The Complexity of 20S Proteasome Assembly

The ubiquitin/proteasome pathway

Healthy living cells should be able to maintain a good quality control of all of their molecules. Since proteins are the main executers of genetic functions, their quality and the balance between their synthesis and destruction have to be controlled. Mutations, errors in transcription or translation and cellular stress can cause alterations in amino acids residues that may prevent proteins from attaining their proper folding, native conformations. The fate of misfolded and unneeded regulatory proteins is proteolysis. Many of the intracellular proteases are found in lysosomes of animal cells and in the vacuoles of fungal and plant cells, but their action is highly unspecific. The main pathway responsible for selective intracellular proteolysis is known as the ***U***biquitin/***P***roteasome ***S***ystem (UPS), in which the central mediator is the proteasome, a large selfcompartmentalized protease complex found in all kingdoms of life. In the cytoplasm and in the nucleus of eukaryotic cells, the UPS provides the main mechanism for a selective and ATP-dependent degradation of short-lived proteins. Intracellular proteolysis contributes to many cellular regulatory mechanisms, including cell cycle control, DNA repair, various stress responses and signal transduction. When this process fails, severe consequences are seen like obesity, neurodegenerative disorders, muscular dystrophies or cancer.

In the ubiquitin/proteasome system, protein substrates are tagged by the attachment of polyubiquityl chains, which target them for degradation by the 26S proteasome. Ubiquitin itself is recycled upon substrate degradation and the peptide products can be converted into free amino acids, which can de degrade or recycled into protein synthesis.

The attachment of a polyubiquitin chain to a protein requires the activity of three classes of enzymes: a ubiquitin-activating enzyme, E1, a ubiquitin-conjugating protein, E2, and a ubiquitin-protein ligase, E3. In some cases, however, it was discovered that an additional ubiquitin chain elongation factor, named E4, is involved in the synthesis of long multiubiquityl chains by catalysing the elongation of the multiubiquityl chain in collaboration with E1, E2 and E3. The C-terminal Gly residue of ubiquitin is activated in an ATP-requiring step by E1, in which a thioester linkage is formed between the ubiquitin and a Cys residue of E1. Activated ubiquitin is next transferred to an active site Cys residue of E2. Finally, ubiquitin is linked by its C-terminus to an internal Lys residue of a substrate protein via an isopeptide linkage. This step is catalyzed by E3 that has the ability to complex with E2 (see figure 1).

Usually there is a single E1, but there are many E2 enzymes and even more E3 proteins (typically over 1,000) encoded in each eukaryotic genome. E3 proteins appear to be responsible mainly for the selectivity of ubiquitin-protein ligation (and thus, protein degradation). They do so by binding, directly or indirectly, to specific protein substrates which contain specific recognition signals. The best characterized degradation signal is known as the N-end rule, in which the *in vivo* half-time of a

protein is determined by the nature of its N-terminal amino acid residue. Phenylalanine, leucine, tryptophan, tyrosine, arginine, lysine, histidine, isoleucine, asparagine, aspartic acid and glutamic acid are the destabilizing N-terminal residues described for *Sacchaomyces cerevisiae*. Sequences rich in proline (***P***), glutamic acid (***E***), serine (***S***) and threonine (***T***), referred to as PEST, have been proposed to function also as degradation signals. Misfolded or abnormal proteins can sometimes expose hydrophobic domains, usually hidden in the correctly folded structure. These domains are also referred to as ***destruction signals*** or ***degrons***. Relate the notion of degrons with the thermal mobility of molecules at temperatures above absolute zero (-273.15 °C or 0 K) and with the low energy difference between the folded and unfolded states of a protein (typically in the range 3.5 to 14 kcal/mol).

