Published in final edited form as:
Structure. 2010 May 12; 18(5): 584–593. doi:10.1016/j.str.2010.02.015.

The Prp19 WD40 Domain Contains a Conserved Protein Interaction Region Essential for its Function

Craig W. Vander Kooi[1,4,6,7,*], **Liping Ren**[2,5], **Ping Xu**[6,7], **Melanie D. Ohi**[2,4], **Kathleen L. Gould**[2,5], and **Walter J. Chazin**[1,3,4,*]

[1] Department of Biochemistry, Vanderbilt University, Nashville, TN 37232; USA

[2] Department of Cell and Developmental Biology, Vanderbilt University, Nashville, TN 37232; USA

[3] Department of Physics, Vanderbilt University, Nashville, TN 37232; USA

[4] Center for Structural Biology, Vanderbilt University, Nashville, TN 37232; USA

[5] Howard Hughes Medical Institute, Vanderbilt University, Nashville, TN 37232; USA

[6] Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY 40536; USA

[7] Center for Structural Biology, University of Kentucky, Lexington, KY 40536; USA

Summary

Prp19 is a member of the WD40-repeat family of E3 ubiquitin ligases and a conserved eukaryotic RNA splicing factor essential for activation and stabilization of the spliceosome. To understand the role of the WD40 repeat domain of Prp19 we have determined its structure using X-ray crystallography. The domain has a distorted seven bladed WD40 architecture with significant asymmetry due to irregular packing of blades one and seven into the core of the WD40 domain. Structure-based mutagenesis identified a highly conserved surface centered around blade five that is required for the physical interaction between Prp19 and Cwc2, another essential splicing factor. This region is found to be required for Prp19 function and yeast viability. Experiments *in vitro* and *in vivo* demonstrate that two molecules of Cwc2 bind to the Prp19 tetramer. These coupled structural and functional studies provide a model for the functional architecture of Prp19.

Highlights

- The Prp19 WD40 domain adopts a distorted seven bladed beta-propeller fold.
- A conserved surface patch is required for interaction with Cwc2 and *in vivo* function.
- Prp19 forms a 4:2 complex with Cwc2 *in vitro* and *in vivo*.

*Corresponding authors: C.W.V.K., Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY 40536, Telephone: (859) 323-8418, Fax: (859) 257-2283, craig.vanderkooi@uky.edu. W.J.C., Center for Structural Biology, Vanderbilt University, 5140 BIOSCI/MRBIII, Nashville, TN 37232-8725, Telephone: (615) 936-2210, Fax: (615) 936-2211, walter.chazin@vanderbilt.edu.

Introduction

Prp19, an essential RNA splicing factor, is a core component of the spliceosomal NinTeen associated-Complex (NTC), a sub-complex required for the activation of the spliceosome (Chan et al., 2003; Makarova et al., 2004). Prp19 plays a direct role in the activation and structural stabilization of the spliceosome (Tarn et al., 1994; Tarn et al., 1993; Chan et al., 2003; Chan and Cheng, 2005). Temperature sensitive mutants of Prp19 lead to destabilization of both the NTC and spliceosome as a whole (Ohi and Gould, 2002). Prp19 is a member of the U-box family of E3 ubiquitin ligases. Accumulating evidence supports the suggestion that Prp19 may regulate spliceosome activity by directly targeting proteins for ubiquitination and progress has been made in identifying putative Prp19 substrates and E3 ubiquitin ligase adaptors (Maeder and Guthrie, 2008; Ohi et al., 2005; Vander Kooi et al., 2006).

Prp19 contains three recognized domains: an N-terminal U-box domain, a central coiled-coil domain, and a C-terminal WD40 (forty residue Trp-Asp containing) β-propeller repeat domain. Prp19 utilizes its N-terminal U-box domain to recruit Ubc3, and possibly other E2 conjugating enzymes (Aravind and Koonin, 2000; Hatakeyama and Nakayama, 2003; Ohi et al., 2003). The central coiled-coil domain of Prp19 mediates tetramerization of the protein, which is important for function in the context of the spliceosome (Ohi et al., 2005). The C-terminal WD40 repeat domain is the least understood both physically and functionally.

In fact, the role of the WD40 repeat domain and whether it is essential for Prp19 function is unknown. E3 ubiquitin ligases function in concert with E1 activating and E2 conjugating enzymes to covalently attach ubiquitin molecules to substrates. E3 ligases often possess a multi-domain architecture and are distinguished by the different domain modules for recruiting E2 enzymes and substrates. E3 ligase proteins are distinguished on the basis of how they interact with E2 enzymes and have been grouped into three families based on the domain responsible for recruiting the E2 ligase: Really Interesting New Gene (RING)-finger (Freemont, 2000), U-box (Patterson, 2002), and Homologous to E6-AP Carboxyl Terminus (HECT) (Pickart, 2001). The substrate binding modules typically consist of well-known protein-protein interaction domains, such as WD40 repeat β-propeller, Leucine-Rich Repeat (LRR), or Tetratric Peptide Repeat (TPR).

The diversity of substrate binding domains used by E3 ligases provides the critical element of substrate specificity. The structural differences in substrate binding domains are functionally significant because E3 ligases differ dramatically in the type and number of substrates they recognize, with certain ligases tuned to bind only one specific partner while others recognize many different substrates. WD40 domains are thought to not only directly interact with targets but may also position them in a catalytically competent orientation. For example, the structure of the β-Transducin Repeat Containing Protein (βTrcP) 1 SCF complex suggests catalytic activity requires precise positioning of the E2 and substrate that is provided by orienting the E2 recruiting RING finger and the substrate recruiting WD40 repeat domains (Wu et al., 2003). We have previously proposed a mechanism for the activity of Prp19 that involves similar positioning of E2 and targeting by the U-box and WD40 domains from two anti-parallel protomers in the tetramer (Vander Kooi et al., 2006).

Investigations into the function of Prp19 in the spliceosome have included the discovery of protein binding partners. Four such proteins have been shown to directly bind to the different domains of *S. cerevisiae* Prp19: the E2 conjugating enzyme Ubc3 to the U-box domain, the NTC components Cef1 and Snt309 to the coiled-coil domain, and Cwc2 to the WD40 repeat domain (Hatakeyama and Nakayama, 2003; Ohi and Gould, 2002; Ohi et al., 2005). A direct interaction with the WD40 domain has been shown for the essential splicing factor Cwc2 (Cwf2 in *S. pombe*) (Ohi and Gould, 2002), but the nature and significance of this interaction has not

been investigated. Since Prp19 exists as a tetramer, it is particularly important to determine the relative stoichiometry of its binding partners because this has implications for its function and the overall organization and stability of the NTC.

The studies reported here use a coupled structural and functional approach to understand the WD40 domain of Prp19. An X-ray crystal structure shows the domain adopts an asymmetric 7-bladed fold. Conserved surface residues are identified and shown to be essential for protein interactions and function. These findings, along with determination of the stoichiometry of Prp19 and Cwc2 provide a model for the architecture of the Prp19 associated spliceosomal subcomplex.

Results

Structure of the WD40 repeat domain of Prp19

The WD40 repeat domain of Prp19, encompassing residues 165-503, was sub-cloned, expressed, purified, and crystallized. The protein crystallized in a tetragonal unit cell and diffracted to 2.6 Å. Seleno-methionine labeled crystals were also produced, which diffracted to 3.2 Å. Phases were derived from selenium MAD experiments and the structure was fully refined to 2.6 Å (Table 1).

While previous sequence analysis identified varying numbers of WD40 repeats in the C-terminus of Prp19, the X-ray crystal structure demonstrates unambiguously that the domain forms a seven bladed β-propeller (Figure 1A). Each WD40 repeat possesses the expected four-sheet blade with the exception of the final beta strand of the seventh blade, which is contributed by residues N-terminal to the first blade (179-186). This is a common feature of WD40 repeats, often termed “Velcro”, that contributes to domain stabilization (Madrona and Wilson, 2004; ter Haar et al., 1998).

