Structural analysis of the yeast exosome Rrp6p-Rrp47p complex by small-angle x-ray scattering

Emil Dedic[a, c], Paulina Seweryn[a, c], Anette Thyssen Jonstrup[a, c], Rasmus Koch Flygaard[a, c], Natalya U. Fedosova[d], Søren Vrønning Hoffmann[e], Thomas Boesen[b, c], and Ditlev Egeskov Brodersen[a, c, *]

[a]Centre for mRNP Biogenesis and Metabolism and [b]Centre for Membrane Pumps in Cells and Disease – PUMPKIN, [c]Department of Molecular Biology and Genetics, Gustav Wieds Vej 10c, [d]Department of Biomedicine, Ole Worms Allé 6, and [e]Institute for Storage Ring Facilities (ISA), Department of Physics and Astronomy, Ny Munkegade 120, Aarhus University, DK-8000 Aarhus C, Denmark.

[*]To whom correspondence should be addressed. Tel: +45 2166 9001; Fax: +45 8612 3178; Email: deb@mb.au.dk.

Keywords

Exosome; 3'-5' nuclease; RNA turnover; RNA processing; snoRNA; ribosomal RNA

Abstract

The RNase D-type 3'-5' exonuclease Rrp6p from *Saccharomyces cerevisiae* is a nuclear-specific cofactor of the RNA exosome and associates *in vivo* with Rrp47p (Lrp1p). Here, we show using biochemistry and small-angle x-ray scattering (SAXS) that Rrp6p and Rrp47p associate into a stable, heterodimeric complex with an elongated shape consistent with binding of Rrp47p to the nuclease domain and opposite of the HRDC domain of Rrp6p. Rrp47p reduces the exonucleolytic activity of Rrp6p on both single-stranded and structured RNA substrates without significantly altering the affinity towards RNA or the ability of Rrp6p to degrade RNA secondary structure.

Abbreviations

SAXS, small-angle x-ray scattering; HRDC, helicase and RNase D C-terminal domain; snoRNP, small nucleolar ribonucleoprotein; SRCD, synchrotron radiation circular dichroism; RALS, right-angle light scattering; MALS, multi-angle light scattering; FAM, 5' carboxy-flourescein; SEC, size-exclusion chromatography; EMSA, electrophoretic mobility shift assay; MW, Molecular weight.

Introduction

The eukaryotic RNA exosome is involved in a multitude of processes including maturation of stable RNAs, turnover of messenger RNA as well as scavenging of aberrant, poly-adenylated nuclear transcripts and short transcripts arising from pervasive transcription [1,2]. Structurally, the exosome consists of a ring of six proteins homologous to bacterial RNase PH and three RNA-binding "cap" proteins, known as the "exosome core" [3,4]. In *Saccharomyces cerevisiae*, the exosome ring has lost catalytic activity and ubiquitously associates with Rrp44p/Dis3p, which functions as both 3'-5' exonuclease and as endonuclease [5,6]. In the yeast nucleus, the complex additionally binds the RNase D-type 3'-5' exonuclease, Rrp6p, which is required for stable RNA processing and turnover specific to this compartment [7,8,9]. Crystal structures currently exist of the human exosome core as well as the core bound to Rrp44p and a C-terminal peptide from Rrp6p showing that Rrp6p binds at the "top" of the core, opposite Rrp44p [3,10]. In addition, crystal structures exist of the isolated exonuclease core domains of yeast, Trypanosoma, and human Rrp6p [11,12,13].

Rrp6p consists of an N-terminal PMC2NT region and a DEDD-type 3'-5' exonuclease domain homologous to bacterial RNase D and at least one Helicase and RNase D C-terminal (HRDC) domain [9]. The available crystal structures of Rrp6p all encompass a part of the N-terminus unique to eukaryotes, the exonuclease core, and one HRDC domain [11,12,13]. These structures have revealed a compact nuclease domain with the N-terminus wrapped around it and a more loosely associated HRDC domain, homologous to the first of two HRDC domains found in *E. coli* RNase D [14]. *In vitro*, isolated Rrp6p is active as a 3'-5' exonuclease on a range of both single-stranded and structured RNA substrates with a subtle preference for poly-A RNA [3,7,12]. Via its N-terminal PMC2NT region, Rrp6p interacts with an additional exosome cofactor, Rrp47p (Lrp1p), which exists as a multimer in isolation,

but appears to associate with Rrp6p in a 1:1 complex [11,15]. Furthermore, yeast Rrp47p has been shown to specifically bind double-stranded nucleic acids *in vitro* but sequence analysis has so far failed to identify any known RNA binding domains [16,17]. Rrp47p is present in pull-down experiments targeting core exosome subunits indicating that it forms a natural part of at least a population of nuclear exosomes, but the details of the interplay between the exosome core, Rrp6p, and Rrp47p are not yet understood [18].

Deletion of either Rrp6p or Rrp47p results in accumulation of 3' extended forms of both 5.8S ribosomal RNA and snoRNAs in yeast [18,19]. Moreover, the PMC2NT region of Rrp6p is required for correct RNA processing as well as for normal levels of Rrp47p strongly supporting that the two proteins are naturally bound to each other *in vivo* [15,17,20]. Studies of the organisation of Rrp47p have revealed that the protein consists of two domains, an N-terminal Sas10/C1D-type domain spanning roughly the first 100 residues, required for binding to Rrp6p, and a C-terminal region shown to be involved in maturation of box C/D small nucleolar ribonucleoprotein (snoRNP) particles [16,21]. Finally, the extreme C-terminus of Rrp47p contains a concentrated patch of basic residues, which appears to be responsible for the majority of the RNA affinity observed *in vitro*. Interestingly, the N-terminal Sas10/C1D domain of Rrp47p is sufficient to complement nearly all of the phenotypes associated with deletion of Rrp47p in yeast [21].

In this paper, we map the interaction between yeast Rrp47p and the Rrp6p PMC2NT domain and show that complex formation leads to structural rearrangements in both proteins. Using small-angle x-ray scattering (SAXS), we determine envelopes of isolated Rrp6p as well as the Rrp6p-Rrp47p complex. Finally, we present functional evidence that Rrp47p finely modulates the activity of Rrp6p on both single-stranded and structured RNA substrates as well as affects degradation through double-stranded regions.

Materials and Methods

Protein expression and purification. Briefly, *S. cerevisiae rrp47/lrp1* and *rrp6* were amplified from yeast genomic DNA and inserted into pET-30 Ek/LIC to create a single construct encoding both full length Rrp47p$^{1-184}$ (residues 1-184) and Rrp6p$^{12-536}$-His$_6$. Rrp47p-Rrp6p$^{12-536, D238A}$ was constructed by site-directed mutagenesis. Expression was carried out in *E. coli* BL21 (DE3) RIL (Novagen) by overnight expression at 18°C and the complex was purified by tandem Ni^{2+}-chelating agarose chromatography (Qiagen) with Tev protease cleavage. Ion exchange chromatography using Source 15Q and Source 15S (GE Healthcare) allowed purification of both the isolated Rrp6p$^{12-536}$ (monomer and dimer) and Rrp47p$^{1-184}$, and the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex. Finally, all samples were purified on a Superdex 200 10/300 GL size-exclusion column (GE Healthcare) in 20 mM Tris-HCl, pH 8.0, 100 mM KCl, and 5 mM BME. See Supplementary Information for further details.

Multi-angle light scattering, SRCD, and SAXS. Multi-angle light scattering data were collected on a Zetasizer µV light scattering (Malvern Instruments) or a MiniDAWN (Wyatt Technology) system connected in-line with gel filtration. MW calculation was achieved using OmniSEC 4.7.0 (Malvern Instruments) or ASTRA (Wyatt Technology) using dn/dc=0.187 in all cases. Synchrotron radiation circular dichroism (SRCD) spectra were collected at beam line CD1 at the ASTRID storage ring (ISA, Aarhus University, Denmark). Samples were measured in triplicate close to 1 mg/ml in 50 mM HPO$_4^{2-}$, pH 7.5, averaged, baseline subtracted using a buffer spectrum, and mildly smoothed using the CDtool software [22]. Synchrotron radiation SAXS data were collected at the X33 at DORIS/DESY (Hamburg, Germany) and I911-4 at MAX-Lab (Lund, Sweden) in 20 mM Tris-HCl, pH 7.5, 100 mM KCl, and 5 mM BME at 10°C. Data were processed using ATSAS [23], *ab-initio* modelling was performed using DAMMIF [24], and the most probable model was identified in

DAMAVER [25]. Docking of the crystal structure was done using the SITUS package [26] and presented using Chimera [27].

In vitro RNA degradation assays and EMSA. In vitro reactions contained 10 nM 5' carboxy-flourescein (FAM) labelled RNA substrate (Invitrogen/Dharmacon) and 30 nM Rrp6p monomer or Rrp6p-Rrp47p complex in 12.5 mM Tris–HCl, pH 8.0, 60 mM KCl, 1.2 mM MgCl$_2$, 2.5 mM BME, and 150 nM BSA. After incubation at 30°C, reactions were quenched with loading dye, resolved on denaturing 18% polyacrylamide gels in 1x TBE, and visualised using a Typhoon Trio Imager (GE Healthcare). Reactions for EMSA contained 10 nM synthetic 5' FAM-labelled RNA oligo in 17.5 mM Tris-HCl, pH 8.0, 50 mM KCl, 10% glycerol, 1 mM MgCl$_2$, and 2.1 mM BME, and varying concentrations of Rrp6p$^{12\text{-}536,\ D238A}$ or its complex with Rrp47p$^{1\text{-}184}$. After incubation at room temperature for 30 min, reactions were mixed with loading dye and separated on native 8% polyacrylamide gels in 1x Tris-glycine buffer. Bands were integrated using TotalLab Quant. See Supplementary Methods for RNA sequences.

