Toward an Integrated Structural Model of the 26S Proteasome*

Friedrich Förster‡§, Keren Lasker¶, Stephan Nickell‡, Andrej Sali¶¶, and Wolfgang Baumeister‡§§

The 26S proteasome is the end point of the ubiquitin-proteasome pathway and degrades ubiquitylated substrates. It is composed of the 20S core particle (CP), where degradation occurs, and the 19S regulatory particle (RP), which ensures substrate specificity of degradation. Whereas the CP is resolved to atomic resolution, the architecture of the RP is largely unknown. We provide a comprehensive analysis of the current structural knowledge on the RP, including structures of the RP subunits, physical protein-protein interactions, and cryo-electron microscopy data. These data allowed us to compute an atomic model for the CP-AAA-ATPase subcomplex. In addition to this atomic model, further subunits can be mapped approximately, which lets us hypothesize on the substrate path during its degradation. *Molecular & Cellular Proteomics* 9:1666–1677, 2010.

The ubiquitin-proteasome pathway is the major route used by eukaryotic cells for the disposal of misfolded or damaged proteins and for controlling the lifespan of proteins (1–3). As a consequence, the ubiquitin-proteasome pathway regulates a plethora of fundamental cellular processes, such as protein quality control, DNA repair, and signal transduction (4). The 26S proteasome is a molecular machine of ~2.5 MDa that targets polyubiquitylated proteins. It comprises two subcomplexes, the 20S core particle (CP) and one or two copies of asymmetric 19S regulatory particles (RPs), which bind to the end(s) of the barrel-shaped CP.

The active sites of the proteasome are located in the CP cavity where proteolytic cleavage of substrates takes place. Electron microscopy (EM) and x-ray crystallography have revealed that the CP is a cylinder consisting of four concentrically stacked rings (5–7): two identical “α”-rings, each assembled of seven homologous proteins, form the outer rings, and two identical “β”-rings, also assembled of seven homologs, form the two inner rings. Proteolysis is confined to the cavity formed by the β-rings, a nanocompartment sequestered from the cytosol.

The RPs regulate substrate degradation by (i) binding polyubiquitylated substrates, (ii) subsequently deubiquitylating them, (iii) substrate unfolding, and (iv) opening the “gate” to the CP (8). The RPs consists of six AAA-ATPase subunits and at least 13 non-ATPase subunits. In contrast to the CP, the architecture of the RP subunits remains largely unknown. The problems that hamper structural characterization of the RP are manifold. It has proven difficult to obtain homogeneous, concentrated preparations of 26S proteasomes or RPs because the RP tends to disassociate into heterogeneous subcomplexes during purification and concentration. Moreover, many of the RP subunits likely exhibit a significant degree of structural variability. As a consequence, x-ray crystallographic analysis of the entire RP has not been accomplished to date, and only a few subunit fragments have been amenable to high resolution structure determination.

For cryo-EM and protein-protein interaction experiments, the requirements for sample homogeneity are less stringent. Recently, the *Drosophila melanogaster* 26S proteasome was resolved to ~20 Å (9). Various proteomics approaches have led to proposals for topological maps of the RP (10–13). The resolution of protein-protein interaction data typically corresponds to the diameters of the proteins or domains found to interact, which are typically far beyond 20 Å. Because of the limited resolution, neither cryo-EM maps nor protein-protein interaction networks are by themselves sufficient to determine the RP architecture (i.e. the localization of the RP subunits in the complex).

The integration of atomic models, cryo-EM maps, and protein-protein interaction data is currently the most promising approach to resolve the architecture of the 26S proteasome (14–18). Here, we provide a comprehensive analysis of the current structural knowledge on the RP, including structures of RP subunits, physical protein-protein interactions, and...
cryo-EM data. Based on these data, we provide a model for the CP-AAA-ATPase subcomplex. Finally, we outline a path toward resolving the architecture of the 26S proteasome by an integrative structure determination approach, which in turn will provide a basis for a mechanistic understanding.