Although it is common to find variability in the overall shape of WD40 repeat domains, most other WD40 repeat domains whose structures have been characterized possess nearly circular geometry with a central pore-like feature. This proves to be the case regardless if there are six, seven, or eight repeats in the WD40 domain. The WD40 domain of Prp19 is atypical in that it has significant deviations from the idealized circular geometry (Figure 1A). Detailed analysis of the Prp19 structure revealed the origin of this unique shape. The packing of the first blade of the β-propeller is central to the structural asymmetry. In particular, blade one possesses an irregular first β-strand (Figure 1A, marked with *). This β-strand contains a conserved proline, P196, which causes a break in the sheet (Figure 1B & C). The N-terminal half of the strand (1A) is rotated approximately 90° and forms an unusual parallel interaction with the first β-strand of blade two. The second half (1A′) forms the expected anti-parallel interaction with the second β-sheet of blade one (1B), and also participates in interactions with the second blade (2A). The seventh blade also shows perturbation in its packing against blade one causing it to be located significantly closer to the domain center than expected in circular geometry. The alterations in blades one and seven result in blades two through six possessing a shallower blade-to-blade angle. These unusual features lead to significant asymmetry in the overall structure, which has dimensions of approximately 35Å X 55Å. In summary, we find that the Prp19 WD40 repeat domain forms an atypical seven bladed beta propeller with significant asymmetry due to altered β-sheet packing, a feature that has not been observed previously in other WD40 repeat structures.

Structure-based prediction of functionally important residues

The overall amino acid conservation of Prp19 is relatively low, with 23% identity and 38% similarity between human and yeast proteins. The N-terminal region comprising the U-box

and coiled-coil is the most conserved with 38% identity and 58% similarity, whereas the C-terminal WD40 repeats are least conserved with 17% identity and 27% similarity. Because it is an essential conserved eukaryotic protein, we hypothesized that the surface residues of the WD40 domain responsible for interaction with partner proteins are the most likely to be conserved. Consequently, multiple sequence alignment of Prp19 from seven species was performed on a diverse range of eukaryotic species including representative members from yeast, plant, and animal. This analysis provided 39 highly conserved WD40 residues, defined on the basis of >85% conservation across species. As expected, many of these residues are core hydrophobic residues. However, 15 of the highly conserved WD40 residues are solvent exposed. When these exposed residues are mapped onto the structure, a striking patch of conserved surface is evident, centered on the fifth blade (Figure 2A). This conserved patch is approximately 250 $Å^2$ and formed by the flat upper surface of the blade.

When the sequence conservation in the fifth blade is analyzed in detail, certain key characteristics are evident (Figure 2B). In particular, residues from the loop between strands one and two of blade five, Asp384-Tyr387, possess a conserved Acidic-N-G-Y sequence, and residues from the loop between sheets three and four, Asp401-Lys404, possess a strictly conserved D-L-R-K sequence.

Physical interaction mediated by blade five

Both Prp19 and Cwc2 are essential for cell viability, but it is not known how they physically interact or whether their interaction is functionally essential. To address this question, a mutant ablating this interaction was needed. The sites selected for mutation were D341 at the edge of the conserved region in the central pore, Y387 near the center of the conserved region, K403 and K404 which contribute to the basic patch in blade five, and the non-conserved K452 from blade six which serves as a control (Figure 3A). These mutations were introduced into the conserved blade five residues and tested for whether the Prp19-Cwc2 interaction was affected in a directed yeast two-hybrid assay (Figure 3B). Differences were observed in the assays with the single site mutants, including no effect in the control experiment, a very weak effect from D341A, the strongest effect from R403A, and intermediate effects for Y387F and K404A. Because mutation of both R403 and K404 show an effect and could be compensatory, a double mutant R403A/K404A was also tested. In this double mutant, the Prp19-Cwc2 interaction was effectively ablated, and indeed, the two-hybrid assay showed this mutant did not interact with Cwc2 over background levels. These results suggest that the site of interaction between Prp19 and Cwc2 is in the WD40 repeat domain in the region centered on the fifth blade and that the double basic combination R403/K404 was likely to be intimately involved in the interaction. Since the relative amount of mutant protein expressed is not determined in these assays, the results provide a qualitative assessment of the interaction and serve as the basis for additional functional and biochemical studies.

Functional significance of conserved residues in blade five

In order to test the functional significance of the identified blade five residues, we assessed the ability of mutant proteins to rescue growth of the *prp19-1* temperature-sensitive strain (Figure 4A). In contrast to the wild- type protein and the single-site mutants, the R403A/K404A double mutant, shown above to be important for interaction with Cwc2, was unable to rescue yeast viability of *prp19-1* at the restrictive temperature. To further determine the physical effect of the R403A/K404A mutations, we developed a co-expression and purification system that allowed pull-down experiments via direct co-purification of Prp19 by his-tagged Cwc2 (Figure 4B). Wild-type Prp19 co-purified robustly with hisCwc2 along with excess free hisCwc2. While there is no apparent difference between the expression and stability of wild-type and the R403A/K404A mutant, co-purification of the R403A/K404A mutant with hisCwc2 was completely lost. Importantly, the WD40 domain containing the R403A/K404A mutations was

well folded, as assessed by circular dichroism (Figure S1), demonstrating that these effects were not due to a general destabilization of the domain.

Thus, mutation of R403A/K404A results in both loss of physical interaction between Prp19 with Cwc2 and loss of function *in vivo*. This shows that, as with the U-box domain (Ohi et. al., 2003) and the coiled-coil region (Ohi et al., 2005), the WD40 domain is essential for Prp19 function and that the conserved blade 5 region centered on R403/K404 is a crucial element.

Architecture of the Prp19/Cwc2 complex

Prp19 exists as a tetramer in solution, with flexibly tethered WD40 repeat domains (Ohi et al., 2005). This suggests that up to four copies of Cwc2 could simultaneously bind to Prp19. Co-purified Prp19 and hisCwc2 (Figure 4B) physically interact and form a stable complex that can be purified by gel filtration (Figure 5A). Interestingly, the complex clearly contains more Prp19 than Cwc2, even though Cwc2 is in excess in the co-purified starting material. Sedimentation velocity analytical ultracentrifugation (AU) was used to quantitatively examine the architecture of the Prp19-Cwc2 complex. Prp19 (58.6 kDa) predominantly exists as an extended tetramer (s = 6.1, 85%, 211 kDa), as previously demonstrated (Figure 5B, Table 2) (Ohi et al., 2005). HisCwc2 (39.4 kDa) was found to exist primarily in a monomeric state (s = 2.67, 60.4%, 41.7 kDa) with a small fraction of higher order species (s = 5.12, 4.1%) (Figure 5C, Table 2). The Prp19-Cwc2 complex purified by gel filtration confirms that Prp19 and Cwc2 form a stable higher molecular weight complex (s = 7.7, 288 kDa) (Figure 5D, Table 2). The molecular weight of the complex is intermediate between what is expected for a complex with Prp19-Cwc2 stiochiometry of 4:1 (265kDa) and 4:2 (305kDa). This is likely due to some level of dissociation of the complex during the course of gel filtration. In order to more carefully examine the stoichiometry of the complex, mixtures of separately purified Prp19 and Cwc2 were mixed and analyzed by AU. Strikingly, mixing the proteins in a 1:1 ratio produced a discrete higher order species (s = 8.04, 301.9 kDa) with a molecular weight consistent with that predicted for a 4:2 complex of Prp19-Cwc2 (305 kDa) along with free Cwc2 (s = 2.7) (Figure 5E, Table 2). Increasing the concentration of Cwc2 two-fold resulted in little change in the molecular weight of the complex (s = 8.13, 306.8kDa) even with a significant excess of Cwc2 (Figure 5F, Table 2). The interacting domains of Prp19 (WD40) and Cwc2 (C-terminus) have been previously identified (Ohi and Gould, 2002). As with the full-length protein, the domains can be co-purified (Figure S2A). While Prp19 exists as a tetramer via the central coiled-coil domain, the isolated WD40 repeat of Prp19 exists as a monomer (Figure S2B). In contrast to the full-length protein, the WD40 domain of Prp19 does not form a high-affinity stoichiometric complex with Cwc2 (Figure S2C), indicating that the stable association of Prp19 and Cwc2 requires the complex quaternary structure of the full-length proteins.

