

The Intrinsically Disordered Sem1 Protein Functions as a Molecular Tether during Proteasome Lid Biogenesis

Robert J. Tomko, Jr.¹ and Mark Hochstrasser^{1,*}

¹Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520-8114, USA

*Correspondence: mark.hochstrasser@yale.edu

<http://dx.doi.org/10.1016/j.molcel.2013.12.009>

SUMMARY

The intrinsically disordered yeast protein Sem1 (DSS1 in mammals) participates in multiple protein complexes, including the proteasome, but its role(s) within these complexes is uncertain. We report that Sem1 enforces the ordered incorporation of subunits Rpn3 and Rpn7 into the assembling proteasome lid. Sem1 uses conserved acidic segments separated by a flexible linker to grasp Rpn3 and Rpn7. The same segments are used for protein binding in other complexes, but in the proteasome lid they are uniquely deployed for recognizing separate polypeptides. We engineered TEV protease-cleavage sites into Sem1 to show that the tethering function of Sem1 is important for the biogenesis and integrity of the Rpn3-Sem1-Rpn7 ternary complex but becomes dispensable once the ternary complex incorporates into larger lid precursors. Thus, although Sem1 is a stoichiometric component of the mature proteasome, it has a distinct, chaperone-like function specific to early stages of proteasome assembly.

INTRODUCTION

Most protein degradation within eukaryotic cells is mediated by the 26S proteasome (Tomko and Hochstrasser, 2013). This 2.5 MDa complex has a proteolytic core particle (CP) capped on its ends by the 19-subunit regulatory particle (RP). The RP recognizes and removes the substrate polyubiquitin targeting signal, unfolds the substrate, and translocates it into the CP. The RP can be subdivided into two complexes, the base and lid. The base contains the substrate recognition and unfoldase activities, whereas the lid provides the major deubiquitylating activity.

Three types of subunits constitute the lid: six α -helical proteasome/cyclosome/initiation complex (PCI) domain subunits, Rpn3, Rpn5, Rpn6, Rpn7, Rpn9, and Rpn12; two Mpr1/Pad1, N-terminal (MPN) subunits, Rpn8 and the deubiquitylase Rpn11; and a small, minimally structured protein, Sem1 (DSS1 in mammals). The PCI subunits arrange via their PCI domains into a horseshoe shape in the order Rpn9-Rpn5-Rpn6-Rpn7-

Rpn3-Rpn12 (Lander et al., 2012; Lasker et al., 2012). Their C-terminal α helices form an intricate helical bundle (Estrin et al., 2013; Beck et al., 2012), while their N-terminal α -helical domains extend outward from this horseshoe like fingers. Rpn8 and Rpn11 are cradled in the PCI-domain horseshoe. The exact position of Sem1 within the lid is not certain but appears to be closely associated with Rpn3 and Rpn7 (Bohn et al., 2013; Gudmundsdottir et al., 2007; Tomko and Hochstrasser, 2011; Wei et al., 2008).

Sem1 remains the most enigmatic of the proteasome subunits. It is also a component of additional complexes. Specifically, it is part of the TREX-2 mRNA-export complex (Thp1-Sac3-Sem1-Sus1-Cdc31) and another complex potentially involved in transcriptional regulation, which contains the Csn12 and Thp3 proteins (Faza et al., 2009; Wilmes et al., 2008). In some species (not *S. cerevisiae*), Sem1 orthologs bind the tumor suppressor BRCA2 (Marston et al., 1999). How Sem1 partitions appropriately between these complexes and whether it serves a similar function in all of them are unknown.

Crystal structures of BRCA2-DSS1 and Thp1-Sac3-Sem1 complexes have been determined (Ellisdon et al., 2012; Yang et al., 2002). Sem1 (DSS1) adopts similar binding configurations in both models. In each structure, it is wrapped around a single protein (BRCA2 or Thp1), burying large surface areas. The buried regions consist primarily of two acidic stretches (here called site 1 and site 2) separated by ~ 20 residues in the Sem1 sequence, which we term the linker (Figure 1A). Both sites interact with basic patches on the surface of the Sem1 binding partner; site 2 also bears two conserved tryptophan residues that occupy hydrophobic pockets on BRCA2 or Thp1. In both crystal structures, large stretches of the Sem1 sequence remain disordered, and the portions of Sem1 that are visible show limited secondary structure. Sem1 has been reported to bind proteasome subunits Rpn3 or Rpn7 in isolation (Bohn et al., 2013; Wei et al., 2008); whether Sem1 can simultaneously bind both subunits in assembly intermediates or within full proteasomes is unknown.

Proteasome assembly must occur with high efficiency and fidelity. Both the CP and RP base rely on dedicated assembly chaperones for efficient assembly (Funakoshi et al., 2009; Hirano et al., 2005, 2006; Kaneko et al., 2009; Kusmierczyk et al., 2008; Le Tallec et al., 2007, 2009; Park et al., 2009; Ramos et al., 1998; Roelofs et al., 2009; Saeki et al., 2009). By contrast, no specific extrinsic factors are required for RP lid assembly. Instead, the lid seems to rely on a hierarchical assembly mechanism mediated in part by C-terminal helical extensions from the PCI and MPN

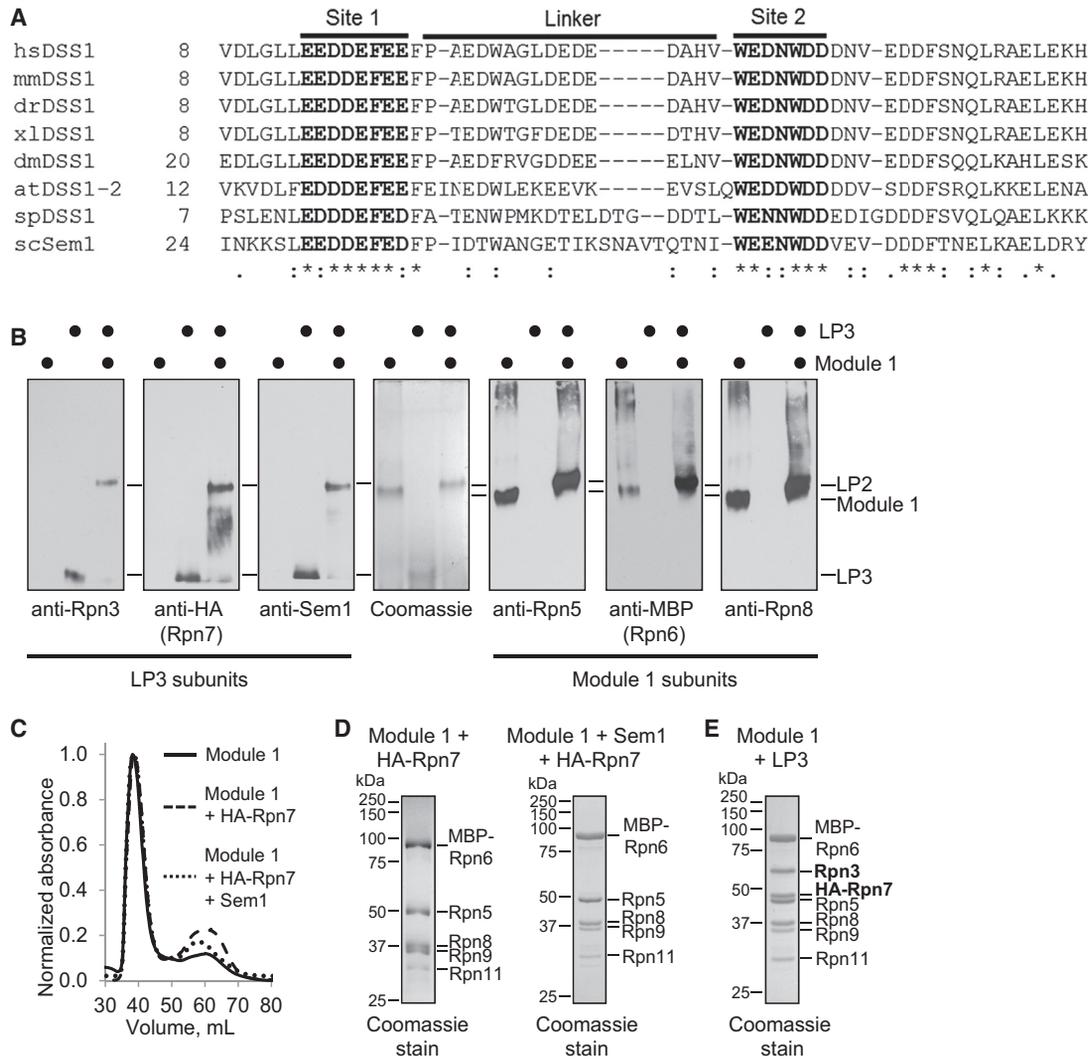


Figure 1. LP3 and Module 1 Are Competent for Assembly

(A) Sequence alignment of Sem1 proteins from the indicated species. Conserved acidic sites 1 and 2 and the poorly conserved linker are indicated.

(B) LP3 and Module 1 form a complex that is indistinguishable from LP2. Purified recombinant LP3 and Module 1 were incubated alone or together for 20 min at 30°C before separation by native PAGE and immunoblotting against the indicated lid subunits or their epitope tags.

(C) HA-Rpn7 does not purify stoichiometrically with Module 1 (Rpn5/6/8/9/11) in the absence of Rpn3. Module 1 subunits were coexpressed with the indicated proteins. Complexes were purified via MBP-Rpn6 and then subjected to Sephacryl S-200 chromatography. Normalized gel filtration traces are shown. Module 1 subunits eluted identically at ~38 ml in all cases.

(D) HA-Rpn7 does not purify stoichiometrically with Module 1 subunits when coexpressed without or with Sem1. SDS-PAGE analysis of fractions from the elution peaks shown in (C) at ~38 ml. HA-Rpn7 was weakly detectable by immunoblot in these fractions (data not shown).

(E) HA-Rpn7 copurifies stoichiometrically with Module 1 subunits when Rpn3 and Sem1 are also present. All lid subunits except Rpn12 were coexpressed in *E. coli*, and complexes were purified via the MBP tag on Rpn6, followed by Superose-6 chromatography. An aliquot of the largest eluted species was resolved on an SDS-PAGE gel and stained with Coomassie Brilliant Blue. No fractions containing a subset of Rpn3, Rpn7, and Sem1 together with Module 1 subunits were recovered (data not shown), supporting the idea that they enter the assembling lid together. See also [Figure S1](#).

domains ([Estrin et al., 2013](#); [Tomko and Hochstrasser, 2011](#)). Formation of Rpn5/Rpn6/Rpn8/Rpn9/Rpn11 and Rpn3/Rpn7/Sem1 complexes (called Module 1 and lid particle 3 [LP3], respectively) appears to occur via parallel pathways ([Estrin et al., 2013](#); [Fukunaga et al., 2010](#); [Tomko and Hochstrasser, 2011](#)). Module 1 and LP3 were proposed to then associate to form the eight-subunit lid particle 2 (LP2) assembly intermediate. Lid assembly culminates with the docking of Rpn12 onto LP2,

which serves as the signal for the subsequent association of the lid and base subcomplexes (and Rpn10) to complete RP assembly ([Tomko and Hochstrasser, 2011](#)). Whether Sem1 contributes to lid assembly is not known; 26S proteasomes purified from *sem1Δ* yeast display no gross structural defects ([Bohn et al., 2013](#)), but mutant *sem1Δ* cells have fewer full proteasomes and accumulate subcomplexes of the lid ([Tomko and Hochstrasser, 2011](#)), suggesting a possible assembly function.