SUBUNITS OF THE REGULATORY PARTICLE

The RP comprises at least 19 different subunits. In addition to six AAA-ATPases (regulatory particle triple A-ATPases Rpt1–6), the proteasome contains 13 integral regulatory particle non-ATPases, Rpn1–15 (19); the proteins Rpn4 and Rpn14 were erroneously noted as integral subunits but later turned out to be a proteasomal transcription factor and an assembly chaperone, respectively (20–22). All integral subunits are conserved from *Saccharomyces cerevisiae* to mammals. In addition, two ubiquitin C-terminal hydrolases (UCHs), Ubp6 and UCH37, are commonly found in preparations of purified 26S proteasomes (23, 24).

AAA-ATPases—AAA-ATPases possess a 250–300-residue ATP-binding domain (AAA domain) and usually assemble into hexameric rings (Fig. 1) (25). Thus, the proteasomal AAA-ATPases Rpt1–6 are expected to form a heterohexamer. The crystal structures of the major domains of the proteasome-activating nucleotidase (PAN), a homohexameric ortholog of the 26S proteasomal AAA-ATPases found in some archaea (26), have led to insights into the structure and mechanism of the AAA-ATPase modules: each PAN monomer consists of coiled coils protruding from an oligonucleotide binding (OB) fold (PAN-N) and an AAA fold, which have been crystallized separately (27, 28). The PAN-N fragments as well as the AAA folds assemble into hexameric rings (N ring and AAA ring, respectively), but the exact spatial relationship between both rings is not yet known (Fig. 2A).

The order of the AAA-ATPase subunits Rpt1–6 in the heterohexamer was suggested to be Rpt1/2/6/4/5/3 based on cross-linking experiments (29). However, the PAN-N structure challenges this model: only Rpt2, Rpt3, and Rpt5 possess an invariant proline (corresponding to Pro-62 of PAN-N) that is required to adopt a cis conformation for coiled coil formation (27). The PAN-N structure suggests strongly that Rpt2, Rpt3, and Rpt5 each pair with an Rpt subunit lacking the cis-proline. Thus, direct adjacencies of Rpt2, Rpt3, or Rpt5, which occur in the previously suggested order (29), are not consistent with the PAN structures. Upon analysis of all available structural and protein-protein interaction data, we concluded that the subunit order in the AAA-ATPase module is most likely Rpt1/Rpt2/Rpt6/Rpt3/Rpt4/Rpt5 (see below) (30).

PAN and the 26S AAA-ATPases promote unfolding of substrates prior to translocation through the pore formed in the hexamer center (31). The C-terminal residues of the AAA-ATPases are responsible for binding of the RP to the CP (32, 33). The C termini probably bind to grooves at the interfaces of the CP subunits in a fashion similar to that of the 11 S activator (34) as demonstrated for the C termini of PAN (35).
that PCI-containing subunits consist of $\alpha$-helical elements, most likely TPRs and coiled coils, besides the PCI modules. Protein-protein interaction data suggest that the PCI-containing subunits are highly interlinked (see Fig. 4 and Table I) through interactions between their PCI modules (45).

Our functional understanding of the PCI-containing subunits is still modest. Winged helix domains are known to be important for the interaction with nucleic acids (47), suggesting that the PCI-containing subunits are important for the role of the proteasome in DNA repair. Indeed, the peptide Rpn15, which associates with the PCI-containing proteins (see Fig. 4), has been shown to mediate proteasome association to DNA double strand breaks (48). The formation of minicomplexes of Rpn15 and the PCI-containing subunits Rpn3 and Rpn7 may be pivotal for regulation of proteins involved in DNA repair (49).

**Fig. 1. Overview of RP subunits.** PC repeat (orange), coiled coils (white, stripes), TPR-like repeats (magenta), PCI module, MPN domain, Sem1 fold, von Willebrand domain (VWA), UIM, PRU, OB fold, AAA-ATPase domain, UCH domain, and ubiquitin-like domain (Ubl) are shown. Secondary structures ($\alpha$-helical; green, $\beta$-strand) were predicted using the MPI bioinformatics toolkit (109) as a front end to PSIPRED (110). Putative coiled coils with probabilities $>0.5$ according to PCOILS (111) are indicated. TPR motifs were assigned using TPRpred (80). ER, endoplasmic reticulum; PIP, proteasome interacting protein.
Proteomics data indicate that Rpn10 is involved in recognition of a ubiquitin receptor (55). It is possible that these two proteins dimerize via their MPN domain forms a dimer in solution. The physical interaction of Rpn8 does not contain the JAMM motif, and its function is not understood. The MPN domain of Rpn11 contains a zinc-binding JAMM (JAB1/MPN/Mov34 metalloenzyme) motif, which is responsible for its function as a deubiquitylating enzyme (DUB) (50, 51). Rpn11 is required for degradation of ubiquitylated substrates (i.e. it controls substrate degradation positively). Because the deubiquitylating activity of Rpn11 is ATP-dependent, Rpn11 is only active in the assembled RP. Recently, the DUB BRCC36, which is homologous to Rpn11, was identified in the BRCA1/BRCA2-interacting complex (BRCC), a ubiquitin E3 ligase complex that is involved in DNA repair (52, 53). The MPN domain of Rpn8 does not contain the JAMM motif, and its function is not known.