These data confirm that Prp19 and Cwc2 stably and directly associate with one another. Further, it surprisingly demonstrates that *in vitro* Prp19 and Cwc2 form a complex with stoichiometry of 4:2. This was unexpected in light of previous evidence that the WD40 domains of Prp19 are flexibly tethered to the core of the tetramer (Ohi et al., 2005) and that Cwc2 possesses no obvious repeat domain allowing a single Cwc2 protein to bind multiple copies of the WD40 repeat domain.

Protein stoichiometry *in vivo*

To validate the unexpected stoichiometry observed *in vitro*, we set out to measure the Cwc2 copy number in the NTC *in vivo*. Members of the NTC complex are highly conserved (Ohi and Gould, 2002) and experiments to determine Cef1 stoichiometry in the NTC were performed previously in *S. pombe* as opposed to *S. cerevisiae* (Ohi et al., 2005). Therefore, the experiments to characterize Prp19 and Cwc2 stiochiometries *in vivo* were also performed in *S. pombe*. Cef1 (Cdc5 in *S. pombe*) binds to the coiled-coil domain of Prp19 (Cwf8 in *S. pombe*) in a 1:4

stoichiometry (Ohi et al., 2005). Building upon these data, it is possible to determine the stoichiometry of Cwc2 (Cwf2 in *S. pombe*) within the NTC relative to Prp19 (Cwf8 in *S. pombe*) using a Cdc5 co-purification strategy. A *cdc5-TAP cwf8-Myc$_{13}$ cwf2-Myc$_{13}$* strain was constructed for this purpose. Cdc5-complexes were purified on IgG sepharose and the Myc-tagged proteins were detected by immunoblotting (Figure 6A). By quantifying the ratio of Myc-tagged proteins, we determined that Cwc2/Cwf2 was 50% as abundant as Prp19/Cwf8 indicating that there are, on average, two copies of Cwc2/Cwf2 for every four copies of Prp19/Cwf8. To confirm that Cwc2/Cwf2 is present in multiple copies in a single spliceosome, a diploid producing two different tagged alleles of Cwf2, Myc and GFP, was produced. Immunoprecipitation with one tag was able to co-immunoprecipitate the alternatively tagged protein, confirming that Cwc2/Cwf2 is indeed multicopy within a single Prp19/Cwf8 containing spliceosomal complex (Figure 6B). Thus, the *in vivo* stoichiometry of Prp19-Cwc2 is consistent with the *in vitro* analysis and refines our understanding of the architecture and network of interactions at the core of the NTC in the spliceosome.

Discussion

WD40 domains are widely distributed protein interaction modules that bind a diverse set of partner proteins, peptides and small molecules. They are distinguished by an ability to mediate protein-protein interactions ranging from small to large surface areas and weak to very strong binding affinities. The location of protein interaction sites on WD40 domains also varies considerably. Important binding sites have been found in the concave region between blades (ter Haar et al., 2000) as well as in the central pore-like region (Wu et al., 2003). Our studies define a conserved face of blade five as an essential interaction region of the Prp19 WD40 domain. The correlation between the essential role for R403 and K404 in the WD40 domain for yeast viability, and for interaction with Cwc2, strongly suggests that the physical coupling of these two proteins represents an essential function of the WD40 domain.

Although data from two-hybrid analysis are qualitative in nature, it is interesting to note that some mutations in the WD40 domain exhibit an intermediate loss of interaction with Cwc2 yet they are still able to rescue *prp19-1* when overexpressed. This suggests a critical threshold in physical interaction must be passed before a functional effect is observed. Such a threshold could arise in this case because Prp19 exists in the large NTC complex with other potential stabilizing interactions *in vivo*. Regardless of the origin of these observations, the loss of direct physical interaction between hisCwc2 and Prp19 does confirm that this site represents either the high-affinity or sole site of interaction between the proteins. However, the available data do not rule out the possibility that the functional deficiency observed for the R403 and K404 mutations may be only partly or indirectly due to loss of Cwc2 binding.

The oligomerization of Prp19 is an overriding feature of its structure and a characteristic common to many E3 ubiquitin ligases. It is also common to find that partner proteins bind to these E3 ligases in sub-stoichiometric ratios relative to the number of copies in the oligomer. The excess of binding sites over target proteins may serve to facilitate polyubiquitination of substrates. For example, half of sites binding of E2 conjugating enzymes to another U-box E3 ligase, CHIP, may contribute to the processivity of the ubiquitination or as a means to accelerate the rate of ubiquitin transfer to substrates (Qian et al., 2009). A recent study suggested that Cwc2 might function by directly binding the U6 snRNA and coupling it to the NTC complex (McGrail et al., 2009). This, combined with our finding that Cwc2 and Prp19 form a stable complex, suggests a constitutive interaction between Prp19 and Cwc2 as part of the NTC. This in turn leads us to speculate that Cwc2 functions as an adaptor for the Prp19 E3 ubiquitin ligase.

The role of Cwc2 as a substrate adaptor is consistent with the inability to demonstrate Cwc2 ubiquitination *in vivo* (data not shown). A number of other spliceosomal proteins that do not

directly contact Prp19 are attractive targets of its E3 ubiquitin ligase activity, including the key spliceosomal protein Prp8. Sontheimer and co-workers demonstrated that ubiquitin mutants or inhibitors of ubiquitin protein-protein interactions reduce U4/U6-U5 tri-snRNP levels (Bellare et al., 2006). It was further demonstrated that Prp8 is ubiquitinated within the tri-snRNP (Bellare et al., 2008). Indeed, the function of Prp8 and Prp19 has previously been linked (Brenner and Guthrie, 2005). Taken together, our data support the suggestion that Prp19 directly targets proteins within the spliceosome for ubiquitination (Maeder and Guthrie, 2008; Ohi et al., 2005) and that the ubiquitination and RNA splicing activities of Prp19 are integrally coupled.

The results reported here enable construction of a refined model of Prp19 (Figure 7). The protein is an antiparallel tetramer with four independent WD40 domains, two at each end of the tetramer available to interact with one or more targets. Direct physical interactions between Prp19 and Cwc2 and other binding partners can be simultaneous or sequentially coupled to the structural and functional rearrangements that are part of the splicing cycle. In complexes of such large modular proteins, direct competition for sites and allosteric coupling represent likely mechanisms for the structural rearrangements required for function. The C-terminus of Cwc2 interacts with the WD40 repeat domain of Prp19 in an extended conformation, yet only two molecules are bound to the Prp19 tetramer. This implies that either a single Cwc2 engages two WD40 domains at once, or that the binding of a Cwc2 molecule sterically precludes the binding of a second Cwc2 on the same side of the tetramer. In either case, two WD40 domains may be available to support additional interactions with other Prp19 target proteins. Our ongoing search for additional binding partners targeted to the WD40 domain is of interest in this context. Studies of Prp19 flexibility and dynamic protein interactions represent an exciting future direction in understanding the assembly/disassembly and changing architecture of the NTC and its role in modulating the activities of the spliceosome.

Experimental Procedures

Protein expression and purification

Prp19(165-503) was expressed and purified as a 6XHis-tag fusion protein from pET15b (Novagen) in BL21(DE3) cells (Novagen). Selenomethionine protein was expressed in 834 (DE3) cells using M9 minimal media supplemented with selenomethionine. Protein was purified using 20 mL Ni^{2+}-nitrilotriacetic acid (Ni^{2+}-NTA) columns (Qiagen). Column equilibration and wash buffer was 20 mM Tris at pH 8.0, 150 mM NaCl, 5 mM BME, and 25 mM Imidazole. Protein was eluted using a seven column volume (CV) linear gradient of imidazole from 25–500 mM. Anion exchange was used to further purify the protein using a Mono-Q 10/10 column (Amersham Pharmacia Biotech). Column equilibration and wash buffer was 20 mM Tris at pH 8.0 and 1 mM DTT. Protein was eluted using a seven CV linear gradient of NaCl from 0–500 mM. Protein was finally purified using a 5 mL HiTrap heparin column (Amersham Pharmacia Biotech). Column equilibration and wash buffer was 20 mM Tris at pH 8.0, 150 mM NaCl, and 1 mM DTT. Protein was eluted using a five CV linear gradient of NaCl from 150–1000 mM. Coexpression of Prp19 (full length and WD40) with hisCwc2 (full length and C-terminus (228-339) was accomplished using the pETDuet system (Novagen). Protein was initially purified using the IMAC protocol above. The complex of full length proteins was further purified by gel filtration using a Superdex200 HiLoad 16/60 column (GE Healthcare) equilibrated and run using a buffer of 20 mM Tris at pH 7.5, 100 mM NaCl, 0.5mM DTT. Gel filtration of two proteins together produced two discreet species: a Prp19:Cwc2 complex at high molecular weight and free Cwc2.

