HHS Public Access

Author manuscript

*J Mol Biol*. Author manuscript; available in PMC 2019 May 11.

Published in final edited form as:

*J Mol Biol*. 2018 May 11; 430(10): 1403–1416. doi:10.1016/j.jmb.2018.03.029.

# Crystal structure of human Rpp20/Rpp25 reveals quaternary level adaptation of the Alba scaffold as structural basis for single-stranded RNA binding

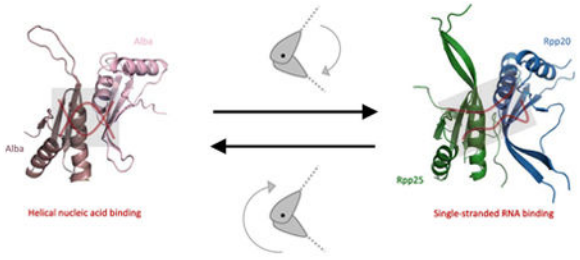
**Clarence W. Chan**[1], **Benjamin R. Kiesel**[1], and **Alfonso Mondragón**[1,*]

[1]Department of Molecular Biosciences, Northwestern University, 2205 Tech Drive, Evanston, IL 60208-3500

## Abstract

RNase P catalyzes the removal of 5′ leaders of tRNA precursors and its central catalytic RNA subunit is highly conserved across all domains of life. In eukaryotes, RNase P and RNase MRP, a closely related ribonucleoprotein enzyme, share several of the same protein subunits, contain a similar catalytic RNA core, and exhibit structural features that do not exist in their bacterial or archaeal counterparts. A unique feature of eukaryotic RNase P/MRP is the presence of two relatively long and unpaired internal loops within the P3 region of their RNA subunit bound by a heterodimeric protein complex, Rpp20/Rpp25. Here we present a crystal structure of the human Rpp20/Rpp25 heterodimer and we propose, using comparative structural analyses, that the evolutionary divergence of the single-stranded and helical nucleic acid binding specificities of eukaryotic Rpp20/Rpp25 and their related archaeal Alba chromatin protein dimers, respectively, originate primarily from quaternary level differences observed in their heterodimerization interface. Our work provides structural insights into how the archaeal Alba protein scaffold was adapted evolutionarily for incorporation into several functionally-independent eukaryotic ribonucleoprotein complexes.

## Graphical abstract




*To whom correspondence should be addressed. Tel: +1-847-491-7726; Fax: +1-847-467-6489; a mondragon@northwestern.edu.

**Accession Numbers:** Atomic coordinates and structure factors for the reported crystal structure have been deposited with the Protein Data Bank under accession number XXXX.

**Conflict of Interest:** None declared.

**Keywords**

X-ray crystallography; ribonuclease P; ribonuclease MRP; MRP RNA; ribonucleoprotein; protein heterodimer; RNA binding; Rpp20; Rpp25; Alba; Th/To autoantigen; molecular evolution

## Introduction

Ribonucleoprotein (RNP) ribonuclease P (RNase P) and protein-only RNase P (PRORP) comprise a group of enzymes whose main function is to catalyze the removal of 5′ leader segments of tRNA precursors [1-3]. The catalytic moiety of RNP RNase P is an RNA subunit that is highly conserved across all domains of life and consists generally of two folded domains. In contrast, the protein content of RNP RNase P varies tremendously among bacterial, archaeal, and eukaryotic organisms [4, 5]. As many as ten protein subunits have been identified as belonging to the human RNP RNase P holoenzyme: Pop1, Pop5, Rpp14, Rpp20, Rpp21, Rpp25, Rpp29, Rpp30, Rpp38, and Rpp40, but how each protein subunit interacts with the others or with the RNA subunit to affect catalytic function, structural organization, regulation, or localization of the entire RNase P complex remains poorly understood [4-7].

To date, Rpp20 and Rpp25 are the best characterized human RNase P subunits, owing in large part to previous biochemical, cellular, and genetic studies [8-16], as well as to extensive work conducted on their yeast homologues Pop7 and Pop6 [17, 18]. The crystal structure of the Pop6/Pop7 heterodimer in complex with the P3 stem loop of yeast RNase MRP, an eukaryotic endoribonuclease structurally and evolutionarily related to RNase P, revealed that these protein subunits interact nearly exclusively with the internal single-stranded regions of the MRP P3 stem loop while demonstrating an apparent specificity for an ACR (where R denotes a purine) nucleotide triad in one of the two strands [19-21]. Rpp20 (Pop7) and Rpp25 (Pop6), as well as the presence of these internal single-stranded P3 regions in both RNase P and MRP RNAs, are conserved features of the nuclear RNase P and RNase MRP enzymes and thus are considered to be essential adaptations specific to eukaryotes [22]. It has been proposed that P3 binding by Pop6/Pop7 helps to protect and to stabilize the overall fold of RNase P and MRP RNAs by specifically replacing tertiary RNA-RNA interactions lost in evolution [20, 21]. Furthermore, the P3/Pop6/Pop7 RNP domain has also been shown to act as a nucleating center for protein-protein interactions by associating with Pop1, the largest protein subunit of the RNase P and MRP holoenzymes, albeit one with relatively undetermined function [22, 23]. Within the current paradigm of RNP structure and function, the Rpp20/Rpp25 or Pop6/Pop7 heterodimer emerges as an example of molecular adaptation whereby novel or additional protein subunits bring about increased structural complexity and functional capacity to an RNP assembly.

Despite the numerous similarities between Rpp20/Rpp25 and its yeast counterpart, Pop6/Pop7, notable differences between the two exist. Rpp20 and Rpp25 exhibit relatively low sequence similarity to each other and to Pop7 and Pop6, respectively, even though they all belong to the Alba superfamily of nucleic acid binding proteins (originally named after a common regulatory property whereby acetylation lowers substrate binding affinity) [24, 25].

On the tertiary level, these proteins exhibit very similar structures. Furthermore, the single-stranded regions of the P3 stem loop of human RNase P and MRP RNAs are not identical in sequence, nor are they similar to the sequences of the corresponding regions in yeast RNase P and MRP RNAs [21], suggesting that these RNA-binding proteins do not abide by a strict sequence consensus. Previous work has also shown that human Rpp20 hydrolyzes ATP, an activity that is seemingly decoupled from tRNA precursor processing by human RNase P, is absent in bacterial RNase P subunits, and so far is unconfirmed in its yeast homologue, Pop7 [10]. Whether this ATPase activity is a unique adaptation in humans remains to be clarified. From a medical perspective, Rpp20/Rpp25 are also the major components of the Th/To autoantigen complex in several autoimmune diseases; however, the molecular basis of their immunogenicity remains completely unknown.

Here we present a crystal structure of the human Rpp20/Rpp25 heterodimer and we examine the structural basis of Rpp20 and Rpp25 heterodimerization. Comparison of the crystal structures of human Rpp20/Rpp25 and yeast Pop6/Pop7 in complex with the P3 stem loop of yeast RNase MRP revealed that although these proteins do not undergo significant conformational change upon RNA binding, the geometry and electrostatic properties of their RNA binding surfaces are conserved and relate to the relative orientation by which the protein subunits heterodimerize. The crystal structure of Rpp20/Rpp25 also provides structural perspective into the previously reported ATPase activity of human Rpp20. Furthermore, we discuss the general implications of subtle molecular adaptations leading to profound functional differences observed in these Alba-related proteins as a potential mechanism of subunit sharing in multi-subunit RNP complexes.

## Materials and Methods

### Expression and purification of human Rpp20, Rpp25, and the Rpp20/Rpp25 heterodimer

Human Rpp20 and Rpp25, encoded in separate pHTT7K vectors and provided by the laboratory of Sidney Altman [8, 16], were expressed with a N-terminal $His_6$ affinity tag that is removable by TEV protease. The proteins were expressed recombinantly in *Escherichia coli* Rosetta(DE3). In brief, for protein production transformed bacterial cultures were grown in Terrific Broth at 37°C under kanamycin (50 μg/ml) and chloramphenicol (34 μg/ml) selection, cold shocked on ice for 30-45 minutes upon reaching an optical density ($OD_{600}$) of approximately 1.5, and induced for protein expression at 16°C for 18 hours with 1 mM isopropyl β-D-1-thiogalactopyranoside (added immediately after cold shocking). Cells were harvested by centrifugation in a Sorvall SLC-6000 rotor spun at 5,500 rpm (6,619 × g) and 4°C for 30 minutes, and then resuspended in 50 mM Tris-HCl, pH 7.5, 400 mM sodium chloride, 20 mM imidazole, and 5 mM β-mercaptoethanol.