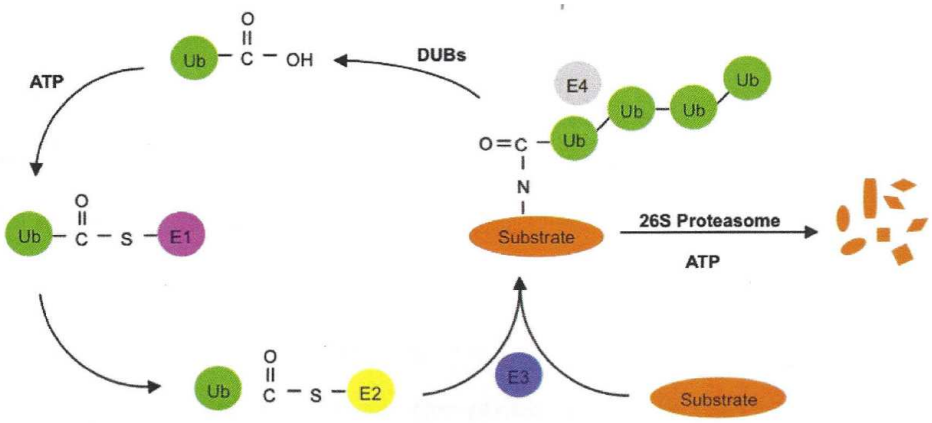

Figure 1. The ubiquitin/proteasome pathway. The formation of a polyubiquityl chain attached to a protein substrate involves the action of three different classes of enzymes in the following order: E1 - ubiquitin-activating enzyme, E2 - ubiquitin-conjugating enzyme and E3 – ubiquitin-protein ligase. E4 can also participate in this process promoting ubiquitin chain elongation, in some cases. Polyubiquitylated substrates are degraded by the 26S proteasome into small peptides and ubiquitin is recycled by deubiquitylation enzymes (DUBs).

The proteasome and its regulatory complexes

The 26S proteasome is a ~2.5 MDa complex that is present ubiquitously in eukaryotes (see figure 2). It is composed of the barrel-shaped catalytic core particle (CP), termed as the "20S proteasome", capped at both ends by the 19S regulatory particle (RP). This complex association requires ATP binding and also some proteins, described to stabilize the 26S proteasome, like Hsp90 and Ecm29 proteins. While the CP is responsible for proteolysis, the RP is responsible for the regulation of the 20S CP catalytic activity. This regulation comprises substrate recognition, deubiquitylation, unfolding and translocation through the channel of the CP. It was described that the 19S RP also participates in nascent 20S proteasomes stabilization and maturation. The RP is composed of at least 19 subunits that include non-ATPases and AAA-ATPases (***A***TPases ***A***ssociated to a variety of cellular ***A***ctivities) proteins. Together, they ensure degradation of ubiquitylated proteins by the 20S proteasome, probably with the help of other proteins like the ubiquitin ligase Hul5 or the deubiquitylating enzyme Ubp6 which were described to interact with 19S subunits.