Crystallization of Prp19 (165-503)

Protein was buffer exchanged into 5 mM Tris at pH 7.5 and 1 mM DTT, then concentrated to 10 mg/mL. Crystals were obtained using hanging drop experiments. Single high quality crystals for native protein were obtained in four days from 2.0 M ammonium sulfate and 0.1 M Tris at pH 7.5 incubated at 18 °C. Selenomethionine labeled protein crystals were obtained in two weeks from 2.1 M ammonium sulfate and 0.1 M Tris at pH 7.5 incubated at 18 °C. Native crystals were, on average, 0.5×0.5×0.4 mm. Mature crystals were serially transferred to mother liquor containing 5, 10, 15% ethylene glycol and then flash frozen.

Crystal structure determination

Data was collected on the Gulf Coast Protein Crystallography Consortium beamline at the Center for Advanced Microstructures and Devices (CamD). Three wavelength selenomethionine MAD data were collected with wavelength selection based on a fluorescence energy scan. Data was processed using DENZO/SCALEPACK (Otwinowski and Minor, 1997). SOLVE was used to locate five of the six expected selenium sites (Terwilliger, 2003). Native crystals were obtained and a complete dataset collected. Phase extension and initial model building was accomplished using RESOLVE (Terwilliger, 2004). Additional model building and refinement was accomplished using Coot (Emsley and Cowtan, 2004) and Refmac5 (Murshudov, 1997). TLS groups for use in refinement were derived from TLSMD (Painter and Merritt, 2006). Three loops with weak electron density were not modeled, encompassing residues 296-298, 417-421, 468-471. Diffraction data and final refinement statistics are summarized in Table 1. Graphical representations of the protein structure were primarily prepared using the program MOLMOL (Koradi et al., 1996). The atomic coordinates and structure factors have been deposited in the Protein Data Bank www.pdb.org, (ID code 3LRV).

Analytical Ultracentrifugation

Sedimentation velocity experiments were run at 30,000 RPM (22 °C) on an Optima XLI (Beckman-Coulter, Fullerton, CA), with a 4-hole An60Ti rotor using double sector cells with charcoal-filled Epon centerpieces (path length 1.2 cm) and quartz windows. The velocity scans were analyzed with the program Sedfit (version 8.7) (Schuck, 2000) using 250 scans collected approximately 2 min apart. Size distributions were determined for a confidence level of p = 0.95, a resolution of n = 200, and sedimentation coefficients between 0.3 and 20 s.

Mutagenesis

Mutations in *PRP19* were made using the Stratagene Quick-Change strategy according to manufacturer s suggestions. DNA sequencing confirmed the presence of only the desired mutations.

Yeast strains and media

Schizosaccharomyces pombe strains were grown in yeast extract medium (Moreno et al., 1991) The genotypes of strains used for this study were: 1) *h− ade6-M210 leu1-32 ura4-D18* (KGY246), 2) *h− cdc5-TAP cwf8-Myc*$_{13}$ *cwf2-Myc*$_{13}$ *ade6-M210 leu1-32 ura4-D18* (KGY7136), 3) *h− cwf2-Myc*$_{13}$ *ade6-M216 leu1-32 ura4-D18 his3-D1* (KGY1431), 4) *h+ cwf2-GFP ade6-M216 leu1-32 ura4-D18* (KGY7145), and 5) *h−/h+ cwf2-GFP/cwf2-Myc*$_{13}$ *ade6-M210/ade6-M216 leu1-32/leu1-32 ura4-D18/ura4-D18* (KGY7220). Crosses were performed on glutamate medium (minimal at pH 5.6 medium lacking ammonium chloride and containing 0.01 M glutamate). The S*accharomyces cerevisiae* strain *MATa prp19-1 ura3-52 lys2-801 ade2-101 leu2-Δ1 trp1-Δ1* was grown either in synthetic minimal medium with the appropriate nutritional supplements or yeast extract-peptone-dextrose (Guthrie and Fink, 1991). Transformations were performed by the lithium acetate method.

Analysis of mutant proteins

The *prp19-1* strain (KGY1811) was transformed with pRS415 plasmids containing wild type *PRP19* and the indicated mutations under the control of the *GAL1* promoter. The ability of *prp19* mutations to rescue growth of the *prp19-1* strain was scored by the ability of transformants recovered at 25 °C to form colonies at 36 °C after 3 days and is indicated by + or −.

Sequences encoding full length Prp19 and Prp19 mutations were incorporated into the pGAD vector to test against amino acids 228-339 of Cwc2 in the pGBT vector (James et al., 1996; Ohi and Gould, 2002) and transformed into *S. cerevisiae* strain PJ69-4A. β-galactosidase reporter enzyme activity in the two-hybrid strains was measured using the Galacto-Star™ chemiluminescent reporter assay system according to the manufacturer s instructions (Applied Biosystems, Foster City, California), with the exception that cells were lysed by glass bead disruption. Each sample was measured in triplicate. Reporter assays were recorded on a Multi-Detection Microplate Reader from Bio-TEK Instruments, Inc., Vermont, USA.

Immunoprecipitations and immunoblotting

Native protein lysates were prepared from approximately 3×10^8 cells as detailed previously (Gould et al., 1991). To the resultant 1.0 ml of protein lysate, 30 μl of 1:1 slurry of IgG beads, 2 μg anti-Myc antibody (9E10) or 2 μg anti-GFP antibody (Roche) was added to precipitate tagged proteins. Immunoprecipitates were resolved on Novex NuPAGE 4–12% bis-Tris gels using NuPAGE MOPS SDS running buffer and transferred by electroblotting to polyvinylidene difluoride membrane (ImmobilonP; Millipore Corp., Bedford, MA). Proteins were detected by incubating with 2 μg/ml anti-Myc antibody (9E10) or 2 μg/ml anti-GFP antibody (Roche) and primary antibodies were detected with secondary antibodies coupled to Alexa Fluor 680 (Invitrogen). Blots were scanned on an Odyssey machine (LI-COR). Quantification of protein intensities was performed using Odyssey version 1.2.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Henry Bellamy for help in collection of x-ray diffraction data and Drs. Michael Oldham and Marcia Newcomer for valuable advice and discussions. This work was supported by NIH training grant T32GM08320, by NIH operating grants P20RR020171 (C.W.V.K.), 1DP2OD004483 (M.D.O.) and R01GM075156 (W.J.C.), and by the Howard Hughes Medical Institute (K.L.G.).

References

Aravind L, Koonin EV. The U box is a modified RING finger - a common domain in ubiquitination. Curr Biol 2000;10:R132–134. [PubMed: 10704423]

Bellare P, Kutach AK, Rines AK, Guthrie C, Sontheimer EJ. Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p. RNA 2006;12:292–302. [PubMed: 16428608]

Bellare P, Small EC, Huang X, Wohlschlegel JA, Staley JP, Sontheimer EJ. A role for ubiquitin in the spliceosome assembly pathway. Nat Struct Mol Biol 2008;15:444–451. [PubMed: 18425143]

Brenner TJ, Guthrie C. Genetic analysis reveals a role for the C terminus of the Saccharomyces cerevisiae GTPase Snu114 during spliceosome activation. Genetics 2005;170:1063–1080. [PubMed: 15911574]

Chan SP, Kao DI, Tsai WY, Cheng SC. The Prp19p-associated complex in spliceosome activation. Science 2003;302:279–282. [PubMed: 12970570]

Chan SP, Cheng SC. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. J Biol Chem 2005;280:31190–31199. [PubMed: 15994330]

Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 2004;60:2126–2132. [PubMed: 15572765]

Freemont PS. RING for destruction? Curr Biol 2000;10:R84–87. [PubMed: 10662664]

Gould KL, Moreno S, Owen DJ, Sazer S, Nurse P. Phosphorylation at Thr167 is required for Schizosaccharomyces pombe p34cdc2 function. Embo J 1991;10:3297–3309. [PubMed: 1655416]

Guthrie, C.; Fink, GR., editors. Guide to yeast genetics and molecular biology. San Diego, CA: Academic Press, Inc.; 1991.