The same purification protocol was utilized to purify Rpp20 and Rpp25 separately, as well as the Rpp20/Rpp25 heterodimer, to apparent homogeneity based on SDS-PAGE (Supplementary Figure S1, Panels a-d). Since Rpp20 and Rpp25 associate readily in solution and appeared to overexpress at comparable levels, purification of the heterodimer simply involved mixing the cell pastes containing each protein at an approximately equal mass ratio prior to cell lysis. Lysozyme, phenylmethylsulfonyl fluoride, leupeptin, and pepstatin A were added to resuspended cells at final working concentrations of 1 mg/ml, 1 mM, 1 μg/ml,

and 1 μg/ml, respectively. Following mechanical disruption by sonication, the cell lysate was clarified first by centrifugation in a Sorvall SS-34 rotor spun at 18,000 rpm (38,724 × g) and 4°C for 60 minutes, and then by filtration through a 0.22 μm polyethersulfone membrane. The supernatant was loaded onto Ni-NTA resin (Qiagen) packed into a gravity flow column, which was subsequently washed with approximately 50 column volumes of 50 mM Tris-HCl, pH 7.5, 400 mM sodium chloride, 50 mM imidazole, and 5 mM β-mercaptoethanol. Protein was eluted from the column with 50 mM Tris-HCl, pH 7.5, 400 mM sodium chloride, 400 mM imidazole, and 5 mM β-mercaptoethanol, and subsequently dialyzed into 50 mM Tris-HCl, pH 7.5, 400 mM sodium chloride, 50 mM imidazole, and 5 mM β-mercaptoethanol at 4°C with the addition of recombinant TEV protease at an approximate ratio of 1:10 (protease:protein) by mass. After the removal of the N-terminal $His_6$ tag of Rpp20 and of Rpp25, the samples were reloaded onto the Ni-NTA resin to obtain tag-free Rpp20, Rpp25, or Rpp20/Rpp25. Size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) was performed subsequently on the Rpp20/Rpp25 samples as an additional step to remove free Rpp20 or Rpp25 from Rpp20/Rpp25 heterotetramers and heterodimers, as well as any small but varying amounts of endogenous soluble aggregates from the expression host. Finally, all protein samples were dialyzed into 50 mM Tris-HCl, pH 7.5, 400 mM sodium chloride, and 5 mM β-mercaptoethanol for subsequent experiments and stored at 4°C. The identities of both Rpp20 and Rpp25 were confirmed by mass spectrometry (Supplementary Figure S1, Panels g-i).

### Crystallization and structure determination of the human Rpp20/Rpp25 heterodimer

Crystallization trials of human Rpp20, Rpp25, and Rpp20/Rpp25 were conducted with a variety of commercial screens, using several equilibration methods including vapor diffusion and capillary counter-diffusion, and in a wide range of sample concentrations and temperatures. The condition resulting in the best diffracting crystals of Rpp20/Rpp25 from this search was as follows: 100 mM potassium formate, pH 7.5, 300 mM magnesium sulfate, 4.5% PEG 300, 4.5% PEG 400, 4.5% PEG 1000, 4.5% PEG 4000, and 4.5% PEG 8000. These crystals were reliably reproduced by vapor diffusion in this and surrounding conditions between a pH range of 7.0 to 8.0, 3% to 7% of each PEG component, and 18°C to 22°C in as quickly as 2 to 3 weeks, while magnesium sulfate was ultimately determined to be dispensable for crystallization. However, the largest (up to 300 μm × 300 μm × 300 μm) and best diffracting crystals took between 6 to 18 months to grow from Rpp20/Rpp25 samples with initial concentrations ranging from 30 to 40 mg/ml.

X-ray diffraction data were collected at 100 K at the Life Sciences - Collaborative Access Team (LS-CAT) beamlines at the Advanced Photon Source, Argonne National Laboratory, using a Rayonix MX300 CCD detector. For data collection, crystals of Rpp20/Rpp25 were cryo-protected by supplementing crystallization solutions with 20-25% glycerol (v/v) and were flash frozen with liquid nitrogen. Processing of diffraction data was carried out using either XDS [26] or MOSFLM [27] for indexing and integration and AIMLESS [28] for scaling. An experimental electron density map was calculated with AutoSHARP [29] by single isomorphous replacement with anomalous scattering (SIRAS) from data obtained from a native and four different crystals derivatized with crystallization solutions containing 10, 20, and 50 mM potassium tetrachloroplatinate (II) (5 to 45-hour soaks in crystallization

solution). The resulting map after solvent flattening was of poor quality, with only a few clearly continuous regions of electron density (Supplementary Figure S3, Panel a). Multiple iterations of automated model building using Buccaneer [30] in CCP4 [31] followed by manual model editing using Coot [32] and Refmac5 [33] allowed us to generate a poly-alanine backbone trace of both Rpp20 and Rpp25. Since the N- and C-terminal regions of both Rpp20 and Rpp25 were disordered in the crystals and our initial datasets were limited in resolution to approximately 3.0 Å, it was difficult to establish the correct amino acid register of the protein models. Only when larger and better diffracting crystals (to 2.25 Å) became available were we able to confirm the exact amino acid sequence for all regions of our model (Supplementary Figure S3, Panel b). All iterative rounds of model building and refinement were performed with Coot [32] and Refmac5 [33], respectively, and the stereochemistry of atomic coordinates of our final model ($R/R_{free}$ = 0.20/0.24 at 2.25 Å resolution) were validated using the MolProbity [34] web server. Additional X-ray crystallographic statistics for Rpp20/Rpp25 are summarized in Supplementary Tables 1 and 2.

Structural alignments were performed in Coot [32] using the Secondary Structure Matching (SSM) superpose feature. Protein sidechain interactions were identified manually, but verified using calculations made by the program “contact” in CCP4 [31]. Secondary structure schematics were adapted from images generated by the Pro-origami web server [35]. The protein electrostatic potential maps were generated using APBS (Adaptive Poisson-Boltzmann Solver) [36-38] with identical parameters to allow for subsequent comparative analyses to be made. All structural models were depicted using PyMOL [39].

### Transcription and purification of the human RNase P RNA P3 stem loop

The P3 stem loop of the human RNase P RNA (H1 RNA) was produced *in vitro* using complementary DNA oligonucleotides purchased from Integrated DNA Technologies (Coralville, IA), which contained the T7 RNA polymerase promoter, as a direct run-off transcription template (Supplementary Table 3). Additional P3 stem loop constructs, in which the internal single-stranded regions were abolished by sequence complementarity in the opposing loop, were generated using similar DNA oligonucleotides (Supplementary Table 3). Standard transcription reaction conditions were used to generate RNA on the milligram scale [40]. To separate product RNA from initial reaction components, the reaction mixture was electrophoresed on a native 10% acrylamide gel in 89 mM Tris-HCl, pH 7.6, 89 mM boric acid (1× TB buffer), in which a principal RNA band was located by UV shadowing and extracted into 50 mM potassium acetate, pH 7.0, 200 mM potassium chloride by passive diffusion at 4°C. The P3 stem loop RNAs were subsequently purified by size exclusion chromatography using a Superdex 200 10/300 GL column (GE Healthcare) in 50 mM Tris-HCl, pH 7.5, 100 mM sodium chloride, and 5 mM β-mercaptoethanol.

### Electrophoretic mobility shift assay with human RNase P RNA P3 stem loop

Electrophoretic mobility shift assays (EMSA) were conducted using purified P3 stem loop RNA and Rpp20, Rpp25, or Rpp20/Rpp25. In brief, 3 μg of each P3 RNA construct were incubated with various stoichiometric ratios of Rpp20, Rpp25, or Rpp20/Rpp25 in 50 mM Tris-HCl, pH 7.5, 100 mM sodium chloride, and 5 mM β-mercaptoethanol (in a reaction

volume of 10 μl per sample) at 22°C for 30 minutes, then supplemented with 17% glycerol (v/v) (thereby increasing the sample volume to 15 μl), and electrophoresed on a native 10% acrylamide gel in 1× TB buffer at 22°C. P3 stem loop RNA was visualized qualitatively by gel staining with Toluidine blue (0.02% w/v) in 1× TB buffer. Since Rpp20, Rpp25, and Rpp20/Rpp25 do not enter the gel under these conditions alone, changes in the electrophoretic migratory behavior of P3 RNAs were interpreted as a direct effect of the presence of increasing amounts of Rpp20, Rpp25, or Rpp20/Rpp25.

## Results

### Purification of the human Rpp20 and Rpp25 heterodimer

In contrast to previous studies in which the apparent insolubility of recombinantly produced human Rpp20 necessitated either its denaturation and subsequent refolding [8, 16], or co-expression of Rpp20 with its binding partner Rpp25 [9], we found that soluble Rpp20 could be obtained simply by low temperature induction combined with using *E. coli* Rosetta(DE3) cells as an expression host to mitigate codon bias. Although Rpp20 derived in this way was still prone to precipitation over time (on the order of hours to days) and particularly at higher concentrations (> 1 mg/ml), its transient stability afforded us a direct means to obtain a relatively homogenous sample of Rpp20 (Supplementary Figure S1, Panels a and g). By comparison, Rpp25 was significantly more stable and was purified to apparent homogeneity using the same purification protocol (Supplementary Figure S1, Panels b and h). Consistent with previous biochemical characterization [9, 18-20], electrophoretic mobility shift assays (Figure 1, Panels b and c) demonstrated that while both Rpp20 and Rpp25 appeared to associate nonspecifically with the P3 region of the human RNase P RNA subunit (or H1 RNA, as depicted schematically in Figure 1, Panel a), neither formed a specific complex with the P3 stem loop RNA, as indicated by the disappearance of free RNA without an accompanying gel shift.