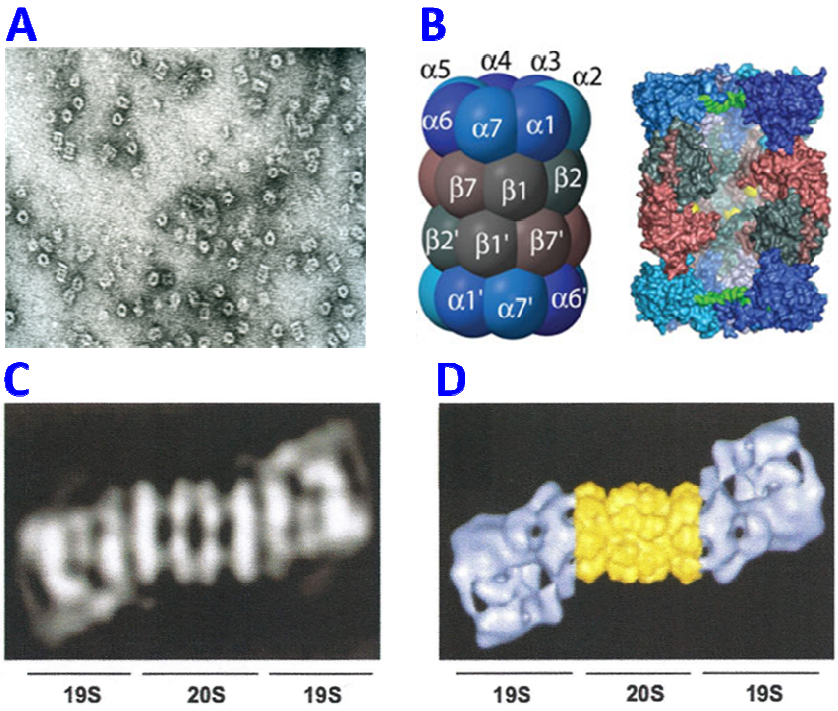

Figure 2. (**A**,**B**) The 20S proteasome from *Lemna minor*. (**C**,**D**) The 26S proteasome from *Drosophila melanogaster.* Electron microscopy images of the 20S (**A**) and 26S (**C**) proteasomes. Three-dimensional model structures based on electron microscopy of 20S (**B**) and 26S (**D**) proteasomes, which is composed of the 20S core particle (shown in yellow) capped at both ends by two 19S regulatory particles (shown in blue).

To gain access to the interior of the CP, where the specific active sites are, the substrate has to pass through a very narrow pore that usually is not accessible to folded substrates. This requires an ATP-dependent mechanism of gate opening and also substrate unfolding to facilitate degradation. These functions are also carried out by 19S RP, more specifically by its AAA-ATPases subunits. While ATP hydrolysis is necessary for protein unfolding the gate opening event requires only ATP binding. The majority of the peptides generated by proteasomes (varying between 3 and 23 amino acids residues in length) are rapidly degraded into amino acids by cytosolic proteases.

Aside from the 19S RP (also known as PA700 - Proteasome Activator of 700 Da), other regulatory particles have been described, that can change 20S CP activities and specifities. In mammalian cells, three more proteasome activators have been identified: PA28 (11 S regulator or REG), PA200 and PI31. ln contrast to the 19S RP, these activators do not recognize ubiquitylated proteins or use ATP. There are two PA28 homologs, called αβ and γ. The αβ forms a heteroheptamer, whereas γ forms a homoheptamer. PA28αβ expression is strongly induced by interferon- γ and infection indicating a role in antigen peptide presentation to the immune system. When bound to 20S proteasomes, PA28 activates cleavage of small peptides but not the degradation of ubiquitylated proteins. While PA28 αβ is predominantly detected in the cytoplasm, PA28γ is mainly found in the nucleus. This complex may have a role in apoptosis. The analysis of proteasomal complexes from mammalian cells led to the detection of hybrid proteasomes, in which the 20S CP is capped by the 19S RP at one end, enabling recognition and unfolding of ubiquitylated substrates, and PA28 at the

other end. These 19S-CP-PA28 hybrid proteasomes may contribute to more efficient proteolysis of some substrates.

The 20S core particle

Haemophilus infuenza 20S proteasome consists of two homohexameric rings built from only one subunit type that is structurally similar to eukaryotic proteasome β-type subunits (figure 3). In most eubacteria and archaebacteria, proteasomes have just two subunit types, α and β, which assemble into a cylinder-shaped complex of four seven-membered rings in an $\alpha_7\beta_7\alpha_7\beta_7$ arrangement. In the eubacteria *Rhodococcus* sp. the proteasome is built of two distinct α (α1 and α2) and two different β (β1 and β2) subunits. Since all β-subunits contain active sites, the prokaryotic proteasomes bear 14 of such sites. In contrast, eukaryotic 20S proteasome are composed of seven distinct α-type (α1 to α7) and seven distinct β-type (β1 to β7) subunits. Only three of the β-subunits are catalytically active (β1, β2 and β5). Despite these differences, subunit arrangement is maintained between prokaryotic and eukaryotic proteasomes (figures 3B and 3C). In mammals, three additional β-type subunits β1i, β2i and β5i) are synthesized after induction by γ-interferon and replaced by their related β-type subunits in newly formed "inmunoproteasomes". Recently, another specialized proteasome subtype was discovered. It contains an alternative subunit (β5t) that is specifically expressed in the thymus, and because of that it is known as "thymoproteasome". Curiously, in *S. cerevisiae* only the *PRE9* gene encoding α3 subunit can be deleted without loss of viability. Proteasomes lacking α3 contain two α4 subunits.

Other important structural feature of eukaryotic proteasomes that distinguishes them from their prokaryotic counterparts includes a complete closure of the α-rings (figures 3B and 3C, top). The N-terminal domains of eukaryotic α-subunits are able to seal the entry port of 20S proteasome, preventing the non-specific degradation of cellular proteins.