Results and Discussion

Rrp6p-Rrp47p complex formation is associated with structural changes. S. cerevisiae Rrp6p$^{12\text{-}536}$ (residues 12-536 of 733) and Rrp47p (full length, residues 1-184) were co-expressed in *E. coli* BL21 (DE3) RIL and purified using a cleavable, C-terminal His-tag on Rrp6p (see Supplementary Methods for details and Figure S1A). The purification also yielded pure samples of the isolated proteins (Figure 1A and 1B). Consistent with previous studies, we find that the level of staining of the two proteins in pure samples of the complex suggests formation of a stoichiometric 1:1 complex between Rrp6p and Rrp47p with a theoretical molar mass of 83.2 kDa (Figure 1A) [11,15]. For molecular mass determination,

we analysed the samples by size exclusion chromatography coupled to a multi-angle static light scattering (SEC/MALS) (Figure 1C and Table 1). The Rrp6p-Rrp47p complex elutes as a monodisperse peak with a MALS MW of 84 kDa, which is significantly smaller than the apparent MW estimated from the elution volume (~130 kDa) suggesting that the complex has a non-globular shape. Isolated Rrp47p elutes as an oligomeric species with a MALS molecular weight of app. 93 kDa, corresponding to a 4-6 mer as previously described [15,17]. Isolated Rrp6p can be separated into two distinct forms, a monomer and a dimer, which elute at two different volumes (Figure 1C and Table 1). N-terminal sequencing reveals that the Rrp6p dimer is intact while the monomer has been cleaved near the end of the PMC2NT domain giving rise to a 85-536 fragment (Figure 1B). This suggests that the Rrp6p PMC2NT domain is not stable in isolation, much like Rrp47p, and that interaction is required for the stability of both proteins. This is consistent with observations *in vivo* that cellular levels of Rrp47p drop when Rrp6p is depleted [17].

Finally, the secondary structure of both the isolated components and the complex was measured using synchrotron radiation circular dichroism (SRCD, Figure 1D) [28]. Isolated Rrp6p and Rrp47p display a high α-helical content (33-35%) and lower amount of β-sheet (16-19%) consistent with crystal structure of Rrp6p [13] and predictions for Rrp47p [17]. The Rrp6p-Rrp47p complex displays a significantly higher α-helical content (42%) than either of the two constituent proteins indicating that structural differences occur upon complex formation. These values are conservative as the isolated proteins are already partially stabilised by oligomerisation and are consistent with the requirement of each protein for their binding partner for stability *in vivo* [20,29].

The Rrp47p Sas10/C1D domain interacts with a conserved N-terminal region of Rrp6p. The N-terminal Sas10/C1D domain of Rrp47p (residues 1-100) is known to interact directly with

the PMC2NT domain of Rrp6p (Figure S1A) [17,21]. To analyse the interaction in more detail, we constructed several truncated versions of both proteins and asked usind His pull-downs of Rrp6p whether the complex was still intact (Figure 1E and 1F). Pull-down of Rrp6p$^{12-536}$ co-expressed with Rrp47p$^{1-100}$ confirmed that the Sas10/C1D domain alone is sufficient for stable complex formation (Figure 1E). Further truncation of Rrp47p, however, caused the complex to break, showing that the intact Sas10/C1D domain is required for complex formation (Figure 1E, lanes 20-100 and 10-89). The N-terminus of Rrp47p is not highly conserved but is predicted to be helical and therefore likely disturbed by the truncations (Figure S1B and S2). At the end of the Sas10/C1D domain, conservation stretches up to residue 93, consistent with the observation that Rrp47p$^{10-89}$ is not able to bind Rrp6p. To map the interaction on the Rrp6p side, truncations were made to produce variants of Rrp6p starting at residue 45, 65, 75, and 85 co-expressed with Rrp47p$^{1-100}$ (Figure 1F). Of these, Rrp47p bound to all except Rrp6p$^{85-536}$ suggesting that the interaction with Rrp47p takes place via the 75-129 region of Rrp6p, which is more conserved than the very N-terminus (Figure S1B and S2). In summary, we conclude that the minimal interaction regions of Rrp6p and Rrp47p involve the PMC2NT domain of Rrp6p starting at residue 75 and the core Sas10/C1D domain of Rrp47p (residues 1-100).

Rrp47p binds on top of the Rrp6p exonuclease core domain. We next analysed Rrp6p on both dimer and monomer forms as well as the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ (full-length Rrp47p), Rrp6p$^{12-536}$-Rrp47p$^{1-100}$ (Sas10/C1D domain), and Rrp6p$^{75-536}$-Rrp47p$^{1-100}$ (minimal interaction region) complexes by synchrotron small-angle x-ray scattering (SAXS) and used the data for *ab initio* modelling using the dummy-atom approach (Figure S3) [24]. The SAXS envelope obtained for the Rrp6p monomer is compact and fits well to the crystal structure of yeast Rrp6p (PDB 2HBJ, Figure 2A) [13]. Extra density on the side and on top of the exonuclease core accounts for the part of the PMC2NT domain not present in the crystal

structure. Likewise, modelling of the Rrp6p dimer (residues 12-536) resulted in a symmetric, flat shape that allows docking of two copies of Rrp6p including PMC2NT domains (Figure S4A) [13]. The envelope calculated for the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex is also elongated and has approximately the same dimensions (D_{max}=141 Å) as the Rrp6p dimer (D_{max} = 136 Å, Figure 2B, S4B and Table T1), consistent with the co-elution of the two samples during gel filtration. However, the complex envelope is not symmetrical and consists of one extended (bottom) and one more globular (top) end (Figure 2B). The exonuclease domain of Rrp6p fits accurately into the envelope with the HRDC domain at the bottom and with density next to and on top of the exonuclease core to account for the PMC2NT domain and Rrp47p. The SAXS envelopes calculated for Rrp6p$^{12-536}$-Rrp47p$^{1-100}$ and Rrp6p$^{75-536}$-Rrp47p$^{1-100}$ have very similar shapes to the more intact complex and the extension assigned to the HRDC domain of Rrp6p is consistently present in all models (Figures 2C-D and S4B). Furthermore, the density assigned to the PMC2NT domain/Rrp47p gets progressively smaller as the complex is truncated, which is also reflected in progressively smaller Porod volumes (165, 147, and 130 nm^3, respectively), while D_{max} remains almost constant (Figure S4B and Table T1). In conclusion, we have shown that the Rrp6p-Rrp47p complex has an elongated shape consistent with gel filtration chromatography, significantly larger than monomeric Rrp6p but with similar dimensions to dimeric Rrp6p. Docking of the crystal structure of Rrp6p suggests that Rrp47p binds on top of the exonuclease domain, in contact with the PMC2NT and opposite of the HRDC domain. Although this position does not preclude a direct interaction of Rrp47p with RNA, Rrp47p is relatively distant from the active site on Rrp6p in this model. As Rrp6p is known to interact with the exosome via its C-terminus, this model further suggests that Rrp47p binds away from the exosome core (Figure 4) [10]. RNA substrates directed to the exosome can then either enter Rrp6p directly or pass through the core for degradation by Rrp44p.

Rrp47p subtly affects the activity of Rrp6p on both single-stranded and structured RNA. The activities of yeast and human Rrp6p have previously been characterised *in vitro* showing that the enzymes are able to degrade a range of RNA substrates in a 3'-5' directed fashion with a subtle preference for poly-A stretches, reflecting the role of Rrp6p in scavenging nuclear poly-adenylated transcripts *in vivo* [3,7,12]. Recent functional studies of Trypanosoma RRP6 bound to the Rrp47p orthologue, EAP3, showed that the co-factor slows down the nuclease marginally, but doesn't affect the specificity or ability of the enzyme to degrade double-stranded RNA [11]. To investigate activity of yeast Rrp6p-Rrp47p, we carried out a series of *in vitro* RNA degradation assays with defined, 5' fluorescently labelled RNA substrates. All comparisons were made between co-purified Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex and isolated Rrp6p monomer originating from the same purification and at conditions where the enzyme concentration (30 nM) exceeded that of RNA (10 nM). First, we compared the activities towards a 30-mer Generic-polyA RNA substrate consisting of 20 nt of constant, generic sequence followed by 10 nt oligo-A. This substrate is rapidly degraded by Rrp6p down to about 2 nucleotides (Figure 3A). Notably, the oligo-A tail (residues 20-30) appears to be degraded much faster than the remaining body of the RNA. Like for Trypanosoma RRP6, degradation is about 5-10 fold slower in the presence of Rrp47p as seen from comparison of the 10 min and 2 min time points [11]. Similarly to isolated Rrp6p, the complex degrades the 10 nt oligo-A tail much more efficiently. In contrast to Trypanosoma RRP6, however, we find that yeast Rrp6p is generally not able to degrade stable, RNA secondary structure. For human RRP6 it was shown that the presence of a 5' extended single-stranded region in front of a stable structured hairpin somehow allows Rrp6p to proceed through the double-stranded region [12]. We therefore asked whether the presence of Rrp47p would affect the ability of Rrp6p to degrade stable RNA secondary structure by analysing degradation of substrates containing a strong hairpin followed by an extended 3' single-stranded region and with

10

(Figure 3B) or without (Figure 3C) a 5' extension. Consistent with the observations for human RRP6, we observe that yeast Rrp6p is able to proceed through the RNA hairpin structure in the presence of a 5' single-stranded extension (Figure 3B, marked with arrows) but not when this element is missing (Figure 3C). In the presence of Rrp47p, the activity of Rrp6p is reduced but appears to eventually reach the same end-point. However, we note that degradation by the complex through the hairpin proceeds with fewer intermediate bands than for isolated Rrp6p (arrows), suggesting that although Rrp47p appears to reduce overall rate of the Rrp6p reaction, degradation through the hairpin is not affected to the same extent. This suggests that the mechanism by which Rrp6p degrades secondary structure is not affected by binding of Rrp47p to the same extent as degradation of single-stranded RNA.