Heterodimerization of Rpp20 and Rpp25 occurred readily upon mixing of the two purified subunits (Supplementary Figure S1, Panel c) and helped to maintain the solubility of both subunits as the heterodimer can remain in solution for many months without significant aggregation or precipitation (data not shown). Since isolated Rpp20, even at low concentrations, precipitated out of solution on the order of hours to days, we combined *E. coli* cells in which Rpp20 and Rpp25 were individually expressed prior to cell lysis and protein extraction as an alternative way to obtain the heterodimer without co-expression. Although Rpp20/Rpp25 obtained in this manner was generally less pure (Supplementary Figure S1, Panel d), Rpp20/Rpp25 purified by both ways behaved identically in electrophoretic mobility shift assays. In particular, the addition of Rpp20/Rpp25 to P3 RNA at a stoichiometric ratio of 1:1 led to a complete and defined gel shift, indicative of the formation of a specific complex. However, the addition of even modestly greater amounts of Rpp20/Rpp25 relative to P3 RNA resulted in higher order aggregation, as indicated by the disappearance of a well-defined, shifted P3 RNA band (Figure 1, Panels d and e). As negative controls, electrophoretic mobility shift assays were performed under the same conditions and using the same RNA and protein concentrations to assess whether Rpp20/Rpp25 binds to P3 RNAs in which the internal single-stranded regions were abolished by

sequence complementarity. As expected, Rpp20/Rpp25 exhibits a clear preference for binding the single-stranded regions within the wild-type P3 RNA construct versus the continuous helical RNA elements comprising the two control constructs (Supplementary Figure S2).

## Crystal structure of the human Rpp20 and Rpp25 heterodimer

Although human Rpp20/Rpp25 crystallized readily after a relatively short period of time and under a variety of conditions, the largest and best diffracting crystals of Rpp20/Rpp25 took about a year to grow (Supplementary Figure S1, Panel e). Structural determination of Rpp20/Rpp25 was accomplished by experimental SIRAS phasing from crystals derivatized with potassium tetrachloroplatinate (II) (Supplementary Figure S1, Panel f; Supplementary Tables 1 and 2). The final model of Rpp20/Rpp25 ($R/R_{free}$ = 0.20/0.24) was obtained from data to 2.25 Å and, as a result of disordered regions in both the N- and C-termini and between the $\beta_3$ and $\beta_4$ antiparallel beta strands, consisted of residues 29 to 59, 63 to 114, and 119 to 140 of Rpp20 and residues 22 to 109 and 125 to 157 of Rpp25 (Figure 2, Panel b).

The overall structure of Rpp20/Rpp25 revealed two very structurally similar polypeptides exhibiting a shared but distinct core fold that is characteristic of the Alba superfamily of nucleic acid binding proteins at both the secondary and tertiary levels (Figure 2), consistent with prior bioinformatics analysis and predictions [24, 25]. Structure-based sequence comparisons and root-mean-square deviation (RMSD) calculations performed on superposed structures further demonstrated that the structures of Rpp20 and Rpp25 are indeed remarkably similar to other previously reported structures of Alba proteins [41, 42] and to their yeast homologues, Pop7 and Pop6 [21], respectively, without significant correlation to sequence conservation (Supplementary Figures S4 and S5). Specifically, we observed that in the context of generally low RMSD values, the sequence similarities between Rpp20 and Rpp25 with other Alba proteins across at least 80 structurally-aligned residues were consistently low [43]. Moreover, visual inspection of the superposed structures (for example, in Supplementary Figure S5) revealed that these differences were not only affected by, but also perhaps increased largely by both conformational variation and flexibility in the distal portions of the $\beta_3$ and $\beta_4$ antiparallel beta strands.

## Structural insights into the heterodimerization interface of Rpp20/Rpp25

Comparisons of the structure of Rpp20/Rpp25 with those of Pop6/Pop7 (Figure 3) and an Alba protein heterodimer [41] (Supplementary Figure S5) demonstrated that the propensity for dimerization between members of the Alba superfamily apparently extends to the P3 binding proteins of RNase P/MRP via a conserved interface [21, 41]. However, the structural alignments also revealed subtle differences in the relative orientation between the subunits of each heterodimer. To characterize better these differences in a systematic manner, we defined a general directional vector for each subunit using residues of its $\alpha_2$ alpha helix to then calculate the angle by which two subunits associate perpendicularly to the heterodimerization interface. Whereas the relative angle between the subunits of Rpp20/Rpp25 and of Pop6/Pop7 was measured to be 106° and 101°, respectively, the angle between

the subunits of an archaeal Alba heterodimer [41] and of another archaeal Alba homodimer [44] was measured to be more obtuse at 128° and 126°, respectively.

Inspection of the heterodimerization interface between Rpp20/Rpp25, Pop6/Pop7, and the two archaeal Alba hetero- and homodimers revealed a similar set of interactions underlying the manner by which each pair of proteins associate (Figure 4). Primarily, numerous nonspecific hydrophobic interactions make up the center of the interface, thereby extending and combining the extensive networks of hydrophobic sidechains that form the core region of each subunit. Flanking the hydrophobic center are also several key electrostatic and additional van der Waals interactions, formed primarily by non-hydrophobic sidechains that are in close physical proximity to one another, which further expand the heterodimerization interface.

A notable difference between the P3 binding proteins of RNase P/MRP, whose subunits are angled between 100° and 110° relative to each other, and the two archaeal Alba hetero- and homodimers, whose subunits are angled approximately 130° relative to each other, is the distribution and positioning of the peripheral electrostatic and van der Waals interactions relative to the central hydrophobic interactions forming the heterodimerization interface. In the archaeal Alba proteins, these electrostatic and van der Waals interactions are more extensively and symmetrically distributed adjacent to the proximal hydrophobic sidechains. Thus, with more intermolecular interactions between the $\alpha_2$ of one subunit and the $\beta_3$ and $\beta_4$ strands of its binding partner, there is increased coaxial alignment and, as a result, a more obtuse engagement between the two subunits along their long edges. In contrast, the flanking non-hydrophobic interactions between both Rpp20/Rpp25 and Pop6/Pop7 are not specifically conserved, are generally fewer in number, and are less evenly distributed on either side of the hydrophobic resides, thereby allowing for the $\alpha_2$ of one subunit to interact with more proximal residues of the $\beta_3$ and $\beta_4$ strands of its binding partner. As a result of these subtle changes in the heterodimerization interface of the P3 binding proteins of RNase P/MRP, their relative orientation is measurably less coplanar (Figure 4).

## Structural insights into the RNA binding interface of Rpp20/Rpp25

Although the crystal structure of Rpp20/Rpp25 was determined in the absence of RNA, a comparison with the structure of P3 RNA bound Pop6/Pop7 [21] revealed that RNA binding would not likely induce any substantial conformational change in either protein subunit or the angle by which Rpp20/Rpp25 associate based on their apparent homology with Pop6/Pop7 (Figure 3). Furthermore, our structure-based sequence alignments also demonstrated that the majority of the residues conserved between Rpp20 and Pop7 and between Rpp25 and Pop6 are those residing within their hydrophobic interior, whereas most of the residues lining the RNA binding interface of Pop6/Pop7 were poorly conserved in Rpp20/Rpp25. Yet, the overall positively charged surface of Pop6/Pop7 was indeed preserved in Rpp20/Rpp25, but by different positively charged residues at the sites where the proteins made contacts with the P3 RNA. This suggested that unless there are undetected covariations between protein and RNA, the predominant factors underlying RNA binding by these RNase P/MRP proteins are charge and surface geometry of the RNA binding surface created as a result of the assembly of the heterodimer.

To compare the RNA binding surface geometry of Rpp20/Rpp25 against that of Pop6/Pop7 and other Alba proteins, we generated an electrostatic potential map of each dimer using APBS [36-38]. To our surprise, we found that despite the striking structural resemblance of the individual subunits, the RNA binding surface of the RNase P/MRP proteins were comparably different than that of the Alba proteins (Figure 5). The electropositive RNA binding surface of both Rpp20/Rpp25 and Pop6/Pop7 form an elongated and curved surface, which appeared to be better suited for binding poorly structured, single-stranded RNAs. In contrast, both RNA binding and DNA binding Alba dimers exhibited a comparatively less extended and flatter electropositive binding surface that is seemingly adapted to binding helical structural elements in a regular and repeating manner, as evidenced by the packing of symmetry related dimers in the Alba crystal structures.