The structure of the Rpn8 MPN domain has been solved by x-ray crystallography (54) (Fig. 2D). Interestingly, the Rpn8 MPN domain forms a dimer in solution. The physical interaction of Rpn8 and Rpn11 is well established (see Fig. 4 and Table I), and it is possible that these two proteins dimerize via their MPN domains, similarly to the Rpn8 crystal structure (54).

Ubiquitin Receptors—Rpn10 was the first proteasome subunit that was shown to function as a ubiquitin receptor (55). Proteomics data indicate that Rpn10 is involved in recognition of ~25% of all proteasomal substrates (56). Rpn10 consists of a von Willebrand fold and a flexible C-terminal ubiquitin interaction motif (UIM) (Fig. 1). Ubiquitin binding is achieved through helices that are connected by flexible linkers and do not possess a distinct tertiary structure (57) (Fig. 2E). Rpn10 seems to be somewhat loosely associated with the RP (58, 59); as a consequence, Rpn10 is detected substoichiometrically in purified proteasomes (9).

Recently, Rpn13 has been shown to be a second integral Ub receptor (60). Rpn13 binds ubiquitin through a pleckstrin-like receptor of ubiquitin (PRU) domain: in contrast to previous structurally characterized UIMs, ubiquitin binds to loops rather than secondary structural elements in PRU (61) (Fig. 2E). In almost all eukaryotes other than S. cerevisiae, Rpn13 possesses a C-terminal extension (Fig. 1), which serves as an anchor of UCH37 (see below). It will now be interesting to reveal which substrates utilize Rpn13 for degradation.

In addition to integral Ub receptors, several “shuttling” Ub receptors, most notably Dsk2 and Rad23, are transiently associated with the proteasome (62). Furthermore, the RP subunits Rpn1, Rpn2, and Rpt5 are candidates for integral proteasomal Ub receptors that await further confirmation (8): Rpn1 has been shown to interact with the ubiquitin-like domains of Rad23 and Dsk2 (63), suggesting that Rpn1 as well as the related Rpn2 might act as Ub receptors. In cross-linking experiments, ubiquitin bound to Rpt5 (64), indicating that Ub binds either directly to Rpt5 or to a subunit in close proximity (see Chemical Cross-linking below).

Associated Deubiquitylating Enzymes—Two additional DUBs are more loosely associated with the RP and are typically found substoichiometrically in preparations of purified 26S proteasomes. The UCHs Uch37 and Usp14/Ubp6 can dissect ubiquitin chains prior to substrate degradation. In contrast to Rpn11, the UCHs control substrate degradation negatively; i.e. deubiquitylation by UCH can save substrate particles from degradation (65). In addition to its deubiquitylating function, Ubp6 was recently found to induce CP gate opening, i.e. to act as a sensor for substrates to be degraded (66).

Whereas Ubp6 is the only proteasome-associated DUB in S. cerevisiae, almost all other eukaryotes possess UCH2/UCH37 as an additional UCH (24). Rpn1 recruits Ubp6 to
the RP (23). UCH37 interacts physically with the ubiquitin receptors Rpn10 (23) and Rpn13 (67); the binding to Rpn13 seems to be more specific than the binding to Rpn10 (67). Rpn13 possesses a C-terminal domain that seems to have evolved for UCH37 binding (Fig. 1) (67). UCH37 is recruited to Rpn13 via its C-terminal domain (67), which contains a coiled coil domain (Fig. 2). Inactivated UCH37 is also part of the nuclear Ino80 complex where Rpn13 can activate it in a hit-and-run fashion (68). It will be interesting to study a possible functional interplay between the RP and the Ino80 complex (69).

