**Solution Structure of Yeast Rpn9**

**INSIGHTS INTO PROTEASOME LID ASSEMBLY**

Received for publication, November 17, 2014, and in revised form, January 25, 2015. Published, JBC Papers in Press, January 28, 2015, DOI 10.1074/jbc.M114.626762

Yunfei Hu*†1, Yujie Wu‡1, Qianwen Li§1, Wenbo Zhang¶1, and Changwen Jin*†¶1

From the †Beijing Nuclear Magnetic Resonance Center, College of Chemistry and Molecular Engineering, §College of Life Sciences, and ¶Beijing National Laboratory for Molecular Sciences, Peking University, Beijing 100871, China

**Background:** Rpn9 is a subunit of the proteasome regulatory particle.

**Results:** Rpn9 interacts with Rpn10 and Rpn5 via its N-terminal α-solenoid and C-terminal proteasome-COP9/CSN-initiation factor (PCI) domains, respectively.

**Conclusion:** The Rpn9-Rpn5 interaction is contributed by a hydrophobic center and surrounding ionic pairs in their PCI domains.

**Significance:** The results provide structural insights into lid assembly regulation via PCI-PCI interactions.

The regulatory particle (RP) of the 26 S proteasome functions in preparing polyubiquitinated substrates for degradation. The lid complex of the RP contains an Rpn8-Rpn11 heterodimer surrounded by a horseshoe-shaped scaffold formed by six proteasome-COP9/CSN-initiation factor (PCI)-containing subunits. The PCI domains are essential for lid assembly, whereas the detailed molecular mechanisms remain elusive. Recent cryo-EM studies at near-atomic resolution provided invaluable information on the RP architecture in different functional states. Nevertheless, atomic resolution structural information on the RP is still limited, and deeper understanding of RP assembly mechanism requires further studies on the structures and interactions of individual subunits or subcomplexes. Herein we report the high-resolution NMR structures of the PCI-containing subunit Rpn9 from *Saccharomyces cerevisiae*. The 45-kDa protein contains an all-helical N-terminal domain and a C-terminal PCI domain linked via a semiflexible hinge. The N-terminal domain mediates interaction with the ubiquitin receptor Rpn10, whereas the PCI domain mediates interaction with the neighboring PCI subunit Rpn5. The Rpn9-Rpn5 interface highlights two structural motifs on the winged helix module forming a hydrophobic center surrounded by ionic pairs, which is a common pattern for all PCI-PCI interactions in the lid. The results suggest that divergence in surface composition among different PCI pairs may contribute to the modulation of lid assembly.

The ubiquitin-proteasome system is the major pathway for programmed protein degradation in eukaryotic cells and is essential for the regulation of various biological processes, including cell cycle, transcription, protein quality control, and antigen presentation (1–3). Substrate proteins destined to be degraded are tagged with polyubiquitin chains and targeted to the 26 S proteasome, the central machinery that carries out the proteolysis steps.

The 26 S proteasome is a 2.5-MDa multiprotein complex comprising a cylindrical 20 S core particle (CP)3 capped by two 19 S regulatory particles (RP) at the two ends (4, 5). The CP harbors the proteolytic chamber, whereas the RP functions in preparing the substrate protein for proteolysis. The structure of the CP has long been well characterized using x-ray crystallography, revealing four heptameric rings stacking on top of each other (6, 7). The RP comprises six AAA-ATPase subunits (Rpt1–6) and 13 non-ATPase subunits (Rpn1–3, Rpn5–13, and Rpn15) and can be further divided into the lid and base complexes (8, 9). The AAA-ATPase subunits form a heterohexamer ring structure in the RP base and contact the outer ring of the CP, functioning in the unfolding and translocation of substrate proteins across the CP gate (10, 11). The non-ATPase subunits constitute the lid and part of the base and are responsible for substrate recognition, interaction, and deubiquitylation. Due to the intrinsic compositional heterogeneity and conformational dynamics of the RP, crystallizations of the RP or the 26 S proteasome holocomplex have proved challenging. Several studies by cryo-EM have emerged in recent years, pushing the elucidation of the RP architecture to near-atomic resolutions (∼7 Å) and revealing different function-associated conformational states (12–18). However, atomic resolution structural information on the RP is still limited. In order to obtain more detailed understanding of the function and assembly of the RP, structures and interactions of the individual subunits or subcomplexes are required.

Among the non-ATPase subunits, Rpn1 and -2 associate with Rpt1–6 and form part of the base; Rpn10 and Rpn13 are both polyubiquitin receptors and are usually assigned to the

---

* This work was supported by Ministry of Science and Technology of China Grant 2006AA02A323 (to C. J.).

