The Ubiquitin-Dependent Protein Degradation Pathway

Literature: Chapter 24, Pollard and Earnshaw

Pickart, C.M. (2001) Mechanisms underlying ubiquitination. *Annu. Rev. Biochem*, **70**, 503-533.
Vale R.D. (2000) AAA proteins: Lord of the Ring. *JCB*, **150**, F13-F19.
Baumeister W. et al. (1998) The proteasome: paradigm of a self-compartementalzing protease. *Cell*, **92**, 367-380.
Weissman, A. (2001) Themes and variations on ubiquitylation. Nature Reviews Molecular Cell Biology 2, 169-178.
Murata S, Yashiroda H, Tanaka K. (2009) Molecular mechanisms of proteasome assembly. Nat Rev Mol Cell Biol. 10(2):104-15.
Ravid T, Hochstrasser M. (2008) Diversity of degradation signals in the ubiquitin-proteasome system. Nat Rev Mol Cell Biol. 9(9):679-90.
Kunjappu, MJ, Hochstrasser M (2013) Assembly of the 20S proteasome. Biochimca Biophysica Acta

1. Introduction

Protein degradation is one additional mechanism how cells regulate gene expression on a posttranscriptional level. In this lecture, I will give an overview about the general principles of ubiquitin-mediated protein degradation. In the lectures of Yves Barral, you will get to know examples for how regulated protein degradation contributes to cell cycle progression

Protein degradation not only serves regulatory functions but it is as essential to the cell as protein synthesis. For example,
➔ to supply amino acids for fresh protein synthesis
➔ to remove excess enzymes
➔ to remove misfolded proteins etc.

Proteases acting in the cytoplasm or the nucleoplasm must be very tightly controlled to ensure that they degrade their targets at the right time and at the right place or that they do not degrade the wrong proteins. Several approaches have been developed during evolution to ensure that intra-cellular protein degradation occurs in a well-controlled manner.

Proteases with very high substrate specificity can ensure that only the right target is hit. Some proteases need to be activated by complex regulatory pathways (e.g. caspases).
Constitutively active proteases are separated from the cytoplasm either through their compartmentalization into an organelle, or through their structural organization.
There are two major intracellular devices following this principle. They represent the main protein degradation facilities in the cell.

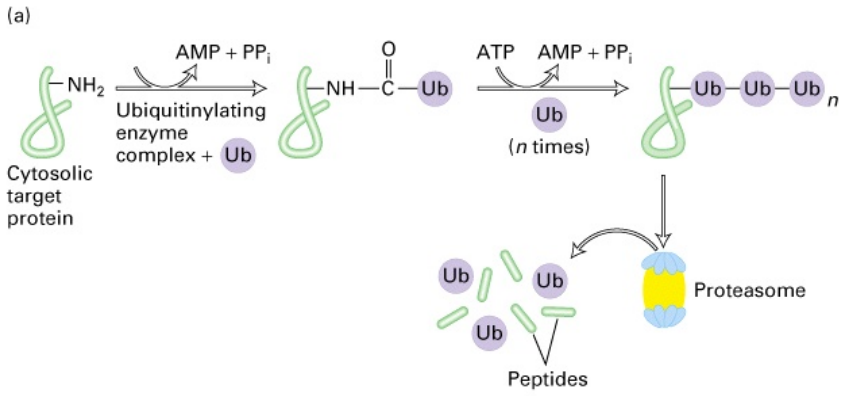
1. **Lysosomes/vacuoles** contain about 50 different kinds of hydrolytic enzymes including proteases, nucleases and polysaccharidases. What concerns protein degradation, they deal with extracellular proteins, e.g., plasma proteins, that are taken into the cell, e.g., by endocytosis or cell-surface membrane proteins that are used in receptor-mediated endocytosis.