Hatakeyama S, Nakayama KI. U-box proteins as a new family of ubiquitin ligases. Biochem Biophys Res Commun 2003;302:635–645. [PubMed: 12646216]

James P, Halladay J, Craig EA. Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. Genetics 1996;144:1425–1436. [PubMed: 8978031]

Koradi R, Billeter M, Wuthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graphics 1996;14:51–55.

Madrona AY, Wilson DK. The structure of Ski8p, a protein regulating mRNA degradation: Implications for WD protein structure. Protein Sci 2004;13:1557–1565. [PubMed: 15152089]

Maeder C, Guthrie C. Modifications target spliceosome dynamics. Nat Struct Mol Biol 2008;15:426–428. [PubMed: 18461042]

Makarova OV, Makarov EM, Urlaub H, Will CL, Gentzel M, Wilm M, Luhrmann R. A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. Embo J 2004;23:2381–2391. [PubMed: 15175653]

McGrail JC, Krause A, O'Keefe RT. The RNA binding protein Cwc2 interacts directly with the U6 snRNA to link the nineteen complex to the spliceosome during pre-mRNA splicing. Nucleic Acids Res 2009;37:4205–4217. [PubMed: 19435883]

Moreno S, Klar A, Nurse P. Molecular genetic analysis of fission yeast Schizosaccharomyces pombe. Methods Enzymol 1991;194:795–823. [PubMed: 2005825]

Murshudov GN. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 1997;53:240–255. [PubMed: 15299926]

Ohi MD, Gould KL. Characterization of interactions among the Cef1p-Prp19p-associated splicing complex. Rna 2002;8:798–815. [PubMed: 12088152]

Ohi MD, Vander Kooi CW, Rosenberg JA, Chazin WJ, Gould KL. Structural insights into the U-box, a domain associated with multi-ubiquitination. Nat Struct Biol 2003;10:250–255. [PubMed: 12627222]

Ohi MD, Vander Kooi CW, Rosenberg JA, Ren L, Hirsch JP, Chazin WJ, Walz T, Gould KL. Structural and functional analysis of essential pre-mRNA splicing factor Prp19p. Mol Cell Biol 2005;25:451–460. [PubMed: 15601865]

Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol 1997;276:307–326.

Painter J, Merritt EA. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. Acta Crystallogr D Biol Crystallogr 2006;62:439–450. [PubMed: 16552146]

Patterson C. A new gun in town: the U box is a ubiquitin ligase domain. Sci STKE 2002;2002:PE4. [PubMed: 11805346]

Pickart CM. Mechanisms underlying ubiquitination. Annu Rev Biochem 2001;70:503–533. [PubMed: 11395416]

Qian SB, Waldron L, Choudhary N, Klevit RE, Chazin WJ, Patterson C. Engineering a ubiquitin ligase reveals conformational flexibility required for ubiquitin transfer. J Biol Chem. 2009

Schuck P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J 2000;78:1606–1619. [PubMed: 10692345]

Tarn WY, Hsu CH, Huang KT, Chen HR, Kao HY, Lee KR, Cheng SC. Functional association of essential splicing factor(s) with PRP19 in a protein complex. Embo J 1994;13:2421–2431. [PubMed: 8194532]

Tarn WY, Lee KR, Cheng SC. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. Proc Natl Acad Sci U S A 1993;90:10821–10825. [PubMed: 8248176]

ter Haar E, Harrison SC, Kirchhausen T. Peptide-in-groove interactions link target proteins to the beta-propeller of clathrin. Proc Natl Acad Sci U S A 2000;97:1096–1100. [PubMed: 10655490]

ter Haar E, Musacchio A, Harrison SC, Kirchhausen T. Atomic structure of clathrin: a beta propeller terminal domain joins an alpha zigzag linker. Cell 1998;95:563–573. [PubMed: 9827808]

Terwilliger T. SOLVE and RESOLVE: automated structure solution, density modification and model building. J Synchrotron Radiat 2004;11:49–52. [PubMed: 14646132]

Terwilliger TC. SOLVE and RESOLVE: automated structure solution and density modification. Methods Enzymol 2003;374:22–37. [PubMed: 14696367]

Vander Kooi CW, Ohi MD, Rosenberg JA, Oldham ML, Newcomer ME, Gould KL, Chazin WJ. The Prp19 U-box crystal structure suggests a common dimeric architecture for a class of oligomeric E3 ubiquitin ligases. Biochemistry 2006;45:121–130. [PubMed: 16388587]

Wu G, Xu G, Schulman BA, Jeffrey PD, Harper JW, Pavletich NP. Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. Mol Cell 2003;11:1445–1456. [PubMed: 12820959]

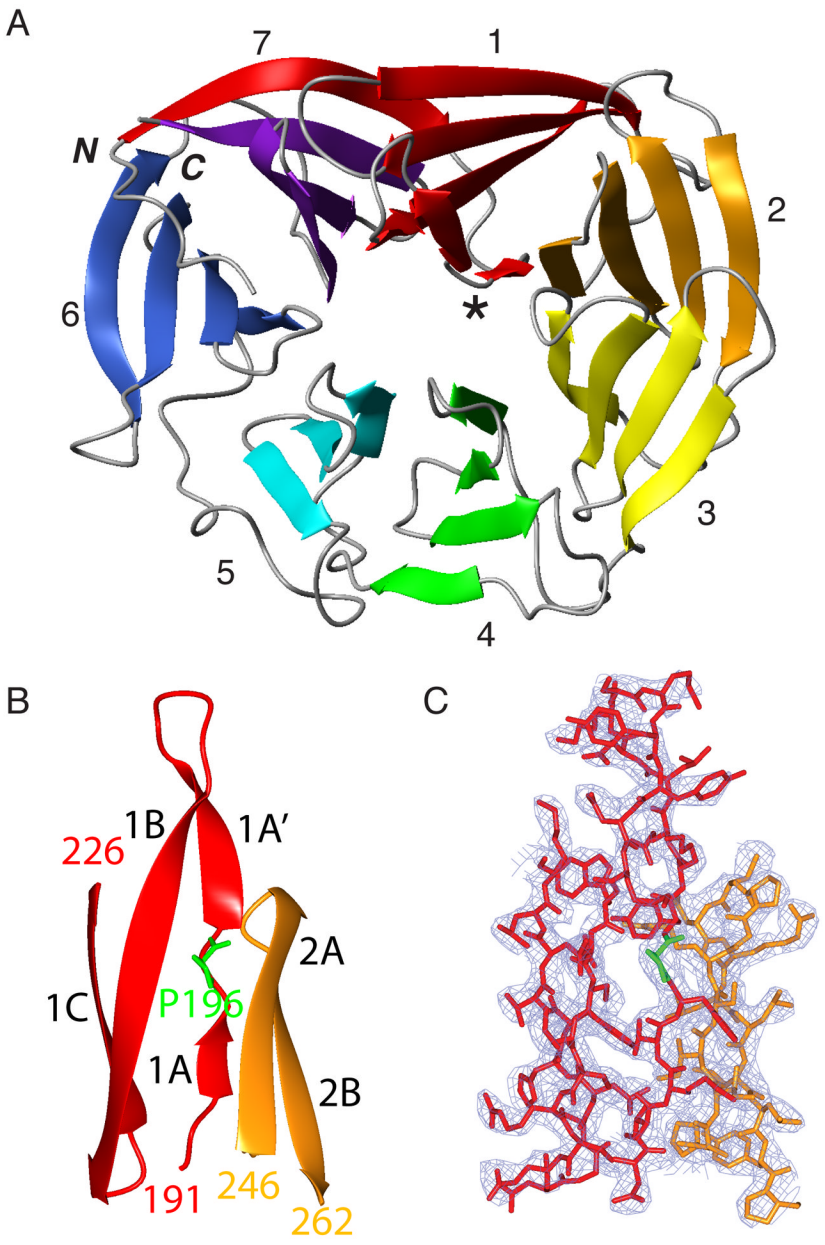

Figure 1.