The atomic coordinates and structure factors (codes 2MQW, 2MRI, and 2MR3) have been deposited in the Protein Data Bank (http://wwpdb.org/).

† Both authors contributed equally to this work.

‡ To whom correspondence should be addressed: Beijing Nuclear Magnetic Resonance Center, Peking University, Beijing 100871, China. Tel.: 86-10-6275-6004; Fax: 86-10-6275-3790; E-mail: changwen@pku.edu.cn.

© 2015 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
base complex; and Rpn3, Rpn5–9, and Rpn11 and -12 together form the eight-subunit lid complex (12–13). Based on sequence homology and domain architecture, the lid subunits can be classified into two groups; Rpn8 and Rpn11 both contain an Mpr1/Pad1 N-terminal (MPN) domain (19–22), whereas the other six subunits all share a C-terminal proteasome-COP9/CSN-initiation factor (PCI) domain. The PCI domain is commonly found in three important multiprotein complexes in cells, namely the proteasome, the COP9/CSN signalosome, and the initiation factor elF3, and is proposed to have essential roles in subunit interactions and complex assembly (23). To date, a number of models for the lid assembly have been suggested, although the detailed molecular mechanisms remain elusive (13, 24–28).

Rpn9 is a PCI domain-containing lid subunit and is necessary for the integrity and efficiency of the 26 S proteasome (29, 30). It interacts with the ubiquitin receptor Rpn10, and the Δrpn9 Saccharomyces cerevisiae strain was reported to accumulate multiubiquitinated proteins at restrictive temperatures (29). Herein, we report the solution structures of the 45-kDa S. cerevisiae Rpn9 protein by a high resolution NMR technique. In addition, interactions of Rpn9 with other subunits in the RP are also investigated. We identified two conserved structural motifs on the WH module of the PCI domains, which are responsible for forming the PCI-PCI interacting surfaces. A hydrophobic center surrounded by charged residue pairs is a common feature for all PCI-PCI interacting surfaces. A hydrophobic center surrounded by charged residue pairs is a common feature for all PCI-PCI interactions in the RP, whereas divergence in surface compositions is also present. The results shed new light on the regulation of PCI-mediated lid assembly.

EXPERIMENTAL PROCEDURES

Sample Preparations—The detailed strategies of gene cloning, protein expression, and purification for Rpn9 and Rpn10 samples were reported previously (31, 32). For study of Rpn9-Rpn10 interactions, an Rpn10 construct containing the segment 1–240 was used. For expression of Rpn5, the rpn5 gene was cloned into pET-28a(+) vector with an N- or C-terminal His6 tag and expressed in Escherichia coli BL21(DE3) strain (Novagen). Isopropyl-β-D-thiogalactoside was added to a final concentration of 0.4 mM when A600 reached 0.8, and cells were harvested after a 12–16-h induction at 25 °C. The protein was purified by nickel-nitrilotriacetic acid affinity chromatography followed by gel filtration (Superdex-75) using an ÄKTA FPLC system (GE Healthcare). The N-terminal His6 tag was removed by thrombin cleavage when necessary.

Trypsin Proteolysis—Full-length Rpn9 protein (1 mg/ml) was subjected to trypsin (0.01 mg/ml) digestion at 4 and 25 °C in a buffer containing 50 mM sodium phosphate (pH 7.0) and 50 mM NaCl. The proteolytic reaction was quenched at different time points by adding phenylmethanesulfonyl fluoride to a final concentration of 1 mM, and the samples were analyzed by SDS-PAGE. The main bands on the gels were subjected to N-terminal protein sequencing, and the molecular weights of the main digestion products were determined by MALDI-TOF mass spectrometry.

Size Exclusion Chromatography—All size exclusion chromatography for analyzing protein interactions was performed in a buffer containing 50 mM sodium phosphate (pH 7.0) and 50 mM NaCl using the Superdex 75 column with an ÄKTA FPLC system (GE Healthcare). For detection of Rpn9-Rpn5 protein complex formation, purified Rpn9 and Rpn5 were incubated together at an ~1:1 molar ratio before loading onto the column.