Much of the observed differences in RNA binding surface geometry directly correlate to the measured differences in the relative subunit orientation of each dimer. Since the subunits of Rpp20/Rpp25 and Pop6/Pop7 associate at an angle that is less obtuse than that of the Alba dimers, the resulting RNA binding surface effectively becomes more extended and less symmetric. This additional twist between the Rpp20/Rpp25 and Pop6/Pop7 subunits also allows for more peripheral residues to engage in substrate binding.

## Discussion

### Minimal induced conformational changes in Rpp20/Rpp25 upon RNase P RNA binding

The internal single-stranded regions of the RNase P/MRP RNA P3 stem loop and their associated binding proteins are eukaryotic specific adaptations to the RNase P/MRP holoenzyme; however, the exact function of this RNP module remains poorly understood even though there has been some suggestion that it assumes many roles, such as regulating subcellular localization of the holoenzymes [45, 46], facilitating proper folding of the RNase P/MRP RNA subunit [14, 47], recruiting Pop1 [23], and affecting enzymatic substrate recognition [47]. While previous genetic, biochemical, and bioinformatic analyses conducted on the human and yeast homologues demonstrated that the general structural and functional properties of Rpp20/Rpp25 and Pop6/Pop7 are indeed conserved, there is substantial sequence variation between both sets of binding proteins and their cognate RNA substrates. The only currently available structural data of the RNP module, which was obtained from a crystal structure of Pop6/Pop7 in complex with a yeast RNase MRP P3 fragment, suggested that a key function of Pop6/Pop7 is to help stabilize the local fold of its cognate RNase P/MRP RNA by replacing peripheral RNA structural elements lost in evolution [21]. Furthermore, there is recent evidence to suggest that Pop6/Pop7 acts additionally to help recruit Pop1, which in turn exerts stabilizing structural effects on the RNase P/MRP RNA on a global level [23]. Nevertheless, the yeast structure also revealed a unique and complex mechanism of association between protein and RNA involving mainly non-sequence specific interactions between an electropositive surface formed by the Pop6/Pop7 heterodimer and various residues of both single-stranded regions of the P3 RNA. Taken together, these findings do not preclude the possibility that significant differences may exist between the P3 modules of eukaryotic RNase P ranging from yeast to human.

In this study, we determined the crystal structure of human Rpp20/Rpp25 (Figure 2). Consistent with primary sequence analysis and secondary structure predictions, both Rpp20 and Rpp25 assume a distinctly Alba protein-like fold and, accordingly, heterodimerize in a manner reminiscent of that seen in Alba dimers (Supplementary Figure S5). Furthermore, superposition of Rpp20/Rpp25 on the MRP P3 bound Pop6/Pop7 demonstrated that in spite of RNA being absent in our structure, both homologous pairs are very similar in structure and the heterodimers overall do not exhibit any significant conformational differences at either the heterodimerization interface or RNA binding surface (Figure 3). The implication of these observations is that Rpp20/Rpp25 and its homologues are likely structurally rigid proteins upon which RNA binding induces minimal conformational change.

## Evolutionary conservation of the heterodimerization interface of Rpp20/Rpp25 and the mechanism of functional divergence of its RNA binding surface

The primary protein-protein interactions facilitating heterodimerization in Rpp20/Rpp25 are hydrophobic and conserved in Pop6/Pop7, as well as in all Alba dimers whose structures have been determined (Figure 4). Additionally, these residues interconnect the hydrophobic centers of each subunit, and in combination with flanking electrostatic and additional van der Waals interactions, allow the formation of an extremely stable heterodimer that is resistant to disassembly even under conditions of high ionic strength and in the presence of nonionic detergents, consistent with previous reports [13-15]. Whereas the core hydrophobic residues are essentially conserved across all dimers, the peripheral electrostatic and van der Waals interactions are not. Alba dimers include more extensive and more symmetrically distributed electrostatic and van der Waals interactions, owing in large part to greater homology shared between its subunits. In contrast, Rpp20/Rpp25 and Pop6/Pop7 contain fewer electrostatic and van der Waals interactions between their subunits and thus assemble in a comparatively twisted manner. As a result of this difference in the relative orientation between subunits, the electropositive RNA binding surface of both Rpp20/Rpp25 and Pop6/Pop7 becomes effectively elongated and less planar.

Remarkably, even though it is apparent that Rpp20/Rpp25 and Pop6/Pop7 are structurally and evolutionarily related to the archaeal Alba proteins, their RNA binding properties are drastically different. The P3 binding proteins of RNase P/MRP bind to poorly structured, intra-helical loops, whereas Alba dimers are archaeal chromatin binding proteins that bind helical segments of DNA or RNA in a manner that allows for higher order compaction. The functional divergence of the RNA binding properties of these two protein subsets is seemingly due to a relatively subtle difference in how their subunits heterodimerize, which results in binding surfaces that are similarly charged but with significantly different geometries. Moreover, this is a clear example of the evolutionary process occurring at the molecular level whereby a parent protein scaffold undergoes subtle structural changes to result in a pronounced alteration of function.

## RNA binding by Rpp20/Rpp25 does not abide by a strict sequence consensus

Despite many structural similarities shared between RNase P and MRP, the two enzymes are physically separate complexes with different and independent functions and substrate specificities [20]. Moreover, sequences of the internal loops of RNase P/MRP RNA P3 stem

loops from the same organism are typically not strictly conserved. In a recent study, it was observed that Pop6/Pop7 and Pop1 form a ternary complex that associates with yeast telomerase by binding to a P3-like region in the telomerase RNA subunit, Tlc1, and thereby converts the telomerase to its active form [48]. Although the activating mechanism of Pop1/Pop6/Pop7 on yeast telomerase remains to be elucidated, these findings illustrate a clear example of Pop6/Pop7 being utilized by an RNP complex other than RNase P or MRP. Additional studies have also demonstrated that not only are the other protein subunits of RNase P shared with RNase MRP, but Pop1, Rpp21, and Rpp29 together help regulate histone H3.3 chromatin formation in the absence of either P or MRP RNA [49]. Taken together, these findings suggest an alternative to the current view that these RNase P/MRP proteins are exclusive subunits of these complexes [50]. In particular, it appears that proteins such as Rpp20 and Rpp25 can have additional cellular and molecular functions by being incorporated into other RNP complexes. Thus, some of the RNase P proteins may not be exclusive to this RNP complex, but rather promiscuous RNA binding proteins that recognize different RNA structural elements and can be used by a variety of RNP complexes in similar or related roles. This observation also suggests that the role of these proteins in RNAs may be structural and not likely directly in catalysis or substrate recognition.

## Rpp20/Rpp25 are major components of the Th/To autoantigen complex in autoimmune diseases

While previous motivations to study Rpp20/Rpp25 have both emanated and developed largely from their established identity as a unique but essential adaptation to eukaryotic RNase P/MRP RNP complexes, human Rpp20/Rpp25 has long been recognized to constitute the major component of the Th/To autoantigen complex in scleroderma (also known as systemic sclerosis), which is a chronic but incurable and potentially devastating autoimmune disease that affects connective tissues [51-53]. A previous report has further showed that sera from patients suffering from systemic lupus erythematosus, polymyositis, and Raynaud disease can also demonstrate reactivity to anti-Th/To autoantibodies [54]; however, it remains unclear whether Th/To autoantigenicity in these specific cases was caused by the manifestation of a scleroderma-like pathology in an overlap syndrome. An often-made assumption in our current understanding of autoimmune diseases is that much of their pathophysiology results from autoantibody targeting of its associated antigen-containing complexes, such as RNase P/MRP in scleroderma. Yet, if Rpp20/Rpp25 is indeed shared by several functionally distinct RNP complexes, it raises the possibility that the pathogenesis of scleroderma might not be caused by autoimmune reactivity directed solely against RNase P/MRP, if at all.

## Locating the putative active site of Rpp20 ATPase activity

In a previous study, human Rpp20 was shown to exhibit an ATPase activity that was decoupled from tRNA precursor processing by RNase P [10]. Based on sequence analysis, it was speculated that this activity was located in a region showing sequence homology with both the C-terminal portion of an ATP binding cassette (ABC) transporter and the DIxxQ box ATPase motif of a RNA helicase subfamily [10]. Although mutations targeting both regions were generated by Li *et al.* [10] for ATPase assays, the location of the active site of Rpp20 ATPase activity is still unknown. For that reason, we also conducted a series of co-

crystallization trials and crystal soaks of Rpp20/Rpp25 supplemented with ATP and various nucleotide analogues but were ultimately unable to obtain structural data for a nucleotide bound Rpp20/Rpp25 complex. However, a superposition of the structures of Rpp20/Rpp25 and Pop6/Pop7 demonstrates that within the same beta strand, $\beta_1$, there is indeed an equivalent element to the human Rpp20 DIxxQ box in Pop6/Pop7 (Supplementary Figure S6, Panels a-c). While structural similarity suggests that these regions may assume a similar role in both Rpp20 and Pop7, there is a notable difference between the two proteins just downstream of the DIxxQ box (Supplementary Figure S6, Panels a, b, and d). In our crystal structure of Rpp20/Rpp25, residues 60A, 61R, and 62G are disordered; however, unlike in Pop6/Pop7, these residues could potentially form a short loop that bends back onto the DIxxQ box to generate an ATPase site and thus explain any differential ATPase activity between Rpp20 and its homologues. Finally, we note that this N-terminal region and putative ATPase site of Rpp20 is also appreciably physically separated from both the heterodimerization interface and RNA binding surface of Rpp20/Rpp25, which is consistent with the functional decoupling of the RNase P and putative ATPase activities [10].

