictorial representation of proposed fusion proteins

Chapter 1 gives general introduction about proteins, protein degradation in prokaryotes and eukaryotes, ubiquitin proteasome system, Ornithine Decarboxylase (ODC), importance of structural and functional characterization of selected peptides and rationale behind this study.

Chapter 2 is focused on structural and functional characterization of Na/β peptide. This chapter describes in detail about the cloning of fragments for structural as well as functional study. Peptide expression and purification is also discussed herewith (Figure 1.15). Structural characterization of the Na/β peptide was done using CD and fluorescence spectra. Under functional study, the Na/β peptide was tagged to a reporter protein and stability of reporter protein was checked. Structural features and functional role of Na/β peptide when it is expressed independently from rest of the protein is discussed.

Chapter 3 covers structural and functional characterization of N50 peptide. Naturally N terminal 50 amino acids stretch of yeast ODC is unstructured region. Any gain of structure by this small unstructured region was tried to resolve by using CD and fluorescence spectra when expressed independently. Under functional study, the N50 peptide was tagged to a reporter protein and stability of reporter protein was checked.
Chapter 1: Introduction

Structural features and functional role of N50 peptide when it is expressed independently from rest of the protein is discussed.

Chapter 4 covers structural characterization of α/β peptide. Standardization of purification conditions and solubility of peptide was discussed here.

Chapter 5 presents the fusion protein UbmODC. Previous reports on fate of ubiquitin and ubiquitin fusion proteins are discussed in the introduction of this chapter. During proteasomal degradation of substrate protein, ubiquitin is recycled. Potential of C37mODC peptide as a portable degradation tag to drag the substrate protein towards proteasome for degradation was already proved. In this chapter C terminal unstructured tail of mouse ODC was tagged at C terminus of ubiquitin and fate of ubiquitin was studied to understand the potential of C terminus as a degron. The construct also brings two different degrons together for increasing the rate of degradation in targeted protein degradation.