NMR Spectroscopy—All NMR experiments were performed on Bruker Avance 500-, 600-, and 800-MHz spectrometers equipped with cryoprobes. A detailed description of sample conditions and NMR experiments used for the chemical shift assignments of Rpn9 N-terminal domain (Rpn9-NTD) and Rpn9-PCI domain as well as the full-length Rpn9 were reported previously (31). For full-length Rpn9, 2H/13C/15N-labeled samples were prepared, and transverse relaxation optimized spectroscopy (TROSY)-based triple resonance experiments were used for backbone assignments (31). For structure calculations of the Rpn9-NTD and Rpn9-PCI, three-dimensional 15N- and 13C-edited NOESY-heteronuclear single quantum coherence (HSQC) spectra (mixing time, 100 ms) were collected at 25 °C to obtain interproton distance restraints. For structure calculation of full-length Rpn9, three-dimensional 1H-15N- and 13C-edited NOESY-HSQC spectra were collected using non-deuterated C terminus-cleaved Rpn9 samples at 30 °C with 50- and 100-ms mixing times to obtain distance restraints.

For paramagnetic relaxation enhancement experiments of Rpn9-NTD (33), we constructed a T9C mutant. 15N-Labeled Rpn9-NTD-T9C (0.1 mM) was mixed with 0.5 mM 1-oxy-2,2,5,5-tetramethyl-D- pyrroline-3-methylmethanethiosulphonate (MTSL; Toronto Research Chemicals, Inc.) and incubated overnight at room temperature. Excess MTSL was removed by buffer exchange to produce the paramagnetic spin-labeled sample. To obtain the diamagnetic reference spectra, MTSL was reduced by the addition of 2 mM ascorbic acid. HSQC spectra were collected for the paramagnetic and the diamagnetic labeled samples, and the signal intensities were compared.

Backbone H-N residual dipolar couplings (RDCs) were measured for Rpn9-NTD and full-length Rpn9 samples. The RDC measurements for Rpn9-NTD were performed using the Pf1 filamentous bacteriophage (34) or the liquid crystalline phase of G-tetrad DNA (35) as the alignment media, and the RDC values were extracted from the differences in 1H-15N splitting measured by 1H-15N IPAP-HSQC (36). The RDC measurements for full-length Rpn9 were performed using 2H/15N-labeled sample and the Pf1 phage as the alignment medium, and the RDC values were determined using the HNCO/TROSY-HNCO pair of experiments.

Structure Calculations—The structure calculations were performed using the program CYANA (37, 38) based on interproton NOE-derived distance restraints and dihedral angle restraints. The program TALOS was used to predict dihedral angle ψ and ϕ restraints (39). For the full-length Rpn9, the three-dimensional NOE-SY-HSQC spectra were compared with those of the individual NTD and PCI domains to verify that the domain structures were identical, and the corresponding intradomain restraints were directly used in the calculation of the full-length Rpn9 structure. Additional NOEs were further assigned for the hinge region. The initial structures were generated using the CANDID module of CYANA (38), and 20
Characterizations of Rpn9 Architecture—The S. cerevisiae Rpn9 is a 393-residue protein with a molecular mass of 45 kDa. Limited trypsin proteolysis shows the presence of a stable core with an apparent molecular mass of ~20 kDa (band 2 in Fig. 1A). When the trypsin proteolysis reaction is quenched quickly (~1–2 min) at room temperature, two major fragments with molecular masses close to 30 kDa (band 1) and 12 kDa (band 3) are observable. At 4 °C, however, the 30-kDa fragment can be stabilized for a longer period of time (data not shown). N-terminal protein sequencing and mass spectroscopy analyses of the fragments (Table 1) suggest the division of Rpn9 protein into four regions: the NTD, the hinge region, the PCI domain, and the C-terminal tail. To facilitate NMR structure determination and interaction studies, we prepared several protein samples corresponding to different fragments (Fig. 1B). Nearly complete chemical shift assignments for the Rpn9-NTD, Rpn9-PCI, and full-length Rpn9 samples were obtained (31).

Solution Structures of NTD and PCI Domains—The solution structures of the NTD and PCI domains of yeast Rpn9 were determined using conventional NMR methods. The representative structure ensembles and ribbon diagrams of the structures are shown in Fig. 2, and the structural statistics are summarized in Table 2.