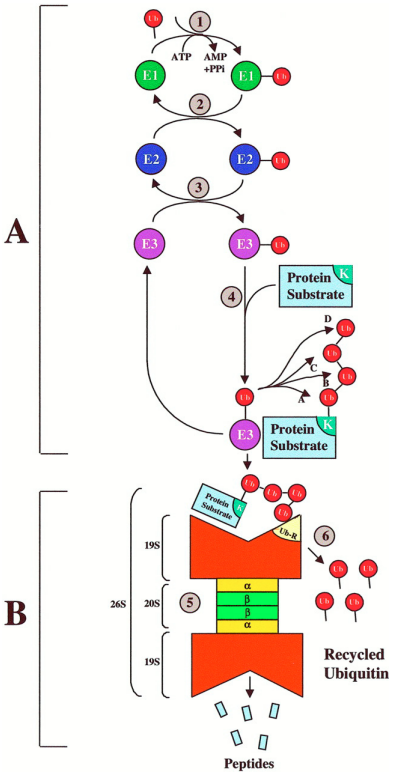
The also contribute to the turnover of organelles and cytosolic components in the process of autophagy.

2. **Proteasomes** deal primarily with endogenous proteins; that is, proteins that were synthesized within the cell such as:
   ➔ transcription factors
   ➔ cyclins (which must be destroyed to prepare for the next step in the cell cycle)
   ➔ proteins encoded by viruses and other intracellular parasites
   ➔ proteins that are folded incorrectly because of
      - translation errors
      - they are encoded by faulty genes (Cystic fibrosis is caused by the accelerated degradation of a mutant version of a chloride transporter)
      - they have been damaged by other molecules in the cytosol.

2. Protein ubiquitination

Proteins are targeted to the proteasome for degradation by multi-ubiquitin chains attached to the target protein. The polyubiquitin chain is attached to e-NH_2 group of a Lysine residue in the degradation substrate.





A - Ubiquitination consists of a cascade of reactions that ultimately lead to the conjugation of ubiquitin onto the substrate. Repeated ubiquitination leads to the attachment of a poly-ubiquitin chain on the substrate.

B- Poly-ubiquitinated proteins are recognized by the proteasome, which recycles the ubiquitin moieties and degrades the substrate into small peptides.

2.1. Ubiquitin

Ubiquitin is a small protein (76 amino acids) with a compact globular structure. The C-terminus extends away from the main body of the protein into the aqueous space. It is the alpha-carboxyl group of the terminal C-terminal glycine, which forms isopeptide bonds with target proteins (which can include another copy of ubiquitin). The isopeptide bond is formed with an (epsilon) amino group in the side chain of a lysine residue of the target protein.
Ubiquitin is generally expressed as a poly-protein, consisting of either several ubiquitin repeats or of a fusion protein between ubiquitin and ribosomal proteins. Specific proteases, called ubiquitin C-terminal hydrolases (or Ubp, for ubiquitin protease), process ubiquitin fusion proteins by cleavage. Ubiquitin is conserved throughout most kingdoms of life – it is virtually identical in sequence in archaebacteria, yeast, or mammals. It is used by all these creatures to target proteins for destruction. However, note that ubiquitin can also serve as a signal for endocytosis of plasma membrane proteins, which are then targeted to the vacuole for degradation.

2.2. The activation reaction

Ubiquitin is activated by a "**ubiquitin activating enzyme**" or "**E1**". The E1 enzyme hydrolyses ATP and forms a complex with the resulting ubiquitin adenylate, reminiscent of amino acyl adenylate formation in protein synthesis. This is followed by transfer of ubiquitin to the active site cysteine of E1 to form a thiol ester between the C-terminus of ubiquitin and the thiol group of the E1. This occurs in concert with adenylation of an additional ubiquitin. Each fully loaded E1 molecule carries two molecules of ubiquitin – one as a thiol ester, the other as an adenylate.

The resulting thiolester is a high-energy intermediate between ubiquitin and an internal Cys residue of the E1.

A E1 + ATP + Ub ----------> E1.Ub-AMP + PPi

B E1.Ub-AMP + Ub + ATP----------> E1-s-co-Ub.AMP-Ub + PPi

2.3. The transfer reaction

The second step starts with the transfer of activated ubiquitin onto an E2 enzyme. This enzyme is responsible for the conjugation of ubiquitin onto the substrate, which in most cases, appear to be the proximal donors of ubiquitin to target proteins. However in many cases, the ubiquitin is further transferred onto an E3 enzyme, prior to conjugation to the substrate.
E2 enzymes, also called UBCs (Ubiquitin Conjugating Enzymes), are highly conserved in their catalytic domain. E2 enzymes, like the E1, also have an active site cysteine, which is absolutely required for the function of the enzyme. Ubiquitin is transferred from E1 to the E2 cysteine to form again a thiol ester. This transfer reaction does not require ATP.