Molecular Cell

Sem1 Chaperones Proteasome Lid Biogenesis

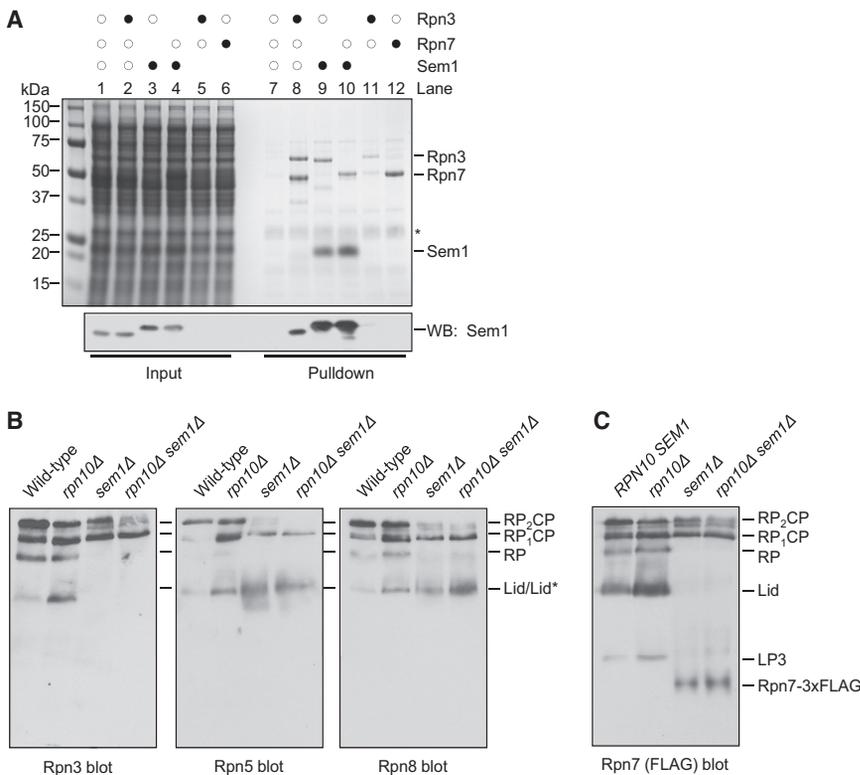


Figure 2. Sem1 Binds Both Rpn3 and Rpn7 and Is Essential for Efficient Lid Assembly In Vivo

(A) Dimeric interaction analysis of LP3 subunits. Dots indicate a subunit was coexpressed; 6His-tagged subunits are marked by filled dots. Rpn3 was largely insoluble unless coexpressed with Sem1. Immunoblotting of the same samples with antibodies against Sem1 is shown at the bottom. (B) Rpn3, but not Rpn5 or Rpn8, is absent from a lid-like particle in *sem1Δ* yeast. Native PAGE immunoblot of extracts from the indicated strains. Lid* indicates a species that migrates similarly to the fully assembled lid but is devoid of Rpn3. (C) Rpn7 is absent from lid*. Immunoblots of the indicated strains as in (B), but each strain contained a chromosomal *RPN7-6xGly-3xFLAG* allele. See also Figures S1 and S2.

Here we describe an assembly factor-like role for Sem1 in proteasome lid biogenesis. We find that Sem1 chaperones proteasomal lid assembly by tethering Rpn3 to Rpn7 during LP3 formation and enforcing their concerted incorporation into the assembling lid. Importantly, this tethering function is no longer required upon formation of LP2. Sem1 also illustrates how a single intrinsically disordered protein can function in disparate protein complexes by deploying the same flexibly linked binding sites in very different ways.

RESULTS

LP3 and Module 1 Can Function as Lid Assembly Intermediates

In yeast, the putative assembly intermediate LP3 had only been observed in lid mutant cells, and it could not be excluded from these data that LP3 is a dead-end complex or a breakdown product of larger species. We therefore sought to directly test whether LP3 can serve as an assembly intermediate by binding Module 1 to form LP2, an established lid precursor. We incubated purified recombinant Module 1 and LP3 (Figure S1A available online) alone or together and then separated the complexes by native polyacrylamide gel electrophoresis (PAGE). Before mixing, Module 1 and LP3 migrated as distinct and slightly diffuse species when evaluated by protein staining (Figure 1B). When the two complexes were mixed, the original complexes nearly disappeared and were replaced by a slower-migrating species that reacted with antibodies to all subunits from LP3 and all tested subunits of Module 1 (Figure 1B). No

additional complexes were evident, suggesting that LP2 forms rapidly and completely from these intermediates via their direct association.

Within the full lid, the Rpn7 PCI domain mediates the majority of contacts of the LP3 subcomplex with Module 1 (Estrin et al., 2013). We investigated whether Rpn7 alone could form a stable complex with Module 1. We coexpressed the five subunits of Module 1 in *E. coli* either alone, with HA-tagged Rpn7, or with both HA-Rpn7 and Sem1. We then purified Module 1 and associated proteins via an N-terminal MBP tag on Rpn6. When expressed alone, Module 1 eluted from a gel filtration column at a position consistent with its predicted size (Figure 1C). Coexpression with HA-Rpn7 or HA-Rpn7 and Sem1 did not affect Module 1 elution. Similarly, HA-Rpn7 did not copurify appreciably with the subunits of Module 1, regardless of whether Sem1 was present (Figure 1D), despite both HA-Rpn7 and Sem1 being abundant and completely soluble (data not shown). When Rpn3 was also coexpressed, Rpn3 and HA-Rpn7 (and Sem1; data not shown) purified stoichiometrically with Module 1 subunits (Figure 1E), forming a complex identical in composition to LP2.

Together, these data strongly suggest that LP3 and Module 1 can serve as assembly intermediates for LP2 and that only the LP3 ternary complex, and not its subunits or subcomplexes, efficiently binds Module 1.

Sem1 Promotes Formation of LP3 In Vitro and In Vivo

To investigate how Sem1, Rpn3, and Rpn7 assemble into LP3, we coexpressed the subunits in various combinations in *E. coli* and tagged individual subunits with a 6His sequence for binding assays using a polyhistidine-binding affinity resin. When Sem1 was fused to 6His, Rpn3 and Rpn7 copurified equally well on the resin (Figure 2A, lanes 9 and 10). Notably, Rpn3 was largely insoluble when expressed alone (data not shown) or with only Rpn7 (Figure 2A, lane 11), and the small pool of soluble Rpn3-6His bound Rpn7 only minimally; correspondingly, very little

Rpn3 copurified with Rpn7-6His (Figure 2A, lanes 11 and 12; Figure S1B). In contrast, Rpn3-6His was soluble when coexpressed with Sem1 and bound Rpn7 stoichiometrically (Figure 2A, lane 8). This indicated that Sem1 stabilizes Rpn3 and promotes LP3 formation. Sem1 did not associate with Rpn9 (data not shown) or Rpn12 when coexpressed alone or with Rpn3 and Rpn7 (Figure S1C). Therefore, Sem1 forms a complex specifically with Rpn3 and Rpn7, and not with other PCI subunits.

We next examined the effect of eliminating Sem1 on lid assembly in vivo. We deleted *SEM1* alone or in combination with *rpn10Δ*. Loss of the Rpn10 RP subunit does not affect lid assembly, but it increases the levels of free lid by reducing lid-base association (Glickman et al., 1998), facilitating detection of lid assembly defects. As expected, *RPN10* deletion increased the levels of free lid in yeast extracts when assayed by native PAGE immunoblotting (Figure 2B), and Rpn12 was absent from the free lid in *sem1Δ* yeast and accumulated as a free subunit (Figure S2A) (Tomko and Hochstrasser, 2011). No base or CP assembly defects were apparent (Figure S2B). However, *SEM1* deletion resulted in loss of detectable Rpn3 in the lid subcomplex and depletion of Rpn3 in doubly capped proteasomes (Figure 2B). The abundance of Module 1 subunits Rpn5 and Rpn8 in this complex (here termed lid*) was unaffected. To test whether Rpn7 was present in lid*, we introduced a sequence encoding a C-terminal 3xFLAG tag at the chromosomal *RPN7* locus. In *sem1Δ* and *rpn10Δ sem1Δ* cells, Rpn7-3xFLAG was reduced in doubly capped 26S proteasomes and was absent from lid* based on native PAGE immunoblotting (Figure 2C). A new species, likely free Rpn7-3xFLAG, accumulated in cells lacking Sem1.

The presence of Rpn5 and Rpn8, but not Rpn3, Rpn7, Rpn12, or Sem1, in lid* suggests that lid* is similar or identical to Module 1. The absence of Rpn3 and Rpn7 in lid* in *sem1Δ* strains implies that both of these subunits rely on Sem1 for stable incorporation into the lid, consistent with our findings in Figure 1. Notably, Rpn3 steady-state levels were modestly reduced in *sem1Δ* cells, and a fraction of Rpn3, likely the extraproteasomal population, was rapidly degraded in *sem1Δ* cells; Rpn7 stability was unaffected (Figure S2C). This suggests that Sem1 helps stabilize the Rpn3 protein in vivo as well.

Sem1 Binds Rpn3 and Rpn7 via Conserved Acidic Sites

In light of the ability of Sem1 to bind independently to Rpn3 and Rpn7, we investigated whether Sem1 binding to the two PCI proteins drives their association. We first determined the elements within Sem1 that recognize Rpn3 and Rpn7. Human Rpn3 contacts Sem1 site 1 (Wei et al., 2008). The site of interaction between Rpn7 and Sem1 is poorly defined, although EM analysis suggests that Rpn7 binds the C-terminal region of Sem1 (Bohn et al., 2013). We confirmed that mutation of the eight residues of site 1 (Figure 1A) to alanines disrupted Sem1 interaction with yeast Rpn3 (Figure 3A, left panel); in contrast, this had no effect on binding to Rpn7 (Figure 3A, right panel). Conversely, mutation of the absolutely conserved Trp60 and Trp64 residues in site 2 disrupted Sem1 binding to Rpn7 while leaving binding to Rpn3 intact (Figure 3A). These data indicate that Sem1 site 1 and site 2 contribute to independent binding sites for Rpn3 and Rpn7.

To determine if these Sem1 elements interact directly with their respective PCI subunits, we produced recombinant LP3 in which Leu29 (abutting site 1) or Trp60 (site 2) in Sem1 was replaced with the photocrosslinkable amino acid *p*-benzoylphenylalanine (Sem1-L29* and Sem1-W60*). Sem1-L29* photocrosslinked only to Rpn3, and Sem1-W60* photocrosslinked only to Rpn7 (Figure 3B). Thus, Sem1 sites 1 and 2 contribute directly to Rpn3 and Rpn7 binding, respectively.

We next tested whether the integrity of both Sem1 sites was required for LP3 formation. When all three subunits were coexpressed in *E. coli*, both Rpn3 and Rpn7 efficiently copurified with WT Sem1-6His (Figure 3C). Mutation of Sem1 site 1 or site 2 as described above abolished formation of the ternary complex (Figure 3C). Therefore, stable association of Rpn3 and Rpn7 does not occur without Sem1 binding to each of them, even when Sem1 (site 2 mutant) maintains Rpn3 in soluble form.

Correspondingly, Sem1 harboring lysine substitutions of four conserved acidic residues in either site 1 or site 2 (*sem1-site1-4K* and *-site2-4K*, respectively) (Figure S3A) failed to suppress the strong growth defects of *rpn4Δ sem1Δ* (Figure 3D) or *rpn10Δ sem1Δ* mutant yeast (Figure S3B), despite protein levels as high as WT Sem1 (Figure S3C). In contrast, substitution of the same four acidic residues in site 2 with uncharged but polar serine residues (*site2-4S*) yielded WT growth (Figure 3D). Mutation of acidic residues in site 1 was less tolerated (Figure S3D). Deletion of *RPN4* or *RPN10* in these experiments genetically linked the observed *sem1*-dependent growth defects to proteasome dysfunction. In agreement, native PAGE immunoblotting of *sem1-site1-4K* and *sem1-site2-4K* cell extracts revealed proteasome assembly defects nearly as severe as those seen with *sem1Δ* (Figure 3E). Taken together, these data strongly suggest that the proteasome assembly defect observed in *sem1Δ* yeast cells derives from a failure to assemble LP3.

Sem1 Tethers Together Rpn3 and Rpn7 within LP3

We considered two models to account for these observations. In the first, Sem1 drives LP3 assembly (and thus lid assembly) by promoting conformational changes in Rpn3 and/or Rpn7 that permit their association with one another and, subsequently, with Module 1 to form LP2. Since Sem1 binds each subunit independently, no physical link between the Rpn3 and Rpn7 binding sites would be necessary in this mechanism. In the second model, Sem1 acts as a molecular clamp, tethering Rpn3 and Rpn7 together within LP3 to promote lid assembly. Here, a physical connection between sites 1 and 2 would be essential.