The NTD adopts an all-helical fold comprising seven anti-parallel α-helices. The helices α2–α7 form a right-handed α-solenoid, whereas the first helix α1 adopts a different configuration and packs on the α2–α4 side. The unique position of α1 is supported by a network of unique NOE signals and is further verified by paramagnetic relaxation enhancement measurement. Paramagnetic spin labeling at position 9 on helix α1 using an Rpn9-NTD T9C mutant results in signal reduction on α4 but not α3, confirming that α1 is indeed packed on the α2–α4 instead of the α2–α3 side (data not shown). The α-solenoid formed by α2–α7 contains three pairs of double-helix repeats and shows structural similarity to the tetratrico peptide repeats (46) as observed in the structure of Rpn6 (47). However, no sequence homology is present for the N-terminal regions of Rpn9 and Rpn6, and no conserved motif for the helical repeats is identified in the Rpn9-NTD sequence. The neighboring helices in the NTD are connected via relative short loops, whereas the loop connecting α6–α7 is considerably longer, comprising a total of 11 residues (Arg115–Gly125). It is therefore not surprising that in the trypsin digestion assays, this loop is immediately reduced on H9251/H11011/H11011 parallel-solenoid formed by α2–α7. The WH module of the Rpn9-PCI structure with other representative PCI structures (e.g. the proteasome subunits Rpn6 (Drosophila melanogaster) (47) and Rpn12 (Saccharomyces pombe) (48), the signalosome subunit CSN7 (Arabidopsis thaliana) (49), and the initiation factor subunit eIF3K (Homo sapiens) (50)) reveals an essentially similar fold with divergence in the relative depositions between the helix bundle and the WH module (Fig. 3). Rpn9-PCI shows the highest similarity with CSN7, with a Z-score of 13.5 by DALI and a root mean square (r.m.s.) deviation of 3.2 Å for 150 aligned backbone Ca atoms (51). This is in accordance with the fact that the two are encoded by paralogous genes possibly originating from gene duplication (52).

Structure of Full-length Rpn9—Based on the structures of the two individual domains, and by extending the NOE assignments in the NOESY–HSQC spectra of the full-length Rpn9, we
Structure of Yeast Proteasome Subunit Rpn9

TABLE 1
Limited trypsin digestion analysis of S. cerevisiae Rpn9

<table>
<thead>
<tr>
<th>Digestion temperature</th>
<th>SDS-PAGE bands</th>
<th>N-terminal sequencing</th>
<th>M1, by mass spectroscopy</th>
<th>Protein segment</th>
<th>M1, based on sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>4°C</td>
<td>~30</td>
<td>N/A</td>
<td>N/A</td>
<td>116–7</td>
<td>N/A</td>
</tr>
<tr>
<td>25°C</td>
<td>~12</td>
<td>N/A</td>
<td>13,876.6</td>
<td>1–115</td>
<td>13,850</td>
</tr>
<tr>
<td></td>
<td>~18</td>
<td>FKNDF</td>
<td>20,269.8</td>
<td>181–356</td>
<td>20,230</td>
</tr>
</tbody>
</table>

* N/A, not available.

TABLE 2
Structural statistics of S. cerevisiae Rpn9

<table>
<thead>
<tr>
<th>Distance restraints</th>
<th>Rpn9-NTD</th>
<th>Rpn9-PCI</th>
<th>Rpn9-ΔC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total unambiguous NOEs</td>
<td>3549</td>
<td>4340</td>
<td>8125</td>
</tr>
<tr>
<td>Intraresidue</td>
<td>1566</td>
<td>2379</td>
<td>3945</td>
</tr>
<tr>
<td>Sequential (i–j = 1)</td>
<td>496</td>
<td>908</td>
<td>1403</td>
</tr>
<tr>
<td>Medium range (1 &lt; i–j &lt; 5)</td>
<td>440</td>
<td>533</td>
<td>1108</td>
</tr>
<tr>
<td>Long range (i–j &gt; 5)</td>
<td>1047</td>
<td>520</td>
<td>1669</td>
</tr>
<tr>
<td>Total ambiguous NOEs</td>
<td>523</td>
<td>1101</td>
<td>1621</td>
</tr>
<tr>
<td>Dihedral angle restraints</td>
<td>82</td>
<td>107</td>
<td>208</td>
</tr>
<tr>
<td>RDC restraint violations</td>
<td>67</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* The largest distance violations are 0.228, 0.256, and 0.299 Å for Rpn9-NTD, Rpn9-PCI, and Rpn9-ΔC, respectively.

* There are no angle violations for both Rpn9-NTD and Rpn9-PCI. For Rpn9-ΔC, there are two angle violations of 2.6 and 2.4.

FIGURE 2. Solution structures of the NTD and PCI domains of S. cerevisiae Rpn9. A and B, ribbon diagrams and 20 lowest energy structure ensembles of the NTD (A) and PCI (B) domains of yeast Rpn9. Secondary structural elements are labeled and numbered according to the full-length structure.

were able to determine the solution structure of the protein as a whole (Table 2). Inspection of the NOESY-HSQC spectra of full-length Rpn9 (the Rpn9 HI5 construct, as shown in Fig. 1) in combination with chemical shift index analysis showed that the 35-residue C-terminal tail (Ile357–Val393) was unstructured under our experimental conditions. Because the C-tail was prone to degradation, we further prepared a C terminus-cleaved sample (Rpn9-ΔC, as shown in Fig. 1) by keeping the full-length protein at room temperature for a few days and subsequently removing the degraded peptides by gel filtration chromatography. The Rpn9-ΔC sample showed higher stability and spectral quality and was used to collect three-dimensional NOESY-HSQC spectra for structure calculation. Therefore, the Rpn9 structures shown hereafter comprise residues Met1–Arg356.