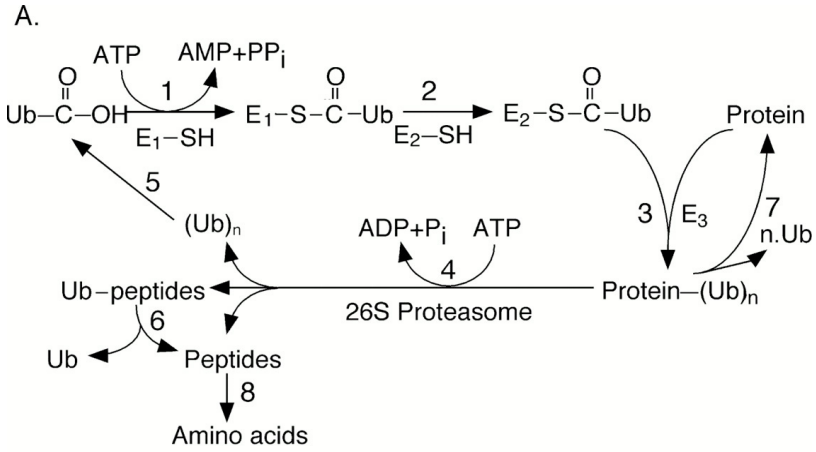
C E1-s-co-Ub.AMP-Ub + E2-SH -----> E2-s-co-Ub + E1.AMP-Ub

Ubiquitin is then transferred to the acceptor lysine of the target protein to form the isopeptide bond.

D E2-s-co-Ub + Protein-NH2 -------> E2-SH + Protein-NH-CO-Ub

While there is generally only one E1 in a given cell type, the number of different E2 expressed is higher. Yeast cells for example express 13 different E2 enzymes, Ubc1 to Ubc13. 11 of them are used for ubiquitin conjugation reactions. The two additional ones, Ubc9 and Ubc12, are involved in conjugation of ubiquitin-related proteins. The sequence differences between Ubcs are thought to play some role in substrate specificity or the intra-cellular distribution of the enzyme. For example, Ubc6 contains a C-terminal trans-membrane domain that anchors the enzyme to the surface of the ER. Ubc3 is the only essential ubiquitin-E2 enzyme (Ubc9 is essential too, but not an ubiquitin-E2), and is also known under the name of Cdc34.

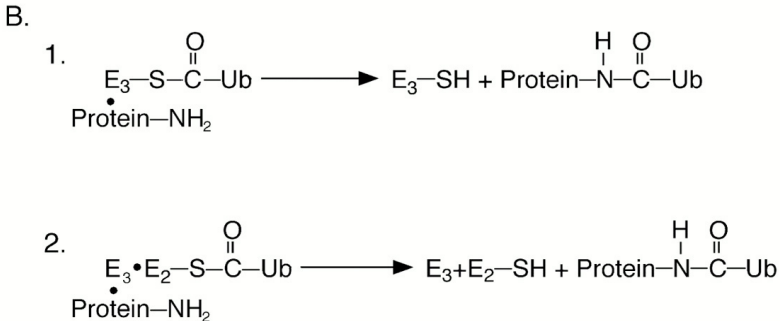
Multi-ubiquitin chains can be built up on a single lysine of the target protein, by isopeptide bond formation between the carboxyl group of gly76 of one ubiquitin and the amino group of the side chain of lys48 of the preceding ubiquitin. Multi-ubiquitin chains may also be pre-assembled and transferred to target proteins intact.



2.4. *Ubiquitin ligases - or E3s*

In vitro, E2s appear not to require other protein factors in vitro to transfer ubiquitin to a suitable target. In vivo, however, additional protein factors, the **E3 ubiquitin ligases** are involved in the ubiquitin transfer reaction. E3s select and bind the target protein, and bring it into the vicinity of the E2s. This implies E3s recognize a motif in the substrate protein that targets it for ubiquitination.