## Conclusions

In our study, we determined the crystal structure of human Rpp20/Rpp25, and our structural model confirmed that Rpp20 and Rpp25 exhibit remarkably similar molecular structures to one another, to their yeast homologues (Pop7 and Pop6, respectively), and to archaeal Alba proteins. A superposition of the crystal structures of Rpp20/Rpp25 and of RNA bound yeast Pop6/Pop7 revealed that these proteins do not appear to undergo significant conformational change upon RNA binding, which suggests that they interact with their cognate RNAs via a pre-formed binding surface. Additionally, our comparative structural analyses show that the overall geometry and charge distribution of this binding surface, conserved between yeast and humans, is dictated largely by quaternary level interactions in the heterodimerization interface, rendering it generally more suitable for associating with unstructured, single-stranded RNA. If RNA binding is indeed governed by complementing surface geometries between protein and RNA, as well as additional specific residue-side chain interactions, this observation provides a physical explanation not only for the differential nucleic acid binding specificities of eukaryotic Rpp20/Rpp25 and their Alba structural homologues, but also for how specific Rpp20/Rpp25 homologues can accommodate sequence variation between at least two different RNA substrates, i.e. the P3 stem loop regions of RNase P/MRP, within the same organism.

In a cellular context, the requirement for RNA substrates to contain single-stranded regions that are both largely unstructured and accessible would reduce substantially the range of RNAs that Rpp20/Rpp25 can bind to, but still represent a large set of molecules. It is possible that additional factors associate with Rpp20/Rpp25 to help limit its binding targets *in vivo* to a subset of all cellular RNAs that conform to its structural requirements. For instance, in the case of RNase P/MRP, association with Pop1 may help augment Rpp20/Rpp25 specificity for the P3 stem loop of the RNase P and MRP RNA subunits over other RNAs that contain a similarly structured domain.

Recent and growing evidence indicate that several RNase P/MRP protein subunits are associated with other functionally-independent RNP complexes, such as the Pop6/Pop7 heterodimer and Pop1 with yeast telomerase [48]. A potential evolutionary basis for the promiscuous mode of complex association displayed by Rpp20/Rpp25 is that a structure-based binding consensus reduces the selective pressure for multiple complexes to simultaneously incorporate Rpp20/Rpp25 into their constitution in a strictly sequence-dependent manner. Although the concept of protein sharing is not a radical departure from what is known of certain specific and related multi-subunit macromolecular complexes [55], it is infrequently invoked since many well-characterized cellular machineries do consist of distinct and specific components. Our study provides an example wherein a simple structural adaptation at the quaternary level leads to a dramatic shift in binding specificity from helical nucleic acid (Alba) to single-stranded RNA (Rpp20/Rpp25). In the case of Rpp20/Rpp25, although specific residue-side chain interactions resulting from co-evolution within a set of protein dimers and cognate RNA can provide an additional means for increasing binding selectivity, this does not preclude the possibility that additional unidentified RNP complexes have also evolved to make use of a Rpp20/Rpp25/P3-like ternary structural module.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Dr. Sidney Altman and Dr. Cecilia Guerrier-Takada for providing us with plasmids encoding the human homologues of Rpp20 and Rpp25. We also acknowledge the beamline scientists at LS-CAT/Sector 21 at the Advanced Photon Source, Argonne National Laboratory, and support from the Structural Biology Facility at Northwestern University.

**Funding:** Funding was provided by the National Institutes of Health (NIH) Grant R01 GM058443 to A. M. and C. W. C. was supported by the Medical Scientist Training Grant 4T32 GM008152, the Molecular Biophysics Training Grant 5T32 GM008382, and an Achievement Rewards for College Scientists Foundation Fellowship. LS-CAT/Sector 21 at the Advanced Photon Source, Argonne National Laboratory was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor. Support from the R.H. Lurie Comprehensive Cancer Center of Northwestern University to the Structural Biology Facility is acknowledged.

## References

1. Liu, F., Altman, S. SpringerLink. Protein Reviews 10. New York, NY: Springe Science+Business Media, LLC; 2010. Ribonuclease P; p. 1Online serviceonline resource
2. Pinker F, Bonnard G, Gobert A, Gutmann B, Hammani K, Sauter C, et al. PPR proteins shed a new light on RNase P biology. RNA Biol. 2013; 10:1457–68. [PubMed: 23925311]
3. Klemm BP, Wu N, Chen Y, Liu X, Kaitany KJ, Howard MJ, et al. The Diversity of Ribonuclease P:Protein and RNA Catalysts with Analogous Biological Functions. Biomolecules. 2016; 6
4. Evans D, Marquez SM, Pace NR. RNase P: interface of the RNA and protein worlds. Trends Biochem Sci. 2006; 31:333–41. [PubMed: 16679018]
5. Esakova O, Krasilnikov AS. Of proteins and RNA: the RNase P/MRP family. RNA. 2010; 16:1725–47. [PubMed: 20627997]
6. Lai LB, Vioque A, Kirsebom LA, Gopalan V. Unexpected diversity of RNase P, an ancient tRNA processing enzyme: challenges and prospects. FEBS Lett. 2010; 584:287–96. [PubMed: 19931535]
7. van Eenennaam H, Jarrous N, van Venrooij WJ, Pruijn GJ. Architecture and function of the human endonucleases RNase P and RNase MRP. IUBMB Life. 2000; 49:265–72. [PubMed: 10995027]

8. Guerrier-Takada C, Eder PS, Gopalan V, Altman S. Purification and characterization of Rpp25, an RNA-binding protein subunit of human ribonuclease P. RNA. 2002; 8:290–5. [PubMed: 12003489]
9. Hands-Taylor KL, Martino L, Tata R, Babon JJ, Bui TT, Drake AF, et al. Heterodimerization of the human RNase P/MRP subunits Rpp20 and Rpp25 is a prerequisite for interaction with the P3 arm of RNase MRP RNA. Nucleic Acids Res. 2010; 38:4052–66. [PubMed: 20215441]
10. Li Y, Altman S. A subunit of human nuclear RNase P has ATPase activity. Proc Natl Acad Sci U S A. 2001; 98:441–4. [PubMed: 11149958]
11. Mahler M, Satoh M, Hudson M, Baron M, Chan JY, Chan EK, et al. Autoantibodies to the Rpp25 component of the Th/To complex are the most common antibodies in patients with systemic sclerosis without antibodies detectable by widely available commercial tests. J Rheumatol. 2014; 41:1334–43. [PubMed: 24931955]
12. Reiner R, Alfiya-Mor N, Berrebi-Demma M, Wesolowski D, Altman S, Jarrous N. RNA binding properties of conserved protein subunits of human RNase P. Nucleic Acids Res. 2011; 39:5704–14. [PubMed: 21450806]
13. Welting TJ, Kikkert BJ, van Venrooij WJ, Pruijn GJ. Differential association of protein subunits with the human RNase MRP and RNase P complexes. RNA. 2006; 12:1373–82. [PubMed: 16723659]
14. Welting TJ, Peters FM, Hensen SM, van Doorn NL, Kikkert BJ, Raats JM, et al. Heterodimerization regulates RNase MRP/RNase P association, localization, and expression of Rpp20 and Rpp25. RNA. 2007; 13:65–75. [PubMed: 17119099]
15. Welting TJ, van Venrooij WJ, Pruijn GJ. Mutual interactions between subunits of the human RNase MRP ribonucleoprotein complex. Nucleic Acids Res. 2004; 32:2138–46. [PubMed: 15096576]
16. Jarrous N, Eder PS, Guerrier-Takada C, Hoog C, Altman S. Autoantigenic properties of some protein subunits of catalytically active complexes of human ribonuclease P. RNA. 1998; 4:407–17. [PubMed: 9630247]
17. Stolc V, Katz A, Altman S. Rpp2, an essential protein subunit of nuclear RNase P, is required for processing of precursor tRNAs and 35S precursor rRNA in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A. 1998; 95:6716–21. [PubMed: 9618478]
18. Perederina A, Esakova O, Koc H, Schmitt ME, Krasilnikov AS. Specific binding of a Pop6/Pop7 heterodimer to the P3 stem of the yeast RNase MRP and RNase P RNAs. RNA. 2007; 13:1648–55. [PubMed: 17717080]
19. Perederina A, Esakova O, Quan C, Khanova E, Krasilnikov AS. Crystallization and preliminary X-ray diffraction analysis of the P3 RNA domain of yeast ribonuclease MRP in a complex with RNase P/MRP protein components Pop6 and Pop7. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2010; 66:76–80.
20. Perederina A, Krasilnikov AS. The P3 domain of eukaryotic RNases P/MRP: making a protein-rich RNA-based enzyme. RNA Biol. 2010; 7:534–9. [PubMed: 20523128]
21. Perederina A, Esakova O, Quan C, Khanova E, Krasilnikov AS. Eukaryotic ribonucleases P/MRP: the crystal structure of the P3 domain. EMBO J. 2010; 29:761–9. [PubMed: 20075859]
22. Khanova E, Esakova O, Perederina A, Berezin I, Krasilnikov AS. Structural organizations of yeast RNase P and RNase MRP holoenzymes as revealed by UV-crosslinking studies of RNA-protein interactions. RNA. 2012; 18:720–8. [PubMed: 22332141]
23. Fagerlund RD, Perederina A, Berezin I, Krasilnikov AS. Footprinting analysis of interactions between the largest eukaryotic RNase P/MRP protein Pop1 and RNase P/MRP RNA components. RNA. 2015; 21:1591–605. [PubMed: 26135751]
24. Aravind L, Iyer LM, Anantharaman V. The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. Genome Biol. 2003; 4:R64. [PubMed: 14519199]
25. Goyal M, Banerjee C, Nag S, Bandyopadhyay U. The Alba protein family: Structure and function. Biochim Biophys Acta. 2016; 1864:570–83. [PubMed: 26900088]
26. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010; 66:125–32. [PubMed: 20124692]
27. Powell HR. The Rossmann Fourier autoindexing algorithm in MOSFLM. Acta Crystallogr D Biol Crystallogr. 1999; 55:1690–5. [PubMed: 10531518]

28. Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013; 69:1204–14. [PubMed: 23793146]
29. Vonrhein C, Blanc E, Roversi P, Bricogne G. Automated structure solution with autoSHARP. Methods Mol Biol. 2007; 364:215–30. [PubMed: 17172768]
30. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006; 62:1002–11. [PubMed: 16929101]
31. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011; 67:235–42. [PubMed: 21460441]
32. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004; 60:2126–32. [PubMed: 15572765]
33. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997; 53:240–55. [PubMed: 15299926]
34. Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr. 2010; 66:12–21. [PubMed: 20057044]
35. Stivala A, Wybrow M, Wirth A, Whisstock JC, Stuckey PJ. Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics. 2011; 27:3315–6. [PubMed: 21994221]
36. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc Natl Acad Sci U S A. 2001; 98:10037–41. [PubMed: 11517324]
37. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G, et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. Nucleic Acids Res. 2007; 35:W522–5. [PubMed: 17488841]
38. Unni S, Huang Y, Hanson RM, Tobias M, Krishnan S, Li WW, et al. Web servers and services for electrostatics calculations with APBS and PDB2PQR. J Comput Chem. 2011; 32:1488–91. [PubMed: 21425296]
39. Schrodinger LLC. The PyMOL Molecular Graphics System. Version 1.8 ed. 2015
40. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic Acids Res. 1987; 15:8783–98. [PubMed: 3684574]
41. Jelinska C, Conroy MJ, Craven CJ, Hounslow AM, Bullough PA, Waltho JP, et al. Obligate heterodimerization of the archaeal Alba2 protein with Alba1 provides a mechanism for control of DNA packaging. Structure. 2005; 13:963–71. [PubMed: 16004869]
42. Tanaka T, Padavattan S, Kumarevel T. Crystal structure of archaeal chromatin protein Alba2-double-stranded DNA complex from Aeropyrum pernix K1. J Biol Chem. 2012; 287:10394–402. [PubMed: 22334696]
43. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. 2005; 33:2302–9. [PubMed: 15849316]
44. Wardleworth BN, Russell RJ, Bell SD, Taylor GL, White MF. Structure of Alba: an archaeal chromatin protein modulated by acetylation. EMBO J. 2002; 21:4654–62. [PubMed: 12198167]
45. Jacobson MR, Cao LG, Wang YL, Pederson T. Dynamic localization of RNase MRP RNA in the nucleolus observed by fluorescent RNA cytochemistry in living cells. J Cell Biol. 1995; 131:1649–58. [PubMed: 8557735]
46. Jacobson MR, Cao LG, Taneja K, Singer RH, Wang YL, Pederson T. Nuclear domains of the RNA subunit of RNase P. J Cell Sci. 1997; 110(Pt 7):829–37. [PubMed: 9133670]
47. Walker SC, Engelke DR. Ribonuclease P: the evolution of an ancient RNA enzyme. Crit Rev Biochem Mol Biol. 2006; 41:77–102. [PubMed: 16595295]
48. Lemieux B, Laterreur N, Perederina A, Noel JF, Dubois ML, Krasilnikov AS, et al. Active Yeast Telomerase Shares Subunits with Ribonucleoproteins RNase P and RNase MRP. Cell. 2016; 165:1171–81. [PubMed: 27156450]

49. Newhart A, Powers SL, Shastrula PK, Sierra I, Joo LM, Hayden JE, et al. RNase P protein subunit Rpp29 represses histone H3.3 nucleosome deposition. Mol Biol Cell. 2016; 27:1154–69. [PubMed: 26842893]
50. Jarrous N. Roles of RNase P and Its Subunits. Trends Genet. 2017; 33:594–603. [PubMed: 28697848]
51. Van Eenennaam H, Vogelzangs JH, Lugtenberg D, Van Den Hoogen FH, Van Venrooij WJ, Pruijn GJ. Identity of the RNase MRP- and RNase P-associated Th/To autoantigen. Arthritis Rheum. 2002; 46:3266–72. [PubMed: 12483731]
52. Welting TJ, Raijmakers R, Pruijn GJ. Autoantigenicity of nucleolar complexes. Autoimmun Rev. 2003; 2:313–21. [PubMed: 14550872]
53. Jee AS, Adelstein S, Bleasel J, Keir GJ, Nguyen M, Sahhar J, et al. Role of Autoantibodies in the Diagnosis of Connective-Tissue Disease ILD (CTD-ILD) and Interstitial Pneumonia with Autoimmune Features (IPAF). J Clin Med. 2017; 6
54. Van Eenennaam H, Vogelzangs JH, Bisschops L, Te Boome LC, Seelig HP, Renz M, et al. Autoantibodies against small nucleolar ribonucleoprotein complexes and their clinical associations. Clin Exp Immunol. 2002; 130:532–40. [PubMed: 12452846]
55. Matalon O, Horovitz A, Levy ED. Different subunits belonging to the same protein complex often exhibit discordant expression levels and evolutionary properties. Curr Opin Struct Biol. 2014; 26:113–20. [PubMed: 24997301]
56. Guo L, Ding J, Guo R, Hou Y, Wang DC, Huang L. Biochemical and structural insights into RNA binding by Ssh10b, a member of the highly conserved Sac10b protein family in Archaea. J Biol Chem. 2014; 289:1478–90. [PubMed: 24307170]


## Abbreviations

| | |
|---|---|
| **RNase** | ribonuclease |
| **SIRAS** | single isomorphous replacement with anomalous scattering |
| **PDB** | Protein Data Bank |
| **EMSA** | electrophoretic mobility shift assay |
| **RNP** | ribonucleoprotein |

## Highlights

- Rpp20/Rpp25 is a heterodimeric protein subunit of eukaryotic RNase P and MRP endonucleases
- Crystal structure of the human Rpp20/Rpp25 was determined at 2.25 Å resolution
- Rpp20/Rpp25 is evolutionarily related and structurally homologous to archaeal Alba chromatin proteins
- Divergent nucleic acid binding properties between Rpp20/Rpp25 and Alba protein dimers correlate primarily with structural differences in their heterodimerization interfaces, resulting in binding surface geometries specific for either single-stranded RNA or helical nucleic acid binding