NOE

TABLE 2
Structural statistics of S. cerevisiae Rpn9

<table>
<thead>
<tr>
<th>Deviations from ideal geometry</th>
<th>Rpn9-NTD</th>
<th>Rpn9-PCI</th>
<th>Rpn9-ΔC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covalent bond lengths (Å)</td>
<td>0.012 ± 0.001</td>
<td>0.012 ± 0.001</td>
<td>0.012 ± 0.001</td>
</tr>
<tr>
<td>Covalent angles (degrees)</td>
<td>1.79 ± 0.03</td>
<td>1.80 ± 0.03</td>
<td>1.80 ± 0.03</td>
</tr>
</tbody>
</table>

* The largest distance violations are 0.228, 0.256, and 0.299 Å for Rpn9-NTD, Rpn9-PCI, and Rpn9-ΔC, respectively.

* There are no angle violations for both Rpn9-NTD and Rpn9-PCI. For Rpn9-ΔC, there are two angle violations of 2.6 and 2.4.

whereas interactions among residues Leu165, Thr168, Asn169, Tyr172, Leu194, Tyr195, and Thr198 bring helix α10 close to α9 (Fig. 4B). Although limited proteolysis divides the Rpn9 sequence into the NTD, hinge, and PCI domains, as shown in Fig. 1, helices α2–α11 actually form a continuous right-handed superhelical solenoid with five pairs of double-helix hairpins, starting from the NTD and extending into the PCI (Fig. 4C). Therefore, the whole protein forms a compact entity, and the N-terminal helical repeats can be viewed as an extension of the PCI domain, as previously suggested for Rpn6 (47). Notably, the Rpn9-NTD construct (residues 1–160) includes the sequence of helix α8. However, helical structures were observed for the three double-helix pairs of α2–α3, α4–α5, and α6–α7, whereas the region corresponding to α8 adopts an unstructured conformation. This suggests that tertiary contact with α9 is required to stabilize the local secondary structure of α8.

Similar to other proteins with helical repeat structures, the trihelical linkage between the NTD and PCI allows a certain...
structural fidelity, with a correlation coefficient of 0.88 for our structure compared with 0.80 for the original model (PDB code 4B4T, chain O) for residues 1–140, as computed using UCSF Chimera (14, 15, 44, 45). The locations of α-helical segments in the two models generally coincide with each other, whereas the positions of α1 and α2 are switched. In our solution structure, all helices are packed anti-parallel, and α1 is the only helix that does not show a right-handed solenoid configuration. This topology is further supported by interaction studies of Rpn9 and Rpn10 and will be discussed below.

The Rpn9-NTD Mediates Interaction with Rpn10—In the structure model of the RP, Rpn9 uses its NTD to interact with the ubiquitin receptor Rpn10. With purified Rpn9 and Rpn10 proteins alone, we were unable to detect tight complex formation by using either pull-down assays or size exclusion chromatography, suggesting that the interaction between the two is not very strong. We therefore performed NMR titration experiments to identify the interaction surface.

When unlabeled Rpn10 was titrated into 15N-labeled Rpn9-NTD sample, we observed a general signal intensity decrease throughout the sequence, which is most probably due to interaction-induced protein aggregation because precipitation was readily visible in the NMR tube. However, certain segments show more profound intensity decrease compared with the average, including Lys36–Glu43 in the α2–α3 loop, Ser67–Val79 in the α4–α5 loop, and residues Glu108, Lys120, and Gly123 in the α6–α7 loop (Fig. 6A). Chemical shift perturbations were also observed, although the changes were slight. Residues showing the most significant peak shifts include Phe69, Ser77, Val79, Lys112, Arg115, and Gly125 (Fig. 6A), all of which locate in the α4–α5 and α6–α7 loops, in accordance with the regions showing an intensity decrease. These observations suggest that Rpn9–Rpn10 interaction may be transient and on the fast-to-intermediate NMR time scale, resulting in line broadening of

An analysis of the surface conservation of Rpn9 is performed and mapped onto the structure using the ConSurf program (Fig. 5D) (55). Three surface regions with high sequence conservation can be identified. Both regions I and II locate in the PCI domain. Region I locates on one side of the WH module formed by helix α18 and strand β2 and contains hydrophobic and negatively charged residues. Region II locates on the other side of the PCI domain, comprising hydrophobic and positively charged residues and forming a groove between the WH module and the α-helical bundle. Region III locates in NTD around the loop connecting helices α2 and α3 and comprises mainly hydrophobic and positively charged residues. In the EM structure of the 26 S proteasome holocomplex, regions I and III are in close contact with the RP subunits Rpn5 and Rpn10, respectively. The conserved region II that forms a large concave surface faces the interior of the RP complex and may be spatially close to the C-terminal helices of Rpn8 (15, 26).