To date, two major classes of E3s have been identified:
1. E3s which do form thiol esters with ubiquitin: ubiquitin is transferred from a charged E2 to a cysteine in the E3, and from there to the target.
2. E3s which do not form thiol esters with ubiquitin but rather target the E2/ubiquitin complex to the substrate. The recognition of some targets for ubiquitinylation relies upon the identity of the N-terminal residue of the target protein.



2.5. Substrate recognition

The substrates of the ubiquitin pathway can be sorted into two main categories: constitutive substrates and occasional substrates.

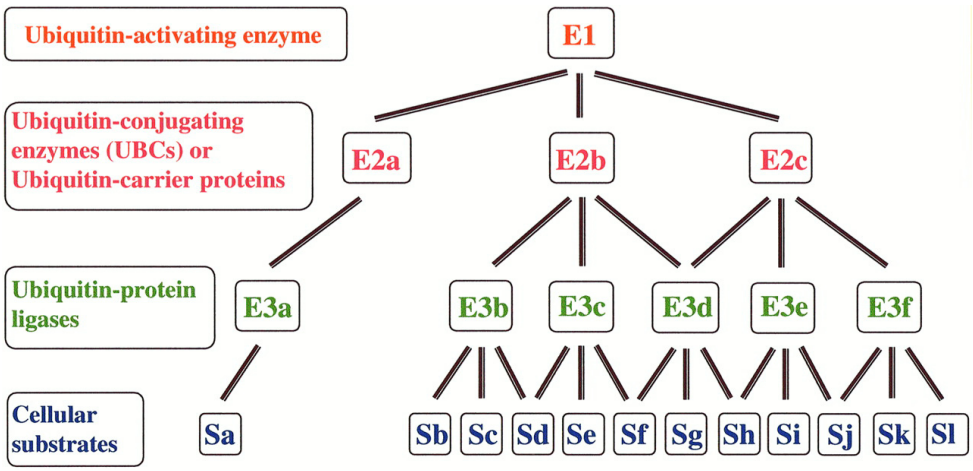
The constitutive substrates are molecules that contain a degradation signal and necessitate for proper function to be short-lived proteins. For these molecules, there must exist specialized recognition E3(s) that recognizes them specifically. This function can be regulated, such that proteins are only ubiquitinated if they are properly modified (e.g. phosphorylation), or only with a specified localization or timing.

Another function of protein ubiquitination is the elimination of misfolded and potentially toxic proteins. In this case, substrate recognition does not involve the specific recognition of specific substrates but the non-specific targeting of all sorts of misfolded proteins. It is unclear, as of now, which are the exact signals recognized by the ubiquitin pathway in this case.

Structural motifs that target proteins for ubiquitination

1. the N-terminal residue (N-end rule)
2. phosphorylation
3. PEST elements (regions rich in P,S,T,E) (often linked to phosphorylation)
4. destruction boxes (e.g. cyclins) **R** (A/T) (A) **L** (G) x (I/V) (G/T) (N)

The structure of the ubiquitin pathway is well suited to ensure a diversity of function and reach a variety of substrates. Indeed, along the pathway, diversity is added at each step. From one E1, ubiquitin is transferred onto at least 10 to 11 E2s, which make use of an even larger variety of E3s. Furthermore, the combinatorial association of different E2s with various E3s is likely to largely amplify this diversity.



:
The ubiquitin pathway and its diversification as ubiquitin progresses through the different steps of the pathway