Rrp47p does not affect the affinity of Rrp6p towards structured RNA. To investigate whether Rrp47p contributes to RNA affinity in the Rrp6p-Rrp47p complex, we analysed the binding of monomeric Rrp6p (85-536) and the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex to defined RNAs including both single-stranded and structured regions by electrophoretic mobility shift assays (EMSA). Our hypothesis was that if both proteins contribute to RNA binding, a significant increase in the measured affinity should distinguish the complex from the isolated nuclease. The RNA substrates used were a range of 32-35 nt mimics of yeast U14 pre-snoRNA, which is naturally processed by Rrp6p in a Rrp47p-dependent manner [18]. Of these, the substrate U14+6 contains both single-stranded and structured elements, including the conserved terminal stem and the box C/D elements characteristic of these RNAs in addition to a 3-6 residue extension mimicking a 3' unprocessed pre-snoRNA (Figure 3D). In both cases, dissociation constants were measured using the Rrp6p^{D238A} active site to prevent RNA degradation during the experiment. For all tested RNAs (some data not shown) isolated Rrp6p and the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex showed very similar dissociation constants (K_D) in the low micro-molar range, suggesting that Rrp47p affects the activity of Rrp6p in a

way not related to binding affinity (Figure 3E). Taken together with the observed changes in activity, our data collectively support a model in which Rrp47p affects Rrp6p activity allosterically, most likely without direct interaction with the substrate. To elucidate the details of the mechanistic interplay between these important cofactors of the nuclear RNA exosome, a high-resolution structural analysis of the Rrp6p-Rrp47p complex, and if possible, their interaction with the exosome core will be required.

Acknowledgements

Figure Legends

Figure 1. Characterisation of the Rrp6p-Rrp47p complex. **A.** Coomassie-stained SDS-PAGE gel showing purified samples of Rrp6p$^{12-536}$ (dimeric form), Rrp47p, and the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex. **B.** As A. showing purified samples of the Rrp6p dimer and monomer (85-536) forms. **C.** UV-absorption (A$_{280}$) profiles showing elution of Rrp6p (monomer, red and dimer, purple), Rrp47p$^{1-184}$ (green), and the Rrp6p$^{12-536}$-Rrp47p (blue) complex during SEC. **D.** SR-CD spectra collected from isolated Rrp6p (red), Rrp47p (green), and the Rrp6p-Rrp47p complex (blue). The sum of the measurements for the individual proteins (purple) is shown for comparison. The inset table shows calculated secondary structure content. **E.** As A.

showing pull-downs of Rrp6p$^{12-536}$-His co-expressed with the Rrp47p Sas10/C1D domain on either the 20-100, 10-89, or 10-100 forms. **F.** As A. showing pull-downs of several forms of Rrp6p-His co-expressed with the Rrp47p$^{1-100}$ Sas10/C1D domain.

Figure 2. Rrp47p binds on top of the Rrp6p exonuclease core domain. **A.** SAXS *ab initio* envelope of Rrp6p (monomer, 85-536) based on dummy-atom modelling in two orthogonal views showing docking of the crystal structure of Rrp6p (129-536, PDB 2HBJ) [13]. "PMC2NT" is the position of the part of Rrp6p missing from the crystal structure. **B.** As A. showing the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex, where "Rrp47p" indicates the position of the combined Rrp47p-PMC2NT interaction domain. **C.** As A. showing the Rrp6p$^{12-536}$-Rrp47p$^{1-100}$ Sas10/C1D complex. **D.** As A. showing the Rrp6p$^{75-536}$-Rrp47p$^{1-100}$ minimal interaction region complex. All figures were made using PyMOL and UCSF Chimera [27].

Figure35. Rrp47p modulates Rrp6p activity but not RNA affinity. **A.** *In vitro* degradation experiments using a defined single-stranded, 5'-labelled RNA substrates with a poly-adenosine (A) 3'-end sequence (Generic-polyA). Experiments were carried out with isolated Rrp6p (monomer) or the Rrp6p$^{12-536}$-Rrp47$^{1-184}$ complex and the resulting 5' RNA fragments separated by denaturing PAGE and visualised by scanning fluorometry. For the lane marked "Rrp6p D238A", the inactive D238A mutant of Rrp6p was used as control for contaminating *E. coli* nucleases. **B.** As A. showing reactions with structured and 5'-labelled RNA substrate with 5' extension as indicated. The star marks the position of the fluorescent label and the arrows indicate where degradation has proceeded through RNA secondary structure. Both the Rrp6p-Rrp47p complex and isolated Rrp6p purified as inactive D238A mutants are included as controls for contamination. **C.** As B, but using a substrate with no 5' extension. **D.** EMSA using a fixed amount (10 nM) of U14 pre-snoRNA substrate and increasing concentrations (0-2.4 μM) of either Rrp6p$^{12-536,\ D238A}$ or the Rrp6p$^{12-536,\ D238A}$-Rrp47p$^{1-184}$ complex. **E.**

Binding curves obtained for Rrp6p (●) and the Rrp6p-Rrp47p complex (▲) calculated as %bound RNA based on integration of the gel bands.

Figure 4. A model for the interaction of Rrp6p-Rrp47 with the exosome core. RNA degradation by the yeast nuclear exosome may proceed either directly by the Rrp6p-Rrp47p complex or by Rrp44p following threading through the exosome core barrel.

Table 1. Measured molecular masses and stoichiometries. Molecular weight (MW) estimates for all observed species of Rrp6p, Rrp47p, and Rrp6p-Rrp47p as measured by SEC, MALS, and SAXS. Oligomer shows the most likely multimer in each case. N.D. = not determined, N.O. = not observed.

References

[1] S. Lykke-Andersen, D.E. Brodersen, T.H. Jensen, Origins and activities of the eukaryotic exosome, J Cell Sci 122 (2009) 1487-1494.

[2] P. Mitchell, E. Petfalski, A. Shevchenko, M. Mann, D. Tollervey, The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'-->5' exoribonucleases, Cell 91 (1997) 457-466.

[3] Q. Liu, J.C. Greimann, C.D. Lima, Reconstitution, activities, and structure of the eukaryotic RNA exosome, Cell 127 (2006) 1223-1237.

[4] E. Lorentzen, J. Basquin, E. Conti, Structural organization of the RNA-degrading exosome, Curr Opin Struct Biol 18 (2008) 709-713.

[5] A. Dziembowski, E. Lorentzen, E. Conti, B. Seraphin, A single subunit, Dis3, is essentially responsible for yeast exosome core activity, Nat Struct Mol Biol 14 (2007) 15-22.

[6] A. Lebreton, R. Tomecki, A. Dziembowski, B. Seraphin, Endonucleolytic RNA cleavage by a eukaryotic exosome, Nature 456 (2008) 993-996.

[7] J. Assenholt, J. Mouaikel, K.R. Andersen, D.E. Brodersen, D. Libri, T.H. Jensen, Exonucleolysis is required for nuclear mRNA quality control in yeast THO mutants, RNA 14 (2008) 2305-2313.

[8] M.W. Briggs, K.T. Burkard, J.S. Butler, Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation, J Biol Chem 273 (1998) 13255-13263.

[9] S. Phillips, J.S. Butler, Contribution of domain structure to the RNA 3' end processing and degradation functions of the nuclear exosome subunit Rrp6p, RNA 9 (2003) 1098-1107.

[10] D.L. Makino, M. Baumgartner, E. Conti, Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex, Nature 495 (2013) 70-75.

[11] R.L. Barbosa, P. Legrand, F. Wien, B. Pineau, A. Thompson, B.G. Guimaraes, RRP6 from Trypanosoma brucei: Crystal Structure of the Catalytic Domain, Association with EAP3 and Activity towards Structured and Non-Structured RNA Substrates, PLoS One 9 (2014) e89138.

[12] K. Januszyk, Q. Liu, C.D. Lima, Activities of human RRP6 and structure of the human RRP6 catalytic domain, RNA 17 (2011) 1566-1577.

[13] S.F. Midtgaard, J. Assenholt, A.T. Jonstrup, L.B. Van, T.H. Jensen, D.E. Brodersen, Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain, Proc Natl Acad Sci U S A 103 (2006) 11898-11903.

[14] Y. Zuo, Y. Wang, A. Malhotra, Crystal structure of Escherichia coli RNase D, an exoribonuclease involved in structured RNA processing, Structure 13 (2005) 973-984.

[15] M. Feigenbutz, R. Jones, T.M. Besong, S.E. Harding, P. Mitchell, Assembly of the yeast exoribonuclease Rrp6 with its associated cofactor Rrp47 occurs in the nucleus and is critical for the controlled expression of Rrp47, J Biol Chem 288 (2013) 15959-15970.

