Assembly and Beyond – The Structure and Functions of Chaperones of the Proteasome

Prasanna Venkatraman

Advanced Centre for Training, Research and Education in Cancer (ACTREC), Tata Memorial Centre, Kharghar, Navi Mumbai – 410210, India

The 26S proteasome is the major protease machinery in eukaryotic cells responsible for spatio-temporally regulated turnover of the proteome and therefore critical to the maintenance of homeostasis. These act not only as efficient garbage disposal units for abnormal/misfolded, denatured and oxidised proteins, but, are crucial for irreversible termination of key cellular process regulation of various checkpoints. Replication, transcription, translation, antigen processing, maintenance of stemness, development and differentiation are several of the functions and processes regulated by the proteasomes. The 2.5-megadalton structure of the eukaryotic proteasome is very complex and is assembled from 66 polypeptides with 33 structurally unique. The biogenesis of the multi subunit architecture is regulated by dedicated chaperones which orchestrate the assembly of intricate and complex structure. The current review focuses on four chaperones PSMD9 (Nas2), PSMD10 (Nas6), PAAF1 (Rpn14) and S5B (Hsm3) responsible for the assembly. The structure of the chaperones, molecular details of interaction of the individual chaperones with subunits of the proteasome during the process of assembly, role in hierarchical steps leading to assembly, recent evidences for the unprecedented role of some of these proteins (PSMD9 and PSMD10) in other physiological processes will be summarized. In addition, potential of two chaperones as targets for development of inhibitors of proteasome function will be explored.

INTRODUCTION

Most of the cellular pathways involved in the biology of a eukaryotic organism and the steady state homoeostasis is regulated by the Ubiquitin Proteasome System (UPS) (Glickman and Ciechanover, 2002; Lecker et al., 2006). UPS comprises of an upstream ubiquitination machinery intimately linked to other post translational modifications such as phosphorylation and acetylation, and a down-stream proteolytic machinery, the 26S proteasome. Ubiquitination is coordinated by the action of several enzymes (E1, E2, E3, E4 proteins) acting...
Figure 1: Components of UPS

(A) Players in the cascade of events involved in ubiquitination of a protein substrate. Mono-ubiquitin is activated by a single E1 enzyme and is transferred in a series of steps via the E2 and E3 enzymes for conjugation to the Lysine residue of a protein substrate. A tetra ubiquitin is sufficient for degradation. However, polyubiquitinated substrates are the norm in physiologically relevant conditions. Ubiquitin is released even before the substrate proteins enter the proteolytic core by the action of deubiquitinating enzymes present in the proteasome. (Adapted from Sorokin et al., 2009)

(B) 26S Proteasome and its various sub complexes. 20S core particle made of the αβ rings and the 19S regulatory particle made of the base and lid components are the two major subcomplexes that assemble to form the functional proteasome. The base component is formed from 6 ATPases and three non-ATPase subunits Rpn1, Rpn2 and Rpn13. The lid is formed by 13 Non-ATPases subunits. The nomenclatures of each of these subunits can be found in Table 1. (Adapted from Sorokin et al., 2009)

(C) Schematic representation of the hierarchical steps involved in substrate degradation by the proteasome. Ubiquitinated substrates are recognized by the ubiquitin binding motifs in the 19S regulatory particle and the substrate is unfolded by the six membered ATPase ring in an ATP dependent manner. The gate at the α-ring is opened to reveal a narrow channel that runs across the proteolytic core. Assembly of the 19S RP and 20S CP is a key event required to open this gate. The unfolded substrate is translocated through the 20S core particle and the catalytic sites in the β-ring cleave the polypeptide and small peptides of 3-25 amino acids are released.

in a cascade to mark and prepare substrates for degradation (Fig. 1A) (Finley, 2009). The downstream degradation unit, the 26S proteasome, also called the 'chamber of doom' (Goldberg et al., 2001), is a specialized self-compartmentalized, nano structure that belongs to AAA+ family of proteolytic machines (Sauer and Baker, 2011) conserved across evolution. Unlike other proteases but similar to the AAA+ family of proteins such as bacterial protease complex (ClpX/ClpP) or the PAN proteasome complex in Archaebacteria, the 26S proteasomes can degrade fully folded proteins and
generate short peptides and amino acids (Akopian et al., 1997; Baumeister et al., 1998; Benaroudj and Goldberg, 2000; Henderson et al., 2011; Zhang et al., 2009). The ATPases of the proteasome helps in unfolding of the folded proteins. Proteasomes also degrade model substrates and physiologically relevant substrates in an ubiquitin independent manner (Asher and Shaul, 2006; Barbour et al., 2013; Coffino, 2001; Hwang et al., 2005; Kahana, 2007; Rosenberg-Hasson et al., 1989; Sdek et al., 2005; Singh et al., 2012; Stewart et al., 2010; Zetter and Mangold, 2005). Based on in vitro studies referred above, it is now well established that substrates require two signals for degradation – a minimum of four ubiquitin and an unstructured region (Schrader et al., 2009). In the absence of ubiquitin, secondary structures within the substrate help in anchoring the protein the 19S regulatory particles (Singh et al., 2012). Impairment in the function of UPS components results in accumulation of proteins leading to cellular stress and apoptosis. Velcade, an active site inhibitor of proteasome is used for treatment of multiple myeloma and mantle cell lymphoma (Adams and Kauffman, 2004; Shahshahan et al., 2011).

The major functional unit of the proteasome is the 26S holo complex

### Table 1: List and nomenclature of the proteasomal subunits involved in proteasomal assembly, structure and function.

A) List of 20S core particle subunit from yeast and mammals. Composition of the immunoproteasome which are involved in antigen processing and presentation are also shown.