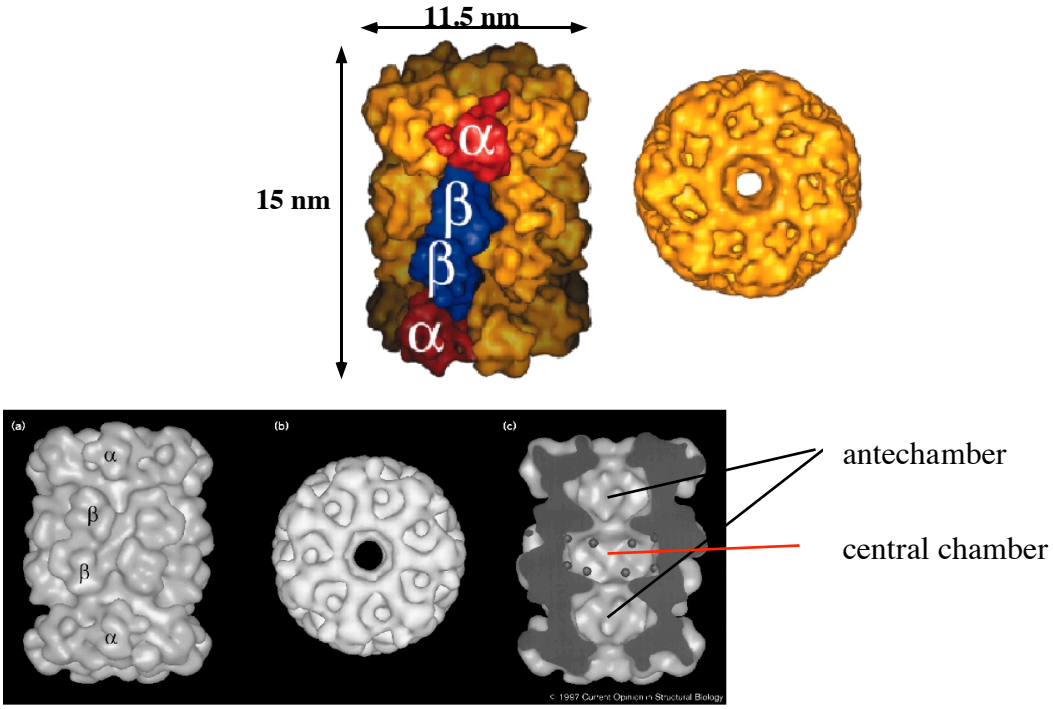
3. The Proteasome

Multi-ubiquitinated proteins are being degraded by the proteasome. The proteasome is a huge macromolecular complex. It exists in pro- and eukaryotes, in the latter it can constitute up to 1% of the total cellular protein. The eukaryotic proteasome can be found both in the cytoplasm and in the cell nucleus.

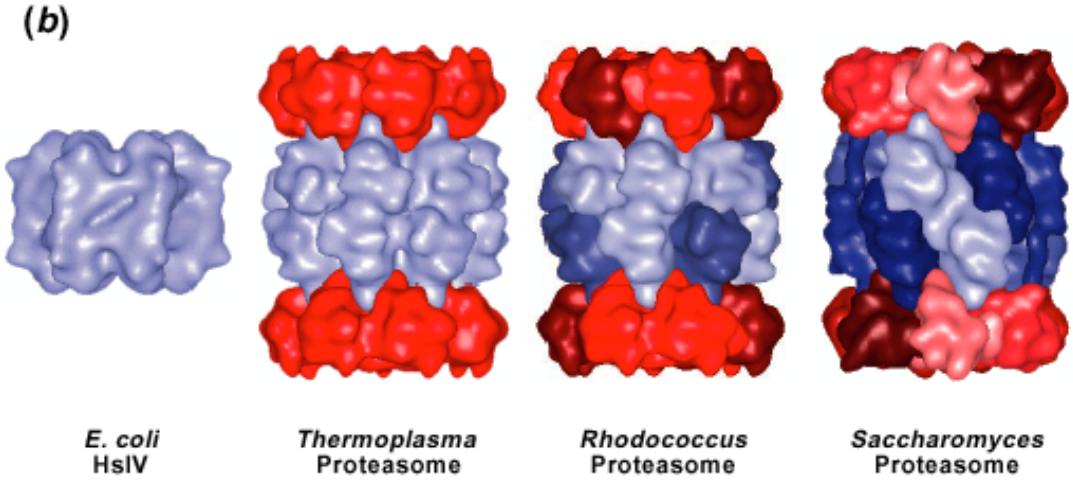
3.1. Structure of the proteasome

20S proteasome

The 20S proteasome (700kDa) is the proteolytic core of the proteasome, which contains the catalytic subunits. It has the structure of a cylinder made of four superimposed rings. Each ring has a seven-fold symmetry and is composed of seven subunits. The two external rings are formed of so-called α-subunits, while the two internal rings are made of β-subunits. This alpha-beta-beta-alpha assembly forms a central channel with three chambers: two antechambers located on opposite sides of a central chamber.