To distinguish between these models, we first tested whether Sem1 could promote LP3 assembly when split into two fragments that uncouple site 1 and site 2. We coexpressed Rpn3 and Rpn7 with one or both Sem1 fragments, namely, FLAG-sem1(1–51) and *sem1*(52–89)-ZZ-6His, which contain site 1 and site 2, respectively. Equal amounts of Rpn3 and Rpn7 copurified with a full-length FLAG-Sem1-ZZ-6His protein, indicating that the two tags did not affect LP3 formation (Figure 4A, lane 7). Expression of FLAG-sem1(1–51) with Rpn3 and Rpn7 followed by FLAG purification resulted in solubilization and copurification of Rpn3, as expected; by contrast, very little Rpn7 was bound (Figure 4A, lane 8). Conversely, Rpn7 readily copurified with

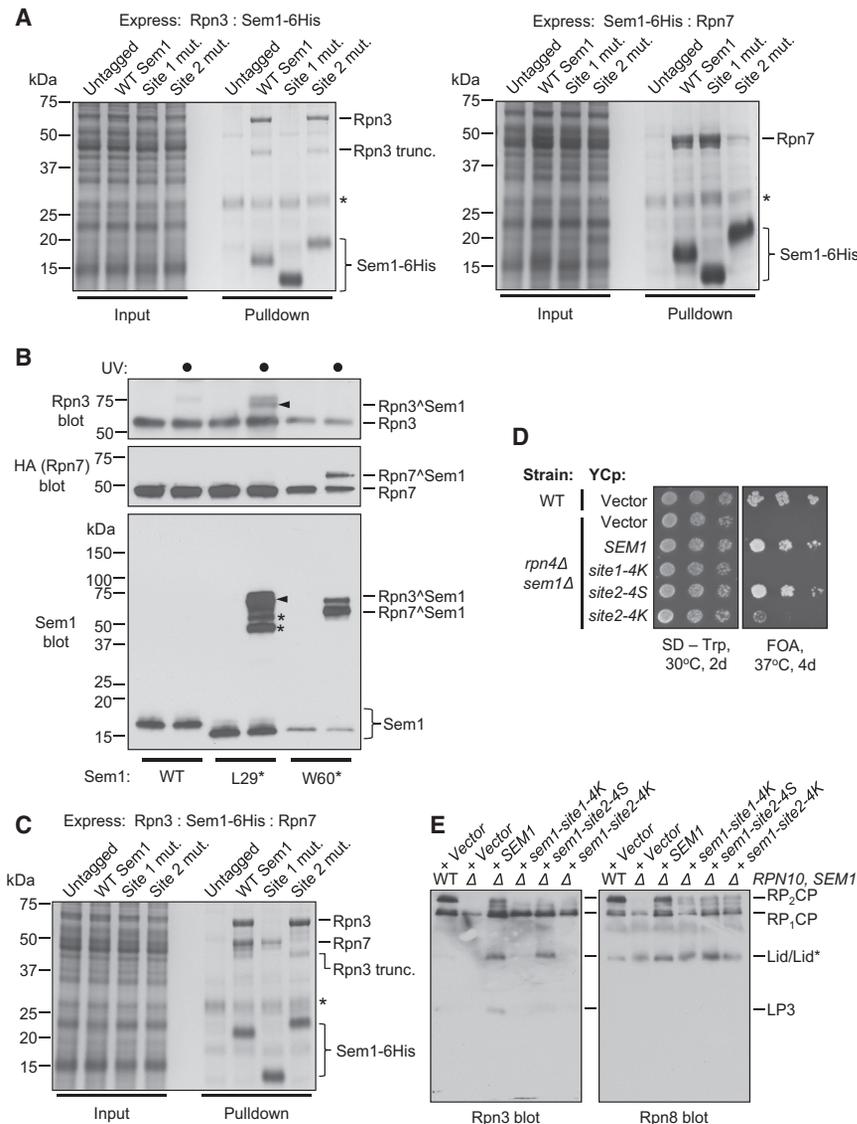


Figure 3. Two Evolutionarily Conserved Sites in Sem1 Are Important for Binding Rpn3 and Rpn7

(A) The indicated Sem1 proteins were coexpressed with Rpn3 (left panel) or Rpn7 (right panel). Sem1 and associated proteins were then isolated via TALON affinity purification. Mutation of residues in Sem1 caused shifts in migration upon SDS-PAGE. Asterisks indicate metal-binding bacterial proteins; Site 1 mut., C-terminally 6His-tagged Sem1 in which residues 30–37 were mutated to alanine; Site 2 mut., C-terminally 6His-tagged Sem1 in which W60 and W64 were mutated to alanine and threonine, respectively.

(B) Site-specific photocrosslinking of Sem1 site 1 and site 2 to Rpn3 and Rpn7, respectively. Sem1-Gly-6His containing *p*-benzoylphenylalanine at the indicated positions and associated proteins were purified on a TALON resin, followed by UV irradiation (indicated by a black dot) to induce crosslinking. The Sem1^{L29*}Rpn3 crosslink is indicated by an arrowhead. Sem1-L29* also appeared to crosslink to Rpn3 truncations (indicated by asterisks in the Sem1 blot).

(C) Both site 1 and site 2 must be intact for LP3 formation. Rpn3, Rpn7, and the indicated forms of Sem1 were coexpressed and purified as in (A). Asterisk indicates bacterial metal-binding protein. (D) The indicated yeast strains were transformed with empty vector or low-copy plasmids encoding the indicated *SEM1* alleles, and spotted in 6-fold serial dilutions on various media and incubated as shown. FOA, 5-fluorouracil.

(E) Native PAGE immunoblots of Rpn3 and Rpn8 in extracts of WT or *rpn10Δ sem1Δ* yeast transformed with the indicated low-copy plasmids. See also Figure S3.

sem1(52–89)-ZZ-6His, but very little Rpn3 copurified (Figure 4A, lane 9). These results indicate that these Sem1 fragments retain their specific binding to Rpn3 and Rpn7, and support the finding that both site 1 and site 2 are necessary for LP3 formation. Importantly, when FLAG-sem1(1–51) and sem1(52–89)-ZZ-6His were both expressed with Rpn3 and Rpn7, again only Rpn3 copurified appreciably with FLAG-sem1(1–51) and only Rpn7 with sem1(52–89)-ZZ-6His (Figure 4A, lane 10). Therefore, Sem1 site 1 and site 2 must be physically linked to one another to promote assembly of LP3.

If Sem1 tethers together Rpn3 and Rpn7 to reinforce their otherwise weak interaction, then overexpressing Rpn3 and Rpn7 in vivo might suppress the *sem1Δ* assembly defect. To test this, we exploited the fact that an *rpn4Δ sem1Δ* double mutant is lethal in the W303 strain background. We introduced *RPN3* and *RPN7* on high-copy plasmids alone, together, or with a plasmid bearing a control lid subunit gene, *RPN5*, into *rpn4Δ sem1Δ* cells kept alive by a WT *SEM1* (*URA3*) plasmid.

The WT *SEM1* plasmid was then evicted from cells grown on 5-fluorouracil, which is toxic to cells expressing *URA3* (Figure 4B). High-copy *RPN3* alone restored viability at 30°C, although growth was much slower than when *SEM1* was present. Both *RPN3* and *RPN7*, but not *RPN5*, rescued lethality at 24°C, but again with much slower growth than seen with reintroduced *SEM1*. These data support the conclusion that Sem1 helps bring or hold together Rpn3 and Rpn7 in vivo, a requirement that is partially overcome by raising the cellular concentrations of these subunits.

Sem1 Requires a Minimal Linker Length between Rpn3 and Rpn7 Binding Sites

If Sem1 tethers Rpn3 and Rpn7 together, then a minimal linker length between the Sem1 site 1 and site 2 would be necessary for their proper positioning on the two PCI subunits. We first tested the effects of shortening or lengthening the linker region (Figure 5A) on proteasome assembly in vivo by generating *sem1* alleles encoding a 20-residue insertion (*sem1-L20*) or linker regions systematically truncated from the middle outward (*sem1-Δ5link*, *sem1-Δ11link*, *sem1-Δ15link*, and *sem1-Δ19link*).

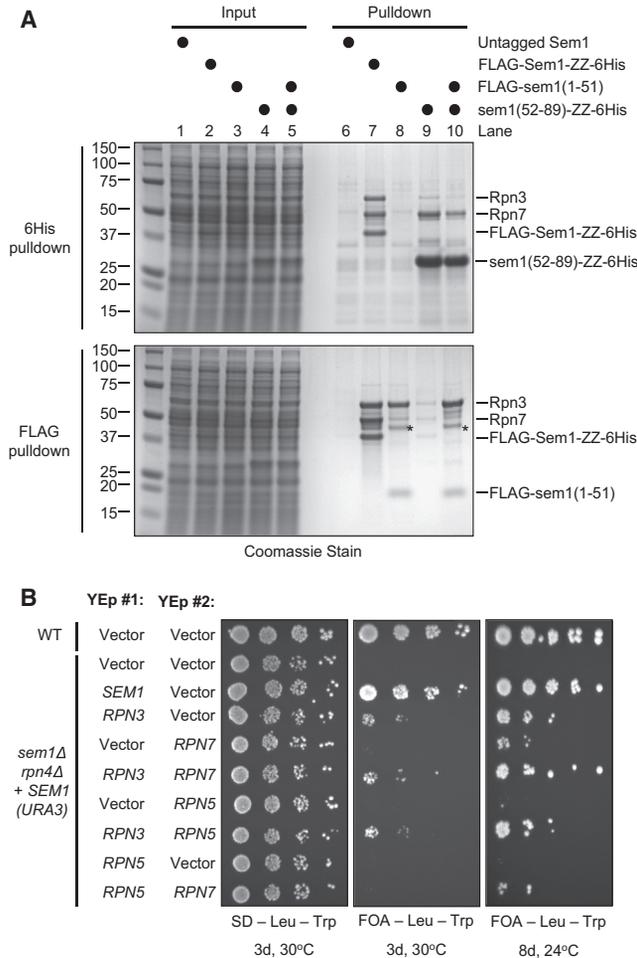


Figure 4. Sem1 Drives Rpn3-Rpn7 Association

(A) The indicated proteins (marked by black dots) were coexpressed in *E. coli*, and FLAG- or 6His-tagged proteins and their binding partners were purified by FLAG immunoprecipitation or TALON resin binding, respectively. Asterisks mark Rpn3 truncation products. We fused *sem1*(52–89) to tandem Z domains of protein A to stabilize it, as the untagged form was rapidly proteolyzed upon bacterial cell lysis (data not shown).

(B) WT or *rpn4Δ sem1Δ* yeast were transformed with empty vector or high-copy plasmids encoding the indicated lid subunits, spotted in 6-fold serial dilutions, and incubated as indicated.

We introduced these alleles into *rpn10Δ sem1Δ* yeast and investigated proteasome formation by native PAGE immunoblotting. The mutant proteins were expressed at levels equal to or near WT Sem1 levels (Figure S4A).

Yeast expressing *sem1-Δ5link* displayed levels of proteasomes and lid subcomplexes equivalent to cells expressing WT SEM1, and they incorporated Rpn3 into the lid and LP3 (Figure 5B). However, yeast expressing *sem1-Δ11link* or *sem1-Δ15link* had decreased levels of proteasomes compared to yeast with WT SEM1, and they did not accumulate detectable levels of LP3. Yeast expressing *sem1-Δ19link* were severely depleted for 26S proteasomes and free lid and had no detectable LP3 (Figure 5B). These *sem1* alleles were then evaluated for their ability to rescue viability of *rpn4Δ sem1Δ* cells, using the

plasmid-shuffle assay described above. All the linker mutants supported viability at 30°C at near-WT growth rates (Figure 5C). At 37°C, *sem1-Δ5link* transformants also grew at WT levels. By contrast, the *sem1-Δ11link* or *Δ15link* transformants displayed a weak but reproducible growth defect compared to those with SEM1, while the *sem1-Δ19link* transformants grew very poorly.

The absence of detectable LP3 in yeast with the more severe Sem1 linker deletions suggested that a minimal distance was necessary for Sem1 to simultaneously bind Rpn3 and Rpn7. To test this, we measured Sem1-dependent copurification of Rpn7 with Rpn3-6His. Assays were performed with extracts made from *E. coli* that coexpressed these proteins along with Sem1 or Sem1 linker deletions (Figure 5D). All deletion constructs bound to Rpn3 based on their ability to maintain Rpn3 solubility. In full agreement with LP3 formation in vivo, *sem1-Δ5link* supported LP3 formation as well as WT Sem1 did. By contrast, *sem1-Δ11link* and *sem1-Δ15link* showed strongly reduced Rpn7-Rpn3-6His association, and *sem1-Δ19link* was nonfunctional.

Lengthening the Sem1 linker by as many as 20 residues fully supported yeast growth even at elevated temperatures (Figure 5C) and supported LP3 assembly in *E. coli* (described below). Simultaneous mutation of nearly every amino acid in the linker did not appreciably affect cell growth, even under proteasome stress conditions (Figures S4B and S4C). This argues against a sequence-dependent function for the linker, consistent with its poor conservation (Figure 1A). In support of a minimal linker length of at least ten residues for full Sem1 activity in LP3 assembly, a comparison of highly divergent species representing all five recognized eukaryotic supergroups showed that the length of the Sem1 linker was at least 11 residues in all species (Figure S4D).

The variability of the linker sequence, together with the requirement for a minimal length for efficient LP3 assembly, leads us to conclude that the linker functions primarily as a flexible spacer between Sem1 site 1 and site 2 that permits their optimal binding to Rpn3 and Rpn7.

Sem1 Tether Function Is Required for Maintaining LP3 Integrity

We next determined whether Rpn3-Rpn7 tethering by Sem1 was required for proteasome assembly after initial formation of LP3. We engineered Sem1 proteins to contain TEV protease (TEVp) cleavage sites at one or more positions within the Sem1 linker region (Figure S4B). Unexpectedly, TEVp did not cleave any of these Sem1 proteins in purified LP3 or 26S proteasomes despite fully cleaving free Sem1 protein (data not shown). To increase TEV site accessibility, we introduced it into the middle of the Sem1-L20 linker extension (*sem1-L20-TEVx*) (Figure S5A). Sem1-L20 or Sem1-L20-TEVx assembled into LP3 (Figure S5B), and *rpn10Δ sem1Δ* yeast harboring either allele had no discernible growth defect compared to the same mutant expressing SEM1 (Figure S5C).