Structure of the WD40 repeat domain of Prp19. A) Ribbon diagram of the WD40 repeat domain with the seven blades numbered. The final blade s fourth sheet is contributed by the region N-terminal to the first blade (red). The atypical first sheet of blade one is highlighted with a *. B) Ribbon diagram of the first three strands of blade one in red (1A–1C, residues 191-226) and the first two of blade two in orange (2A–2B, residues 246-262). P196 (green) causes a break in the first sheet of blade one allowing interaction with both blades one and two via the 1A′ and 1A sections of the first sheet, respectively. C) Stick model and 2Fo-Fc electron density map contoured at 0.9σ for the same region as in 1B.

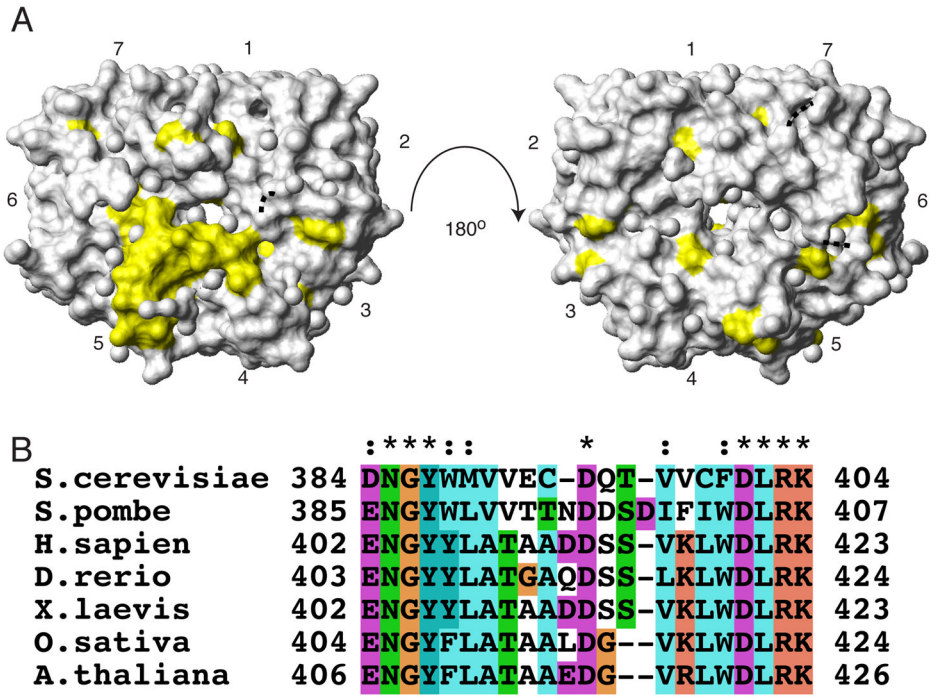

Figure 2.

Identification of conserved surface residues. A) Mapping of conserved residues on a surface representation of the structure (yellow) reveals a distinct patch formed primarily by residues from the fifth blade. The three short loops not modeled in the structure are represented by dashed black lines. B) Multiple sequence alignment of the conserved residues found in the fifth blade of the WD40 domain.

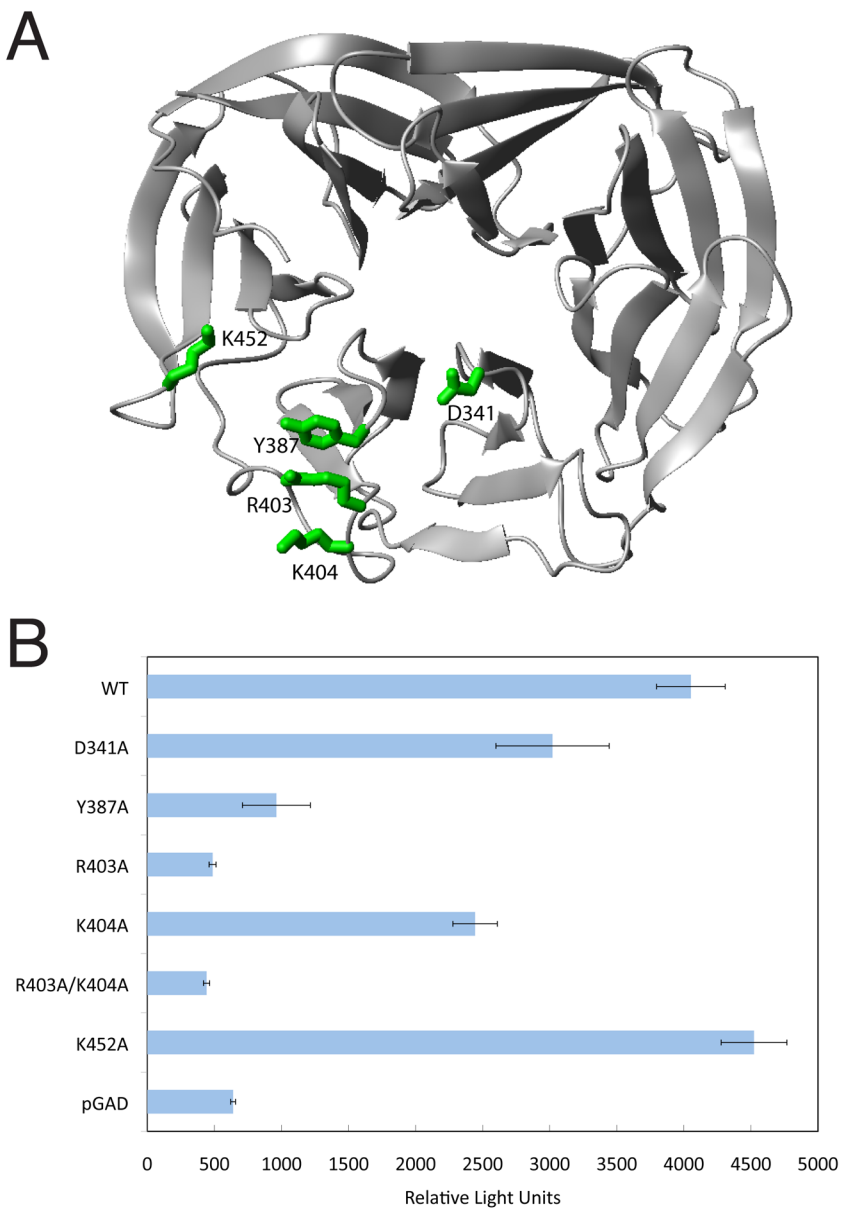

Figure 3.

Location of functionally and physically important residues in the WD40 repeat domain of Prp19. A) Mapping of mutated residues onto the structure of Prp19. B) Interaction between Prp19 mutant proteins and Cwc2 assessed in β-galactosidase activity (relative light units).