[16] P. Mitchell, Rrp47 and the function of the Sas10/C1D domain, Biochem Soc Trans 38 (2010) 1088-1092.

[17] J.A. Stead, J.L. Costello, M.J. Livingstone, P. Mitchell, The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein, Nucleic Acids Res 35 (2007) 5556-5567.

[18] P. Mitchell, E. Petfalski, R. Houalla, A. Podtelejnikov, M. Mann, D. Tollervey, Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs, Mol Cell Biol 23 (2003) 6982-6992.

[19] A. van Hoof, P. Lennertz, R. Parker, Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs, Mol Cell Biol 20 (2000) 441-452.

[20] M. Feigenbutz, W. Garland, M. Turner, P. Mitchell, The Exosome Cofactor Rrp47 Is Critical for the Stability and Normal Expression of Its Associated Exoribonuclease Rrp6 in Saccharomyces cerevisiae, PLoS One 8 (2013) e80752.

[21] J.L. Costello, J.A. Stead, M. Feigenbutz, R.M. Jones, P. Mitchell, The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation, J Biol Chem 286 (2011) 4535-4543.

[22] J.G. Lees, B.R. Smith, F. Wien, A.J. Miles, B.A. Wallace, CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving, Anal Biochem 332 (2004) 285-289.

[23] P.V. Konarev, M.V. Petoukhov, V.V. Volkov, D.I. Svergun, ATSAS 2.1, a program package for small-angle scattering data analysis, J Appl Cryst 39 (2006) 277-286.

[24] D. Franke, D.I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering, J Appl Cryst 42 (2009) 342-346.

[25] V.V. Volkov, D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering, J Appl Cryst 36 (2003) 860-864.

[26] W. Wriggers, Using Situs for the integration of multi-resolution structures, Biophysical Rev 2 (2010) 21-27.

[27] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF chimera - A visualization system for exploratory research and analysis, J Comp Chem 25 (2004) 1605-1612.

[28] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, Biopolymers 89 (2008) 392-400.

[29] I. Stuparevic, C. Mosrin-Huaman, N. Hervouet-Coste, M. Remenaric, A.R. Rahmouni, Cotranscriptional recruitment of RNA exosome cofactors Rrp47p and Mpp6p and two distinct Trf-Air-Mtr4 polyadenylation (TRAMP) complexes assists the exonuclease Rrp6p in the targeting and degradation of an aberrant messenger ribonucleoprotein particle (mRNP) in yeast, J Biol Chem 288 (2013) 31816-31829.

Table 1

Table 1. Measured molecular masses and stoichiometries. Molecular weight (MW) estimates

	MW^{SEC} (kDa)	MW^{MALS} (kDa)	MW^{SAXS} (kDa)	Theoretical MW (kDa)	Oligomer
Rrp6p	82	67	56	55	Monomer
	116	94	114	124.4	Dimer
Rrp47p	N.O.	N.O.	N.O.	21	Monomer
	116	93 ± 16	N.D.	84 - 126	4 - 6mer
Rrp6p-Rrp47p	130	84	97	83.2	1:1 complex

Figure 1

Dedic et al. Figure 1. Characterisation of the Rrp6p-Rrp47p complex

A

B

C

	Alpha	Beta	Turn	Disordered
Rrp6p	35%	19%	13%	33%
Rrp47p	33%	16%	16%	35%
Rrp6p-Rrp47p	42%	13%	14%	31%

D

E

F
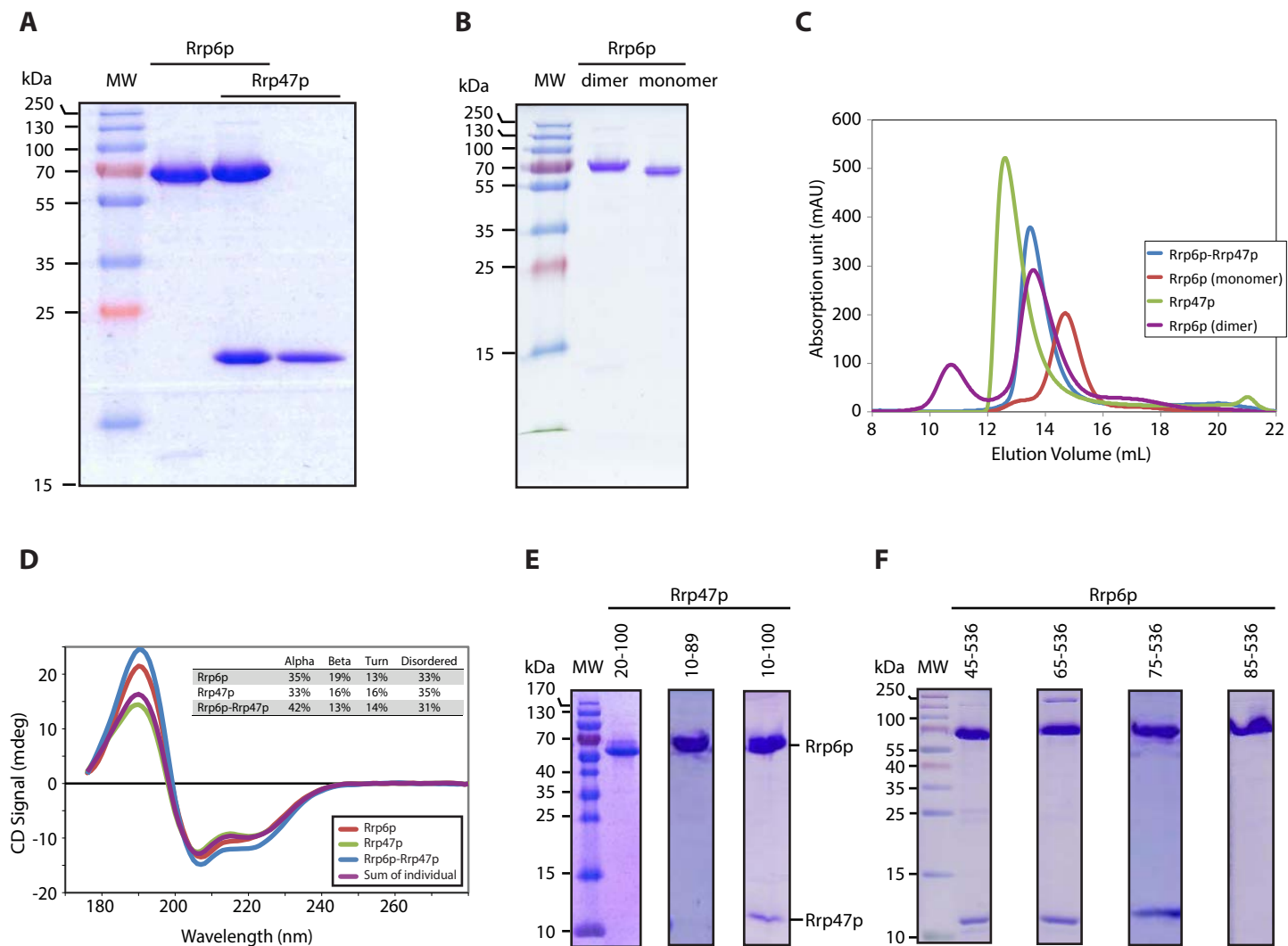
Figure 2

Dedic et al. Figure 2. Rrp47p binds on top of the Rrp6p exonuclease core domain

A — Rrp6p 85-536

B — Rrp6p 12-536 / Rrp47p 1-184

C — Rrp6p 12-536 / Rrp47p 1-100

D — Rrp6p 75-536 / Rrp47p 1-100
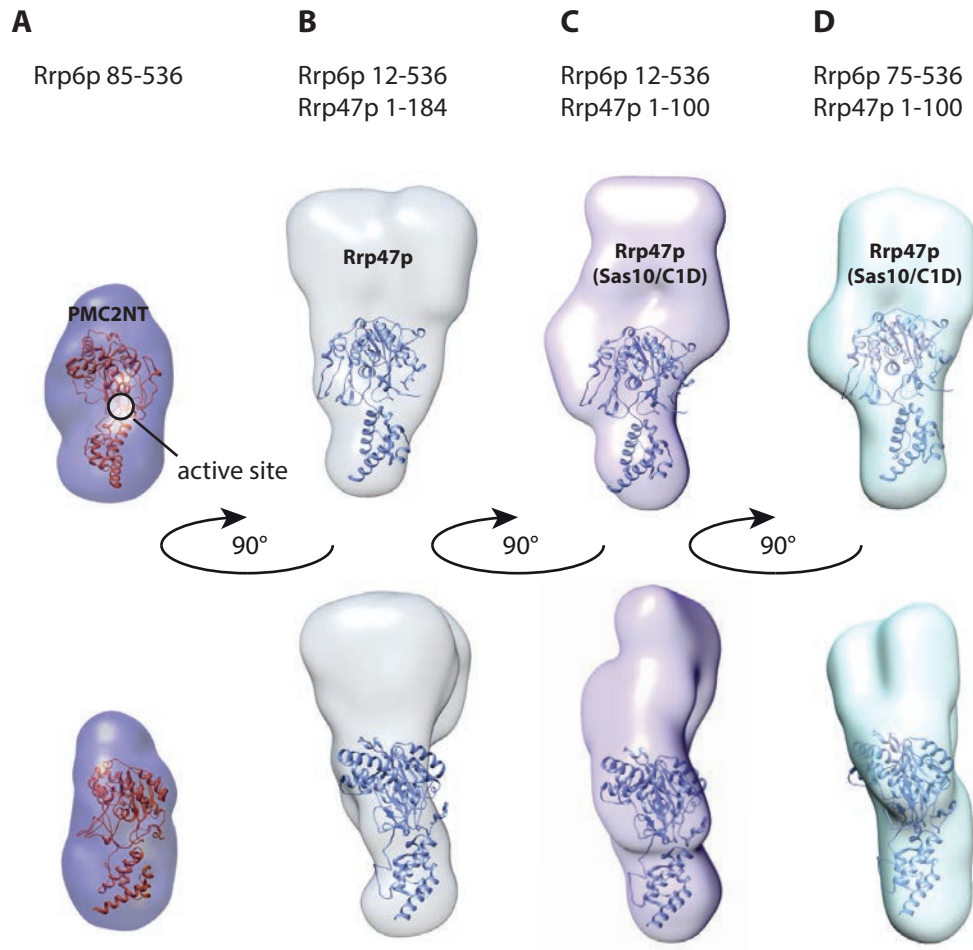
Figure 3