We followed the copurification of Rpn3 with HA-Rpn7 following TEVp-mediated scission of Sem1 in isolated LP3. Sem1-L20-TEVx was completely cleaved, as indicated by loss of full-length protein and appearance of a smaller species with a mass expected for the N-terminal cleavage fragment

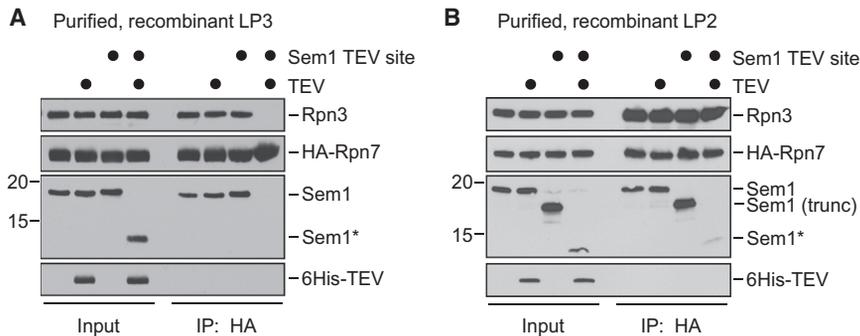


Figure 6. Sem1 Serves as a Molecular Clamp during Lid Assembly

(A) Site 1 and site 2 must remain covalently tethered for the integrity of LP3. Purified LP3 containing Sem1 with the L20 linker extension with or without an engineered TEV protease cleavage site was incubated with TEV protease before immunoprecipitation of HA-Rpn7 and associated proteins. Black dots indicate the presence of a component. Sem1* indicates the N-terminal cleavage fragment of Sem1-L20-TEVx, which is lost during HA immunoprecipitation, presumably because it remains bound to Rpn3.

(B) The tethering function of Sem1 is dispensable in the context of LP2. As in (A), but with recombinant

purified LP2. In the context of recombinant LP2, the Sem1-L20-TEVx protein appears to be partially proteolyzed (Sem1-trunc) compared to Sem1-L20, resulting in a more rapid migration. A weak full-length band is apparent in these lanes. Note that the Sem1* N-terminal cleavage fragment copurifies with HA-Rpn7 in LP2, but not in LP3 (A), presumably because it is bound to Rpn3. See also [Figure S5](#).

Sem1 Functions as a Molecular Clamp during Lid Assembly

Our finding that Sem1 drives the association of Rpn3 and Rpn7 and, by extension, their ordered incorporation into the assembling lid fits well with our previously proposed model of a hierarchical lid assembly pathway (Tomko and Hochstrasser, 2011). In this model, intrinsic properties of lid subunits determine their ordered, stepwise assembly via avid interactions among multiple subunits or conformational changes caused by subunit binding (or both). Although Rpn7 makes substantial direct contact with Rpn6 in the fully assembled lid (Lander et al., 2012; Lasker et al., 2012), the lack of stable binding of Rpn7 to Module 1 in the absence of Rpn3 suggests that these contacts are either insufficient or unavailable. The C-terminal α helices of Rpn3 and Rpn7 are important for their incorporation into the assembling lid, even though both are dispensable for Rpn3-Rpn7 association (our unpublished data; Estrin et al., 2013). Within the lid helical bundle, the Rpn7 helix occupies a peripheral position in which it primarily contacts the Rpn3 helix (Estrin et al., 2013). The Rpn3 helix in turn makes extensive contacts with the Rpn8 C-terminal helix. Thus, in the absence of Rpn3, the interaction between the Rpn6 and Rpn7 PCI domains may not be sufficient to anchor Rpn7 within the assembling lid, and likewise, the interaction between the C-terminal helices of Rpn3 and Rpn8 may not be sufficient in the absence of Rpn7. Thus, the simultaneous entry of Rpn7 and Rpn3 imposed by Sem1 may overcome the weak affinity of either subunit in isolation for Module 1. These additional contacts with Rpn8 may further stabilize the interface between Rpn3 and Rpn7, rendering the tethering function of Sem1 dispensable in LP2 and lid-containing complexes.

Functional Diversity of Sem1 in Multiprotein Complexes

Our study provides a framework to evaluate the function of Sem1 in proteasome assembly compared to its participation in the BRCA2 and Sac3-Thp1 complexes. One emerging role of Sem1 is the stabilization of its binding partners. Rpn3 requires Sem1 for its folding stability or solubility when expressed in bacteria (Figure 2A), and a subpopulation of Rpn3 is rapidly degraded in *sem1* Δ yeast (Figure S2C). Similarly, both Thp1 and BRCA2 depend on Sem1 coexpression for their solubility

in heterologous expression systems (Ellisdon et al., 2012; Yang et al., 2002). Depletion of Sem1 (DSS1) by siRNA in human cells dramatically decreases the half-life of BRCA2 (Li et al., 2006). We propose that a general function of Sem1 is to stabilize its binding partners, although the mechanism remains obscure. In yeast, a subpopulation of Rpn3 in *sem1* Δ yeast appears metabolically stable (Figure S2C). We imagine that any Rpn3 that successfully incorporates into the lid in *sem1* Δ cells is stabilized due to binding by adjacent subunits.

Rpn3-Rpn7 tethering by Sem1 requires physical linkage between Sem1 site 1 and site 2, and the linker must be ≥ 10 residues in length to function normally in proteasome biogenesis (Figures 4A and 6A). Both site 1 and site 2 also mediate Sem1 binding to Thp1 and BRCA2, whereas much of the linker region is disordered (Ellisdon et al., 2012; Yang et al., 2002). In the Thp1-Sem1-Sac3 and BRCA2-Sem1(-ssDNA) crystal structures, the unresolved 11–14 residues of the linker need to bridge only ~ 18 – 22 Å to connect the resolved segments of Sem1. This is much less than the theoretical ~ 39 Å length of 11 residues in extended conformation. Therefore, it is the proteasome that probably imposes the longer Sem1 linker-length requirement.

In yeast, the proteasome lid, TREX-2, and Csn12-Sem1-Thp3 complexes each contain one evolutionarily conserved PAM (PCI domain-associated module) protein (Rpn3, Thp1, and Csn12, respectively) and one atypical PCI (aPCI) domain protein (Rpn12, Sac3, and Thp3). Based on this, Sem1 was originally proposed to bind PAM-aPCI pairs (Faza et al., 2009; Wilmes et al., 2008). We show instead that Sem1 forms a stoichiometric complex with Rpn3 and Rpn7, a standard PCI subunit. In further opposition to the PAM-aPCI-binding model, Rpn12 failed to join Sem1 complexes, even when coexpressed with both Rpn3 and Rpn7 (Figure S1C). We propose that the critical shared feature of these complexes for Sem1 binding is the PAM subunit rather than a PAM-aPCI subunit pair.

A Role for Intrinsic Protein Disorder in Protein Complex Assembly

Sem1 is predicted to be an intrinsically disordered protein (IDP) (<http://dis.embl.de/>) and displays several characteristics of IDPs (Dyson and Wright, 2005), including a very hydrophilic sequence, little secondary and tertiary structure along with

Molecular Cell

Sem1 Chaperones Proteasome Lid Biogenesis

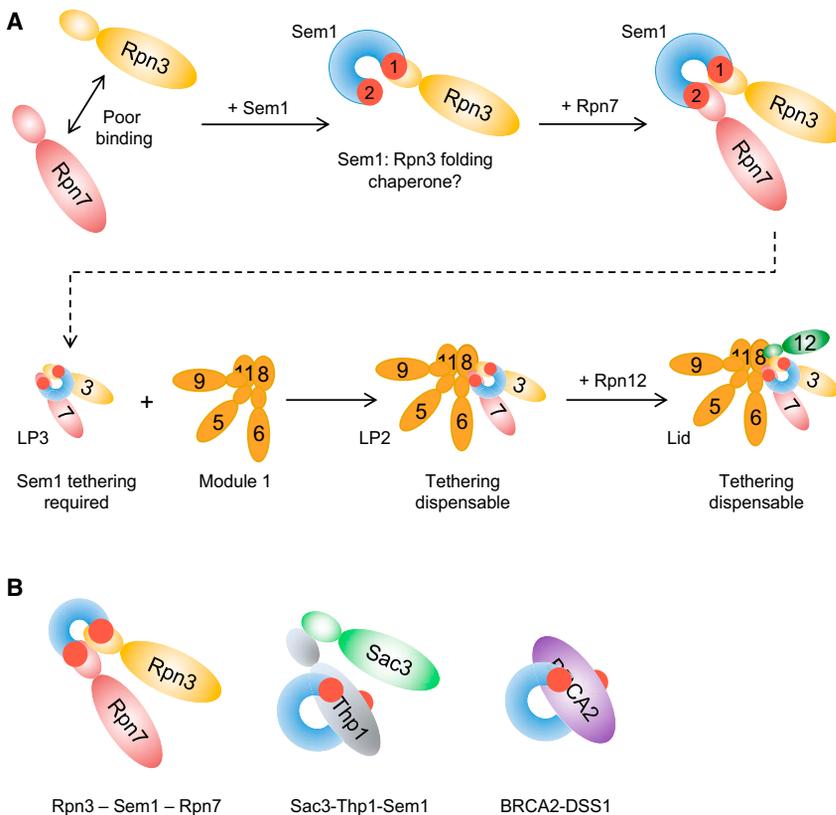


Figure 7. Model for The Role of Sem1 in Proteasomal Lid Assembly

(A) Rpn3 and Rpn7 have poor affinity for one another in the absence of Sem1. Red circles on Sem1 indicate acidic sites 1 and 2 as shown. Sem1 is depicted as first binding Rpn3 and then Rpn7, but the reverse is also possible. We hypothesize that Sem1 associates first with Rpn3 because Rpn3 requires Sem1 for stability when produced in *E. coli* and probably also in yeast. Sem1 tethers Rpn3 and Rpn7 together in LP3 until it associates with Module 1, forming LP2. The tethering role of Sem1 is dispensable in LP2 and in the fully formed lid and 26S proteasome.

(B) Cartoon schematic comparing known binding configurations of Sem1 within protein complexes. In the lid, Sem1 recognizes a distinct protein (Rpn3 or Rpn7) with each acidic patch. In the crystal structures of Sac3-Thp1-Sem1 (Ellisdon et al., 2012) and BRCA2-DSS1 (Yang et al., 2002), the two acidic patches recognize a single protein (Thp1 or BRCA2, respectively). Although both the lid and the Sac3-Thp1 complexes consist of PCI subunits, Sem1 adopts distinct binding configurations in each of them.

regions of disorder in crystal structures, and in our hands a large Stokes radius and high sensitivity to proteolysis in isolation (Figure S1D). An important contribution from protein disorder has been suggested for the function of many protein and RNA folding chaperones (Kovacs et al., 2013). Analogously, Sem1 disorder may be important for its ability to stabilize Rpn3. We expect that it will also contribute to the tethering function of Sem1 in proteasome assembly. Intriguingly, the dedicated CP chaperone Ump1 has also recently been shown to be an intrinsically disordered protein (Sá-Moura et al., 2013; Uekusa et al., 2013). Whether Ump1 serves a function similar to Sem1 during CP assembly remains to be tested.

The configuration of functional domains within Sem1 recalls that of the kinase inhibitory protein (KIP) family of IDPs, which function in part to promote the assembly of cyclin-dependent kinase (Cdk)-cyclin complexes (Cheng et al., 1999). KIP proteins, specifically p21 and p27, contain an N-terminal kinase inhibitory domain consisting of Cdk- and cyclin-binding sites separated by a spacer element (Yoon et al., 2012). The flexibility of this spacer allows a KIP protein to reach and recognize Cdk and cyclin binding sites that are separated by different distances in distinct Cdk-cyclin pairs (Wang et al., 2011). Similarly, Sem1 linker flexibility may enable its two key acidic binding sites to recognize the appropriate surfaces on its multiple binding partners.

The mechanism by which KIP proteins drive Cdk-cyclin assembly has not been explicitly demonstrated, but it is likely that these domains must be physically connected and sufficiently spaced for function, analogous to our results with Sem1 and LP3 assembly. In support of this, shortening of the p21

spacer by three amino acids results in Cdk-cyclin complexes with reduced thermal stability in vitro and reduced Cdk-cyclin complex formation in vivo (Wang et al., 2011). Indeed, other proteins with

important biomedical implications, such as the oncoprotein BRCA1 (Mark et al., 2005) and the tumor suppressor axin (Noutsou et al., 2011), contain protein- or DNA-binding sites interspersed with disordered regions; these motifs may also function to drive assembly of multicomponent complexes via disorder-dependent mechanisms.

EXPERIMENTAL PROCEDURES

Yeast Strains and Media

Yeast manipulations were carried out according to standard protocols. Strains used in this study are listed in Table S1.