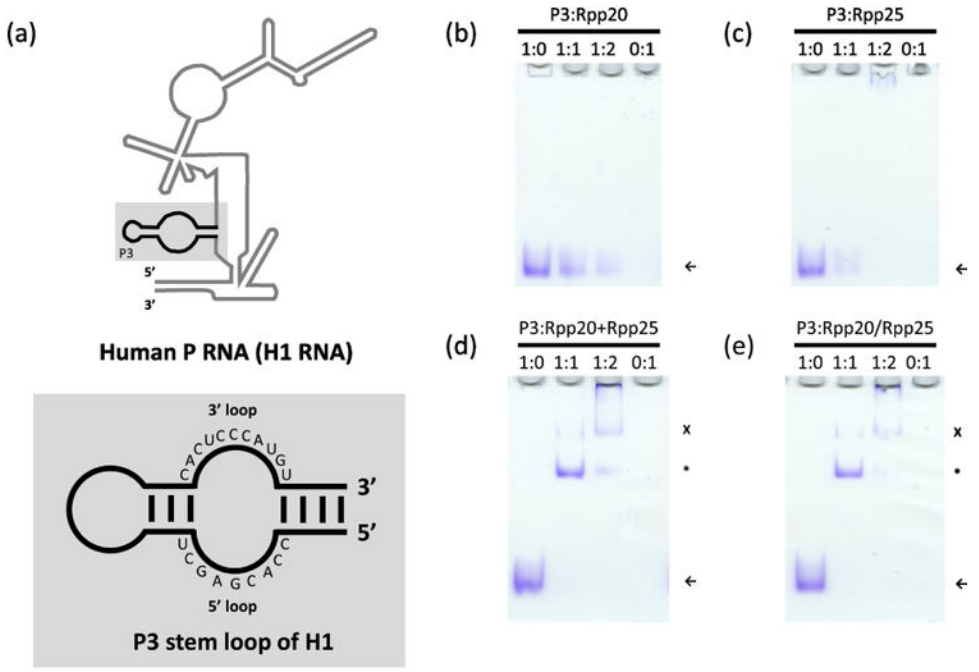



**Figure 1. Binding of human Rpp20, Rpp25, and the Rpp20/Rpp25 heterodimer to the P3 stem loop**

(a) Top: schematic diagram of the RNA component of human RNase P (H1 RNA). The position of the P3 loop is shown in black. Bottom: schematic diagram of the P3 stem loop. Electrophoretic mobility shift assay of *in vitro* transcribed human RNase P RNA P3 stem loop by recombinantly purified human (b) Rpp20, (c) Rpp25, and (d, e) the Rpp20/Rpp25 heterodimer. Arrows show the position of free P3 RNA, whereas asterisks show the position of gel shifted, protein bound P3. No significant differences were observed as a result of forming the Rpp20/Rpp25 heterodimer by (d) post purification mixing of Rpp20 to Rpp25 or by (e) co-purification of both subunits. Despite the disappearance of free P3 RNA, Rpp20 and Rpp25 alone do not produce a clear gel shift, whereas Rpp20/Rpp25 associates readily with the P3 RNA to produce a defined complex. Even at a modest stoichiometric excess of Rpp20/Rpp25, higher order oligomers or assemblies appear to form (crosses).

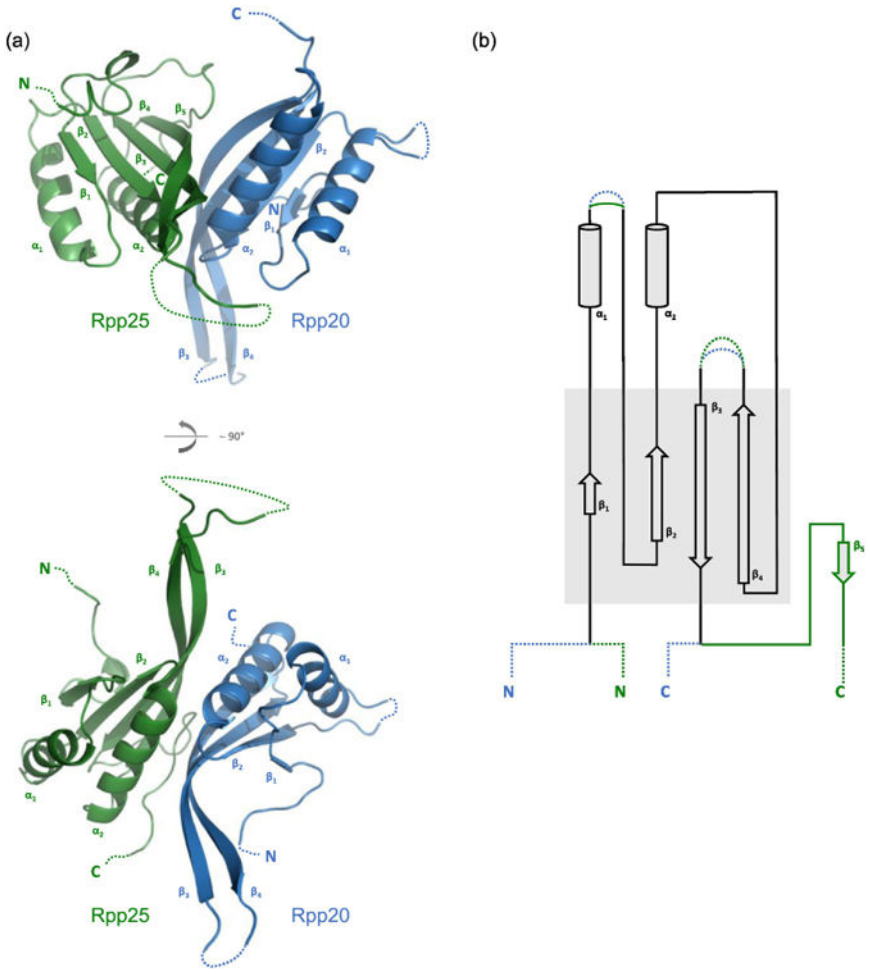



**Figure 2. Crystal structure of the human Rpp20/Rpp25 heterodimer**

(a) Cartoons showing the structure of the Rpp20/Rpp25 heterodimer in orthogonal views. Rpp20 and Rpp25 are shown in blue and green, respectively. Disordered regions are indicated by dotted lines. (b) Schematic diagram of the secondary structures of Rpp20 and Rpp25. The diagrams reveal a common βαβαββ fold (black) identical to that of their yeast homologues, Pop7 and Pop6, respectively, and of Alba proteins in general [21, 41, 42, 44, 56]. The core of Alba proteins consists of four beta strands ($\beta_1$-$\beta_4$) forming a sheet through mostly hydrophobic interactions. Alpha helices $\alpha_1$ and $\alpha_2$ are roughly coplanar and point in generally the same direction as $\beta_1$-$\beta_4$. In addition to the Alba fold, Rpp25 contains an additional short beta strand ($\beta_5$) that lies coplanar to $\beta_1$-$\beta_4$ and is present in yeast Pop6 as well. Disordered N- and C-terminal and internal loop regions in the crystal structure are shown as dotted lines in the secondary structure diagram [35].

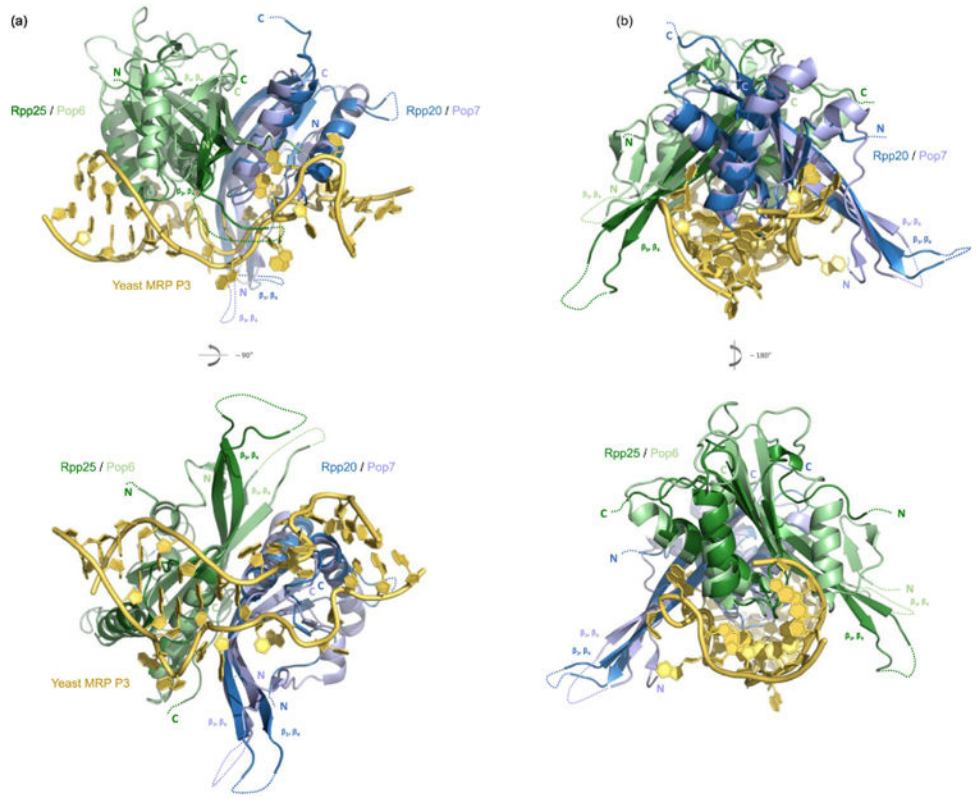



**Figure 3. Superposition of human Rpp20/Rpp25 and yeast MRP P3 RNA bound Pop6/Pop7**

The superposition of the human Rpp20/Rpp25 heterodimer structure on the yeast MRP P3 RNA bound Pop6/Pop7 structure (PDB ID: 3IAB) [21] reveals minimal overall conformational differences between the RNA free and RNA bound structures of the homologous heterodimers. Despite exhibiting relatively low sequence similarity, Rpp20 and Pop7, as well as Rpp25 and Pop6, exhibit significant homology at the tertiary level. Notable structural differences between the two homologous pairs occur in terminal regions that do not facilitate protein-protein or RNA-RNA interactions [9], whereas the heterodimerization interface and RNA binding surface are both conserved. Regardless of whether RNA is present, the regions that make up the electropositive RNA binding surfaces align closely, even though there is divergence between the two surfaces at the sequence level. The long antiparallel beta strands $\beta_3$ and $\beta_4$ of all four polypeptides shown here are disordered to different extents in both crystal structures and represent one of the more flexible, but still conserved regions of these proteins. (a) Orthogonal views of the superposed structures and protein-RNA interaction. (b) End-on views of the superposed structures show that the P3 stem loop binds primarily to proximal regions of Rpp20/Pop7 and Rpp25/Pop6 and does not make contacts with distal regions of the beta strands $\beta_3$ and $\beta_4$ of either polypeptide.