The least complex proteasome of 20S is found in archaebacteria of the genus Thermoplasma and consists of four stacked rings, each a homoheptamer (two outer rings consist of subunit alpha, two inner rings of beta). In the eukaryotic 20 S proteasome, each ring consists of seven different, but closely related subunits. Despite the difference in subunit composition, the proteasome from archaebacteria to eukaryotes have the same basic architecture, which under the electron microscope appears as a cylinder shaped particle, made up of four stacked rings with dimensions of approximately 15 nm in height and 11 nm in diameter.



3.2. Proteolytic mechanism – The Proteasome: a Threonine protease

Only the β-subunits are proteolytically active and the internal space of the ring is dedicated to peptidase activity. The catalytic site of the protease is formed by the amino-termini of the b-subunits, which are pointing toward the lumen of the cylinder. Thereby, any peptide introduced in the lumen is rapidly degraded by the added activity of the reactive groups with very low sequence specificity. Protease activity results in the liberation of very short peptides and amino acids, which presumably escape through the interstices of the cylinder wall. While degradation by the proteasome requires ATP in vivo, the 20S proteasome does not require energy to function and does neither bind nor hydrolyze ATP.
Eukaryotic proteasomes are more complex than their prokaryotic counterparts. In higher eukaryotes there are 7 different β-subunits in one ring but only 3 of them are catalytically active. The proteasome is a fairly unspecific protease, however, some preferences can be assigned to the individual β type subunits.

β1 - cleavage behind acidic aa like **Glu** (Caspase-like)
β2 - cleavage behind basic aa like **Lys, Arg** (Trypsin-like)
β5 - cleavage behind bulky hydrophobic aa like **Phe, Tyr, Trp** (Chymotrypsin-like)

Binding studies with the inhibitory peptide aldehyde acetyl-Leu-Leu-norleucinal (Calpain inhibitor I) reveal 14 catalytic sites within the central chamber and suggest a novel proteolytic mechanism in which the hydroxyl group of a **threonine acts as the nucleophilic group in the peptide hydrolysis**. The role of this threonine located at the N-terminus of the beta-subunit in the catalytic process is further supported by its covalent modification by lactacystin, a natural inhibitor of the proteasome. The specificity of the proteasome seems to be rather unspecific but the size of the hydrolysis products is mostly between 6 and 9 residues. This corresponds to the length between adjacent catalytic sites in the central chamber, which probably means that the substrate must be channeled into a single 20S molecule during the hydrolysis process. This generation of peptides of defined length is of biological relevance in the context of the implication of the eukaryotic proteasome in the antigen presentation by MHC molecules during the T-cell immune response.

3.3. Assembly of the 20S proteasome

The assembly of the proteasome occurs in a well-defined manner to exclude that active β subunits would ever be exposed to the cytoplasm. The β subunits of the proteasome are synthesized as inactive precursors with an N-terminal extension that needs to be cleaved off to expose the N-terminal proteolytically active threonine. In the proteasome assembly, this cleavage occurs during a late step in assembly, thus deferring activation until the β subunits are confined to the inner cavity of the proteasome. This process takes place as two double rings close into a single 20S cylinder. The activation of the β subunits does not require any additional factor suggesting that they ensure their own cleavage.