Dedic et al. Figure 3. Rrp47p modulates Rrp6p activity but not RNA affinity
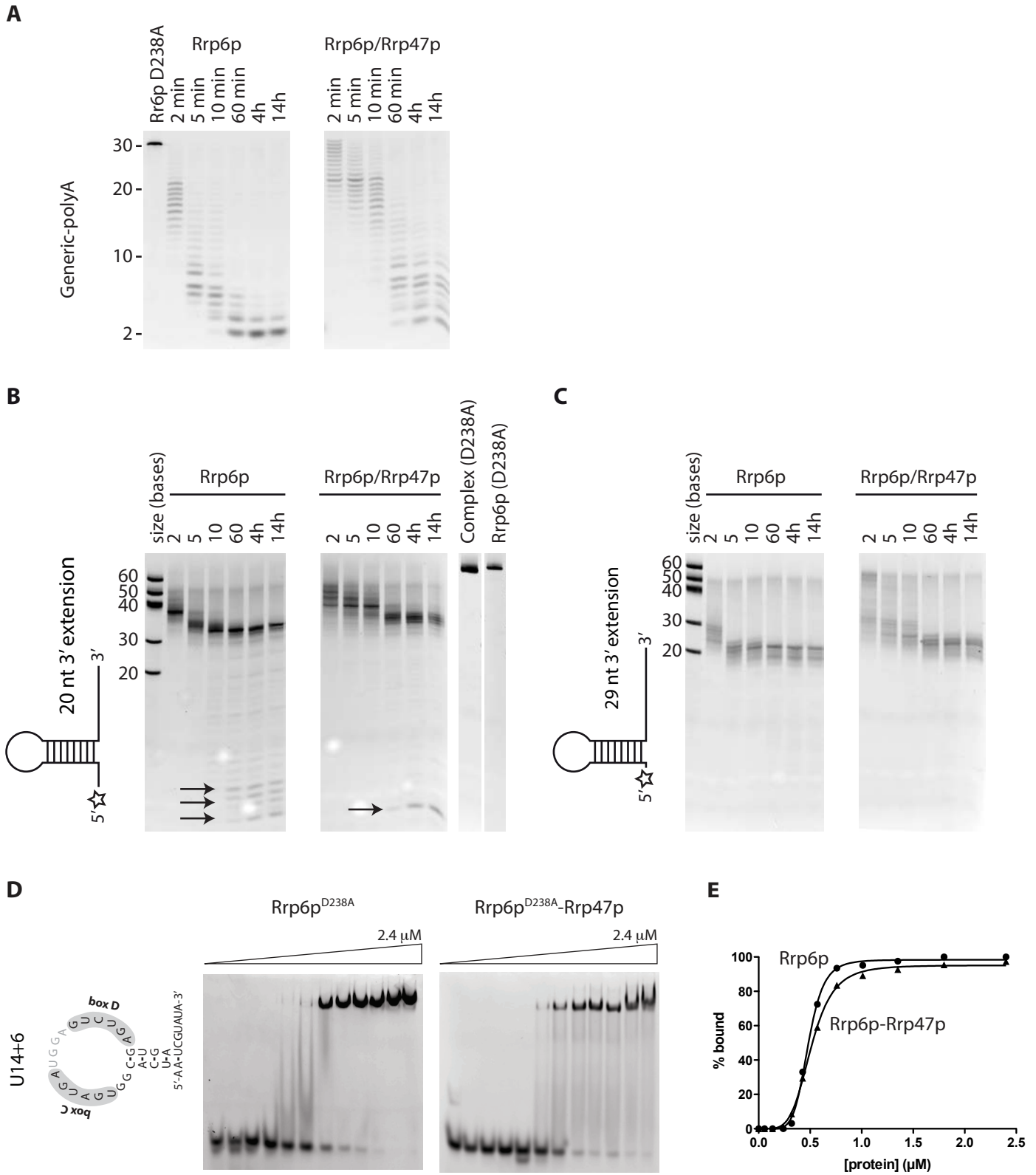
Figure 4

Dedic et al. Figure 4. A model for the interaction of Rrp6p-Rrp47 with the exosome core
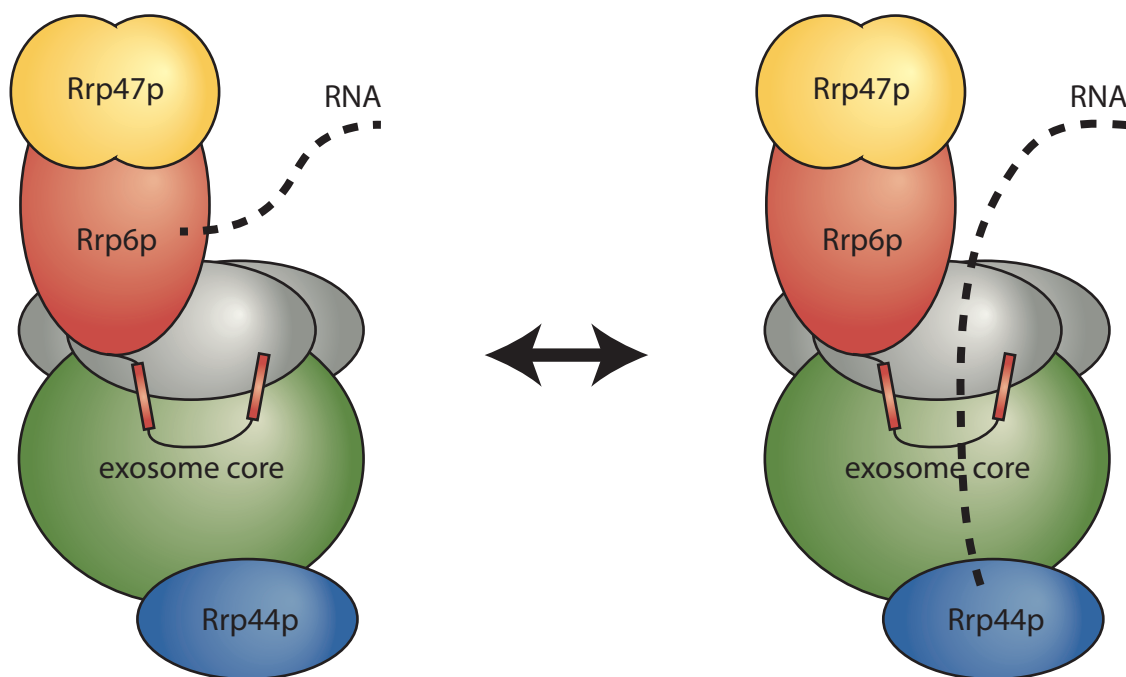
Dedic *et al*. "Structural analysis of the yeast exosome Rrp6p-Rrp47p complex by small-angle x-ray scattering"

METHODS

Protein expression and purification

S. cerevisiae rrp47/lrp1 and *rrp6* were amplified from genomic DNA and inserted into pET-30 Ek/LIC by the ligation-independent cloning method using a LIC Duet Minimal Adaptor (Novagen) to create a single construct coding for Rrp47p and Rrp6p$^{12-536}$-Tev-His$_6$ (residues 12-536 of Rrp6p followed by a Tev-cleavable His-tag). Rrp47p-Rrp6p^{D238A} was subsequently constructed by site-directed mutagenesis and was used as an template for introduction of a stop codon at position 100 in Rrp47 by site-directed mutagenesis, using primers 5'-AAATCATACATGGATAAGTAAAAACAATACGATAATAGG-3' and 5'-CCTATTATCGTATTGTTTTTACTTATCCATGTATGATTT-3'. The resulting Rrp47p$^{1-100}$-Rrp6p$^{12-536}$ construct harbouring the D238A mutation was subsequently used to generate the Rrp47p$^{1-100}$-Rrp6p$^{75-536}$ construct by inverse PCR (iPCR), using primers 5'-TCTGACTTGTGGAACAACTTTG-3' and 5'-CATGGTATATCTCCTTCTTAAAG-3'. *S. cerevisiae rrp47/lrp1* was likewise cloned into pET-30 Ek/LIC vector to encode full-length Rrp47p with an N-terminal, Tev-cleavable His-tag. Protein expression was carried out in *E. coli* BL21 (DE3) RIL (Novagen) at 37°C in 2xTY medium containing 50 μg/ml kanamycin and 34 μg/ml chloramphenicol by growing the cells to A$_{600}$=1.1, cold shocked for 45 min on ice, and induced with 0.5 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) followed by overnight expression at 18°C. Cells were resuspended in 50 mM Tris-HCl pH 8.0, 300 mM KCl, 10 mM imidazole, 5 mM