<table>
<thead>
<tr>
<th>Subunit</th>
<th>\textit{T. acidophilum}</th>
<th>\textit{S. cerevisiae}</th>
<th>Constitutive 20S</th>
<th>Immunoproteasome</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1)</td>
<td>(\alpha)</td>
<td>Sc(1/\text{Prc2/Prs2/C7})</td>
<td>PSMA6/ProS27/Iota</td>
<td>PSMA6/ProS27/Iota</td>
</tr>
<tr>
<td>(\alpha_2)</td>
<td>(\alpha)</td>
<td>Pre8/Prs4/Y17</td>
<td>PSMA2/C3/LmpC3</td>
<td>PSMA2/C3/LmpC3</td>
</tr>
<tr>
<td>(\alpha_3)</td>
<td>(\alpha)</td>
<td>Pre9/Prs5/Y13</td>
<td>PSMA4/C9</td>
<td>PSMA4/C9</td>
</tr>
<tr>
<td>(\alpha_4)</td>
<td>(\alpha)</td>
<td>Pre6</td>
<td>PSMA7/C7/XAPC7</td>
<td>PSMA7/C7/XAPC7</td>
</tr>
<tr>
<td>(\alpha_5)</td>
<td>(\alpha)</td>
<td>Pup2/Doa5</td>
<td>PSMA5/Zeta</td>
<td>PSMA5/Zeta</td>
</tr>
<tr>
<td>(\alpha_6)</td>
<td>(\alpha)</td>
<td>Pre5</td>
<td>PSMA1/C2/Pros30</td>
<td>PSMA1/C2/Pros30</td>
</tr>
<tr>
<td>(\alpha_7)</td>
<td>(\alpha)</td>
<td>Pre10/Prc1/Prs1/C1</td>
<td>PSMA3/C8</td>
<td>PSMA3/C8</td>
</tr>
<tr>
<td>(\beta_1)</td>
<td>(\beta)</td>
<td>Pre3</td>
<td>PSMB6/Y/delta/LMPY/LMP19</td>
<td>PSMB6/B1/LMP2/Ring12</td>
</tr>
<tr>
<td>(\beta_2)</td>
<td>(\beta)</td>
<td>Pup1</td>
<td>PSMB7/Z/Mmc14</td>
<td>PSMB10/B2/LMP10/MECL1</td>
</tr>
<tr>
<td>(\beta_3)</td>
<td>(\beta)</td>
<td>Pup3</td>
<td>PSMB3/C10</td>
<td>PSMB3/C10</td>
</tr>
<tr>
<td>(\beta_4)</td>
<td>(\beta)</td>
<td>Pre1/C11</td>
<td>PSMB2/C7</td>
<td>PSMB2/C7</td>
</tr>
<tr>
<td>(\beta_5)</td>
<td>(\beta)</td>
<td>Pre2/Doa3/Prg1</td>
<td>PSMB5/X/MBl</td>
<td>PSMB5/(\beta5/\text{Ring10/Y2/C13/LMP7})</td>
</tr>
<tr>
<td>(\beta_6)</td>
<td>(\beta)</td>
<td>Pre7/Prs3/Pts1/C5</td>
<td>PSMB1/C5</td>
<td>PSMB1/C5</td>
</tr>
<tr>
<td>(\beta_7)</td>
<td>(\beta)</td>
<td>Pre4</td>
<td>PSMB4/N3/beta/LMP3</td>
<td>PSMB4/N3/beta/LMP3</td>
</tr>
</tbody>
</table>
made up of two modules – the 19S regulatory particle (RP) and the 20S proteolytic core (CP) (da Fonseca and Morris, 2008; Finley et al., 2016; Lander et al., 2012; Rubin and Finley, 1995; Walz et al., 1998). The 20S proteolytic core is a central four ringed cylindrical barrel made up of two types of seven member ring structures, the α and the β. The inner two β-rings carry three different types of catalytic sites, the trypsin-like (β2), caspase-like (β1) and the chymotrypsin-like (β5) sites responsible for the degradation of a variety of protein molecules of diverse sequences with the exception of repeat

Table 1: List and nomenclature of the proteasomal subunits involved in proteasomal assembly, structure and function.

<table>
<thead>
<tr>
<th>Subunit</th>
<th>S. cerevisiae</th>
<th>Mammals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rpt1/S</td>
<td>Rpt1/Cim5/Yta3</td>
<td>PSMC2/Mss1</td>
</tr>
<tr>
<td>Rpt2/S4</td>
<td>Rpt2/Yhs4/Yta5</td>
<td>PSMC1</td>
</tr>
<tr>
<td>Rpt3/S6b</td>
<td>Rpt3/Ynt1/Yta2</td>
<td>PSMC4/Mip224/Tbp7</td>
</tr>
<tr>
<td>Rpt4/S10b</td>
<td>Rpt4/Crl13/Pcs1/Sug2</td>
<td>PSMC/Sug2/p42</td>
</tr>
<tr>
<td>Rpt5/S6a</td>
<td>Rpt5/Yta1</td>
<td>PSMC3/Tbp1</td>
</tr>
<tr>
<td>Rpt6/S8</td>
<td>Rpt6/Cim3/Crl3/Sug1/TbpY/Tby1</td>
<td>PSMC5/p45/Sug1/Trip1</td>
</tr>
</tbody>
</table>