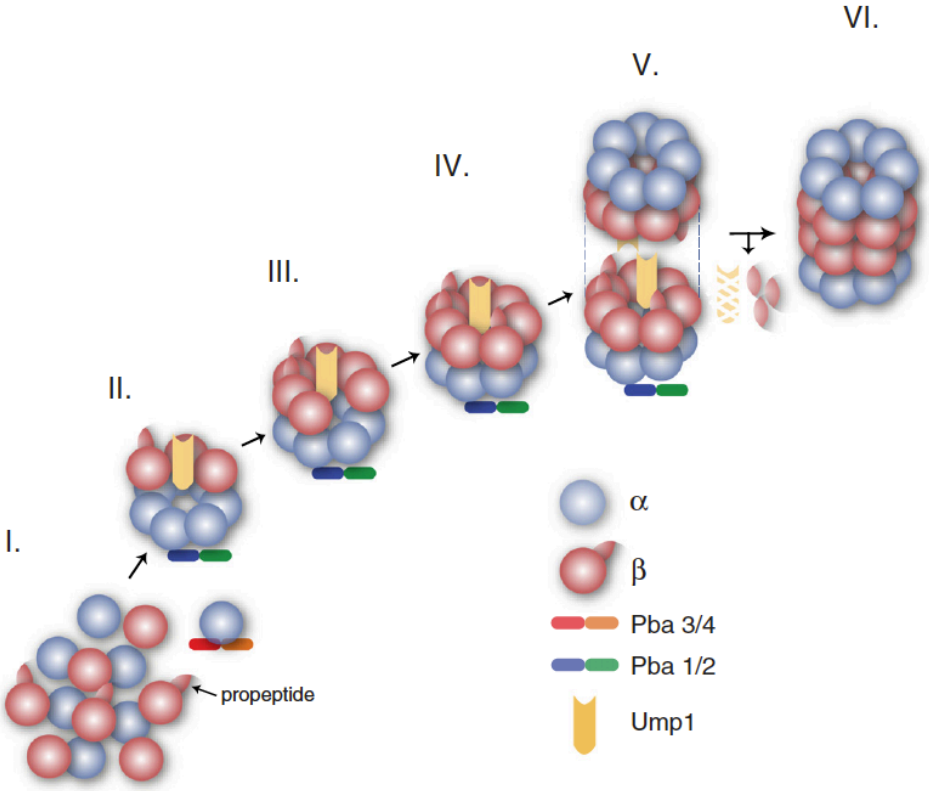


Fig. 2. Model for the assembly of the eukaryotic CP. This cartoon depicts assembly precursors that were isolated, or shown to form, in yeast. I. α (blue) and β subunits with propeptides (red) are synthesized as free polypeptides. An α ring is formed initially with the aid of the Pba3/4 (red and orange) chaperone that interacts specifically with the α5 subunit. II. The isolated 15S intermediate contains a full α ring, the β2, β3 and β4 subunits and two assembly chaperones: Pba1/2 (blue and green) and Ump1 (yellow). III. The –β7 intermediate is composed of a full complement of subunits except the β7 subunit and Pba1/2 and Ump1. IV. The half-proteasome has a full α ring and full β ring, but is still associated with the assembly chaperones Pba1/2 and Ump1. V. The dimerization of the half-proteasomes forms the preholoproteasome that is still immature. VI. CP maturation is achieved by the autocatalytic processing of the β subunit propeptides and the degradation of the Ump1 chaperone. The Pba1/2 chaperone is also released upon maturation. This process yields a functional CP competent for protein degradation.

3.4. The Proteasome and its regulatory complexes

The access to the central cavity of the proteasome is restricted to unfolded proteins. Isolated proteasomes can degrade unfolded proteins in an ATP-independent manner. However, as unfolding is required to funnel a protein into the antechambers, a mechanism must exist to ensure delivery of the degradation substrate into the central cavity. This requires recognition and unfolding of the degradation substrates. The proteasome cooperates with additional factors, which serve these functions.

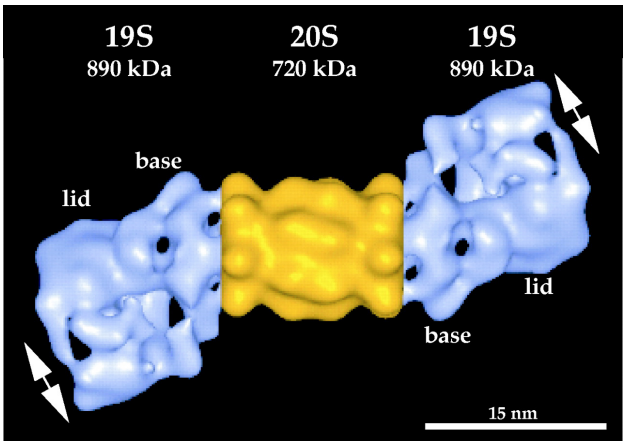
3.4.1. AAA-family (tripleA) and substrate unfolding

AAA-family (**A**TPases **a**ssociated with various cellular **a**ctivities) members take part in many different biological processes throughout the kingdoms. They all possess a conserved ATPase domain of about 250 aa. Very importantly, many of them have been shown to serve the unfolding of proteins or the assembly/ disassembly of protein complexes. They form characteristic ring like oligomers (often hexameric). Proteasomes from all organisms are assisted by AAA family members. The ATPase component is likely to function as a reverse chaperone to unfold and disaggregate protein substrates and to facilitate their entry into the chamber, which harbors the proteolytic active site. Unfolding of the degradation substrates renders proteasomal degradation into an ATP-dependent process.