While the α-ring function seems to be binding to the different proteasome regulatory particles and also regulation of the gating pore, β-subunits are involved in protein hydrolysis, proteasome assembly and stability. The proteasome active β-subunits are members of the Ntn (N-terminal nucleophile) hydrolases. Common to this family is the ability to hydrolyse amide bonds. All Ntn-hydrolases are synthesized as inactive precursors and are converted to an active form by an autocatalytic internal cleavage, which exposes a threonine, serine or cysteine residue as the new N-terminus. In the proteasome, a threonine (Thr1) serves as the N-terminal nucleophile that performs the nucleophilic attack onto the carbonyl carbon of a peptide bond. Other residues were found close to the N-terminal Thr, such as Lys33, Asp17, Asp166, Ser129 or Ser169 (numbering according to the *Thermoplasma* proteasome), which appear to be required for structural integrity of the proteolytic site. Beyond these residues, a water molecule was also found in the vicinity of Thr, which was described to be involved both in intramolecular autolysis and substrate proteolysis.

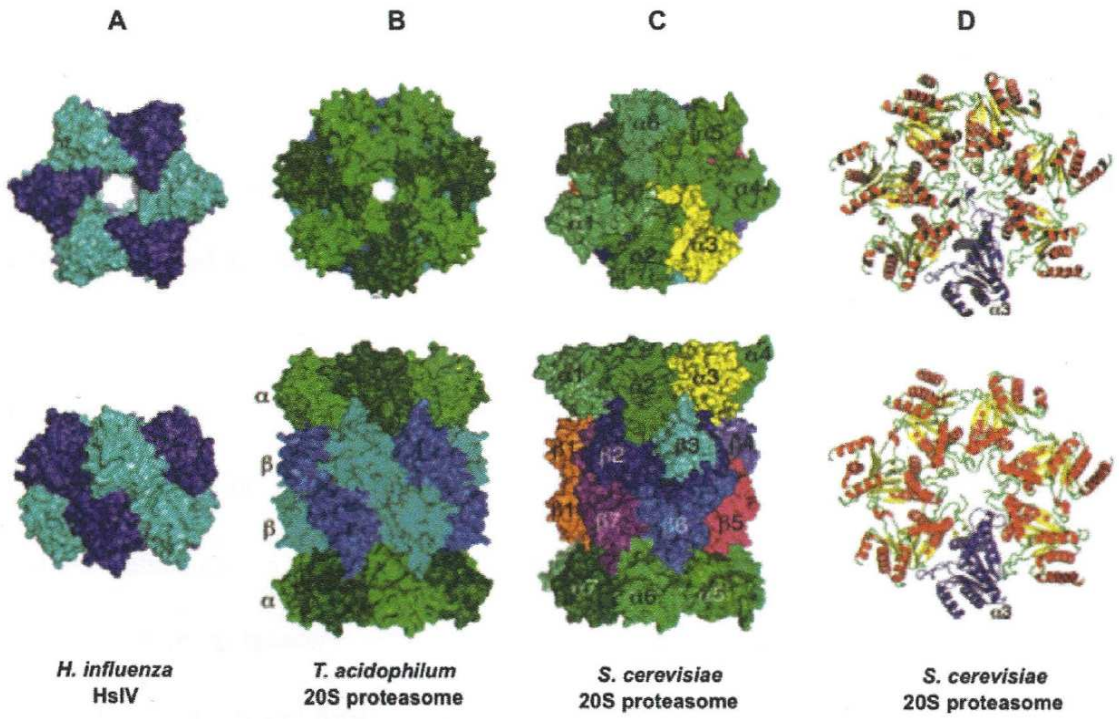

Figure 3. Structure of prokaryotic and eukaryotic proteasomes. Side and top view surface representation of the crystal structures of the (**A**) protease HsIV from the eubacterium *Haemophilus influenza;* (**B**) proteasome from the archaebacterium *Thermoplasma acidophilum;* (**C**) proteasome from the eukaryote *Saccharomyes cerevisiae.* The identical β-type subunits of Hs1V and the *Thermoplasma* proteasome are each shown in two different shades of blue. The identical α-subunits of the *Thermoplasma* proteasome are shown in different shades of green. The 14 individual subunits of the yeast proteasome are shown in distinct colours. (**D**) Ribbon diagrams of two different top view configurations of the *S. cerevisiae* 20S proteasomal α-rings: wild-type (top) and *a3*ΔN mutant (bottom).