**ATPases Unfolding of protein, gate opening and translocation**

- Rpn1/S2: Rpn1/Hrd2/Nas3/Rpd1 - PSM2/p97/Trap2
- Rpn2/S1: Rpn2/Sen3 - PSM1/p112
- Rpn3/S3: Rpn3/Sun2 - PSM2/p58
- Rpn4: Rpn4/Son2/Ufd5 - PSM12/p55
- Rpn5/S5: Rpn5/Nas5 - PSM11/p44.5
- Rpn6/S9: Rpn6/Nas4 - PSM12/p44.5
- Rpn7/S10: Rpn7 - PSM2/p42a
- Rpn8/S11: Rpn8 - PSM7/p40/Mow34
- Rpn9/S12: Rpn9/Nas7 - PSM13/p40.5
- Rpn10/S5a: Rpn10/Sun1/Mcb1 - PSM4/S5a/Mcb1
- Rpn11/S13: Rpn11/Mpr1 - PSM14/Pad1/Poh1
- Rpn12/S14: Rpn12/Nin1 - PSM8/p31

**Chaperones of 19S regulatory particle assembly**

- Rpn13: Rpn13/Daq1 - PSM10/p28/Gankyrin
- p28: Nas6 - PSM9/p27
- p27: Nas2 - PSM5/KIAA0072

sequences such as poly Gly-Ala, alanine or glutamine repeats (Levitskaya et al., 1997; Venkatraman et al., 2004). The two \( \alpha \)-rings sandwich the inner two \( \beta \)-rings to form the 20S proteolytic core. A single 20S core particle is decorated on either side by a 19S regulatory particle. Each 19S regulatory particle is made up of at least 13 non-identical subunits, 6 of which are ATPases (Table 1; Fig. 1B). Some of these subunits are responsible for substrate recognition via ubiquitin. The non-ATPase subunit is also a deubiquitinating enzyme (Rpn11) which releases ubiquitin from the polyubiquitin chain before the substrate enters the proteolytic core. Access to the 20S core particle is restricted by a closed gate guarded by the loops in the \( \alpha \)-ring which restricts entry of small peptides. Assembly of the 19S regulatory particles with ATPases directly in contact with the \( \alpha \)-ring, helps in opening of gate, allowing access to the active site chamber (Groll et al., 2000). The diameter of the channel remains small measuring about 13Å ensuring that only unfolded proteins enter the chamber even when the gate is open (Finley et al., 2016). The ATPases are presumed to unfold and translocate the polypeptide chain into 20S particles (Benaroudj et al., 2003; Navon and Goldberg, 2001) where proteolysis occurs place (Fig. 1C) (Bar-Nun and Glickman, 2012).

The biogenesis of the well-organized, compartmentalized, modular architecture of the proteasome is assisted by a dedicated group of molecular chaperones (Bai et al., 2014; Bedford et al., 2010; Kusmierczyk and Hochstrasser, 2008; Park et al., 2010; Roelofs et al., 2009). Since function is intricately linked to structure, guided assembly of such large heterologous complexes by chaperones is not surprising and is relevant to the proteasome, since no single subunit in the complex is identical to each other. The registry/position of the subunits within each assembly module/subcomplex and subsequently in the fully assembled complex must be programmed and coded in the structure of the proteins to prevent formation of non-productive complexes.

These chaperones dissociate at various stages during their assembly (Tomko et al., 2010). Under specific conditions, however, these chaperones can be found in association with functional 26S proteasome (Dawson et al., 2002; Jorgensen et al., 2006; Sahu et al., 2014). The reversibility of the process, mechanism of association/dissociation,
structural basis of protein-protein interaction, thermodynamic and kinetic parameters of the interactions are several of the active areas of research in the field. Other less well addressed albeit key questions that demand answers are fate of the chaperones after assembly and their role in other cellular functions.

The review focuses on four known chaperones of the 19S assembly – PSMD9 (Nas2), PSMD10 (Nas6), PAAF1 (Rpn14) and S5B (Hsm3). Several of the key points of discussion include the role of protein structure, domain motif interactions in assembly, bandwidth of protein interactions that permits association of some of the chaperones with the assembled proteasomes and other proteins. Special reference to the new 'avatar' of PSMD9/Nas2 and PSMD10/Gankyrin/ Nas6 in cellular functions, new avenues and strategies for possible intervention of function in diseases will be highlighted.

Historical Perspective on the Discovery, Interactions and Structure of 19S Chaperones

Discovery and the role of 20S assembly chaperones precede that of the 19S chaperones. Aided by the crystal structure of 20S proteasome, role of the PAC family of the 20S chaperones in the core particle assembly and other molecular events needed to form the mature 20S particles were worked out in great detail (Hirano *et al*., 2005; Le Tallec *et al*., 2007; Murata, 2006; Ramos and Dohmen, 2008; Yashiroda *et al*., 2008). A number of authors identified chaperones of the 19S regulatory particle assembly (Funakoshi *et al*., 2009; Roelofs *et al*., 2009; Tomko *et al*., 2010). Three of the chaperones Nas6, Hsm3 and Rpn14 (yeast nomenclature) were identified as 19S regulatory particle associated proteins in pull down assays, whereas 20S subunit was not found in the complex. Nas2 (yeast)/PSMD9, was discovered in a genetic screen as a suppressor of growth defect with a mutant of Rpt4 (Funakoshi *et al*., 2009). These proteins did not seem to be involved in proteolytic independent functions of the proteasome such as transcription (Roelofs *et al*., 2009). Individual silencing of the chaperones causes RP assembly defects but do not affect viability in mammalian cells (Kaneko *et al*., 2009) and in yeast (Funakoshi *et al*., 2009; Le Tallec *et al*., 2007; Roelofs *et al*., 2009; Saeki *et al*., 2009). Single deletion of the chaperones (nas6Δ, rpn14Δ, nas2Δ, and hsm3Δ) do
not adversely affect growth, but when combined, exhibit severe temperature sensitive growth defect paralleled by impairment in proteasome assembly in the following order: \( \text{nas6}\Delta \ rpn14\Delta > \text{nas2}\Delta \ hsm3\Delta > \text{nas6}\Delta \ hsm3\Delta \cong \text{hsm3}\Delta \ \text{nas6}\Delta > \text{hsm3}\Delta \ rpn14\Delta > \text{rpn14}\Delta \ \text{nas2}\Delta \cong \text{WT} \). Loss of Hsm3 was found to cause DNA repair deficiency in slow growing cells, although the mechanism has not been established (Fedorova et al., 1998). These genetic interaction studies indicate, unique but partially overlapping functions for the four chaperones. Since a) mutations of these genes mimic loss of (proteasome) function phenotype, b) deletion of these proteins in combination resulted in assembly defects of the proteasome and accumulation of ubiquitinated proteins and c) in most cases not found in association with the assembled 26S proteasome, these proteins were tested for their chaperone function in RP assembly.