Plasmids

Routine cloning was performed in *E. coli* strain TOP10 F'. Plasmids used in this study are listed in Table S2. Yeast genes were amplified by PCR using *S. cerevisiae* genomic DNA as template and included sequences extending 500 bp upstream and downstream of the start and stop codons, respectively. Yeast SEM1 plasmids contained a 1.3 kb HindIII fragment from *S. cerevisiae* genomic DNA containing the SEM1 coding sequence. Generation of multigene operons for bacterial expression was done as described previously (Kusmierczyk et al., 2008). All lid subunit mutations were first introduced into host plasmids containing individual subunit expression cassettes, confirmed by DNA sequencing, and then transferred by subcloning into the appropriate operons. Point mutations were introduced by QuikChange mutagenesis (Stratagene). Sem1 linker-length mutants and intein-tagged subunit mutants were made by recombineering with bacterial strain MC1061 as host.

Native PAGE Immunoblot Analysis

Yeast cell extracts were prepared essentially as described (Tomko and Hochstrasser, 2011). Mid to late log phase cells (OD₆₀₀ = 1.5–2.0) grown in YPD or

the appropriate selective minimal medium were washed with ice-cold water and frozen in liquid nitrogen. The frozen cells were ground with mortar and pestle, and the resulting cell powder was thawed in 26S buffer (50 mM Tris·HCl [pH 7.5], 5 mM MgCl₂, 10% glycerol, 1 mM ATP). Extracts were centrifuged for 10 min at 21,000 × g to remove cell debris. Protein concentrations were determined by BCA assay (Pierce). Native PAGE was performed at 100 V at 4°C on 50 µg of protein per sample. Purified proteins and proteasomal subcomplexes were diluted into 26S buffer and separated as described for yeast extracts. Native PAGE-separated proteins were then transferred to PVDF membranes and subjected to immunoblot analysis.

Protein Expression and Purification

BL21-STAR (DE3) cells (Invitrogen) containing pRARE2 or pRARE2LysS (Novagen) were used for all bacterial protein expression experiments except for expression of proteins containing *p*-benzoylphenylalanine (see below). Cells were grown to mid-log phase in LB medium and the appropriate antibiotics before induction with 0.5 mM IPTG for 4 hr at 30°C or overnight at 16°C. Copurification assays were performed with TALON polyhistidine affinity resin (Clontech). Cells were lysed in TALON buffer (50 mM Tris·HCl [pH 7.5], 500 mM NaCl, 0.2% Tween-20, 10% glycerol, 0.5 mM TCEP) containing protease inhibitors using an M-110EH microfluidizer (Microfluidics Corp.) and cleared by centrifugation for 20 min at 30,000 × g at 4°C. Supernatants were bound to TALON resin, washed extensively in buffer HA (50 mM HEPES·NaOH [pH 7.5], 150 mM NaCl, 5 mM MgCl₂, 10% glycerol, 0.5 mM TCEP) supplemented with 10 mM imidazole, and eluted with buffer HA containing 150 mM imidazole. LP3 for protease cleavage experiments was purified on TALON resin as above and concentrated by centrifugal filtration using a 100 kDa cutoff filter; the concentrated eluate was separated isocratically in buffer HA by Superose-12 chromatography on an AKTA FPLC at 4°C. Recombinant Lid, LP2, and Module 1 were purified via lysis of bacterial cells in lid buffer (50 mM HEPES·NaOH [pH 7.5], 100 mM NaCl, 100 mM KCl, 5% glycerol, 1 mM DTT) plus protease inhibitors, bound to amylose resin (NEB), washed extensively with lid buffer, and eluted in lid buffer containing 10 mM maltose. Eluates were further purified via AKTA FPLC at 4°C using Superose-6 or Sephacryl S-200 chromatography as appropriate based on the size of the complex. We found that coexpression of Hsc82 (Lander et al., 2012) was not required for assembly of the lid or its subcomplexes in *E. coli*. To ensure that the C termini of Rpn3 and Rpn7 were intact in purified LP3 used for assembly assays, the proteins were expressed from pRT978, which encodes a C-terminal 6His-tagged Rpn7 and a C-terminal *Mxe* intein-chitin binding-domain fusion-tagged Rpn3. LP3 was purified under nonreducing conditions (to avoid spurious intein cleavage) via sequential TALON affinity chromatography as above followed by chitin affinity chromatography. The protein was eluted from the TALON resin in TALON-intein buffer (50 mM Tris·HCl [pH 8.0], 500 mM NaCl, 10% glycerol, 0.2% Tween-20, 150 mM imidazole), and applied directly to pre-equilibrated chitin resin (NEB). After washing extensively with TALON-intein buffer, intein cleavage was initiated by incubation with 50 mM DTT overnight at 4°C. The flowthrough was then collected, concentrated, and fractionated isocratically in lid buffer using a Sephacryl S-200 column. LP2 and 26S proteasomes of yeast origin for TEVp cleavage assays were purified exactly as described previously (Tomko and Hochstrasser, 2011) from MHY7810 (LP2) or MHY7773 (26S) carrying pRT837 or pRT902. All proteins were concentrated via centrifugal filtration (Amicon), flash frozen in liquid nitrogen, and stored at –80°C until use.

Lid Subcomplex Assembly Assays

Protein complexes were diluted together to 5 µM in 26S buffer containing 10 mg/mL BSA and lacking ATP, incubated at 30°C for 20 min, and then analyzed by native PAGE as described above.

Site-Directed Photocrosslinking Assays

A BL21-STAR (DE3) strain containing plasmid pRT59/pEVOL-pBpF (Young et al., 2010) was transformed with a plasmid encoding Rpn3, HA-Rpn7, and Sem1-Gly-6His with the indicated amber suppressor codons. Transformants were grown in terrific broth to OD₆₀₀ ≈ 1.0. At that time, IPTG, *D*-arabinose, and *p*-benzoylphenylalanine (Bachem) were added to 1 mM, 0.2% (w/v), and 500 µM, respectively, to induce expression. Cultures were shaken at

30°C for 20 hr before harvest. After purification by TALON affinity chromatography, protein complexes were diluted to equal concentrations (based on Bradford assay) with buffer HA and irradiated for 30 min on ice with a Blak-Ray B-100AP High-Intensity UV Lamp (100 W, 365 nm; UVP, LLC). Samples were then subjected to immunoblotting.

TEV Protease Cleavage

The indicated complexes were diluted to 100–400 nM in buffer HA containing 5 mg/mL BSA and 1 mM DTT instead of TCEP and were treated with either buffer control or 10 µM recombinant TEV protease overnight at 4°C. Reactions containing 26S proteasomes also contained an ATP-regenerating system (1 mM ATP, 50 µg/mL creatine kinase, 5 mM creatine phosphate). For recombinant LP3 and LP2, HA-Rpn7 and associated proteins were immunoprecipitated using anti-HA affinity matrix (Roche) before separation by SDS-PAGE and immunoblotting for the indicated proteins. For yeast-derived LP2 and 26S proteasomes, reactions were instead separated by native PAGE and subjected to immunoblotting as described above.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two tables and five figures and can be found with this article at <http://dx.doi.org/10.1016/j.molcel.2013.12.009>.

ACKNOWLEDGMENTS

We thank Yanjie Li for helpful comments, Peter G. Schultz for pEVOL-pBpF, and Dan Finley for antibodies. This work was supported by NIH grant GM083050 (to M.H.). R.J.T. was supported in part by an American Cancer Society postdoctoral fellowship.

Received: October 24, 2013

Revised: November 19, 2013

Accepted: December 5, 2013

Published: January 9, 2014

REFERENCES

- Beck, F., Unverdorben, P., Bohn, S., Schweitzer, A., Pfeifer, G., Sakata, E., Nickell, S., Plitzko, J.M., Villa, E., Baumeister, W., and Förster, F. (2012). Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl. Acad. Sci. USA* 109, 14870–14875.
- Bohn, S., Sakata, E., Beck, F., Pathare, G.R., Schnitger, J., Nagy, I., Baumeister, W., and Förster, F. (2013). Localization of the regulatory particle subunit Sem1 in the 26S proteasome. *Biochem. Biophys. Res. Commun.* 435, 250–254.
- Cheng, M., Olivier, P., Diehl, J.A., Fero, M., Roussel, M.F., Roberts, J.M., and Sherr, C.J. (1999). The p21(Cip1) and p27(Kip1) CDK ‘inhibitors’ are essential activators of cyclin D-dependent kinases in murine fibroblasts. *EMBO J.* 18, 1571–1583.
- Dyson, H.J., and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6, 197–208.
- Ellisdon, A.M., Dimitrova, L., Hurt, E., and Stewart, M. (2012). Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. *Nat. Struct. Mol. Biol.* 19, 328–336.
- Estrin, E., Lopez-Blanco, J.R., Chacón, P., and Martin, A. (2013). Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* 21, 1624–1635.
- Faza, M.B., Kemmler, S., Jimeno, S., González-Aguilera, C., Aguilera, A., Hurt, E., and Panse, V.G. (2009). Sem1 is a functional component of the nuclear pore complex-associated messenger RNA export machinery. *J. Cell Biol.* 184, 833–846.
- Fukunaga, K., Kudo, T., Toh-e, A., Tanaka, K., and Saeki, Y. (2010). Dissection of the assembly pathway of the proteasome lid in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* 396, 1048–1053.

- Funakoshi, M., Tomko, R.J., Jr., Kobayashi, H., and Hochstrasser, M. (2009). Multiple assembly chaperones govern biogenesis of the proteasome regulatory particle base. *Cell* 137, 887–899.
- Glickman, M.H., Rubin, D.M., Coux, O., Wefes, I., Pfeifer, G., Cjeka, Z., Baumeister, W., Fried, V.A., and Finley, D. (1998). A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* 94, 615–623.
- Gudmundsdottir, K., Lord, C.J., and Ashworth, A. (2007). The proteasome is involved in determining differential utilization of double-strand break repair pathways. *Oncogene* 26, 7601–7606.
- Hirano, Y., Hendil, K.B., Yashiroda, H., Iemura, S., Nagane, R., Hioki, Y., Natsume, T., Tanaka, K., and Murata, S. (2005). A heterodimeric complex that promotes the assembly of mammalian 20S proteasomes. *Nature* 437, 1381–1385.
- Hirano, Y., Hayashi, H., Iemura, S., Hendil, K.B., Niwa, S., Kishimoto, T., Kasahara, M., Natsume, T., Tanaka, K., and Murata, S. (2006). Cooperation of multiple chaperones required for the assembly of mammalian 20S proteasomes. *Mol. Cell* 24, 977–984.
- Kaneko, T., Hamazaki, J., Iemura, S., Sasaki, K., Furuyama, K., Natsume, T., Tanaka, K., and Murata, S. (2009). Assembly pathway of the Mammalian proteasome base subcomplex is mediated by multiple specific chaperones. *Cell* 137, 914–925.
- Kovacs, D., Szabo, B., Pancsa, R., and Tompa, P. (2013). Intrinsically disordered proteins undergo and assist folding transitions in the proteome. *Arch. Biochem. Biophys.* 531, 80–89.
- Kusmierczyk, A.R., Kunjappu, M.J., Funakoshi, M., and Hochstrasser, M. (2008). A multimeric assembly factor controls the formation of alternative 20S proteasomes. *Nat. Struct. Mol. Biol.* 15, 237–244.
- Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., and Martin, A. (2012). Complete subunit architecture of the proteasome regulatory particle. *Nature* 482, 186–191.
- Lasker, K., Förster, F., Bohn, S., Walzthoeni, T., Villa, E., Unverdorben, P., Beck, F., Aebersold, R., Sali, A., and Baumeister, W. (2012). Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc. Natl. Acad. Sci. USA* 109, 1380–1387.
- Le Tallec, B., Barrault, M.B., Courbeyrette, R., Guérois, R., Marsolier-Kergoat, M.C., and Peyroche, A. (2007). 20S proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and in mammals. *Mol. Cell* 27, 660–674.
- Le Tallec, B., Barrault, M.B., Guérois, R., Carré, T., and Peyroche, A. (2009). Hsm3/S5b participates in the assembly pathway of the 19S regulatory particle of the proteasome. *Mol. Cell* 33, 389–399.
- Li, J., Zou, C., Bai, Y., Wazer, D.E., Band, V., and Gao, Q. (2006). DSS1 is required for the stability of BRCA2. *Oncogene* 25, 1186–1194.
- Mark, W.Y., Liao, J.C., Lu, Y., Ayed, A., Laister, R., Szymczynska, B., Chakrabarty, A., and Arrowsmith, C.H. (2005). Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J. Mol. Biol.* 345, 275–287.
- Marston, N.J., Richards, W.J., Hughes, D., Bertwistle, D., Marshall, C.J., and Ashworth, A. (1999). Interaction between the product of the breast cancer susceptibility gene BRCA2 and DSS1, a protein functionally conserved from yeast to mammals. *Mol. Cell. Biol.* 19, 4633–4642.
- Noutsou, M., Duarte, A.M., Anvarian, Z., Didenko, T., Minde, D.P., Kuper, I., de Ridder, I., Oikonomou, C., Friedler, A., Boelens, R., et al. (2011). Critical scaffolding regions of the tumor suppressor Axin1 are natively unfolded. *J. Mol. Biol.* 405, 773–786.
- Park, S., Roelofs, J., Kim, W., Robert, J., Schmidt, M., Gygi, S.P., and Finley, D. (2009). Hexameric assembly of the proteasomal ATPases is templated through their C termini. *Nature* 459, 866–870.
- Ramos, P.C., Höckendorff, J., Johnson, E.S., Varshavsky, A., and Dohmen, R.J. (1998). Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* 92, 489–499.
- Roelofs, J., Park, S., Haas, W., Tian, G., McAllister, F.E., Huo, Y., Lee, B.H., Zhang, F., Shi, Y., Gygi, S.P., and Finley, D. (2009). Chaperone-mediated pathway of proteasome regulatory particle assembly. *Nature* 459, 861–865.
- Sá-Moura, B., Simões, A.M., Fraga, J., Fernandes, H., Abreu, I.A., Botelho, H.M., Gomes, C.M., Marques, A.J., Dohmen, R.J., Ramos, P.C., et al. (2013). Biochemical and Biophysical Characterization of Recombinant Yeast Proteasome Maturation Factor Ump1. *Comp. Struct. Biotech. J.* 7, 1–12.
- Saeki, Y., Toh-E, A., Kudo, T., Kawamura, H., and Tanaka, K. (2009). Multiple proteasome-interacting proteins assist the assembly of the yeast 19S regulatory particle. *Cell* 137, 900–913.
- Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004). Sem1p is a novel subunit of the 26 S proteasome from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 279, 28807–28816.
- Tomko, R.J., Jr., and Hochstrasser, M. (2011). Incorporation of the Rpn12 subunit couples completion of proteasome regulatory particle lid assembly to lid-base joining. *Mol. Cell* 44, 907–917.
- Tomko, R.J., Jr., and Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. *Annu. Rev. Biochem.* 82, 415–445.
- Uekusa, Y., Okawa, K., Yagi-Utsumi, M., Serve, O., Nakagawa, Y., Mizushima, T., Yagi, H., Saeki, Y., Tanaka, K., and Kato, K. (2013). Backbone H, C, and N assignments of yeast Ump1, an intrinsically disordered protein that functions as a proteasome assembly chaperone. *Biomol. NMR Assign.* Published online September 25, 2013. <http://dx.doi.org/10.1007/s12104-013-9523-1>.
- Wang, Y., Fisher, J.C., Mathew, R., Ou, L., Otieno, S., Sublet, J., Xiao, L., Chen, J., Roussel, M.F., and Kriwacki, R.W. (2011). Intrinsic disorder mediates the diverse regulatory functions of the Cdk inhibitor p21. *Nat. Chem. Biol.* 7, 214–221.
- Wei, S.J., Williams, J.G., Dang, H., Darden, T.A., Betz, B.L., Humble, M.M., Chang, F.M., Trempus, C.S., Johnson, K., Cannon, R.E., and Tennant, R.W. (2008). Identification of a specific motif of the DSS1 protein required for proteasome interaction and p53 protein degradation. *J. Mol. Biol.* 383, 693–712.
- Wilmes, G.M., Bergkessel, M., Bandyopadhyay, S., Shales, M., Braberg, H., Cagney, G., Collins, S.R., Whitworth, G.B., Kress, T.L., Weissman, J.S., et al. (2008). A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Mol. Cell* 32, 735–746.
- Yang, H., Jeffrey, P.D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N.H., Zheng, N., Chen, P.L., Lee, W.H., and Pavletich, N.P. (2002). BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. *Science* 297, 1837–1848.
- Yoon, M.K., Mitrea, D.M., Ou, L., and Kriwacki, R.W. (2012). Cell cycle regulation by the intrinsically disordered proteins p21 and p27. *Biochem. Soc. Trans.* 40, 981–988.
- Young, T.S., Ahmad, I., Yin, J.A., and Schultz, P.G. (2010). An enhanced system for unnatural amino acid mutagenesis in *E. coli*. *J. Mol. Biol.* 395, 361–374.