Subcomplexes of the RP—The RP easily disassociates into two subcomplexes, a distal “lid” consisting of the non-ATPases Rpn3–12 and the “base” comprising the remaining non-ATPases and the AAA-ATPases (70). Recently, the discovery of numerous smaller subcomplexes has provided significant insights into the assembly process of the RP, such as the scaffold function of Rpn1 and Rpn2 for AAA-ATPase

### Table I
Protein interactions in RP involving six or fewer subunits

Protein-protein interactions were compiled from yeast two-hybrid screens (Y2h), chemical cross-linking (x-link), in vitro binding assays (IVB), different in vivo pulldown experiments (IVPD), and co-expression (co-exp). The data were obtained from 26S proteasomes of *H. sapiens* (h), *S. cerevisiae* (y), *Schizomycetes pombe* (sp), *Sus barbatus* (p), *D. melanogaster* (dm), and *Caenorhabditis elegans* (w).

<table>
<thead>
<tr>
<th>Subcomplex</th>
<th>Technique</th>
<th>Refs.</th>
<th>Species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rpt1, Rpt2</td>
<td>Y2h</td>
<td>29, 76, 77, 92</td>
<td>y, h</td>
</tr>
<tr>
<td>Rpt1, Rpt3</td>
<td>Y2h, x-link</td>
<td>29</td>
<td>h</td>
</tr>
<tr>
<td>Rpt1, Rpt2, Rpt3, Rpt6</td>
<td>IVB</td>
<td>76</td>
<td>h</td>
</tr>
<tr>
<td>Rpt1, Rpt2, Rpt5, Rpn1</td>
<td>IVPD</td>
<td>79</td>
<td>y</td>
</tr>
<tr>
<td>Rpt1, Rpn1</td>
<td>Filter, IVB</td>
<td>20, 77</td>
<td>h</td>
</tr>
<tr>
<td>Rpt1, Rpt2, Rpn1</td>
<td>Filter, IVB</td>
<td>41, 77</td>
<td>h</td>
</tr>
<tr>
<td>Rpt2, Rpt6</td>
<td>x-link</td>
<td>29</td>
<td>h</td>
</tr>
<tr>
<td>Rpt2, Rpn1</td>
<td>IVPD, y2h, IVB</td>
<td>10, 77, 88, 93</td>
<td>h</td>
</tr>
<tr>
<td>Rpt2, α4</td>
<td>IVPD, y2h</td>
<td>87, 88</td>
<td>h</td>
</tr>
<tr>
<td>Rpt2, Rpn12, Rpn2, Rpt1, Rpn13, Rpn8</td>
<td>IVPD</td>
<td>94</td>
<td>y</td>
</tr>
<tr>
<td>Rpt3, Rpt4</td>
<td>Y2h</td>
<td>88, 95, 96</td>
<td>y</td>
</tr>
<tr>
<td>Rpt3, Rpt5</td>
<td>Y2h, x-link</td>
<td>13, 29, 29–95–97</td>
<td>y, h</td>
</tr>
<tr>
<td>Rpt3, Rpt6</td>
<td>Y2h, IVB, IVPD</td>
<td>13, 22, 76, 88, 95</td>
<td>y, h</td>
</tr>
<tr>
<td>Rpt3, Rpn11, Rpt6, Rpn8</td>
<td>IVPD</td>
<td>98</td>
<td>y</td>
</tr>
<tr>
<td>Rpt4, Rpt5</td>
<td>Y2h, IVB, x-link, IVPD</td>
<td>10, 13, 29, 41, 76, 97</td>
<td>w, y, h</td>
</tr>
<tr>
<td>Rpt4, Rpt6</td>
<td>Y2h, x-link</td>
<td>29, 88, 99</td>
<td>y, h</td>
</tr>
<tr>
<td>Rpt4, α4</td>
<td>Y2h</td>
<td>10, 88</td>
<td>y, w</td>
</tr>
<tr>
<td>Rpt4, α6</td>
<td>x-link</td>
<td>29</td>
<td>h</td>
</tr>
<tr>
<td>Rpt6, Rpn1</td>
<td>IVPD</td>
<td>100</td>
<td>p</td>
</tr>
<tr>
<td>Rpt6, Rpn2</td>
<td>x-link</td>
<td>29</td>
<td>h</td>
</tr>
<tr>
<td>Rpn1, Rpn10</td>
<td>IVB</td>
<td>20, 102</td>
<td>y</td>
</tr>
<tr>
<td>Rpn1, Ubp6</td>
<td>IVB</td>
<td>23</td>
<td>sp</td>
</tr>
<tr>
<td>Rpn2, Rpn13</td>
<td>Y2h, IVPD</td>
<td>97, 103, 104</td>
<td>dm, y, sp</td>
</tr>
<tr>