Initially many of these chaperones were found to bind to only one of the six unique Rpt subunits: Nas2 binds to Rpt5, Nas6 to Rpt3, and Rpn14 to Rpt6. However, Hsm3/S5b binds to Rpt1 and Rpt2 subunits (Barrault et al., 2012). PAAF1 (mammalian Rpn14) binds most strongly to Rpt6 but is also associated with other five proteasomal ATPases, Hsm3 is credited with the scaffolding function since it can bind to both Rpt1 and Rpt2 proteins. Hsm3 enhances the affinity of interaction between Rpt1 and Rpt2. Pull-down experiments and yeast two hybrid studies further supported these interactions. Over the years, structures of the chaperones in their apo form or in complex with cognate Rpts were unfolded (Barrault et al., 2012; Kim et al., 2010; Nakamura et al., 2007).

It was evident that the four different chaperones have different fold and domain architecture (Park et al., 2010). Nevertheless, these bind to the C-domain of the cognate ATPase subunits. There are seven ankyrin repeats in Nas6; armadillo/heat-like repeats form Hsm3; Rpn14 contains a WD40 domain and Nas2 is predicted to contain a PDZ-like domain (Fig. 2A). While all the chaperones bind to the C-domain of the cognate ATPase subunit, the interaction of Nas2 involves the C-terminal residues of Rpt5 (Lee et al., 2011). Rpn14 binds close to the C-terminus ~8 residues away from the last residue. Gankyrin/Nas6 interacts with the sequence 29 residues inside the C-terminus of Rpt6/S6 ATPase. Hsm3 interacts with the last 90 residues of the C-terminus of Rpt1.
Figure 2: Structures of the chaperones, ATPases and the docking of ATPases to the 20S proteasome. (A) 4 major chaperones of the 19S regulatory particle assembly. Gankyrin/Nas6 contains 7 ankyrin repeats; HSM3 is made up of armadillo/heat like repeats; RPN14 contains WD40 domain and Nas2/PSMD9 harbors a PDZ-like domain. From the cryo-EM structure of the 26S proteasome, the group of the three ‘active’ ATPases (c-type Rpt 2, 5 and 3) with the HbYX motif, seen here as an extended stub (B) and group of three ATPases with the tail tucked away from the α-ring (C) are independently represented for clarity. Notice the two major domains, a C-terminal compact structure and the N-terminal coiled coil domain that extends out. The C-terminal domain docks to the 20S α-ring via specific HbYX motif present in the tails of the RPT subunits. Two ATPases from the two different groups (B and C) along with the bound α-ring shows how the hydrophobic tail of the ATPase (white) is docked into the α-ring (purple) while the other ATPase (violet) is not making any extensive contact with the α-ring (D). The structures were drawn from available PDB deposited at http://www.rcsb.org/ except for Nas2 which was modelled (Sangith et al., 2014).
Binding of these chaperones to the C-terminal domains of ATPases can lead to steric clashes with the core particle mandating their dissociation for stable assembly of the 19S to the CP. This is evident in the case of Nas2 which directly binds to the tail of Rpt5. In the final assembled structure the ATPase subunits Rpt2, Rpt5 and Rpt3 dock into specific α-ring pocket via the HbYX motif (hydrophobic-Tyr-any amino acid) at the C-terminus (Fig. 2B and C). A main chain carboxylate in the tail engages ε-amino group of the pocket lysine in a salt bridge interaction enforcing the docking interaction between RP and CP (Park et al., 2011). The docking is expected to open the gate and activate the proteasome (Fig. 2D) and disengage the chaperones from the assembled complex.

Assembly mechanism

Two hypotheses prevail to explain the mechanism of assembly; a template assisted model and the RP base assembly model (Fig. 3A and 3B). In the template assisted model, an assembly intermediate called BP1 binds to 20S core particle followed by binding of BP2. Further organization of the 19S regulatory complex occurs on the 20S core particle. According to the RP base model, reaction intermediates formed from twin ATP subunits with the cognate chaperones and Rpn subunits of the base, merge in a sequential manner to complete the formation of the base complex (ATPase hexameric ring plus Rpn subunits). The structure then combines with the lid (all other non-ATPase subunits) and this 19S regulatory complex arranges on the 20S core particle.