Molecular Cell, Volume 53

Supplemental Information

The Intrinsically Disordered Sem1 Protein Functions as a Molecular Tether during Proteasome Lid Biogenesis

Robert J. Tomko, Jr. and Mark Hochstrasser

Supplemental Figure Legends

Supplemental Figure S1 pertaining to Figures 1 and 2. Purification of lid subcomplexes, Sem1-Rpn12 interaction analysis, and elution profiles of Rpn3 when coexpressed with combinations of Rpn7 and Sem1. (a) Coomassie-stained SDS-PAGE separation of purified recombinant Module 1 and LP3. The asterisk indicates a truncation product of MBP-Rpn6. (b) Normalized gel filtration UV traces of TALON-purified Rpn3-6His expressed in the presence of the indicated proteins. Arrowheads mark the position of the Rpn3-6His elution peak as determined by SDS-PAGE. Note that Rpn3-6His, when coexpressed with Rpn7, elutes with an M_r as expected for free Rpn3 and with no apparent copurification of Rpn7 (not shown), indicating that Rpn7 does not appreciably bind Rpn3 in the absence of Sem1. (c) Rpn12 does not copurify with Sem1. Sem-6His was coexpressed with Rpn12 alone or in combination with Rpn3 and Rpn7. Sem1 and associated proteins were then purified by TALON affinity and separated by SDS-PAGE. An arrowhead marks the Rpn12 band in the input lanes, and asterisks indicate metal-binding *E. coli* proteins. Note that the input and pulldown panels are from nonadjacent lanes of the same gel. (d) Sem1 elutes from a gel filtration column with a relative migration much larger than its calculated molecular weight. Rpn7 and Sem-Gly-6His were coexpressed as described in the Experimental Procedures, purified via the 6His tag on Sem1, and the 6His eluate was separated on a Superose 12 column. An arrowhead indicates the elution peak position of Sem1 that is unbound by Rpn7. Proteolytic fragments of the free proteins are indicated.

Supplemental Figure S2 pertaining to Figure 2. Additional characterization of *sem1A* proteasome assembly and Rpn3 stability. (a) Rpn12 immunoblot of extracts from the

indicated yeast strains showing the accumulation of the full lid in *rpn10Δ* yeast, lack of Rpn12 in lid* (compare to main **Fig. 2b**), and the accumulation of free Rpn12 in *sem1Δ* mutants. **(b)** Immunoblots against RP base and CP subunits from extracts of the indicated yeast strains. No obvious defects in base or CP assembly were evident. B₂CP, CP capped on each end with base. **(c)** The half-life of Rpn3, but not Rpn7, is decreased in *sem1Δ* yeast. The indicated strains were grown to mid-log phase in YPD, treated with 250 μg/mL cycloheximide (CHX), and samples were taken for immunoblotting at each of the indicated timepoints. PGK, phosphoglycerate kinase.

Supplemental Figure S3 pertaining to Figure 3. Expression levels of *sem1* site mutants and associated growth defects. **(a)** The amino acid sequence of each allele is shown in an alignment with WT Sem1. Mutated amino acids are colored blue. **(b)** WT or *rpn10Δ sem1Δ* yeast were transformed with empty vector or low copy plasmids encoding the indicated *SEM1* alleles before spotting on synthetic medium lacking tryptophan and incubation for 2 days at the indicated temperatures. **(c)** WT and *rpn10Δ sem1Δ* yeast were transformed with empty vector or low copy plasmids encoding the indicated *SEM1* alleles. After growth to mid-log phase, transformants were harvested and lysed under denaturing conditions, followed by immunoblotting for Sem1 or PGK. The migration of mutant Sem1 by SDS-PAGE was enhanced upon mutation of acidic amino acids. **(d)** Mutation of Sem1 site 1 to noncharged residues ablates interaction with Rpn3, but not Rpn7. Mutant forms of Sem1-Gly-6His in which site 1 residues 30-37 were changed to the indicated amino acids were coexpressed with Rpn3 (left panel) or Rpn7 (right panel), and copurification of Rpn3 or Rpn7 with Sem1 was assessed by SDS-PAGE. The input and pulldown lanes for each panel are from non-adjacent lanes of the same gel.

Asterisks indicate contaminating *E. coli* proteins.

Supplemental Figure S4 pertaining to Figure 5. Expression levels of *sem1* linker mutants and associated growth defects. (a) WT and *rpn10Δ sem1Δ* yeast were transformed with empty vector or low copy plasmids encoding the indicated *SEM1* alleles. After growth to mid-log phase, transformants were harvested and lysed under denaturing conditions, followed by immunoblotting for Sem1. (b) The amino acid sequence of Sem1 TEV site mutants are shown in an alignment with WT Sem1. Mutated or inserted amino acids are colored blue. (c) WT or *rpn10Δ sem1Δ* yeast were transformed with empty vector or low copy plasmids encoding the indicated *SEM1* alleles before spotting on synthetic medium lacking tryptophan and incubation for 2 days at the indicated temperatures. (d) An alignment of representative Sem1 orthologs from all five superfamilies of Eukarya is shown, with the change in linker length compared to scSem1 indicated.

Supplemental Figure S5 pertaining to Figure 6. TEV-cleavable proteasomes and subcomplexes. (a) An alignment of the protein sequences of WT Sem1, Sem1-L20, and Sem1-L20-TEVx. Inserted amino acids are colored blue. (b) SDS-PAGE analysis of LP3-L20 and LP3-L20-TEVx. The asterisk indicates a truncation product of Rpn3. Note that the N-terminal FLAG tag was rapidly cleaved from a portion of Sem1 and thus was not used for comparative analyses by Western blotting. (c) WT (1) or *rpn10Δ sem1Δ* (2 through 5) yeast were transformed with the plasmids encoding WT or mutant *SEM1* alleles, struck onto SD-Trp plates, and incubated for 2 days at the indicated temperatures. (1) empty vector; (2) empty vector; (3) *SEM1*; (4) *sem1-L20*; (5) *sem1-L20-TEVx*. (d) SDS-PAGE analysis of purified recombinant LP2

containing sem1-L20 or sem1-L20-TEV_x. The dot marks the presence of the TEV cleavage site.

(e) SDS-PAGE analysis as in **(d)** of 26S proteasomes and LP2 containing sem1-L20 or sem1-L20-TEV_x, purified from yeast. **(f)** Sem1-L20-TEV_x within yeast-purified LP2 is fully cleaved by TEV protease. Immunoblot analysis of SDS-PAGE-separated proteasomes from **(e)**; the dot indicates the presence of a component. Sem1* indicates the sem1 N-terminal cleavage product.

(g) The tethering function of Sem1 is dispensable within yeast LP2. Purified LP2 from **(e)** was treated with TEV protease or buffer before resolving by native PAGE and immunoblotting for the indicated lid subunits. **(h)** Sem1-L20-TEV_x within 26S proteasomes is fully cleaved by TEV protease. Immunoblot analysis of SDS-PAGE-separated LP2 from **(e)**, performed as in **(f)**. **(i)** The tethering function of Sem1 is dispensable within 26S proteasomes. Purified proteasomes from **(e)** were treated with TEV protease or buffer before resolving by native PAGE and immunoblotting for the indicated lid subunits.

Supplementary Table S1: Yeast strains used in this study

Name	Genotype	Source
MHY500	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1</i>	Chen <i>et al.</i> (1993)
MHY960	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 rpn10Δ::HIS3</i>	Tomko and Hochstrasser (2011)
MHY1396	<i>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 (W303)</i>	Thomas and Rothstein (1989)
MHY4785	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 sem1Δ::HIS3</i>	Tomko and Hochstrasser (2011)
MHY7773	<i>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 RPN5-6xGly-3xFLAG:hphMX4 sem1Δ::kanMX6</i>	This study
MHY7789	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 rpn10Δ::HIS3 sem1Δ::HIS3</i>	This study
MHY7810	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 RPN5-6xGly-3xFLAG:hphMX4 rpn12-234Δ:hphMX4 sem1Δ::scHIS3</i>	This study
MHY7882	<i>MATa leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3- 11,15 rpn4Δ::natMX4 sem1Δ::kanMX6 [YCplac33-SEM1]</i>	This study
MHY8051	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 RPN7-6xGly-3xFLAG:kanMX6 sem1Δ::HIS3</i>	This study
MHY8297	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 RPN7-6xGly-3xFLAG:kanMX6 rpn10Δ::HIS3</i>	This study
MHY8287	<i>MATa his3-Δ200 leu2-3,112 ura3-52 lys2-801 trp1-1 RPN7-6xGly-3xFLAG:kanMX6 rpn10Δ::HIS3 sem1Δ::HIS3</i>	This study

Supplementary Table S2: Plasmids used in this study.