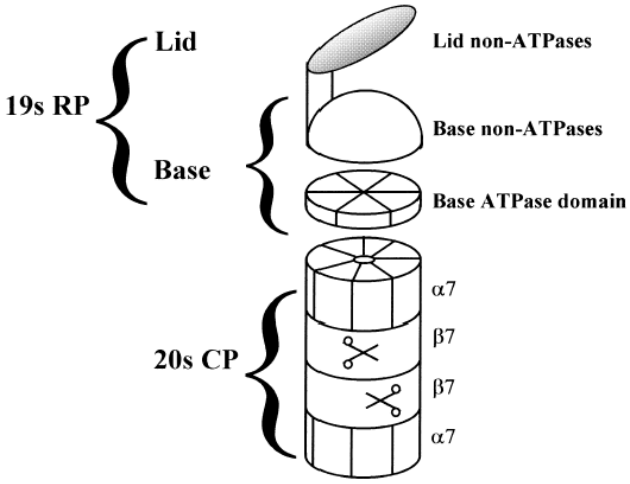
The 20 cylinder forms the core of the 26S proteasome, which contains two additional cap structures, which as isolated particles have a sedimentation coefficient of 19 S ("19 S cap complex" or "PA700"). Purified 20 S proteasomes cleave proteins in an ATP independent manner; 26S proteasomes need ATP for proteolytic activity, their major substrates are polyubiquitinated proteins. While most eubacteria lack a 20S or 26S proteasome, 20S proteasomes consisting of two different subunits each of alpha and beta type) have been purified from Rhodococcus spec. and other Actinomycetales. All eubacteria appear to contain Clp proteases, which like the proteasome, consist of a proteolytic core structure (two heptameric rings of subunit ClpP) and ATPase subunits (ClpA), which are distant relatives of the AAA superfamily.

3.4.2. The 19S Cap Complex

As mentioned above, in the eukaryotic cell, the AAA proteins are part of the regulatory complex (19S cap). The addition of two such caps to each side of the 20S cylinder results in the formation of the so-called 26S proteasome.



Structure of the 26S proteasome.



The 19S cap can be separated into a Base and a Lid domain. While the Base is the energy-dependent component responsible for unfolding and feeding protein substrates into the proteolytic chamber, the Lid is needed for the recognition of ubiquitinated proteins and does not directly interact with the 20S core. It helps in determining the substrate specificity of the enzyme. For example the ubiquitin-dependent degradation of protein by the proteasome requires the presence of the cap. However, how the cap recognizes ubiquitinated proteins is not clear yet. Other regulatory roles of the cap may lie in determining the subcellular localization of the complex.

Finally, additional functions of the caps comprise cleaving off ubiquitin moieties conjugated to the substrates to recycle them. This isopeptidase function is ensured by so-called ubiquitin C-terminal hydrolases, which are very specific enzymes conserved in evolution.

3.4.3. The Immunoproteasome

Proteasomes plays also role in the processing of proteins for presentation by the MHC class I pathway.
During an immune response to pathogens, the proinflammatory cytokine interferon (IFN)-gamma and tumor necrosis factor (TNF)-alpha are released, which induce the proteasome subunits LMP2, LMP7, and MECL-1. These replace the constitutively expressed active site β subunits of the proteasome leading to a marked change in the cleavage preference of the proteasome and the production of T-cell epitopes.
Proteasome activity is further changed by the IFN-gamma-mediated induction of a different type of regulator, the PA28 activator (or 11S regulatory complex). The PA28 activator consists of PA28 α and β subunits, which form a 200kDa hexamer. The PA28 complexes can bind to the ends of the 20S proteasome to form PA28-proteasome complexes. It enhances the degradation of small peptides but not denatured or ubiquitinated proteins.

It is believed that these changes stimulate the production of antigenic peptides which can be bound by MHC class I molecules in the ER and after delivery to the cell surface being presented to cytotoxic T cells.