Despite these differences, both models are derived from early intermediates which may be similar as per reports from various groups. Three Rpt heterodimeric assembly intermediates, containing the four chaperones form the following modules: Nas2 module, Nas2-Rpt5-Rpt4, Nas6 module, Nas6-Rpt3-Rpt6-Rpn14 and Hsm3 module, Hsm3-Rpt1-Rpt2-Rpn1 (Funakoshi et al. 2009; Kaneko et al., 2009; Park et al., 2009; Saeki et al., 2009). In the template assisted model, the Hsm3 module including Rpt5 is called the BP2 assembly intermediate. The composition was later modified to include the fact that Rpt5 does not always associate with BP2. Therefore, BP2 intermediate would primarily be of the same composition as the RP assembly model i.e., Hsm3 module =
BP2 (Fig. 3A and 3B).
The other intermediate in the template assisted model BP1 whose composition was not clear but predicted to contain Rpt3, Rpt4, Rpt6, Rpn14, Rpn2 and Nas6. The intermediate very closely resembles the Nas2-Nas6-Rpn14 module barring presence of Nas2 and Rpn13 and Rpt5. However exclusion of Rpt5 from BP2 and from binding of Nas2 to Rpt5, and therefore in the same complex as Rpt5, BP1 intermediate may take the form of Intermediate 1 of the RP base model (Fig. 3A and 3B).

**Template assisted Model**
The two models differ in the manner in which the assembly takes place and whether or not 20S core structure was identified in the assembly intermediates. In the template assisted model, BP2 would bind to 20S to form the BP2-CP complex. BP1 will bind to the BP2-CP complex leading to the release of Hsm3 to form the Base CP structure. The lid, which exists in isolation and can be stripped off from purified 26S proteasome by salt washes as a separate entity, now binds to the Base CP to form the RP-CP module.
The mechanism is reinforced by three observations: a) The 20S proteasome assembly defects perpetuate the 19S assembly defects as a secondary manifestation (Kusmierczyk et al., 2008); b) The hydrophobic tails of the Rpt subunits when truncated by single amino acid show more severe assembly defects than any previously seen assembly/growth defects from chaperone deletion (Park et al., 2013); c) ∆1Rpt6 and a mutant with a single amino acid insertion in Rpt6 near the C-terminal tail, remain trapped with CP. Since Rpt6 tail is not a part of the fully assembled proteasome, the assembly intermediate BP1 with CP was established. Of all the pair wise deletion studies carried out, Rpt4 and Rpt6 showed the strongest assembly defective phenotype. Therefore it is presumed that Rpt4 and Rpt6 form a nucleating complex on the 20S to initiate base assembly, and that the complex is subsequently joined by BP1 to complete the Rpt ring (Fig. 3A).
The template assisted assembly process is further supported by the recently proposed 'affinity switch' model (Wani et al., 2015). Taking cue from the fact that among the short synthetic tail peptides of the ATPases only the Rpt6 tail inserts into the cognate α-ring pocket and the Rpt6 is not engaged in the core particle interaction in the fully assembled
Figure 3. Mechanism of assembly and the role of four chaperones.

(A) The template assisted Model. The intermediates are common to the RP-Base assembly model. However BP1 intermediate binds to the CP which is then bound by the Hsm3 module to form the assembled 26S proteasome and the chaperone such as Nas2 and Rpn14 dissociate at this stage.

(B) The RP-Base assembly model. Notice that Nas2/PSMD9 leaves the assembly early during the process and before RPT1 and RPT2 are recruited to the base complex. 20S joins the assembled base and lid to form the holo proteasome (please see text under Assembly mechanism for details). (Adapted from: Tomko and Hochstrasser, 2013)
complex, the authors tested the role of the tail regions in the context of the full length protein in the assembly process. Similar to the peptide, Rpt6 tail in the full length protein inserts into the cognate $\alpha$-ring pocket. The binding, induces the Rpt3 tail to engage in interaction with its own cognate $\alpha$-ring pocket. This relay of interactions has three consequences – as it reduces the affinity of Rpt6 tail interaction with the $\alpha$-ring allowing Rpt6 to disengage, enhances affinity of interaction of Rpt3 with the cognate $\alpha$-ring, and enables the disassociation of PSMD10/gankyrin from Rpt3. The key steps in ring opening and activation of the 20S core particle are orchestrated in a coordinated manner upon addition of the lid to Rpt-Base complex (Park et al., 2013).

NMR studies that captured the structural details of Rpt6 in different conformation add additional dimension to the assembly process. Distinct helices in the Rpt6 interact with Rpn14. The helices are in a conformational equilibrium with the coiled structure and it's the compact helical structure is preferentially bound by Rpn14 (Ehlinger et al., 2013). It is hypothesized that insertion of Rpt6 tail into the $\alpha$-ring pocket will unwind the helix resulting in the disengagement of Rpn14 from the assembled complex.

**RP base assembly model**

The key emphasis of this model is the formation of RP base structure by discrete chaperone-assisted base subunit complexes. The assembled base binds to the lid and the bound chaperones are released prior to or during holo 26S formation. The mechanism supports a preassembled ATPase ring base complex from elegant studies of Hochstrasser group at Yale (Tomko et al., 2010). The authors aimed to determine the correct order of the ATPase ring on the assembled proteasome, to represent a large fraction of the 26S particles. An indication was the structure of PAN AAA ATPase, a hexa homologmeric structure that assembles on the heptameric $\alpha$-ring of the archaeabacterial 20S proteasome. The PAN ATPase hexamer is a trimer of dimers. Presence of a proline in an alternate cis-trans configuration within each dimeric pair is proposed to be responsible for the observed trimer or dimer configuration. The proline residue (Pro 91) is present at the junction of C-terminus and N-terminus domain, and the cis-trans pairing allows formation of the coiled-coil domain to fold into register, extending them out and placing the compact C-terminal ATPase domain at the base and in contact with the $\alpha$-ring.
The Hochstrasser group observed that such a proline residue is conserved in three of the ATPase subunits in yeast, and Proline exists as a cis-isomer in three of the ATPase subunits Rpt2, Rpt5 and Rpt3 (c-type, c for cis), and speculated that the registry of the Rpt subunits in eukaryotes may follow the pattern as PAN ATPase complex. Homology modeling, a series of protein engineering experiments and cross linking studies were reported by the group. The three cis-type ATPases were part of three distinct dimeric units and the following dimers were obtained Rpt2-Rpt1; Rpt5-Rpt4 and Rpt3-Rpt6. The disulfide cross linking experiments conducted on assembled proteasome, allowed the authors to position the ATPase in the following order: Rpt1-Rpt2-Rpt6-Rpt3-Rpt4-Rpt5 and Rpt5-Rpt1 closing the ring. These biochemical experiments describing the order of assembly were confirmed by later higher resolution cryoEM studies (Unverdorben et al., 2014).