Because β3, β4, β6 and β7 lack the three critical residues (Thr1, Asp17 and Lys33) of the catalytic centre, they are inactive. These four subunits are found in the mature proteasome either in a unprocessed form (as is the case for β3 and β4) or in a intermediated processed form (as is the case for β6 and β7). In contrast to autolysis of the active β-subunits (β1, β2 and β5), processing of β6 and β7 seems to be due to proteolytic action of the neighbouring active subunits. The N-terminal active sites of β-subunits are on the inner surface of the β-rings, whereas the C-terminal of β-subunits are on the outer surface of the 20S proteasome. Maturation of β-subunits occurs during proteasome assembly. Prokaryotic proteasomes, like the complexes from *Thermoplasma* and *Rhodococcus*, harbour only one type of activity, which is termed chymotrypsin-like because cleavage occurs preferentially after bulky hydrophobic residues. Eukaryotic proteasomes show two additional activities, termed trypsin-like, because cleavage occurs after basic residues, and postacidic, because cleavage occurs after acidic residues: β1 mediates the postacidic, β2 the tryptic and β5 the chymotryptic activity (figure 4). These different enzymatic activities can be modulated by inter-subunit contacts and also by substrate binding either to noncatalytic sites or to the active sites in the proteasome.

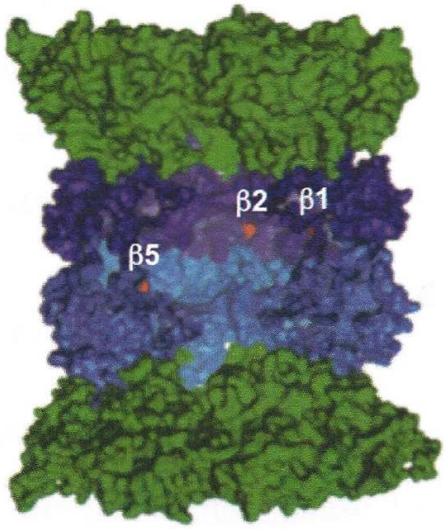

Figure 4. *Saccharomyces cerevisiae* 20S proteasome active sites. Surface representation with a cut open view of the proteasome catalytic chamber. The localization of the catalytic triad: Thr1, Asp17 and Lys33 in the β active proteasome subunits (β1, β2 and β5) are represented in red.

Assembly of the 20S proteasome

The understanding of how eukaryotic 20S proteasomes are assembled is a research area that gained some advances in the last few years. Archaeal, eubacterial and eukaryotic proteasomes do not seem to follow exactly the same assembly pathways, which is not surprising if one considers the increase in subunit complexity. The relatively low complexity with respect to subunit composition of prokaryotic 20S proteasomes has been a precious tool to elucidate the basic mechanism of 20S proteasome assembly. Figure 5 illustrates the proposed assembly pathway of the yeast 20S proteasome.

Molecular chaperones comprise a large and diverse group of proteins that share the property of assisting the folding, unfolding, assembly and disassembly of other macromolecular structures, but are not permanent components of these structures when they are performing their normal biological functions. The accumulationof misfolded proteins (aggregation) is recognized as a common characteristic of a number of degenerative diseases like Alzheimer or Parkinson. Therefore, the importance of cellular protein quality control is essential for cell viability and this task is taken care of by both chaperones and proteasomes. Some chaperones, but not all, are stress or heat shock proteins, because the requirement for chaperone function increases under stress conditions since stress causes protein unfolding and misassembly. Among these heat shock chaperones, the members of the Hsp70 (***H***eat-***s***hock ***p***rotein 70) family are the most abundant and best studied. The members of this essential super-family, with an approximate molecular mass of 70 kDa, assist a wide range of processes, including protein synthesis, folding, stabilization,

translocation, degradation, assembly and disassembly of protein complexes, both in times of cell stress and under normal growth conditions.