<td>Rpn3, Rpn5</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn3, Rpn5, Rpn8</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn3, Rpn5, Rpn8, Rpn9</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn3, Rpn5, Rpn7, Rpn9, Rpn11</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn3, Rpn7</td>
<td>Y2h</td>
<td>10, 13, 97</td>
<td>y, w, h</td>
</tr>
<tr>
<td>Rpn3, Rpn12</td>
<td>Y2h</td>
<td>13, 88, 105</td>
<td>y</td>
</tr>
<tr>
<td>Rpn3, Rpn15</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn5, Rpn6</td>
<td>Y2h</td>
<td>88, 97, 106</td>
<td>y</td>
</tr>
<tr>
<td>Rpn5, Rpn6, Rpn8</td>
<td>MS/MS</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn5, Rpn6, Rpn8, Rpn9</td>
<td>MS/MS</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn5, Rpn8, Rpn9</td>
<td>MS/MS</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn5, Rpn8</td>
<td>IVPD</td>
<td>107</td>
<td>h</td>
</tr>
<tr>
<td>Rpn5, Rpn9</td>
<td>Y2h</td>
<td>13, 103</td>
<td>y, dm</td>
</tr>
<tr>
<td>Rpn6, Rpn8, Rpn9</td>
<td>MS/MS</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn7, Rpn15</td>
<td>x-link</td>
<td>12</td>
<td>y</td>
</tr>
<tr>
<td>Rpn8, Rpn9</td>
<td>Y2h</td>
<td>10, 88</td>
<td>y</td>
</tr>
<tr>
<td>Rpn8, Rpn11</td>
<td>Y2h</td>
<td>10, 13, 95, 103</td>
<td>y, dm</td>
</tr>
<tr>
<td>Rpn9, Rpn11</td>
<td>Y2h</td>
<td>10, 13</td>
<td>y</td>
</tr>
<tr>
<td>Rpn10, UCH37</td>
<td>IVPD</td>
<td>23</td>
<td>sp</td>
</tr>
<tr>
<td>Rpn12, UCH37</td>
<td>IVB, y2h</td>
<td>108</td>
<td>h</td>
</tr>
<tr>
<td>Rpn13, UCH37</td>
<td>IVPD, IVB</td>
<td>104, 107</td>
<td>sp</td>
</tr>
</tbody>
</table>
Fig. 3. Cryo-EM density map of D. melanogaster 26S proteasome. A, image classification revealed that ~50% of all analyzed particles lacked an additional density (red circle). B, the remaining 50% possessed an additional density (red circle). C, side cut of density depicted in B. D, atomic models of the CP (gold) and the AAA-ATPases (Rpt1/Rpt2/ Rpt6/Rpt3/Rpt4/Rpt5, red/orange/yellow/green/blue/purple) fitted into map B. E, segmentation of B according to the fitted atomic models. F, hybrid representation of the 26S proteasome by atomic models and the cryo-EM density where interpretation on the atomic level is not possible yet. G, enlarged view of F.

assembly (71). The assembly process of the RP and binding of the RP to the CP is facilitated by the assembly chaperones Nas2, Nas6, Hsm3, and Rpn14 in yeast (p27, p28, S5b, and Rpn14 in human) (21, 22, 41, 42).

**CRYO-EM ANALYSIS OF THE 26S PROTEASOME**

Because of the aforementioned difficulties to obtain homogeneous, concentrated 26S proteasome preparation, most information on the structure of the complete RP to date has been revealed by single particle EM analysis. Reconstructions of air-dried and negatively stained 26S proteasomes from D. melanogaster (72) and Homo sapiens (73) revealed that the double-capped proteasome measured ~45 nm in length and that the C2 symmetry of the CP is approximately conserved in the two RPs (Fig. 3, A and B).