MgCl$_2$, 5 mM β-mercaptoethanol (BME), 10% glycerol and Complete Protease Inhibitor Cocktail tablets (Roche) and lysed by sonication and subsequent high-pressure homogenization at 15,000 psi. The cleared lysate was then purified on a 5 ml Ni^{2+}-chelating agarose column (Qiagen) pre-equilibrated in lysis buffer, washed, and eluted using 200 mM imidazole. The His-tag in the C terminus of Rrp6p was then cleaved off by incubation with 1:50 molar recombinant, His-tagged Tobacco Etch Virus (Tev) protease overnight at 4°C during dialysis against 50 mM Tris-HCl pH 8.0, 150 mM KCl, 5 mM MgCl$_2$, 5 mM β-mercaptoethanol (BME), and 10% glycerol. The Tev protease and any uncleaved protein or non-specifically binding proteins were then removed by reapplying the eluate to the Ni^{2+} agarose column and collecting the run-through. Sequential ion exchange purification using custom 5 ml Source 15Q and 15S columns (GE Healthcare) running in 50 mM Tris-HCl pH 8.0, 75 to 450 mM KCl, 5 mM MgCl$_2$, and 5 mM BME allowed purification of both isolated Rrp6p$^{12-536}$ and Rrp47p as well as the Rrp6p$^{12-536}$-Rrp47p complex. Isolated Rrp6p$^{12-536}$ that was not retained during the Source S ion-exchange step was additionally purified on a 1.5 mL Mono Q column (GE Healthcare) with identical buffers to resolve Rrp6p monomer and dimer populations. The purification of Rrp47p was similar to Rrp6p and Rrp6p-Rrp47p but did not involve a cation exchange step. The purification steps for the truncated Rrp6p-Rrp47p complexes (Rrp47p$^{1-100}$-Rrp6p$^{12-536}$ and the minimal Rrp47p$^{1-100}$-Rrp6p$^{75-536}$) were essentially identical as for the Rrp47$^{1-184}$-Rrp6p$^{12-536}$ complex. Finally, all samples were individually purified on a Superdex 200 10/300 GL size-exclusion column (GE Healthcare) equilibrated in 20 mM Tris-HCl pH 8.0, 100 mM KCl, and 5 mM BME, concentrated by spin filtration, and stored at -80°C in 40% glycerol. All purifications were carried out with dual A$_{280}$/A$_{260}$ monitoring to ensure that no RNA was co-purified with the proteins.

SR-CD analysis

CD spectra were collected at beam line CD1 at the ASTRID storage ring (ISA, Aarhus University, Denmark) detected by a photo multiplier tube (type 9406B, ETL, UK) [1,2]. Spectra were measured from Rrp6p, Rrp47p and Rrp6p-Rrp47p individually at concentrations close to 2 mg/ml in 50 mM KH_2PO_4/K_2HPO_4 at pH 7.5. Three independent spectra of the buffer (for baseline correction) and samples were measured in 100 μm path-length Suprasil cell (Hellma GmbH), spectra were averaged, baseline subtracted, and mildly smoothed with a Savitzky-Golay filter using the CDtool software [3]. Secondary structure analysis was carried out using the CDSSTR method and the SP175 reference data set in DichroWeb [4].

SAXS analysis

Synchrotron radiation SAXS scattering data were collected at the EMBL SAXS-WAXS beam line X33 at DORIS/DESY (Hamburg, Germany) as 8*15 sec exposures at 1.5 Å using a MAR345 image plate detector, and scattering profiles for the 8 passes were compared to detect radiation damage. Measurements were conducted at 10 °C using 75 μL sample in 20 mM Tris-HCl pH 7.5, 100 mM KCl, and 5 mM BME. Measurements were carried out at three different concentrations of $Rrp6p^{12-536}$-$Rrp47p^{1-184}$ complex (3.5, 1.8 and 0.9 mg/mL), $Rrp6p^{12-536}$-$Rrp47p^{1-100}$ (4.3 and 3.9 mg/mL) and the $Rrp6p^{85-536}$ monomer (4.0, 2.0 and 1.0 mg/ml). For the $Rrp6p^{12-536}$ dimer and the $Rrp6p^{75-536}$-$Rrp47p^{1-100}$ complex, scattering at a single concentration of 2.2 mg/mL and 2.75 mg/mL was measured, respectively. Background scattering was subtracted using PRIMUS [5] in the ATSAS package [6]. For the $Rrp6p^{12-536}$-$Rrp47p^{1-100}$ and $Rrp6p^{75-536}$-$Rrp47p^{1-100}$ complexes, the synchrotron radiation SAXS scattering data were collected at the I911-4 beam line at MAX-Lab (Lund, Sweden) and processed essentially the same as for the full-length Rrp47 complex ($Rrp47p^{1-184}$-$Rrp6p^{12-536}$). Pair distance distribution functions of the particles, P(r), and the maximum particle dimensions, D_{max}, were computed using GNOM [7] and molecular weights were estimated by the

calculated Porod volumes. Porod volumes were calculated using ATSAS AUTOPOROD [8] and *ab initio* shapes were determined using DAMMIF [9] and GASBOR [10], in the case of the Rrp6p dimer, both with and without a P2 symmetry constraint. After 12 DAMMIF runs, DAMAVER [11] was used to analyse the normalized spatial discrepancy (NSD) between the 12 models and the lowest NSD model was used as representative. The known crystal structure of Rrp6p (residues 129-536) was then docked manually into the representative GASBOR models and for the Rrp6p monomer and the three Rrp6p-Rrp47p complexes, *ab initio* envelopes were computed from the dummy atom bead models using SITUS pdb2vol based on the GASBOR and DAMMIF models, respectively [12]. For the Rrp6p$^{12-536}$-Rrp47p$^{1-184}$ complex, the molecular envelope of the Rrp6p crystal structure was docked automatically using SITUS colores and the best fits were inspected manually.

Electrophoretic mobility shift assay

The 5'-labelled FAM-labelled U14+6 RNA substrate (Microsynth), designed based on *S. cerevisiae* U14 snoRNA sequence, was used for EMSA and containts the sequence (5'-AAUCACGGUG AUGAUGGAGU CUGAGUGAUC GUAUA-3'). As it contains a stuctured stem loop, the substrate was re-annealed by incubation at 96°C for 2 min followed by cooling at RT for 60 min. Concentrations were verified by A$_{260}$ measurements using a Nanodrop ND-1000 spectrophotometer. Data were quantified using TotalLab Quant.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure S1.The Rrp47p SAS10 domain interacts with a conserved N-terminal region of Rrp6. **A.** Overview of Rrp6p and Rrp47p and the constructs used in this paper. Left, the purified constructs of Rrp6p, Rrp6p$^{129-536}$ (crystal structure, orange), Rrp6p$^{85-536}$ (monomer, red), and Rrp6p^{12-}

^{536} (dimer, purple). Of these, only Rrp6p$^{12\text{-}536}$ includes the full N-terminal PMC2NT domain (residues 12-102, magenta). The exonuclease core consisting of the 3'-5' exonuclease domain is shown in brown, the Helicase and RNase D C-terminal (HRDC) domain in orange, and the putative, C-terminal nuclear localisation signal (NLS) in white. Right, Rrp47p consisting of the Sas10/C1D domain (residues 10-100, purple) and a basic patch of residues near the C-terminus (brown). **B.** Sequence alignments of the N and C terminal regions of the Sas10/C1D domain of Rrp47p (top) and the N-terminal PMC2NT domain of Rrp6p (bottom).

Supplementary Figure S2. Sequence alignments of S.cerevisiae Rrp6p and Rrp47p. **A.** Alignment of the N-terminal region of Rrp6p up till the beginning of the 3'-5' exonuclease domain with relevant, eukaryotic homologues from *S. pombe*, *C. albicans*, *S. stipites*, *S. japonicus*, *H. sapiens*, *M. musculus*, and *D. melanogaster* as indicated. **B.** Alignment of the full sequence of Rrp47p with relevant, eukaryotic homologues from *C. albicans*, *D. bruxellensis*, *H. sapiens*, *B. taurus*, and *D. melanogaster*. Alignmnent was carried out using CLC bio (http://www.clcbio.com).

Supplementary Figure S3. SAXS scattering curves, Guinier, and Kratky plots. **A.** SAXS scattering curves for the Rrp6p monomer and dimer forms shown as relative log(intensity) versus inverse scattering angle (s). **B.** As A. for the Rrp6p-Rrp47p complexes. **C.** SAXS scattering curves for the Rrp6p monomer and dimer forms shown as log-log plots with Guinier plots (ln(I) versus s^2) inserted. **D.** As C. for the Rrp6p-Rrp47p complexes. **E.** Kratky plots (l(s)*s^2 versus s) showing the peak characteristic of folded proteins for the Rrp6p monomer and dimer forms. Functions were calculated using AutoGNOM [8]. **F.** As E for the Rrp6p-Rrp47p complexes.

Supplementary Figure S3. Dimensions of protein complexes. **A.** Left, the pair distance distribution

function, P(r) versus r, for the Rrp6p monomer and dimer forms with the calculated radii of gyration (R_g) and maximum dimensions (D_{max}) inserted; right, SAXS *ab initio* envelopes illustrating the D_{max} dimension. **B.** As A. for the Rrp6p-Rrp47p complexes.