Plasmid	Genotype
pRT59	pEVOL- <i>pBpF</i>
pRT375	pET42b-Rpn3-6His : Rpn7
pRT514	pET42b-Rpn3-6His : Sem1 : Rpn7
pRT515	pET42b-Rpn3-6His : Sem1
pRT562	pET42b-Sem1-Gly-6His : Rpn7
pRT568	pET42b-Rpn3 : Sem1-Gly-6His : Rpn7
pRT585	pET42b-Rpn3 : Sem1 : Rpn7
pRT588	pET42b-Rpn3 : Rpn7-Gly-6His
pRT589	pET42b-Rpn3 : Sem1-Gly-6His
pRT562	pET42b-Sem1-Gly-6His : Rpn7
pRT620	YCplac33- <i>SEM1</i>
pRT644	pET42b-Rpn3 : sem1(E30A/E31A/D32A/D33A/E34A/F35A/E36A/D37A)-Gly-6His
pRT645	pET42b-Rpn3 : sem1(W60T/W64A)-Gly-6His
pRT649	pET42b-sem1(E30A/E31A/D32A/D33A/E34A/F35A/E36A/D37A)-Gly-6His : Rpn7
pRT650	pET42b-sem1(W60T/W64A)-Gly-6His : Rpn7
pRT664	pET42b-Rpn3 : sem1(E30A/E31A/D32A/D33A/E34A/F35A/E36A/D37A)-Gly-6His : Rpn7
pRT665	pET42b-Rpn3 : sem1(W60T/W64A)-Gly-6His : Rpn7
pRT681	pET42b-Rpn3 : FLAG-sem1(1-51)
pRT712	pET42b-HA-Rpn7
pRT746	YCplac22- <i>SEM1</i>
pRT754	YCplac22- <i>sem1(E61S/E62S/D65S/D66S)</i>
pRT755	YCplac22- <i>sem1(E61K/E62K/D65K/D66K)</i>
pRT769	pET42b-Rpn3 : FLAG-Sem1-TEVx-ZZ-6His : Rpn7
pRT772	pET42b-Rpn3 : sem1(52-89)-TEVx-ZZ-6His : Rpn7
pRT773	pET42b-Rpn3 : sem1(52-89)-TEVx-ZZ-6His : Rpn7 : FLAG-sem1(1-51)
pRT796	pRS424- <i>RPN3</i>

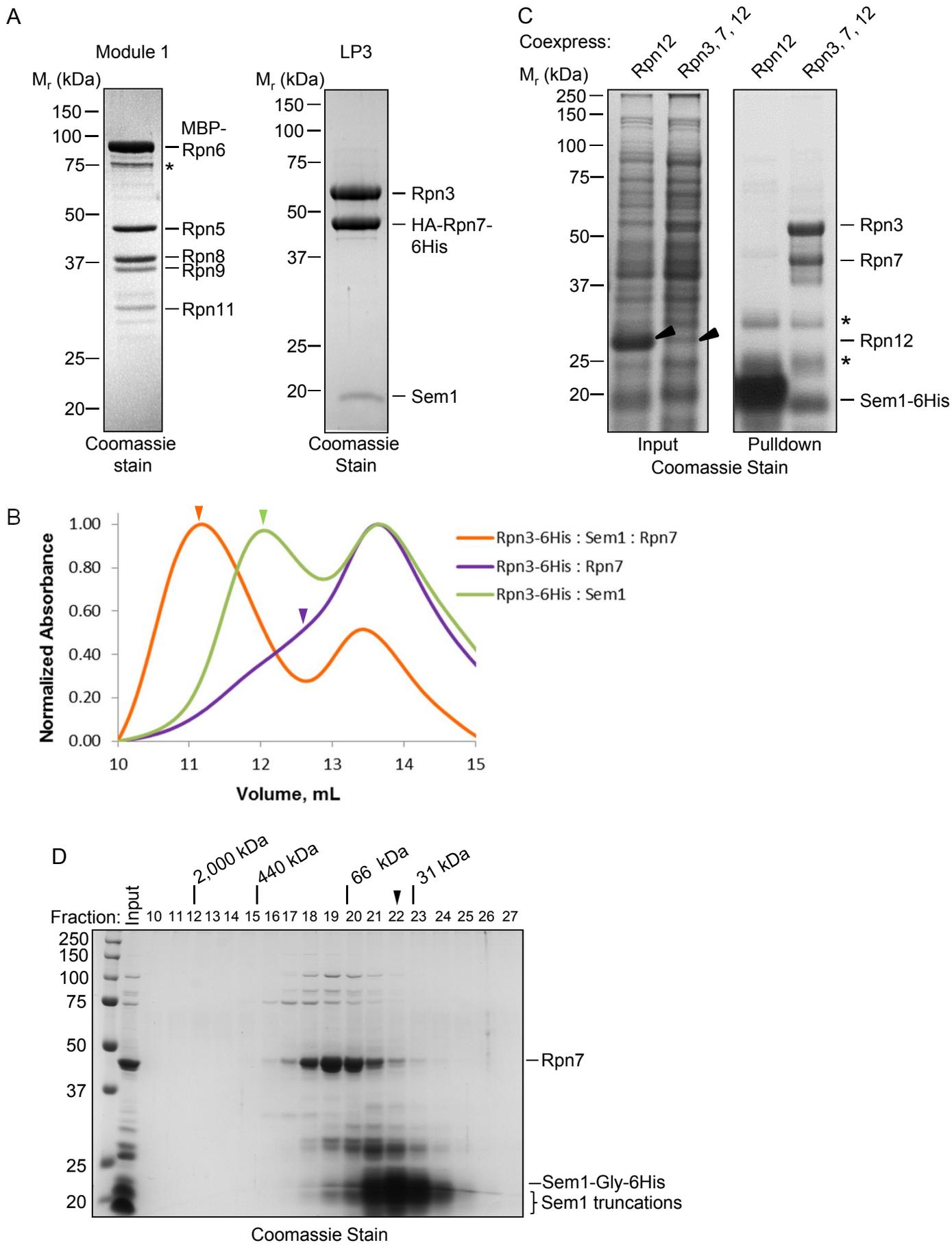
pRT815	pET42b-Rpn3 : sem1-W60X-Gly-6His : HA-Rpn7
pRT818	YCplac22- <i>sem1-Δ5link</i>
pRT819	YCplac22- <i>sem1-Δ11link</i>
pRT820	YCplac22-- <i>sem1-Δ15link</i>
pRT821	YCplac22-- <i>sem1-Δ19link</i>
pRT834	pET42b-Rpn3-6His : HA-Rpn7
pRT837	YCplac22- <i>sem1-L20</i>
pRT843	pET42b-Rpn3-6His : FLAG-Sem1 : HA-Rpn7
pRT847	pRS424- <i>RPN5</i>
pRT851	pRS425- <i>RPN5</i>
pRT854	pRS425- <i>RPN7</i>
pRT859	pET42b-Rpn3-6His : FLAG-Sem1-L20 : HA-Rpn7
pRT869	YCplac22- <i>sem1(E30K/E31K/E34K/E36K)</i>
pRT897	pET42b-Sem1-Gly-6His : HA-Rpn7
pRT902	YCplac22- <i>sem1-L20-TEVx</i>
pRT905	pET42b-Rpn3-6His : FLAG- <i>sem1-L20-TEVx</i> : HA-Rpn7
pRT937	pET42b-Sem1 : HA-Rpn7
pRT939	pET42b-Rpn3 : Sem1
pRT945	pCDF42-6His-MBP-Rpn6 : Rpn9 : Rpn11 : Rpn5 : Rpn8
pRT958	pET42b-Rpn3-6His : FLAG- <i>sem1-Δ5link</i> : HA-Rpn7
pRT959	pET42b-Rpn3-6His : FLAG- <i>sem1-Δ11link</i> : HA-Rpn7
pRT960	pET42b-Rpn3-6His : FLAG- <i>sem1-Δ15link</i> : HA-Rpn7
pRT961	pET42b-Rpn3-6His : FLAG- <i>sem1-Δ19link</i> : HA-Rpn7
pRT962	pET42b-Rpn3 : FLAG-Sem1 : HA-Rpn7
pRT978	pET42b-Rpn3-GyrA-CBP : Sem1 : HA-Rpn7-Gly-6His
pRT1000	pET42b-Rpn3 : Sem1-Gly-6His : HA-Rpn7
pRT1001	pET42b-Rpn3 : <i>sem1-L29X-Gly-6His</i> : HA-Rpn7

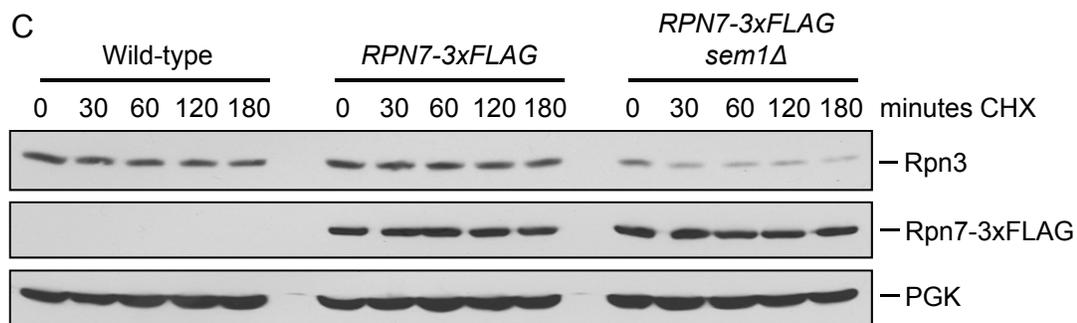
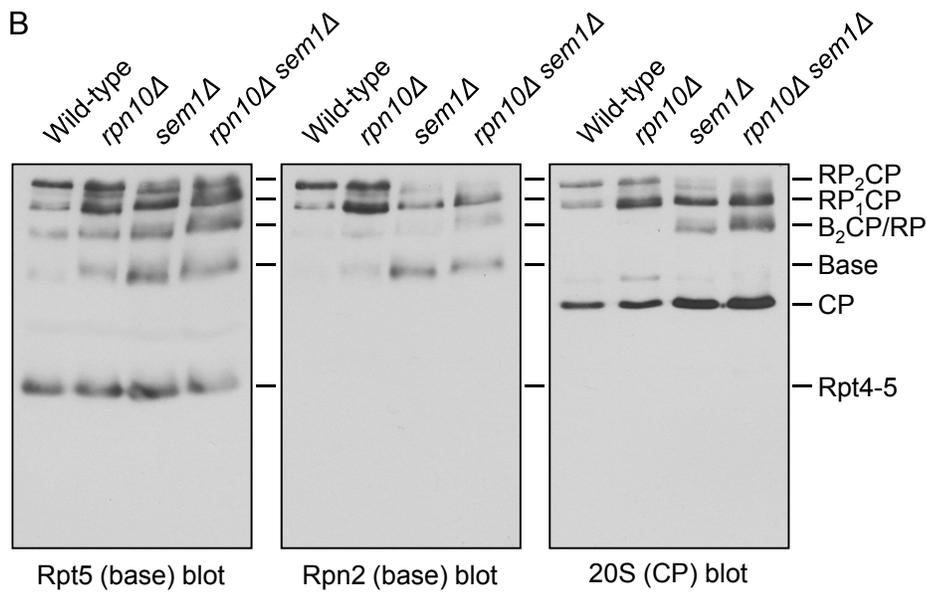
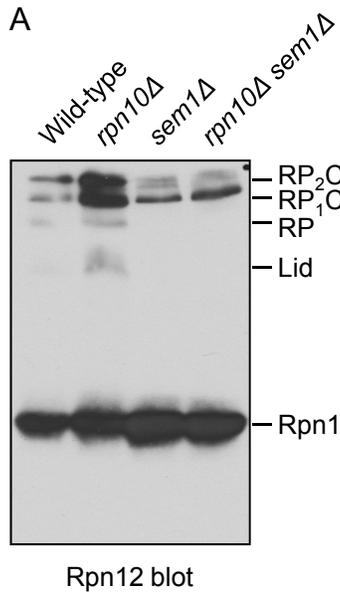
Note: all plasmids except pRT59 (described in J Mol Biol. 2010; 395:361-74) were produced in this study. 3Cx and TEVx indicate human rhinovirus 3C protease and tobacco etch protease

cleavage sites, respectively. For pRT815 and pRT1001, the “X” indicates this codon was mutated to TAG for amber suppression.