More detailed structural insights were obtained from cryo-EM reconstructions of D. melanogaster 26S proteasomes at ~20-Å resolution (9). The higher resolution allowed segmentation of the AAA-ATPase hexamer in the 26S proteasome. The (pseudo-) 6-fold symmetry axes of both ATPase hexamers do not coincide with the (pseudo-) 7-fold symmetry of the CP; the ATPase axes are shifted by ~3 nm. In addition, the axes of the ATPases are tilted by ~10° compared with the CP axis.

Advanced image classification methods revealed that the RPs possess a substantial degree of structural heterogeneity (9). Most significantly, a density of ~60 kDa can be found at one of the two RPs in ~50% of the analyzed particles (Fig. 3, A and B). Quantitative mass spectrometry suggests that Rpn10 is present in only 25% of the RPs, indicating that the variable mass may correspond to Rpn10. However, this hypothesis needs to be corroborated for example by cryo-EM imaging of 26S proteasomes with labeled Rpn10.

**INTRAPROTEASOME INTERACTOME**

**Measuring Protein-Protein Interactions in the RP**

A variety of molecular biology techniques have revealed physical interactions among the RP subunits. Here, we summarize the methods that have been applied to the RP along with the information they provide.

**Two-hybrid Assay**—One candidate protein (“bait”) is fused to a transcription factor DNA binding domain, and another is fused to the activation domain (“prey”), typically using yeast as the host organism. Activation of a reporter gene implies a physical interaction between the two proteins. However, the assay is prone to false negatives as well as false positives (74, 75). To decrease the number of false-positive two-hybrid interactions in the RP interactome, we considered only those interactions that were reported in two or more independent publications (Table I).

**In Vivo Pulldown**—Co-immunoprecipitation or tandem affinity purification can be used to purify complexes that include a bait protein. Subsequent analysis, traditionally by Western blotting or more recently by MS, reveals the identity of affinity-purified complexes. Variation of elution buffers often allows purification of different subcomplexes for the same bait. The experiment indicates physical interactions of the detected subunits, but physical contacts between pairs of proteins typically cannot be directly deduced from the data because subcomplexes often comprise more than two proteins.

**In Vitro Binding Assays**—A bait protein is recombinantly expressed and attached to GST beads. The beads are incubated in cell lysate, again separated from the lysates, and analyzed, typically by Western blotting. Alternatively, the in vitro expressed bait protein may be attached to a nitrocellulose matrix, which is then exposed to cell lysate, washed, and analyzed (76, 77). In vitro binding assays reveal binary physical interactions.

**Chemical Cross-linking**—The protein complex is chemically cross-linked, typically using relatively extended cross-linking agents, such as bis(sulfosuccinimidyl) suberate, that cross-link specific residues (e.g., lysines) in spatial proximity (for a review, see Leitner et al. (112) in this issue). The cross-linked samples are then separated in a denaturing two-dimensional gel and subsequently analyzed by Western blotting or mass spectrometry. The experiment reveals spatial proximity of two
or more proteins. Direct physical interactions cannot be inferred with certainty as the typical cross-linking agents are more than 12 Å long.

Co-expression—Expression of two or more proteins in a heterologous expression system or an in vitro translation system and subsequent purification of a complex of these proteins implies a stable subcomplex of the respective protein.

Mass Spectrometry of Whole Complexes—MS of whole, intact complexes isolated from cells reveals the mass of specific complexes (78). When the subunits are identified by conventional shotgun MS, the stoichiometry of the complex under scrutiny can be determined. In addition, subunits that are peripherally located in the complex can be identified by collision of the complex in the gas phase.

**Interactome Topology**

The proteasomal interaction map clearly reveals two distinct clusters (Fig. 4). The first consists of the PCI- and MPN-containing non-ATPases as well as Rpn15. The second cluster consists of the AAA-ATPases, the PC repeat-containing proteins Rpn1 and Rpn2, the Ub receptors Rpn10 and Rpn13, and the UCHs. This clustering corresponds to the lid and base subcomplex, which can be purified independently (70).