Comparison of our Rpn9 structure with the previous reported cryo-EM-based model (PDB code 4B4T or 4CR2) (15) reveals a topology difference in the NTD, particularly for the first two helices. Both models show good correlation with the EM density, whereas our solution structure shows a slightly improved result. The cross-correlation coefficient, which reports on the model accuracy, is 0.88 for our structure compared with 0.80 for the original model (PDB code 4B4T, chain O) for residues 1–140, as computed using UCSF Chimera (14, 15, 44, 45). The locations of α-helical segments in the two models generally coincide with each other, whereas the positions of α1 and α2 are switched. In our solution structure, all helices are packed anti-parallel, and α1 is the only helix that does not show a right-handed solenoid configuration. This topology is further supported by interaction studies of Rpn9 and Rpn10 and will be discussed below.
most residues on the interaction surface. Segments showing the most significant intensity decrease or chemical shift changes are mapped onto the structure, as shown in Fig. 6B. These residues are clustered on the highly conserved surface region III, as shown in Fig. 5D, and generally coincide with the Rpn9-Rpn10 contacting site based on the fitting of the Rpn9 NMR structure into the EM density (Fig. 6C).

In particular, the $\alpha_2-\alpha_3$ loop of Rpn9 shows the highest sequence conservation and forms a small hydrophobic patch by the tripeptide Leu$^{37}$-Trp$^{38}$-Phe$^{39}$ capped by a highly conserved Lys$^{36}$ on one side (Fig. 6, C and D). Based on this structure model, the side chains of Rpn9-Trp$^{38}$ and Phe$^{39}$ residues possibly interact with Rpn10-Tyr$^{15}$, which is restricted to aromatic residues among Rpn10 proteins from different species. The Rpn9-Lys$^{36}$ is in close proximity to the invariant Rpn10-Asp$^{20}$, suggesting possible involvement of electrostatic interactions. In contrast, the $\alpha_4-\alpha_5$ loop of Rpn9 is less conserved, and the $\alpha_6-\alpha_7$ loop is highly variable. We subsequently prepared three Rpn9-NTD mutants, including two single-site mutants (K36E and F39A) and a K36E/F39A double mutant, the backbone resonances of which were assigned by $^{15}$N-edited NOESY-HSQC spectra. The ability of these mutants to interact with Rpn10 was investigated by two-dimensional $^1$H-$^1$N HSQC spectra using $^{15}$N-labeled Rpn9-NTD mutant samples mixed with unlabeled Rpn10 at a 1:2 molar ratio. All mutants appear to affect the interaction because the interaction-induced precipitation phenomenon was alleviated. Signal intensity reduction was still observable in the three specific regions. Both K36E and F39A mutants caused a slight decrease in reduction level, suggesting a weakened interaction, whereas signal reduction was significantly suppressed when using the K36E/F39A double mutant (Fig. 6E). The results demonstrate the role of the conserved $\alpha_2-\alpha_3$ loop in mediating Rpn9-Rpn10 interaction.

Comparison of our results with previously published EM-based models (14, 18) shows that the position of the segment Lys$^{36}$–Glu$^{43}$ is different between the two. In the original EM-based Rpn9 model, this segment is inserted into the protein structure core and is solvent-inaccessible, which is inconsistent with the NMR titration results, and the helices $\alpha_1-\alpha_2$ are probably misassigned.

The WH Module of the Rpn9-PCI Domain Mediates Interaction with Rpn5—The structure model of the RP reveals that Rpn9 contacts Rpn5 via the PCI domain. Size exclusion chromatography assays demonstrate that the Rpn9 and Rpn5 subunits are able to form a stable heterodimeric complex (Fig. 7A). The strong interactions between the two proteins can be mediated by Rpn9-PCI domain alone, whereas the Rpn9-NTD is not involved (data not shown). By incubating unlabeled Rpn5 with
15N-labeled Rpn9-PCI samples followed by size exclusion chromatography, we were able to obtain a heterodimeric Rpn5-[15N]Rpn9-PCI complex. An overlay of the HSQC spectra of the Rpn9-PCI domain alone and in complex with Rpn5 identifies significant signal disappearance or chemical shift changes clustering in the segment Glu325-Asn345 (Fig. 7E). This segment maps onto the WH module in the Rpn9-PCI domain, in particular the highly conserved helix α18 and strand β2 (Fig. 7C). Notably, this segment corresponds to the conserved region I as depicted in Fig. 5D.