The various assembly intermediates indicated that prior to assembly, the Rpt subunits are segregated to form different modules with the chaperones demonstrating the right specificity: Rpt2-Rpt1 (Hsm3 module); Rpt5-Rpt4 (Nas2 module) and Rpt3-Rpt6 (Nas6-Rpn14 module). The assembly process was based on key observations: a) Nas2 binds to Rpt4 which binds to Rpt5 and the trimodular complex can be independently isolated from both yeast and mammalian cells (Kaneko et al., 2009); and b) Nas2 is not a stable complex of the Hsm3 module but is part of the Nas6 and Rpn14 modular complexes. As per the assumptions, the dimer pair of Rpt3, part of the Nas2 module and Rpt4 part of the Nas6-Rpn14 module would be trapped. The disulfide cross linking experiments confirmed the presence of an Rpt3-Rpt4 heterodimer which is possible if the two ATPases were brought in register by association of the Nas6 and Nas2 modules in an assembly intermediate. Hsm3 was not obtained in the immunoprecipitation experiments conducted with Flag Nas2, but Nas6 and Rpn14 were observed. The proposed model is depicted in Fig. 3B.

In mammalian cells, same assembly intermediates have been detected, however, the Nas2 module joins last and binds to Nas6-Hsm3 complex to form the RP base complex. The 20S CP is added after the lid has formed the RP base complex (Kaneko et al., 2009).

To summarize, specific interaction of chaperones with the ATPase subunits to
form assembly intermediates, formation of paired dimers of Rpts, arrangement of Rpt pairs as a trimer of dimers that sit atop α-ring, propagation of correct registry, may contribute to chaperone assisted assembly process and ultimate dissociation from the holo 26S complex.

**Unique Role of Nas2 and association with the assembled proteasome**

Role of Nas2 in ATPase assembly are intriguing. The chaperone leaves the assembly intermediates early during the process. Nas2/PSMD9 is a PDZ-like domain containing proteins, and typically PDZ domains recognize the last four to six C-terminal residues in proteins. Deletion of the last three residues and not the last C-terminal residue in Rpt5 abolishes interaction with Nas2 (Lee et al., 2011). The direct involvement of PDZ domain in the interaction with Rpt5 was not demonstrated. Nevertheless the non-requirement of the terminal residues for interaction was observed. Direct experiments to demonstrate the ability of isolated PDZ domain in Nas2 to recognize the specific peptides and the recent structure determination of the hybrid complex between Nas2 and PAN ATPase-Rpt5 chimera were not conclusive (Satoh et al., 2014). For example the Rpt C-terminal peptide does not seem to bind to the isolated PDZ domain as seen by Surface Plasmon Resonance experiments, but binds to the full length protein and the N-terminal 1-120 segment. In NMR experiments where the ligand and the protein are taken at 200 μM in 1:1 ratio, chemical shifts were seen at the tail region of Nas2, a structure not defined by conventional PDZ domain boundaries. Therefore, even with these structural (Singh et al., 2014) and genetic studies, the precise role of PDZ domain of Nas2/PSMD9 in the interaction and assembly remains unanswered. Whether the PDZ domain followed the general rules of PDZ interactions was debatable. Our recent findings on the structure, binding specificity and functions of PSMD9 brings a perspective and provides various alternate avenues. We observed that purified human PSMD9 recognizes C-terminal residues in proteins similar to the typical PDZ domains that recognize a hydrophobic residue such as Phe at the C-terminus of the interacting partners (Sangith et al., 2014). By modelling, docking, simulation and site directed mutagenesis, we established that PDZ domain was specifically interacted with the peptides.
derived from the C-terminal four residues of the proteins (Fig. 4A). These interactions were confirmed using recombinant proteins. Mutagenesis of a single key residue in the PDZ domain and a single amino acid deletion or substitution at the C-terminus of the protein disrupted the interaction. In addition we observed other atypical interactions. For example, in growth hormone (GH), we noted that deletion of the last two residues (ΔGF) within SCGF (Kd of peptide ~10 µM) had no effect on binding. However when Cysteine at P-2 was also deleted (ΔCGF) or mutated to a glycine in the peptide (SC(G)GF), the interaction between mutant GH and PSMD9 was abolished, and the mutant peptide did not inhibit interaction with the full length protein. This mechanism can be extrapolated to explain the above mentioned Nas2-Rpt5 interaction, where deletion of last three residues, FYA in Rpt5, resulted in loss of binding (ΔFYA), and deletion of the last C-terminal residue was not observed. Our studies established that Nas2/PSMD9 carries an atypical PDZ domain which mediates both canonical and non-canonical interactions expanding the scope of its function. The direct interaction with C-terminal tail of the Rpts which insert into the 20S α-ring during the formation of the holocomplex, necessitates removal of the chaperone early in the assembly process. The occupancy in the final assembled complex precludes interaction with 20S proteasome in a competitive manner.