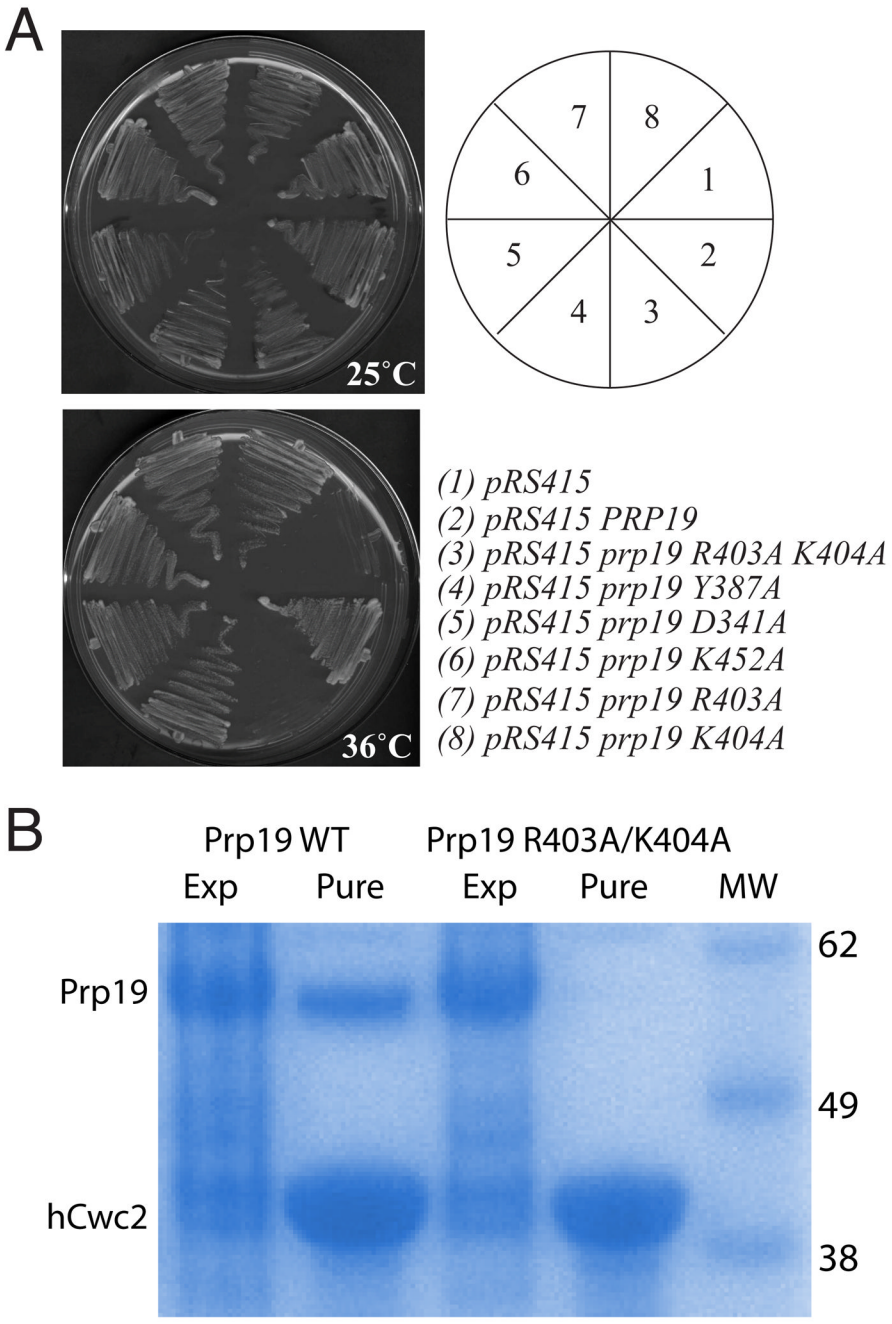

Figure 4.

Identification of functionally important residues in the WD40 repeat domain of Prp19. A) Ability of Prp19 protein with mutations in the WD40 repeat domain to rescue growth of the *prp19-1* temperature sensitive strain. The prp19-1 strain (KGY1811) was transformed with pRS415 plasmids containing wild type PRP19 and the indicated mutations under the control of the GAL1 promoter. Transformants were struck to selective plates and grown at 25 °C or 36 °C to determine their ability to rescue growth of the prp19-1 strain. B) Co-expression and purification of Prp19 and hisCwc2 demonstrate that the R403A/K404A mutant protein is expressed as wild-type but results in loss of physical interaction with Cwc2. See also Figure S1.

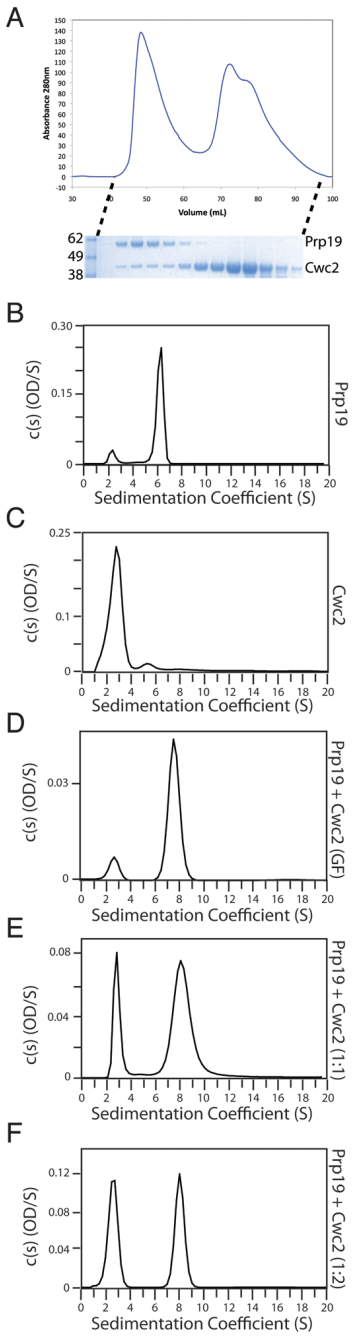

Figure 5.

Prp19:hisCwc2 forms a stable complex with 4:2 stoichiometry. A) IMAC co-purified Prp19/hisCwc2 run on gel filtration resolves a complex of the two proteins along with free hisCwc2. Sedimentation velocity data of B) Prp19, C) Cwc2, D) The complex of Prp19 and Cwc2 purified by gel filtration, E) Mixture of Prp19:Cwc2 in a 1:1 molar ratio, F) Mixture of Prp19:Cwc2 in a 1:2 molar ratio. Prp19 concentrations were constant, 10μM, with Cwc2 concentrations varied to the indicated molar ratio. AU experiments were conducted at 22 °C at a speed of 30,000 rpm and concentration profiles measured at 280 nm. See also Figure S2.