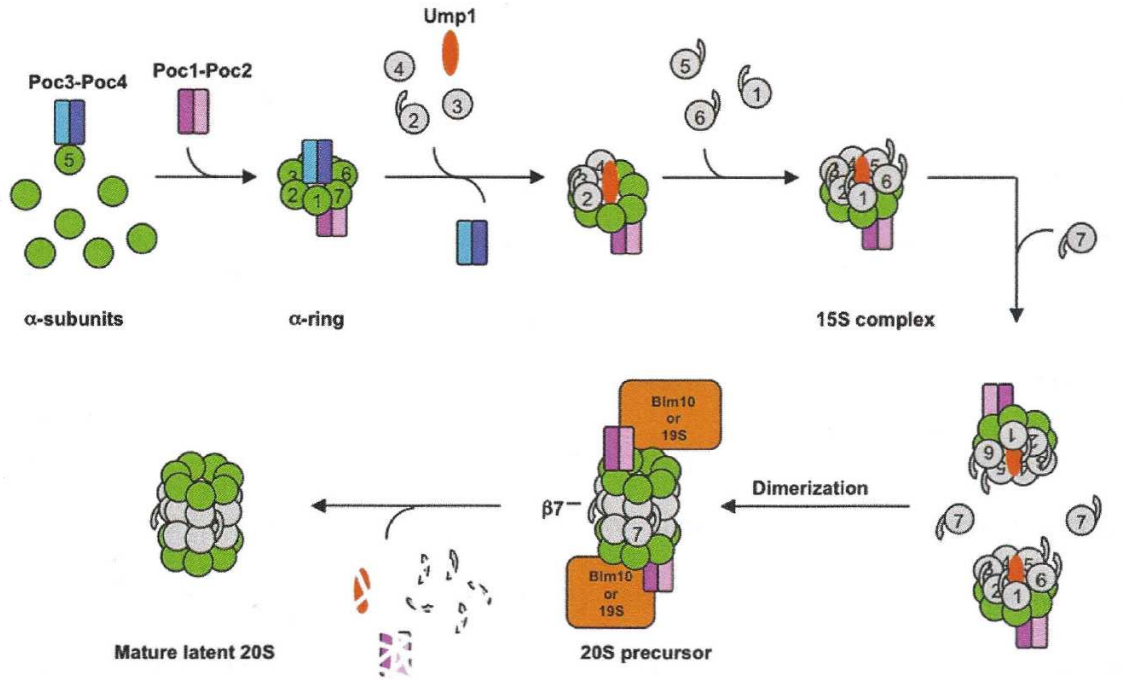

Figure 5. Assembly pathway of the yeast 20S proteasome. The two chaperone complexes Poc1-Poc2 (pink) and Poc3-Poc4 (blue) promote the assembly of α-subunits α1- α7 (green) into heptameric rings. The resulting α-ring serves as a template for assembly of Ump1 (red) and β-subunits (grey): first β2, followed by β3 and β4. At this stage, the Poc3-Poc4 complex dissociates and β5, β6 and β1 join the assembly event to form the 15S complex or half-proteasome. Dimerization of such precursor complexes is triggered by the binding of β7, which stabilizes nascent proteasomes via its long C-terminal extension. The short-lived precursor is activated by autocatalytic maturation of its active β-subunits. The activated proteasome then degrades Ump1 and Poc1-Poc2 chaperone complex. The mature latent 20S proteasome is ready to be activated by the binding of Blm10 or 19S RP to its exposed closed α-ring. Some β-subunits are drawn with their N-terminal propeptides, which are cleaved upon 20S proteasome formation.

The essential role of chaperones in yeast proteasome assembly

The 26S proteasome is composed by thirty-three distinct subunits that occupy specific positions within the structure. The assembly of this complex requires additional factors known as molecular chaperones, which promote the process and prevent incorrect subunit associations. Recent studies have revealed an ordered pathway for the assembly of the 19S component of the 26S proteasome, In this assembly pathway, four chaperones called Nas2, Hsm3, Nas6 and Rpn14 in yeast ensure efficient formation of this complex structure. Assembly of the 20S component requires three different molecular chaperones: Ump1, and Pocl-Poc2 and Poc3-Poc4 heterodimers. Other factors like Hsp90 and Hsc73 have been also implicated in 26S proteasome assembly. The participation of these factors in proteasome assembly

provided new mechanistic insight into the cooperative interactions between molecular chaperones and the proteolysis systems. Hsp110 and Hsp70 molecular chaperones have also been identified together with α4 proteasome subunit in precursor complexes.