Supplementary Table T1. SAXS parameters

Protein/Complex	R_g (nm)	D_{max} (Å)	Porod volume (nm^3)	MW_{calc} (kDa)	$MW_{theor.}$ (kDa)
Rrp6p monomer	3.47	100.5	95.4	56.1	62.2
Rrp6p dimer	4.04	141.4	193.0	113.5	124.4
Rrp6p$^{12-536}$-Rrp47p$^{1-184}$	3.89	136.1	165.4	97.3	83.2
Rrp6p$^{12-536}$-Rrp47p$^{1-100}$	3.97	139.0	147.4	86.7	74.1
Rrp6p$^{75-536}$-Rrp47p$^{1-100}$	3.89	136.4	129.7	76.31	65.7

References

[1] A.J. Miles, S.V. Hoffmann, Y. Tao, R.W. Janes, B.A. Wallace, Synchrotron radiation circular dichroism (SRCD) spectroscopy: New beamlines and new applications in biology, Spectroscopy 21 (2007) 245-255.

[2] A.J. Miles, R.W. Janes, A. Brown, D.T. Clarke, J.C. Sutherland, Y. Tao, B.A. Wallace, S.V. Hoffmann, Light flux density threshold at which protein denaturation is induced by synchrotron radiation circular dichroism beamlines, J Sync Radiation 15 (2008) 420-422.

[3] J.G. Lees, B.R. Smith, F. Wien, A.J. Miles, B.A. Wallace, CDtool-an integrated software package for circular dichroism spectroscopic data processing, analysis, and archiving, Anal Biochem 332 (2004) 285-289.

[4] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, Biopolymers 89 (2008) 392-400.

[5] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H.J. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, J Appl Cryst 36 (2003) 1277-1282.

[6] P.V. Konarev, M.V. Petoukhov, V.V. Volkov, D.I. Svergun, ATSAS 2.1, a program package for small-angle scattering data analysis, J Appl Cryst 39 (2006) 277-286.

[7] D.I. Svergun, Determination of the Regularization Parameter in Indirect-Transform Methods Using Perceptual Criteria, J Appl Cryst 25 (1992) 495-503.

[8] M.V. Petoukhov, P.V. Konarev, A.G. Kikhney, D.I. Svergun, ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis, J Appl Cryst 40 (2007) S223-228.

[9] D. Franke, D.I. Svergun, DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering, J Appl Cryst 42 (2009) 342-346.

[10] D.I. Svergun, M.V. Petoukhov, M.H.J. Koch, Determination of domain structure of proteins from X-ray solution scattering, Biophys J 80 (2001) 2946-2953.

[11] V.V. Volkov, D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering, J Appl Cryst 36 (2003) 860-864.

[12] W. Wriggers, Using Situs for the integration of multi-resolution structures, Biophysical Rev 2 (2010) 21-27.
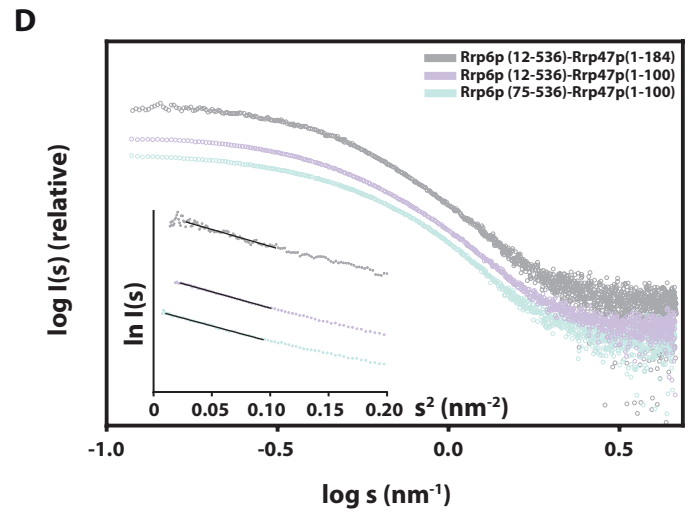
A

B
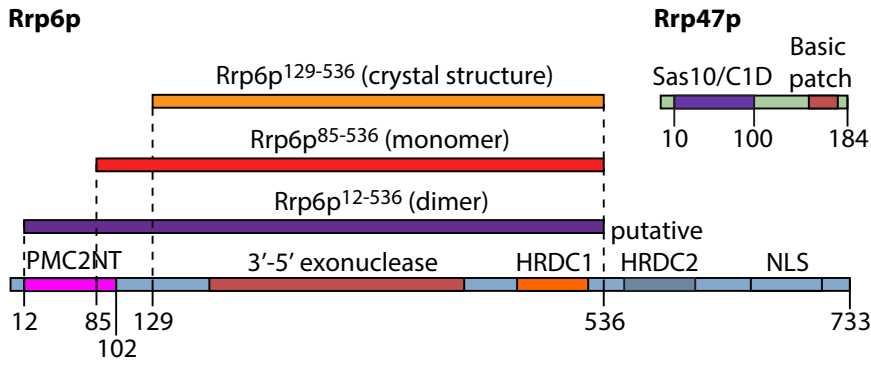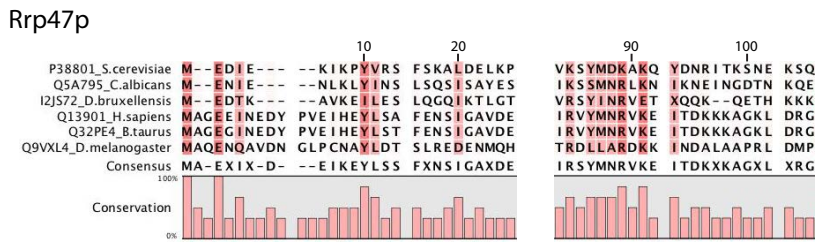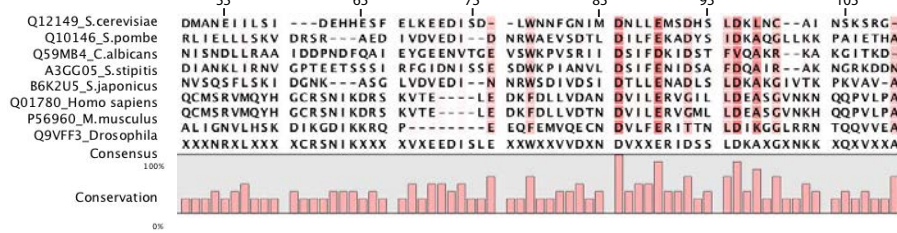
Rrp47p