Supplemental References

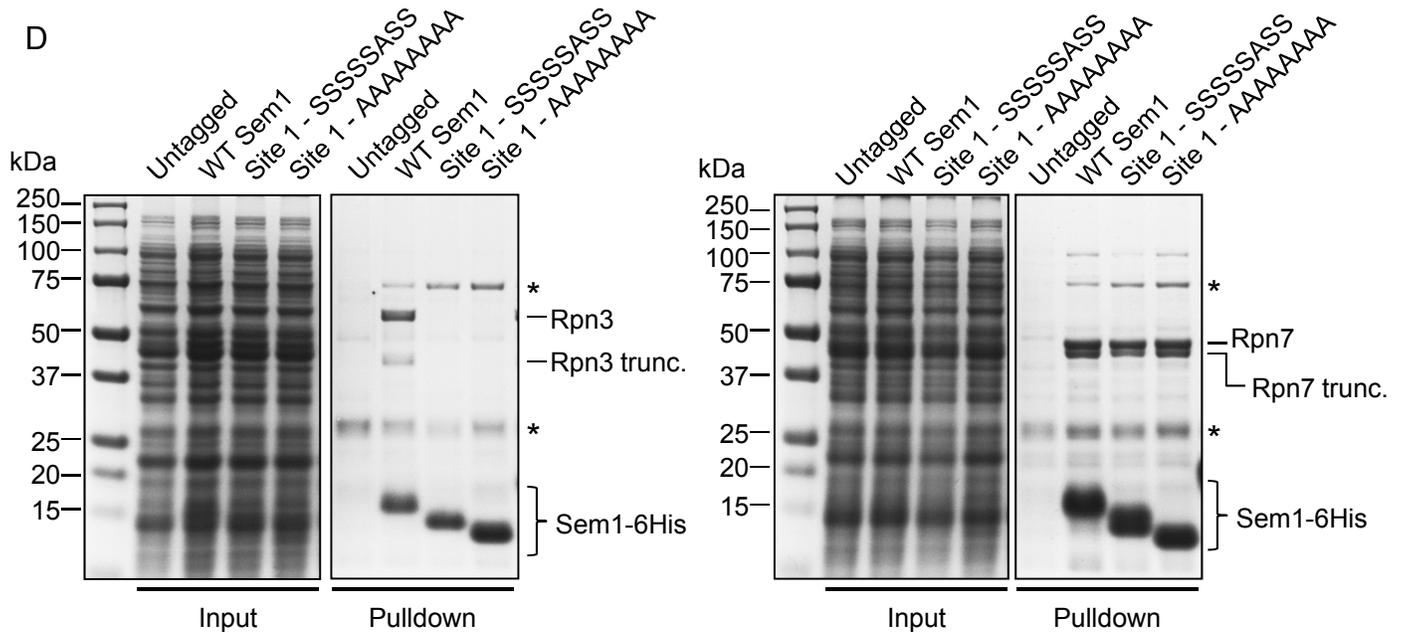
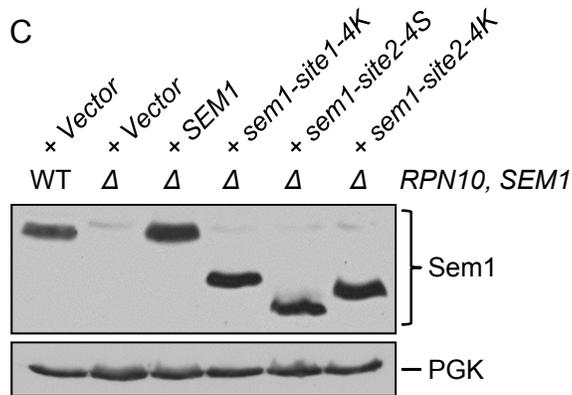
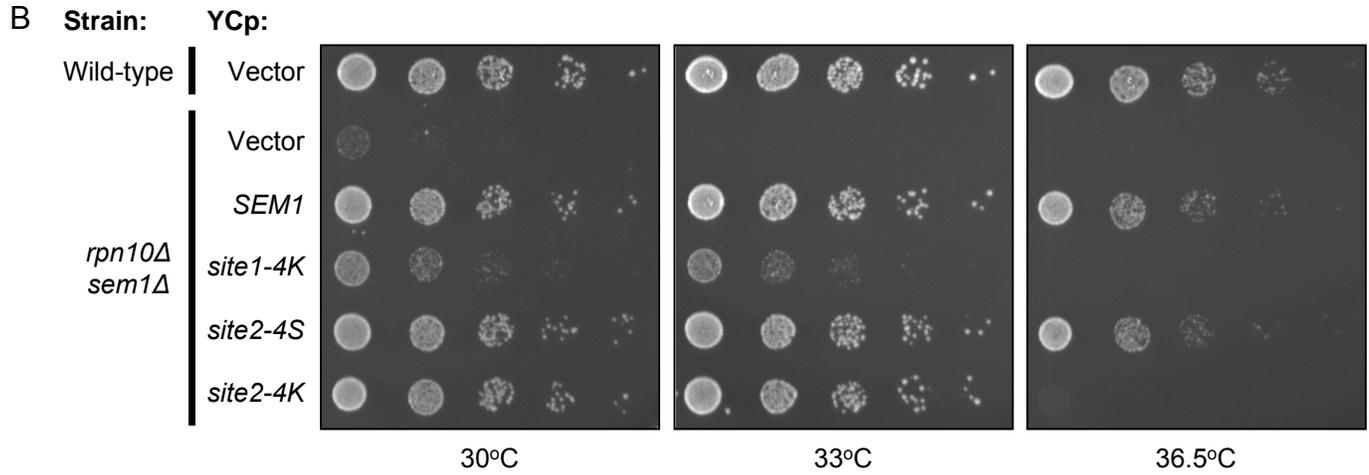
1. Chen, P., Johnson, P., Sommer, T., Jentsch, S., and Hochstrasser, M. (1993). Multiple ubiquitin-conjugating enzymes participate in the in vivo degradation of the yeast MAT alpha 2 repressor. *Cell* 74, 357-369.
2. Thomas, B.J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* 56, 619-630.

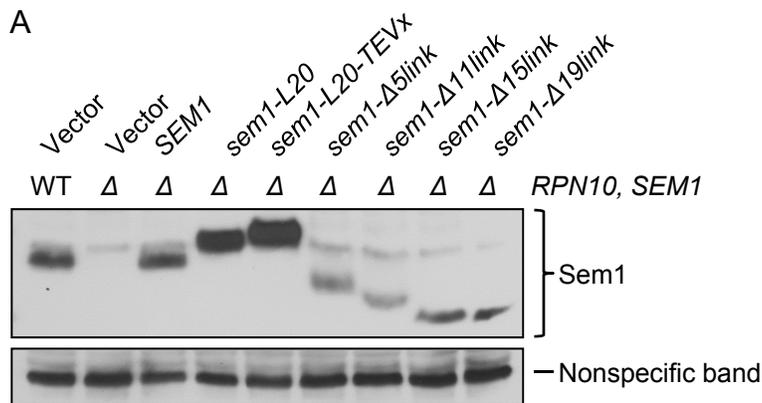




A

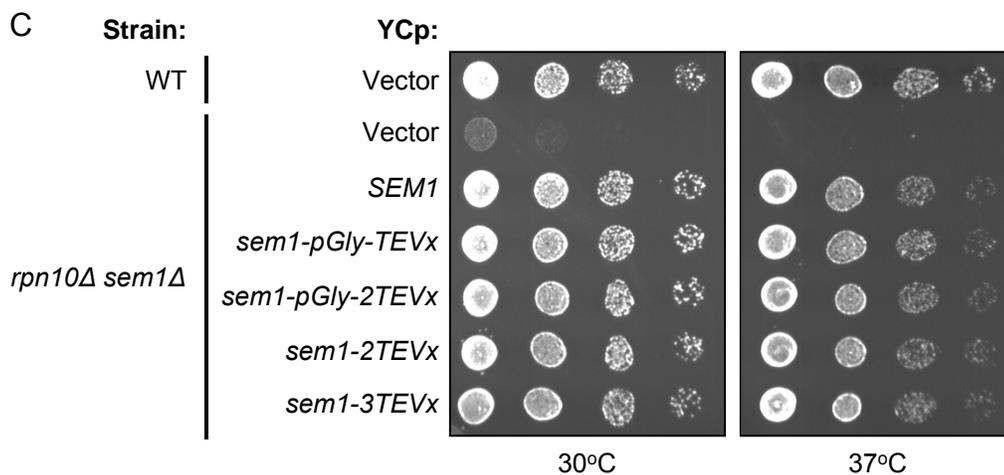
	Site 1	Linker	Site 2
Sem1	MSTDVAAAQAQSKIDLTKKKNEEINKKSL EEDEFE DFPIDTWANGETIKSNAVTQTNI WEENW DDVEVDDDFTNELKAELDRYKRENQ		
sem1-site1-4K	MSTDVAAAQAQSKIDLTKKKNEEINKKSL KKDDKFKD FPIDTWANGETIKSNAVTQTNI WEENW DDVEVDDDFTNELKAELDRYKRENQ		
sem1-site2-4S	MSTDVAAAQAQSKIDLTKKKNEEINKKSL EEDEFE DFPIDTWANGETIKSNAVTQTNI WSSNW SSVEVDDDFTNELKAELDRYKRENQ		
sem1-site2-4K	MSTDVAAAQAQSKIDLTKKKNEEINKKSL EEDEFE DFPIDTWANGETIKSNAVTQTNI WKKNW KKVEVDDDFTNELKAELDRYKRENQ		





B

	Site 1	Linker	Site 2	
Sem1	MSTDVAAAQAQSKIDLTKKKNEEINKKSL	LEEDDEFED	FPIDTWANGETIKSNAV-TQN-I	WEENWDD VEVDDDFTELKAEIDRYKRENQ
sem1-pGly-TEVx	MSTDVAAAQAQSKIDLTKKKNEEINKKSL	LEEDDEFED	FPIDT GGGGENLYFQGGG QTN-I	WEENWDD VEVDDDFTELKAEIDRYKRENQ
sem1-pGly-2TEVx	MSTDVAAAQAQSKIDLTKKKNEEINKKSL	LEEDDEFED	FPIDT ENLYFQGENLYFQGGG -I	WEENWDD VEVDDDFTELKAEIDRYKRENQ
sem1-pGly-2TEVx	MSTDVAAAQAQSKIDLTKKKNEEINKKSL	LEEDDEFED	FPIDT ENLYFQGENLYFQGG QTN-I	WEENWDD VEVDDDFTELKAEIDRYKRENQ
sem1-pGly-3TEVx	MSTDVAAAQAQSKIDLTKKKNEEINKKSL	LEEDDEFED	FP ENLYFQGENLYFQGENLYFQGI	WEENWDD VEVDDDFTELKAEIDRYKRENQ



D

	Site 1	Linker	Site 2	Δ (sc)
hs_NP_006295.1	13 LEEDDEFEEFF-AEDWAGL-DED	----	EDAHVWEDN-----	WDDDNV-EDDFSNQLRAE -5
mm_NP_033195.1	13 LEEDDEFEEFF-AEDWAGL-DED	----	EDAHVWEDN-----	WDDDNV-EDDFSNQLRAE -5
dr_NP_955887.1	13 LEEDDEFEEFF-AEDWTGL-DED	----	EDAHVWEDN-----	WDDDNV-EDDFSNQLRAE -5
xl_NP_001078914	14 LEEDDEFEEFF-TEDWTGF-DED	----	EDTHVWEDN-----	WDDDNV-EDDFSNQLRAE -5
dm_Q9VM46.1	25 LEEDDEFEEFF-AEDFRVG-DDE	----	EELNVWEDN-----	WDDDNV-EDDFSQQLKAH -5
bn_145938	61 LDEDDEFEEFF-RENWEKK-DEQ	----	QDEKLWEDN-----	WDDDDV-DAQFSEQLRKE -4
at_Q9FL96.1	17 FEDDDEFEEFF-EINEDWLEK-EVK	----	EVSLQWEDD-----	WDDDDV-SDDFSRQLRKE -3
mR_ACO62697.1	16 LECDDDEFEEFF-NEEWGAE-BEDA	----	EDVNQWEDD-----	WDDSEK-SDDFTROLRAE -4
sc_NP_010651.3	29 LEEDDEFEEFF-IDTWANG-ETIKSNAV	TQTN	IWEEN-----	WDDVEV-DDDFTNELKAE
sp_NP_594968.1	12 LEEDDEFEEFFA-TENWPMK-DTELDT	-GDDTL	WENN-----	WDDDEDIGDDFSVQLQAE -4
cm_BAM80515.1	49 FEEDDDFEFF-LHNWEPDLAAVE	----	EDELEWQDD-----	WEMIDQ-DEDFAAKLAAE -5
eh_EMS14886.1	17 LEEDDIFEEFF--DFKTK-DVD	----	DSTLHWQDN-----	WEDEGA--DEFTNHIRSS -9
tg_XP_002369043	37 DEPDDELEFFD--EIGGG-DGVVD	----	TEVAQWEDWDAAGWDD	EDV--NDDFCKRLQQE -9
gl_XP_001704618	13 LMY-DEFEEFGVLQDPEKALTEAI	----	LSKDTWQEA-----	WNKEPT--DAYIQMLQ-R -4

A

	Site 1	Linker	Site 2
Sem1	M...KKSLE EDDEFED FPIDTWANGET-----IKSNAVITQTN WEENWDD VEVDDDFTNELKAELDRYKRENQ		
sem1-L20	M...KKSLE EDDEFED FPIDTWANGET ASGAGGSEGGGSEGGTSGAT IKSNAVITQTN WEENWDD VEVDDDFTNELKAELDRYKRENQ		
sem1-L20-TEVx	M...KKSLE EDDEFED FPIDTWANGET ASGAGGENLYFQGGGTS GATIKSNAVITQTN WEENWDD VEVDDDFTNELKAELDRYKRENQ		