The above result is in good accordance with the Rpn9-Rpn5 binding surface revealed by the RP structure model. Briefly, the helix α18 of Rpn9 WH module docks into a shallow groove of the Rpn5 PCI domain, whereas the strand β2 makes additional contacts on the side. The groove on the Rpn5-PCI surface is mainly formed by the C-terminal tip of its central helix, the β1 strand in its WH module, and the following short helix. The contacting surface has a hydrophobic center comprising residues Met329, Ile332, and Ile341 from Rpn9 and residues Tyr356, Tyr357, and Ile368 from Rpn5. In addition, three pairs of oppositely charged residues are present on the periphery of the contacting surface, namely Arg330(Rpn9)-Glu307(Rpn5), Glu325(Rpn9)-Arg364(Rpn5), and Asp342(Rpn9)-Arg359(Rpn5), suggesting electrostatic contribution to the PCI-PCI interactions. Many of these hydrophobic or charged residues are highly conserved (Fig. 7D).

To gain further information on the contributions of hydrophobic and charged residues in the Rpn9-Rpn5 interactions, single-point mutations, including Rpn9-M329A, Rpn9-I332A, Rpn9-E325K, and Rpn5-R364E, were prepared. Residue Met329 of Rpn9 lies in the hydrophobic center of the contacting surface and is surrounded by three hydrophobic residues, Tyr356, Tyr357, and Ile368, from Rpn5. Alanine substitution of Met329(Rpn9) results in weakening of Rpn9-Rpn5 interaction, as shown by size exclusion chromatography analysis, whereas the mutation of the sideways-pointing residue Ile332(Rpn9) to alanine has no effect (Fig. 7E). The Glu325(Rpn9)-Arg364(Rpn5) ionic pair also locates in the center of the interaction surface, and both residues are highly conserved. Single mutation of either Rpn9-E325K or Rpn5-R364E can fully disrupt Rpn9-Rpn5 interaction, as shown by size exclusion chromatography (Fig. 7E). These observations strongly establish that both hydrophobic and electrostatic interactions are essential for the interactions between the PCI domains of Rpn9 and Rpn5, whereas different residues on the interface have differential contributions.

**DISCUSSION**

The six PCI-containing subunits form a horseshoe-shaped complex via sequential interactions of Rpn9-Rpn5-Rpn6-Rpn7-Rpn3-Rpn12 (Fig. 5A). The contacts between neighboring subunits are mediated by the WH module using an

**FIGURE 5. Fitting of S. cerevisiae Rpn9 structure into EM density.** A, location of Rpn9 in the 7.4 Å cryo-EM density of the S. cerevisiae 26 S proteasome (EMDB code 2165). The 20 S CP is clipped off, and the 19 S RP subunits are colored and labeled. The horseshoe-shaped structure formed by six PCI-containing subunits, Rpn9-Rpn5-Rpn6-Rpn7-Rpn3-Rpn12, is shown on the right. B, detailed view of the Rpn9 fitted into the EM envelope. Density assigned for Rpn9 was segmented from the map. C, comparison of the Rpn9 structure fitted to the EM density (green) with two models from the 20-conformer NMR structure ensemble. The conformer showing the lowest backbone r.m.s. deviation (3.9 Å) with the fitted structure is shown in blue, and the other showing the highest backbone r.m.s. deviation (7.7 Å) with the fitted structure is shown in red. D, surface conservation of Rpn9 with a cyan-white-magenta color gradient representing increasing conservation. The conservation score is calculated using the multiple-sequence alignment of 21 representative sequences from S. cerevisiae, Neurospora crassa, Talaramyces stipitatus, Yarrowia lipolytica, S. pombe, Candida albicans, H. sapiens, Bos taurus, Mus musculus, Rattus norvegicus, Gallus gallus, D. melanogaster, Caenorhabditis elegans, A. thaliana, Physcomitrella patens, Volvax carteri, Naegleria gruberi, Dictyostelium discoideum, Paramecium tetraurelia, Trichomonas vaginalis, and Trypanosoma brucei. The structure of Rpn9 fitted to the EM map is used, and three highly conserved regions are indicated.
interface showing essentially similar characteristics as shown in the Rpn9-Rpn5 interactions (Fig. 8). All interactions involve two structure motifs contributed from neighboring WH modules, which we designate motifs A and B (Fig. 9). Motif A comprises the third α-helix and the second β-strand of the WH module (helix α18 and strand β2 in Rpn9), abbreviated as WH-αIII,βII hereafter. Motif B mainly comprises the second α-helix and first β-strand of the WH module (helix α17 and strand β1 in Rpn9), abbreviated as WH-αII,βI hereafter. The two motifs form a contacting interface with a hydrophobic center surrounded by ionic pairs, which appears to be a common pattern for the PCI subunits in the lid. For example, on the Rpn6-Rpn7 interface, residues Ile381(Rpn6) and Leu382(Rpn6) in the WH-αIII helix (motif A) of Rpn6, together with residues Leu341(Rpn7) and Tyr345(Rpn7) in the WH-αII helix (motif B) of Rpn7, form the hydrophobic center of the interface (Fig. 8). In the recent structural study of Drosophila Rpn6, the equivalent residues Ile369(Rpn6) and Leu370(Rpn6) were identified as essential for interaction with Rpn7 by mutagenesis (47).