Rrp6p

Dedic et al. Supplementary Figure S2. Multiple alignments of Rrp6p and Rrp47p

A

Rrp6p

B

Rrp47p
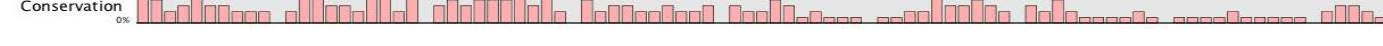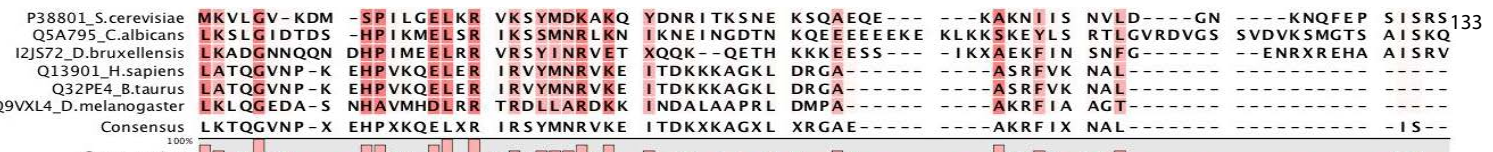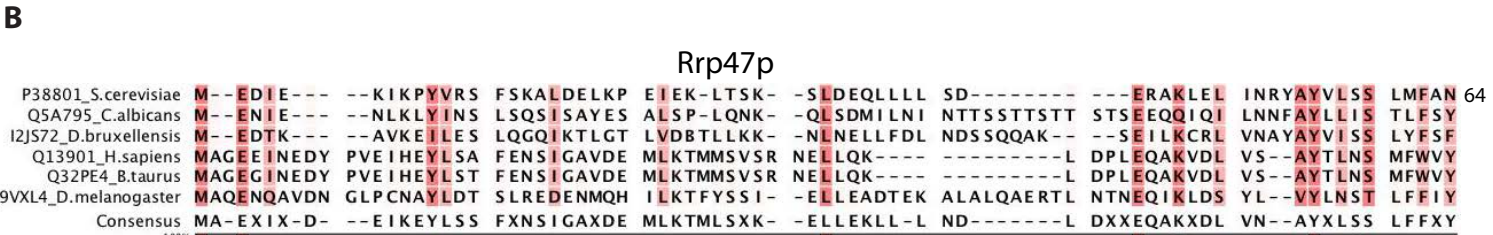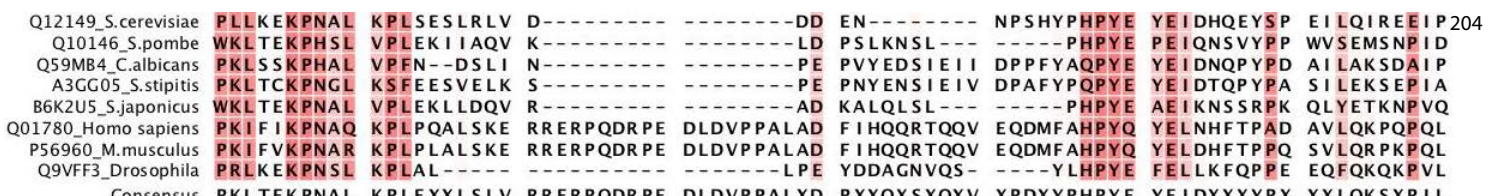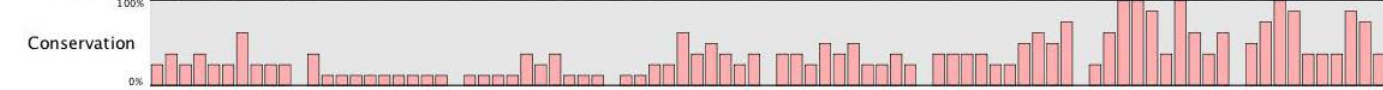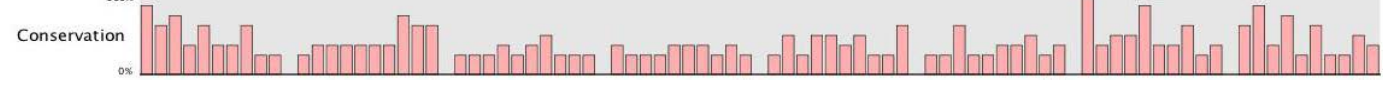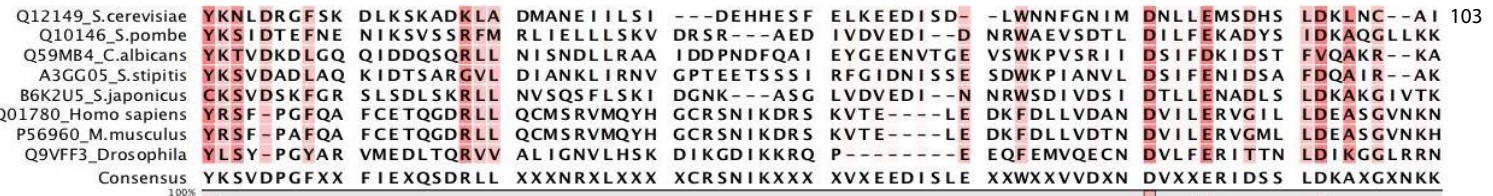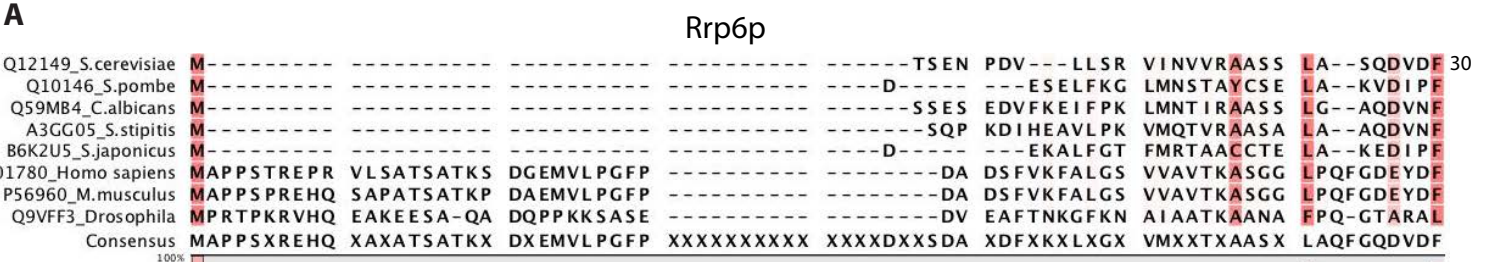
Dedic et al. Supplementary Figure S3. SAXS scattering curves, Guinier and Kratky plots
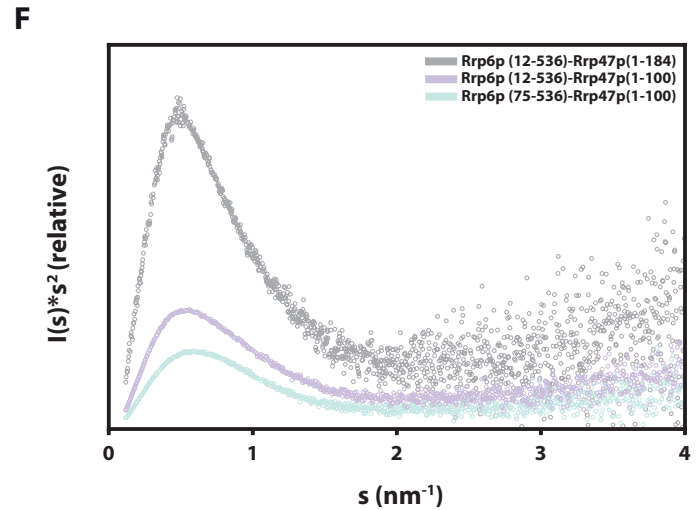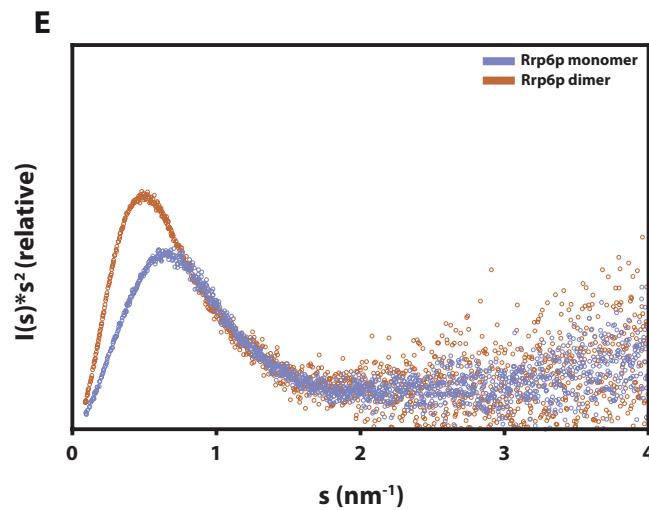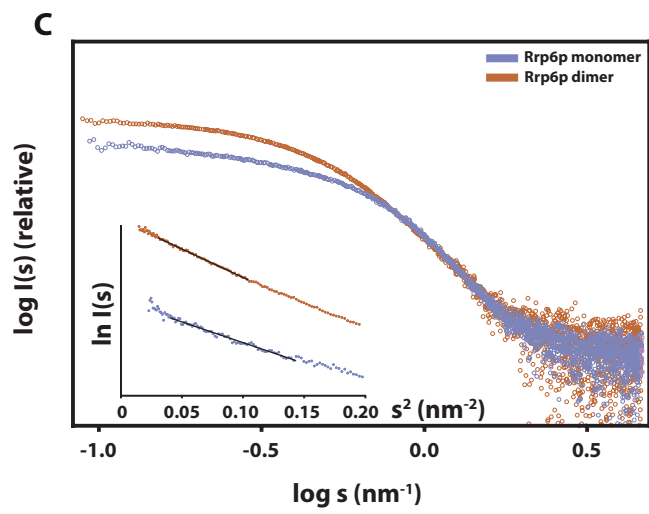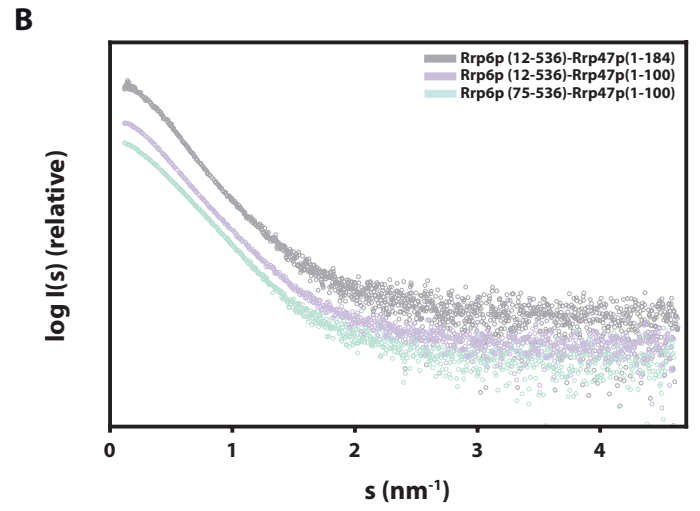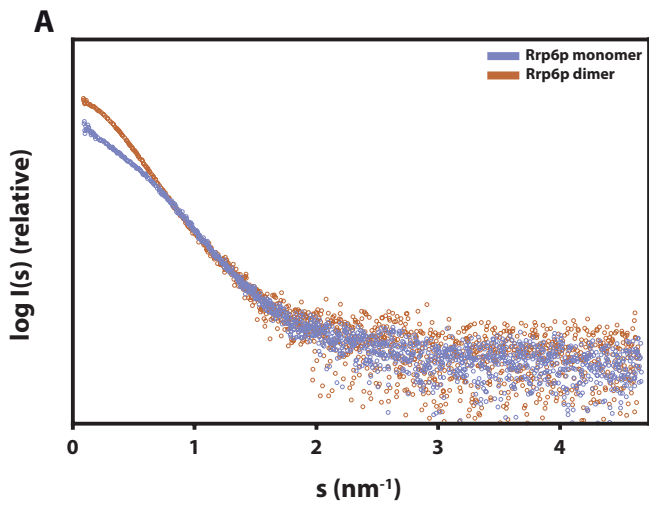
Dedic et al. Supplementary Figure S4. Dimensions of protein complexes