**Base**—Numerous protein-protein interactions have been reported among the AAA-ATPases, strongly supporting the notion that the AAA-ATPases form a hexameric subcomplex. The PC repeat-containing subunits Rpn1 and Rpn2 interact with many of the AAA-ATPases, which is in agreement with their likely function as scaffolding proteins for base assembly (79). The Ub receptors Rpn10 and Rpn13 bind to the PC subunits Rpn1 and Rpn2, respectively. Because UCH37 is probably associated with the Ub receptors in substoichiometric amounts, it cannot be excluded that the UCH37 binds to multiple sites on the RP. Thus, Rpn10 and Rpn13 are not necessarily in spatial proximity.

**Lid**—The PCI subunits form an extensively connected network where only the evolutionary distant Rpn12 is somewhat peripherally bound (12). The peptide Rpn15 binds to Rpn3 and Rpn7. The dimer of the MPN subunits Rpn8 and Rpn11 is integrated into the lid via Rpn5 and Rpn9. The MPN subunits directly interact with the base complex via the AAA-ATPase dimer Rpt3/Rpt6.

**INTEGRATIVE STRUCTURE DETERMINATION OF THE RP**

Integrative methods for structure determination combine input data of different kinds to obtain models of assemblies with a substantially improved accuracy and completeness compared with the individual input data (also see the review by Lasker et al. (113) in this issue). For example, a comprehensive model of the NPC has been obtained based on protein-protein interaction data, the NPC envelope derived by EM, and ultracentrifugation data about the shapes of the individual proteins (17).

**CP-AAA-ATPase Subcomplex**—One option to use protein-protein interaction data for structural research is to discriminate between different candidate models. We have used this strategy to build a model of the subcomplex of the CP and the AAA-ATPases (30). First, we built comparative models of the hexameric N ring and AAA ring based on the PAN crystal structures in all possible subunit orders and fitted the models into the density corresponding to the AAA-ATPases from the 26S density determined by cryo-EM at 20-Å resolution (9).
Indeed, the volume of the density (Rpt1/Rpt2 corresponds to Rpn1 and Rpn10 (Fig. 3, model. The interactome suggests that the density adjacent to can be localized approximately using the CP-AAA-ATPase not bind to additional subunits. (Fig. 3, consists of the variable mass detected by image classification (Fig. 3, E and F). The coiled coils of Rpt4/Rpt5 apparently do not bind to additional subunits.

Taking into account the interactome, additional subunits can be localized approximately using the CP-AAA-ATPase model. The interactome suggests that the density adjacent to Rpt1/Rpt2 corresponds to Rpn1 and Rpn10 (Fig. 3, E and F). Indeed, the volume of the density (~2 × 10^5Å^3) is in good agreement with the expected volume of the two proteins (~2.2 × 10^5Å^3). Moreover, Rpn8 and Rpn11 should be in close proximity to the Rpt6/Rpt3 pair. Thus, substrates of the Rpn10 pathway get recognized at the peripherally located Rpn10 and are subsequently deubiquitylated by Rpn11, which is presumably located in the RP cavity near Rpt6/Rpt3.

Solving the Proteasomal Puzzle—To facilitate the approximate localization of all RP subunits ("molecular architecture") by integration of cryo-EM maps, protein-protein interaction data, and atomic models of subunits, additional data are required. Some subunits can be localized by means of cryo-EM single particle analysis of 26S proteasomes with appropriate genetic tags or deletions of the corresponding genes. Moreover, high throughput data acquisition and image processing will yield higher resolution density maps of the 26S proteasome (14, 82).

An increase in the resolution of protein-protein interaction data will be pivotal for accurate model building. Toward this goal, identification of cross-linked peptides appears to be the most promising technology (Ref. 83; for a review, see Leitner et al. (112) in this issue). Whereas the resolution of the current RP interactome is determined approximately by the diameters of the constituent proteins or, at best, their domains, the resolution of a residue-specific cross-linking restraint is much higher. In addition, the cross-linking/MS technology promises to be less prone to false positives than some established protein-protein interaction methods, such as two-hybrid assays.

Most RP non-ATPase subunits are still structurally poorly characterized (Fig. 1); further structural analysis of these subunits will be required to position them accurately. Given that most structural data available today have only been obtained in the last 5 years, structural coverage can be expected to increase rapidly in the coming years. Structural genomics initiatives may also play a significant role in expanding structural data on the RP subunits as suggested by the recently deposited structure of UCH37 (Protein Data Bank code 3IHR).