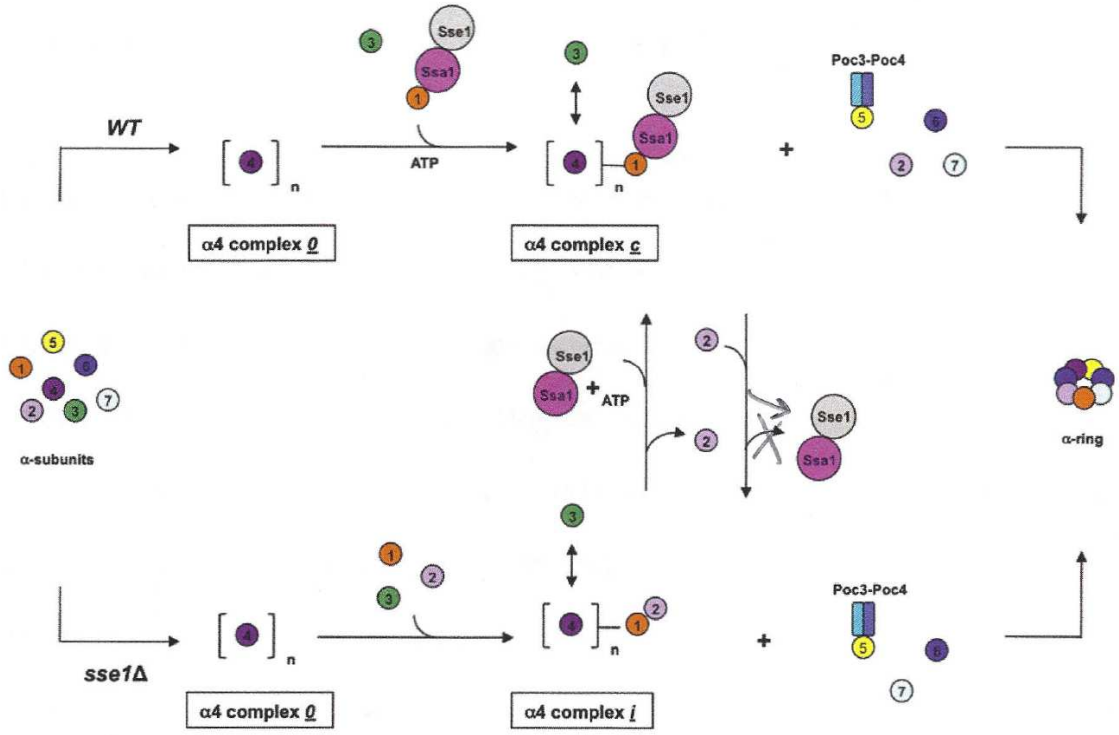

Figure 6. α-Ring assembly model. In the wild-type (WT), α-ring assembly might be initiated by α4 complex ***0***, formed by multiple (n) α4 subunits. In the presence of ATP, Sse1 and Ssa1 chaperones, perhaps in association with α1, would help the replacement of α4 by α1 and α3. This event will avoid premature incorporation of α2 into the complex named α4 complex ***c***. In *sse1Δ* mutant, α2 is incorporated and an intermediary complex is formed (α4 complex ***i***). The addition of Sse1-Ssa1 and ATP promote the formation of α4 complex ***c*** with simultaneous dissociation of α2. Later stages of α-ring assembly involve the association of α2, α6, α7 and α5-Poc3-Poc4. Sse1-Ssa1 should be released upon α2 incorporation into the α-ring. Poc3-Poc4 chaperone should be bound to the α-ring but it is not represented. Poc1-Poc2 chaperone complex is also not represented.

Maio de 2011

Ricardo Boavida Ferreira

Adapted and modified from:

Ana Catarina Parreira Peres Matias

Tese de Doutoramento em Bioquímica

Especialidade em Biologia Celular e Molecular

Universidade do Algarve, 2011