**Beyond Assembly – Functions of the chaperones**

By definition, the chaperones must dissociate upon assembly of the 19S/26S proteasome, with many of the proteins not observed in fully assembled complex or in the cryo EM snap shots. Our group demonstrated unequivocal evidence for the presence of PSMD9 in the 26S holocomplex by several techniques. Mutation of the PDZ domain which disrupted interaction with hnRNPA1 did not affect association with the proteasome, indicating other regions of PSMD9 such as the coiled coil N-domain may be involved interaction with the proteasome. Further support was provided as Nas2 1–120 N-terminal region rescued the growth defect on specific genetic background (Satoh et al., 2014) more strongly than the PDZ domain/C-terminus of the protein. Several groups have indicated that the role of the chaperone is not limited to
preventing premature engagement of the C terminal tail of Rpt to the core particle. The Rpt deletion, rpt5Δ1, prevents tail docking and therefore the assembly, showed a temperature sensitive phenotype when Nas2 was deleted in the Hsm3 background (Lee et al., 2011). The assembly details have been reported in yeast, with minimal information on the functions of the chaperones outside the context of assembly. In the following section we discuss the functions of PSMD9 and PSMD10.
**Functions of PSMD9**

Bridge 1, the rat homolog of PSMD9 interacts with transcription factors E12 and histone acetyl transferase, p300 (Lee *et al.*, 2005; Thomas *et al.*, 1999) via its PDZ-like domain and activates insulin gene transcription. Bridge-1 modulates PDX-1 functions and regulates glucose metabolism (Stanojevic *et al.*, 2005). Overexpression of Bridge-1 increases pancreatic apoptosis with a reduction in the number of insulin-expressing beta cells leading to insulin deficiency and diabetes (Thomas *et al.*, 2009; Volinic *et al.*, 2006). Hence, it is noteworthy that the initial investigations in yeast on the non-proteolytic role of the chaperone, did not support a role in transcription. It is nevertheless unclear of Bridge 1 role in insulin gene transcription is dependent or independent of the proteolytic role of the proteasome.

In ovarian cells, changes in the levels of PSMD9 alters activin signaling (Banz *et al.*, 2010). In human breast carcinoma cell line MCF-7, activin-A stimulation enhances RNA and protein levels of PSMD9/Bridge-1 (Banz-Jansen *et al.*, 2011). The increased expression of Bridge-1 influences activin-A signaling by affecting expression of Smad 2, 3 and 4. In mouse melanoma cells p27/PSMD9 (not be confused with p27/kip, the cell cycle protein) apparently negatively regulates the activity and protein levels of Tyrosinase (Tyr) enzyme and Tyrosinase related protein 1 (Trp1) (Godbole *et al.*, 2006) and may be directly involved in regulation of melanin biosynthesis. A recent review highlights the role of PSMD9 in several diseases including diabetes, mental disorders, neurodegenerative disease and polycystic ovarian syndrome (Gragnoli, 2014; Hopper, 2015).

We established that the PDZ domain of PSMD9 interacts specifically with a variety of C-terminal residues. The P0 position is occupied by a Phe/Cys/Arg/Met (Fig. 4A). Cys and Met rank along with the long aliphatic and aromatic amino acids in hydrophobicity and Arg is routinely used to solubilize inclusions or elute proteins from a hydrophobic column (Sangith *et al.*, 2014). We proposed a broader functional repertoire of the chaperone. We established that PSMD9-hnRNPA1 interaction is a rate limiting edge in IkBα degradation and NF-kB activation. We provided evidence for hnRNPA1 as a shuttle receptor that recruits IkBα for degradation; and PSMD9 as a subunit acceptor that anchors hnRNPA1 to...
facilitate degradation of IkBα by the proteasome (Fig. 4B). We further demonstrated that PSMD9 also interacts with S14, UPF2, soluble IL6 receptor, GH and E12 proteins via the C-termini. Since these proteins are involved in ribosome biogenesis, nonsense mediated decay, signaling and transcription, we speculate that PSMD9 may regulate these processes. The mechanism is likely to involve adaptor protein like response observed with hnRNPA1. It is also likely that the functions may be independent of the whole 26S proteolytic machinery.

Gankyrin the Oncoprotein
Gankyrin/Nas6/PSMD10 acts as an oncoprotein and is overexpressed in many epithelial cancers (Fu et al., 2002; Higashitsuji et al., 2000; Li et al., 2011; Man et al., 2010; Meng et al., 2010; Tang et al., 2010; Zhen et al., 2013). Gankyrin oncoprotein functions in deregulating key signaling networks and/or influences degradation of crucial regulatory molecules by the proteasome. By binding to MDM2, gankyrin facilitates degradation of p53 (Higashitsuji et al., 2005), and on binding to Rb, gankyrin increases Rb phosphorylation and degradation by the proteasome (Higashitsuji et al., 2000) resulting in the release of E2F transcription factor responsible for cellular proliferation. Gankyrin also interacts with CDK4 kinase and cell cycle regulation (Dawson et al., 2002; Krzywda et al., 2004). Interaction of gankyrin with MAGE-A4 suppresses tumor formation in athymic mice overexpressing gankyrin (Nagao et al., 2003). Gankyrin also binds to ankyrin repeats in NF-κB and inhibits its activity (Chen et al., 2007; Higashitsuji et al., 2007). Ras mediated oncogenic signaling is dependent on the presence of gankyrin (Man et al., 2010). NIH3T3 cells overexpressing gankyrin form tumors in nude mice (Higashitsuji et al., 2000). On silencing gankyrin expression, cells undergo reduced proliferation and colony formation on soft agar assay (Man et al., 2010). When gankyrin silenced pancreatic cancer cells were injected into nude mice, the tumors formed were of reduced size. In contrast, when gankyrin was overexpressed, the tumors formed were large in size (Meng et al., 2010). Recently, Gankyrin overexpression in cholangiocarcinoma tissue and cell lines was found to promote metastasis and tumor growth. Gankyrin is an independent prognostic indicator for overall survival in cholangio-carcinoma.
The tumorigenic properties of Gankyrin overexpression are mediated via IL6 and STAT3 pathway (Zheng et al., 2014). These findings suggest that gankyrin connects multiple oncogenic pathways and hence shows characteristics of a key hub protein in a protein interaction network with highly connected nodes. The hub protein interacts with multiple proteins by virtue of different domains or by sharing a common interface. Hub proteins that are overexpressed, are estimated to be three times more essential than the non-hub counterparts (Ekman et al., 2006). Therefore, a network may be deregulated by inhibiting the hub protein or knocking the protein down. Being part of the ubiquitin-proteasome pathway gankyrin is uniquely positioned to link some of these pathways to the ubiquitin proteasome network with a decisive role in disease progression. Hence, gankyrin may be an essential hub protein in cancers when over expressed and the cancer cells may be addicted to their function.