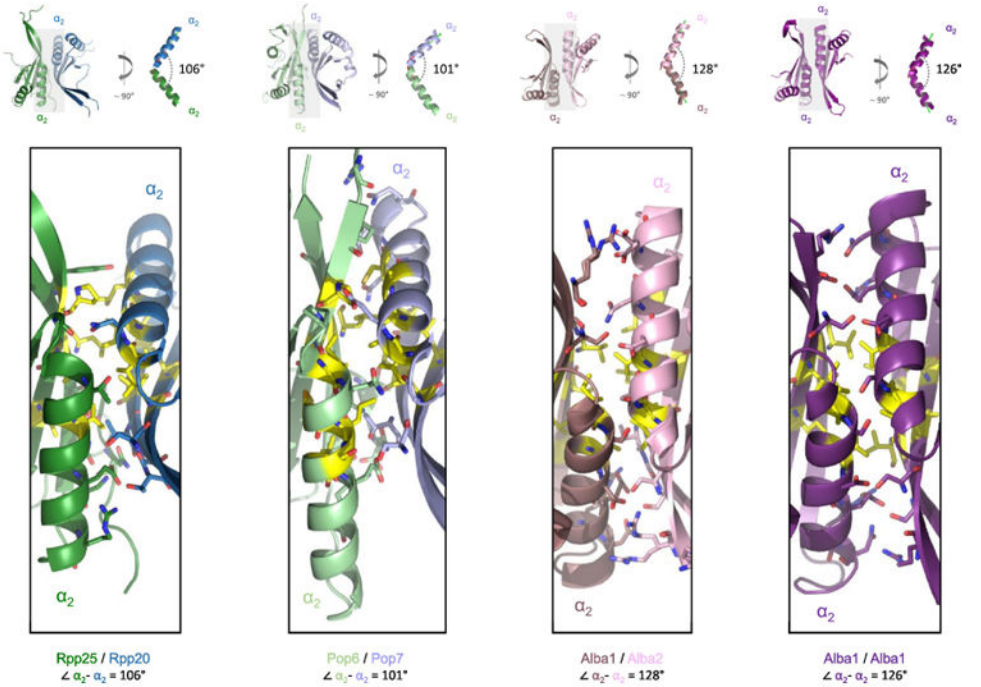



**Figure 4. Comparison of the heterodimerization interface of the human Rpp20/Rpp25 heterodimer, yeast Pop6/Pop7 heterodimer, an *S. solfataricus* Alba1/Alba2 heterodimer, and an *S. solfataricus* Alba1 homodimer reveals a conserved mechanism of association between their respective subunits**

Although each protein pair associate in a generally similar fashion (top), a notable difference between the eukaryotic heterodimers (Rpp20/Rpp25 and Pop6/Pop7) and their archaeal counterparts is the relative orientation of the subunits. We quantified this difference by calculating the angle between vectors through the helical axis of $\alpha_2$ of each subunit, which we took to represent the general direction of each subunit given the apparent rigidity of the core of each protein. Whereas such a measure of inter-subunit orientation resulted in 106° and 101° for Rpp20/Rpp25 and Pop6/Pop7, respectively, the angles between the archaeal Alba heterodimer and homodimer were more obtuse at 128° and 126°, respectively. The structural basis for this variation in subunit orientation appears to originate from differences in the distribution of additional electrostatic and van der Waals interactions flanking the central hydrophobic interactions at the heterodimerization interface (top, boxes shaded in grey). Insets show a close-up view of the heterodimerization interfaces. Every heterodimerization interface consists of an extensive network of hydrophobic residues (colored in yellow) that are in turn flanked by additional electrostatic and van der Waals interactions (by side chains shown in stick representation). Together, these interactions enable the formation of a stable homo/heterodimer, consistent with other observations [9, 14, 15, 41]. Between the subunits of Rpp20/Rpp25 and Pop6/Pop7, the peripheral protein-protein interactions are fewer in number and less symmetrically distributed along the long edge of each subunit. In contrast, the corresponding peripheral protein-protein interactions between the archaeal proteins are far more extensive and more evenly distributed along the long edge of each subunit, thus creating a larger interaction surface and a more obtuse angle of association. The atomic coordinates of Pop6/Pop7, Alba1/Alba2, and Alba1/Alba1 correspond to PDB ID: 3IAB [21], 2BKY [41], and 1H0X [44], respectively.

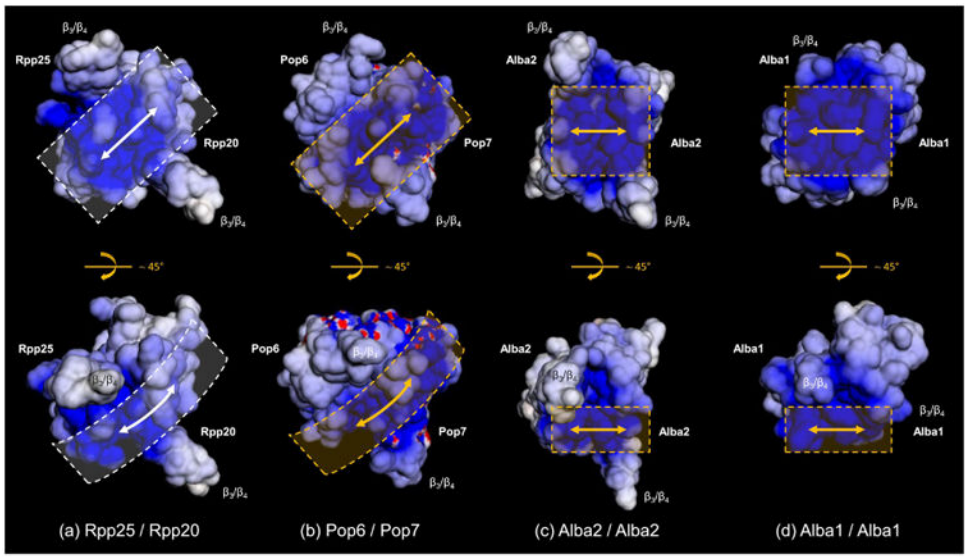



**Figure 5. Differences in the nucleic acid binding surfaces of the human Rpp20/Rpp25 heterodimer, yeast Pop6/Pop7 heterodimer, an *S. shibitae* Alba1/Alba1 heterodimer, and an *A. pernix* Alba2 homodimer**

Electrostatic potential of the nucleic acid binding surface of (a) human Rpp20/Rpp25, (b) yeast Pop6/Pop7 (PDB ID: 3IAB) [21], (c) an archaeal homodimer of Alba2/Alba2 from *A. pernix* (PDB ID: 3U6Y) [42], and (d) an archaeal homodimer of Alba1/Alba1 from *Sulfolobus shibatae* (PDB ID: 3WBM) [56] calculated using APBS [36-38] and rendered with a range of +/- 8 kT/e. Whereas the crystal structure of Rpp20/Rpp25 was solved in the absence of nucleic acid, the crystal structures of Pop6/Pop7 [21, 42], *A. pernix* Alba2/Alba2 [42], and *S. shibatae* Alba1/Alba1 [56] were solved in complex with RNA, DNA, and RNA, respectively. The two archaeal Alba homodimers are purposefully shown here as representative examples to compare a DNA binding Alba dimer and a helical RNA binding Alba dimer. Rpp20/Rpp25 and Pop6/Pop7 contain a relatively curved and elongated electropositive RNA binding surface (represented by an arrow drawn in a lightly shaded boxed region), which bind to less structured, single-stranded RNAs rather than helical elements. In contrast, and in spite of the type of nucleic acid to which they bind, both archaeal Alba homodimers exhibit a more compact and flatter electropositive surface, which facilitates binding to nucleic acid helices. The distal end of $\beta_3/\beta_4$ of each polypeptide is labeled to provide landmarks for the orientations shown of each dimer. The yellow shaded regions are drawn according to where the proteins contact RNA or DNA in the respective crystal structures, whereas the white shaded region depicted on Rpp20/Rpp25 is extrapolated from the crystal structure of Pop6/Pop7 bound to P3 RNA (PDB ID: 3IAB) [21].