Based on our CP-AAA-ATPase model and the RP interactome, we can already hypothesize on the mechanism of the RP in substrate degradation (Fig. 5). Substrates using the Rpn10 pathway are probably recruited to the RP near Rpt1/ Rpt2. The interactome suggests that the Ub receptor Rpn13 is located in proximity to the Rpt4/Rpt5 pair. Thus, Rpn10 and Rpn13 substrates may enter the proteasome through different portals (84). The gate to the CP may then be opened differently for both types of substrates: Rpn10 substrates might use Rpt2 as a "gate opener" to the CP, possibly facilitated by Ubp6 (66), whereas Rpt5 might open the CP gate for Rpn13 substrates. Subsequently, substrates are deubiquitylated at Rpn11, which appears to be centrally located in the RP cavity before being unfolded in the upper AAA-ATPase cavity and translocated to the CP where degradation occurs.

The mathematical framework of "modeling by satisfaction of spatial restraints" allows a systematic integration of all available data to obtain, in principle, all subunit configuration(s) consistent with the input data (18). Specifically, the RP requires a platform that can represent subunits at different
levels of detail, reflecting the varying structural characterization of subunits. The Integrative Modeling Platform (IMP) may provide the required functionality for the exciting endeavor of localizing the RP subunits (for a review, see Lasker et al. (113) in this issue; see also Refs. 18 and 85). However, we emphasize that the quality of any model is ultimately limited by the quality and amount of input data. To obtain a most accurate model, it will be important to remove erroneous protein-protein interaction data. For example, recent in vitro redesign of the CP-AAA-ATPase interface suggests that Rpt5 binds to the α3/α4 CP pocket (86), whereas the data underlying our CP-AAA-ATPase model suggest that Rpt2 binds to α4 (87, 88), and genetic data indicate interaction of Rpt2 with α3 (89). Thus, taking into account the data from Yu et al. (86), two conflicting models are possible where either Rpt2 or Rpt5 interact with the α3/α4 CP pocket. Again, the identification of cross-linked peptides is probably the most promising technology to identify only those protein-protein interactions that occur in the fully assembled complex and in turn filter out erroneous data (for a review, see Leitner et al. (112) in this issue). In addition, the CP-AAA-ATPase model can be validated using cryo-EM of the 26S proteasome with specifically labeled AAA-ATPase subunits.

Elucidation of proteasomal dynamics will undoubtedly be the next challenge after obtaining a draft of the static architecture of the 26S proteasome. In particular, cryo-EM may be uniquely suitable for unraveling the sequence of events during substrate degradation similar to the role of cryo-EM in elucidating ribosomal protein synthesis (90). More detailed insights into the temporal behavior of proteasomal degradation are then to be expected from systematically incorporating such temporal data into the molecular modeling approach (91).

Acknowledgment—We thank Eni Sakata for insightful discussions.

* This work was supported by a 3D Repertoire grant and a PROSPECTS grant within the Research Framework Programs 6 and 7 (FP6 and FP7) of the European Commission, respectively, and by the Deutsche Forschungsgemeinschaft Cluster of Excellence “Munich-Centre for Advanced Photonics.” This work was also supported, in part, by National Institutes of Health Grants R01 GM54762, U54 RR022220, PN2 EY016525, and R01 GM038960 (to A. S.).

§ Supported by a Human Frontier Science Project Organization career development award. To whom correspondence may be addressed. Fax: 49-89-8578-2641; E-mail: foerster@biochem.mpg.de.

** Supported by the Clore Foundation Ph.D. scholars program and carried out in partial fulfillment of the requirements for the Ph.D. degree at Tel Aviv University.

†† Supported by the Sandler Family Supporting Foundation, National Science Foundation Grant IIS-0705196, Ron Conway, Mike Homer, Hewlett-Packard, NetApp, IBM, and Intel.

§§ To whom correspondence may be addressed. Fax: 49-89-8578-2641; E-mail: baumeist@biochem.mpg.de.

REFERENCES


sional arrangement in proteasomes from Thermoplasma acidophilum. EMBO J. 11, 1607–1616


some and the COP9 signalosome. EMBO J. 20, 7096–7107


ology of the cell. Nature 450, 973–982


molecular assemblies. Nature 450, 683–694


(2006) hRpn13/ADR1M1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37. *EMBO J.* 25, 5742–5753