Notably, the A and B motifs are located on two opposite sides of the WH module, and surface conservations are generally high for each PCI subunit. However, the Rpn9 and Rpn12 subunits locate on the two distal ends of the horseshoe-shaped
complex, and each has one unengaged motif. Intriguingly, the center of Rpn9 motif B is less conserved compared with others, and the Rpn12 motif A is highly variable. These observations suggest that evolution selectively retains the residues essential for PCI-PCI assembly while allowing random mutations for the unengaged surfaces.

Based on previously reported biochemical and structural data, an ordered self-assembly process has been proposed for the 26 S lid complex (13, 24–26). Briefly, an Rpn5/8/9/11 subassembly is first formed and recruits the Rpn6 subunit. An Rpn3/Rpn7 complex is stabilized with the help of a small protein, Rpn15 (also known as Sem1), and subsequently
incorporated into Rpn5-Rpn6-Rpn8-Rpn9-Rpn11 complex. The assembly is completed by the addition of the Rpn12 subunit. The C-terminal tails of all PCI and MPN subunits were shown to form a helical bundle and govern the assembly process (26), whereas the contributions of the PCI and MPN domains in regulating the assembly are less well understood.

Despite the similar structural motifs for PCI-PCI interactions, different PCI pairs exhibit divergent amino acid composition on the binding interface, which may correlate with the relative different binding affinities (Fig. 8). For example, the Rpn9-Rpn5 interaction surface contains a total of six hydrophobic residues and three pairs of possible ionic contacts, all of which are highly conserved (Fig. 7, C and D). In this study, we observe that the Rpn9-PCI domain alone can mediate strong interaction with Rpn5 interaction without the presence of the C-terminal tail, which is in accordance with a previous reported observation that deletion of the Rpn9 C-helix does not affect its assembly into the lid complex (26). The Rpn6-Rpn7 interaction surface contains four highly conserved hydrophobic residues (Ile^{381}(Rpn6), Leu^{382}(Rpn6), Leu^{341}(Rpn7), and Tyr^{345}(Rpn7)) and one pair of relatively conserved ionic contact (Asp^{291}(Rpn6)-Lys^{346}(Rpn7)). A previous report on Drosophila Rpn6 showed Rpn6-Rpn7 complex formation in pull-down experiments with requirement of the presence of the C-terminal helix (47). For the Rpn5-Rpn6 pair, on the other hand, the hydrophobic area is much smaller, and the interaction appears to be mainly contributed by the highly conserved Arg^{395}(Rpn5)-Glu^{353}(Rpn6) charged pair. No tight complex formation could be detected for the Rpn5-Rpn6 pair either by size exclusion chromatography in our study (data not shown) or pull-down experiments in a previous report on Drosophila Rpn6 (47). We therefore speculate that the differences in
hydrophobicity and charges on the interaction surfaces may help to modulate the binding affinities of the PCI-containing subunits, preventing the premature assembly of unwanted or poisonous subcomplexes and functioning in the regulation of the lid assembly hierarchy.

Acknowledgments—We thank Professor Xianzhong Yan (National Center of Biomedical Analysis, Beijing, China) for help with MS analysis. All NMR experiments were performed at the Beijing NMR Center and the NMR facility of the National Center for Protein Sciences at Peking University.

REFERENCES

tion and related to the COP9-signalsome and elf3. Cell 94, 615–623
some lid. Structure 21, 1624–1635
teasome subunit RPN9 inhibits viral systemic transport and alters plant