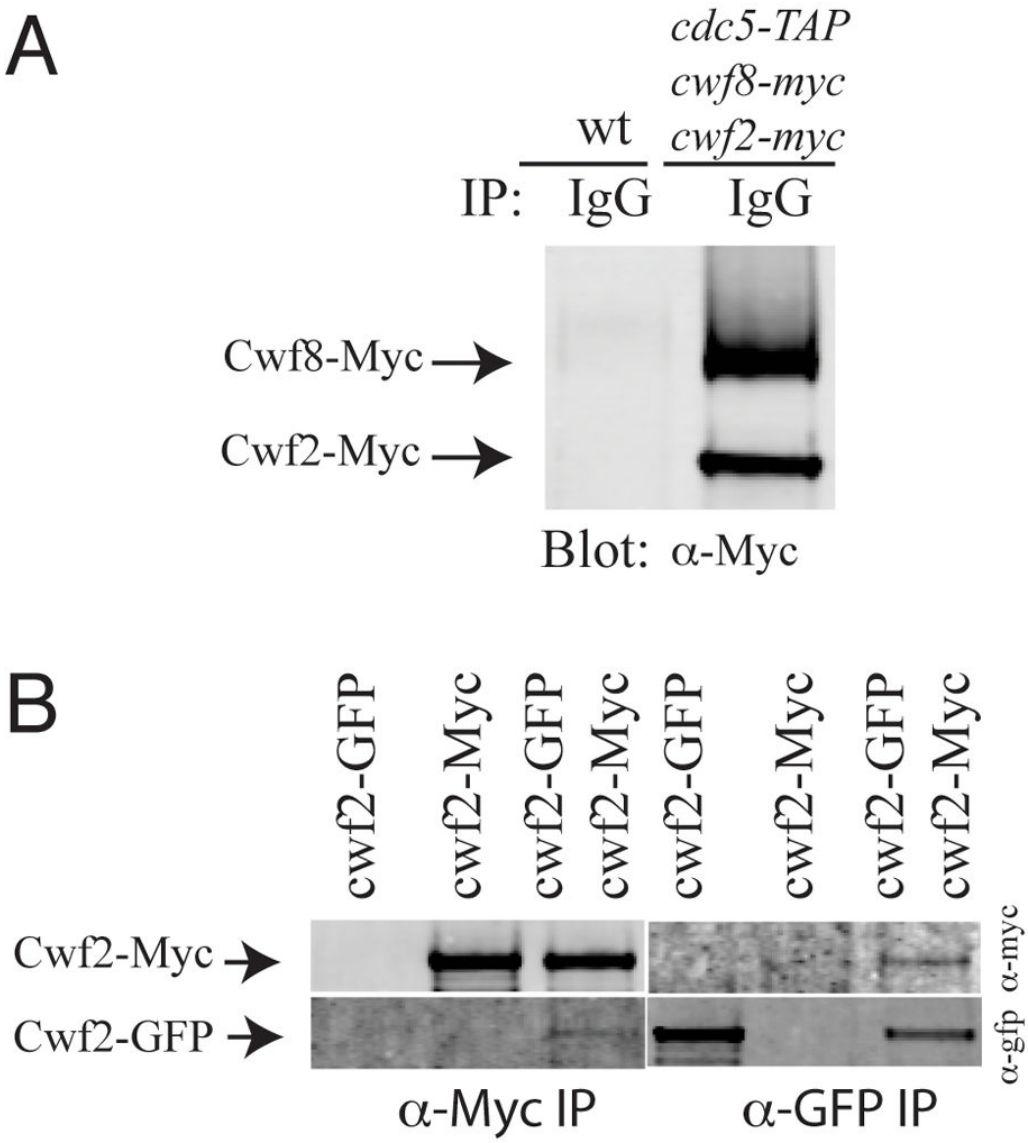

Figure 6.

Cwc2/Cwf2 is present in the NTC in a 1:2 ratio with Prp19/Cwf8. A) An anti-myc immunoblot of an IgG immunoprecipitate from either a wild-type *S. pombe* strain (KGY246) or from *h− cdc5-TAP cwf8-Myc*$_{13}$ *cwf2-Myc*$_{13}$ *ade6-M210 leu1-32 ura4-D18* (KGY7136). B) Anti-Myc (upper panel) and anti-GFP immunoblots of immunoprecipitates (IP) from *cwf2-Myc*$_{13}$, *cwf2-GFP*, and *cwf2-Myc*$_{13}$ *cwf2-GFP* strains. Immunoprecipitations were performed with anti-Myc or anti-GFP antibodies.

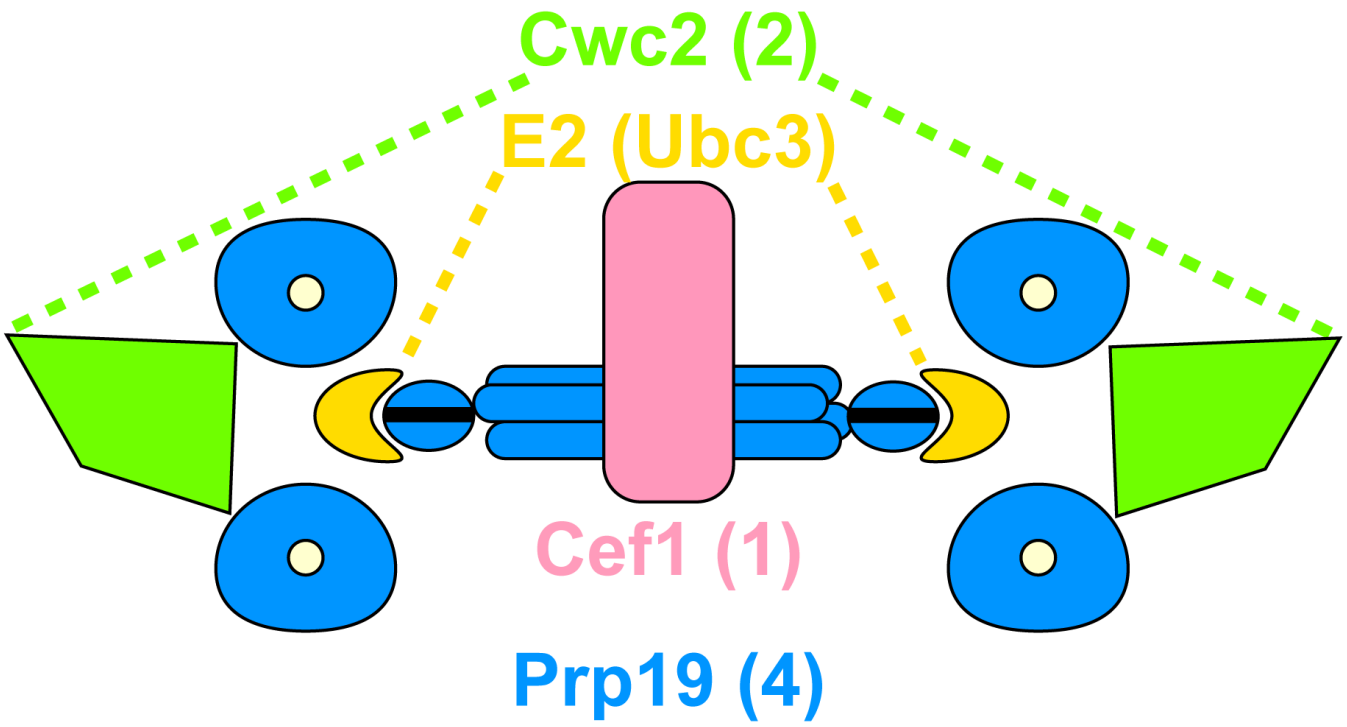

Figure 7.
Model of the Prp19 tetramer bound to Cwc2 showing the overall architecture and stoichiometry of the proteins in the complex.

Table 1

Crystallographic data and refinement statistics

	Native	**SeMet**		
		Edge	**Peak**	**Remote**
Beamline	PDX	PDX		
Space group	$P4_22_12$	$P4_22_12$		
Wavelength	1.381	0.9797	0.9793	0.9465
Unit-cell parameters				
a (Å)	83.110	83.426		
b (Å)	83.110	83.426		
c (Å)	126.641	126.307		
Unique reflections	14158	14092	14059	14063
Completeness (%)	99.2 (99.6)	99.9 (100.0)	99.8 (100.0)	99.8 (100.0)
Resolution (Å)	2.6 (2.60–2.69)	3.2 (3.20–3.31)	3.2 (3.20–3.31)	3.2 (3.20–3.31)
R_{merge} (%)	8.0 (44.5)	12.2 (43.6)	8.4 (27.1)	8.7 (29.0)
Redundancy	6.3 (5.5)	5.5 (5.0)	5.5 (5.3)	5.6 (5.3)
I/σ (I)	16 (3.5)	12.1 (3.6)	18.9 (6.8)	21.0 (7.3)
FOM (after density modification)		0.48 (0.62)		
Resolution Limits (Å)	25.0–2.60		Ramachandran	
Number of reflections used in refinement	13408		Most Favored	91.9
			Add. Allowed	8.1
Number of reflections used to compute R_{free}	712		Additionally	0.0
			Disallowed	0.0
R (R_{free})	21.9 (26.9)			
			RMS deviation	
# solvent molecules	66		Bond, Å	0.011
# sulfate molecules	3		Angle, °	1.397

Table 2

Sedimentation velocity AU Data summary

	S values	Predicted MW (kDa)	Frictional Ratio	r.m.s.d.
Prp19:			1.8	0.0057
Peak 1	6.10 (85%)	211.6		
Peak 2	2.13 (14%)	41.7		
hisCwc2			1.65	0.0053
Peak 1	5.12 (4.1%)	118.4		
Peak 2	2.67 (60.4%)	41.7		
Prp19: hisCwc2 (Gel Filtration pure)			2.03	0.0050
Peak 1	7.75 (30.8%)	287.0		
Peak 2	2.77 (9%)	36.3		
Prp19: hisCwc2 (1:1 molar concentration)			1.96	0.0049
Peak 1	8.04 (54.2%)	301.9		
Peak 2	2.7 (26.7%)	43.8		
Prp19: hisCwc2 (1:2 molar concentration)			1. 94	0.0048
Peak 1	8.13 (42%)	306.8		
Peak 2	2.63 (42.5%)	44.8		