The above observations prompted us to investigate the PSMD10/Nas6/ gankyrin role in carcinogenesis by describing the protein interaction network. To identify new interactions that may be crucial in consolidating the role of an oncogenic hub, we used atomic details of a known interaction between gankyrin-and proteasome ATPase complex. EEVD, a four amino acid linear sequence may form a hot spot site at the interface. Since hot spot sites are conserved in evolution, we searched for other proteins in the human proteome with EEVD in exposed regions. From the PDB database, we predicted 34 novel interactions. Eight proteins were tested and seven of these were found to interact with gankyrin (Fig. 5). Affinity of four interactions is high enough for endogenous detection. Others require gankyrin overexpression in HEK 293 cells or occur endogenously in breast cancer cell line- MDA-MB-435, reflecting lower affinity or presence of a deregulated network. Mutagenesis and peptide inhibition confirm that EEVD is the common hot spot site at the interfaces and therefore a potential polypharmacological drug target. In MDA-MB-231 cells with the endogenous CLIC1 silenced, trans-expression of Wt protein (CLIC1_EEVD) and not the hot spot site mutant (CLIC1_AA VA) resulted in significant rescue of the migratory potential (Nanaware et al., 2014).
studies established the functional significance of the predicted interactions. We believe that other reported interactions are likely to be functionally relevant and will help better understand the many functions of gankyrin in normal and disease condition.

Assembly chaperones as putative targets for intervention
There seems to be an immense scope to target the assembly pathway of proteasome for intervention in diseases (Funakoshi et al., 2009; Gaczynska and Osmulski, 2015). Recently a clinical sample based study indicated that PSMD9 expression may predict radiotherapy benefit in human breast cancers, with low expression indicative of relative radio-sensitivity (Langlands et al., 2014). The study showed a reduction of PSMD9 expression using siRNA in breast cancer cell lines (MCF7 and MDA-MB-231), sensitized the cells to...
radiotherapy and evaluated colony forming assays after irradiation. This is likely due to the mechanisms associated with PSMD9 in NF-kB activation via the PSMD9-hnRNPA1 interaction edge (Sahu et al., 2014). Higher levels of PSMD9 resulted in enhanced degradation of Ik-Bα and therefore enhanced induction of NF-kB when challenged with radiation, rendering cells resistant to therapy and disease relapse. Small molecule inhibitors that target the interaction and perhaps those that stabilize the interaction may be important in dealing with induced NF-kB with a dual role in cancer. The strategy may benefit from the interactions that govern association with the proteasome differ from those that target the hnRNPA1-PDZ domain interaction. The role of PSMD9 in human diseases has been recently reviewed (Hopper et al., 2015).

Gankyrin as an attractive anti-cancer drug target

We proposed that gankyrin over expression results in oncogenesis because a) it results in a rewired network with new nodes (proteins); b) network formed under normal condition are probably long lived leading to deregulated or constitutive aberrant signaling. As the cancer cells are likely to be addicted to gankyrin and its interaction, an inhibitor based on the EEVD motif may be designed to disrupt and collapse the network. This may be a better therapeutic strategy because inhibitors designed against a single target do not work effectively and are subject to resistance by mutagenesis. Some of the key interactions that may be mediated by this 'hot spot site' motif are MDM2 and Rb, master regulators of oncogenesis. MDM2 carries EEND and Rb harbors EEPD. Recently a small molecule inhibitor that interacts with Gankyrin was found to stabilize p53 (Chattopadhyay et al., 2016). Interestingly the molecule occupies an overlapping site occupied by EEVD of the S6 ATPase. Designer proteins that bind to gankyrin and prevent S6ATPase interaction, stabilize p53 have been designed (Chapman and McNaughton, 2015). These results and the relevance of physical interaction in in vitro wound healing and invasion assays hold promise for development of PPI inhibitors of gankyrin interaction.

Conclusion

The review provides a comprehensive literature on the current status of our
understanding of the physiological role of these chaperones. The convergent and the divergent evolutionary aspects of structure and function are other anticipated areas of growth. Several dogmas are likely to be challenged and we are likely to see emergence of new mechanisms and discover novel roles of the chaperones, dependent on or independent of proteasome.

REFERENCES
Benaroudj N, Goldberg AL. PAN, the proteasome-activating nucleotidase from archaeabacteria, is a protein-unfolding
Godbole D, Mojamdar M, Pal JK. Increased level of p27 subunit of proteasomes and its co-


Lee SY, De la Mota-Peynado A, Roelofs J. Loss of Rpt5 protein interactions with the core particle and Nas2 protein causes the formation of faulty proteasomes that are inhibited by Ecm29 protein. *J Biol Chem* 2011;286:36641–36651.


Park S, Kim W, Tian G, Gygi SP, Finley D. Structural defects in the regulatory particle-core particle interface of the proteasome.


Singh CR, Lovell S, Mehzabeen N, Chowdhury


