

The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA

CRYSTAL L. YOUNG¹ and KATRIN KARBSTEIN¹

Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055, USA

ABSTRACT

RNA-binding proteins mediate the function of all RNAs. Since few distinct RNA-binding domains (RBDs) exist, with most RBDs contacting only a few nucleotides, RNA-binding proteins often combine multiple RNA-binding motifs to achieve a higher affinity and selectivity for their targets. Rrp5, a ribosome assembly factor essential for both 40S and 60S ribosome maturation, is an extreme example as it contains 12 tandem S1 RNA-binding domains. In this study, we use a combination of RNA binding and DMS probing experiments to probe interactions of Rrp5 with pre-rRNA mimics. Our data localize Rrp5's binding site to three distinct regions within internal transcribed spacer 1 (ITS1), the sequence between 18S and 5.8S rRNAs. One of these regions is directly adjacent to a recently uncovered helical structure, which prevents premature cleavage at the 3'-end of 18S rRNA. This finding, together with previous results, suggests a role for Rrp5 in regulating the above-mentioned helical element. Furthermore, we have produced two truncated forms of the protein, Rrp5N and Rrp5C, which together encompass the entire protein and fully restore growth. Quantitative analysis of the RNA affinity of these Rrp5 fragments indicates that the first nine S1 motifs contribute much of Rrp5's RNA affinity, while the last three domains alone provide its specificity for the pre-rRNA. This surprising division of labor is unique, as it suggests that S1 domains can bind RNA both specifically as well as nonspecifically with high affinity; this has important implications for the molecular details of the Rrp5•pre-rRNA complex.

Keywords: RNA-binding protein; S1 domain; ribosome assembly

INTRODUCTION

All functions of RNA, including cellular localization, translation, degradation, splicing, and catalysis, are mediated by RNA-binding proteins. Even though their roles are diverse, most RNA-binding proteins comprise a basic set of RNA-binding modules (Lunde et al. 2007). Examples include RNA-recognition motifs (RRMs), K-Homology (KH) domains, Zn-fingers, double-stranded RNA-binding domains (dsRBDs), as well as S1 motifs. While the repertoire of RNA-binding domains is limited, the large variety of cellular functions of RNA-binding proteins as well as their specificity for distinct RNA targets often arises because of the presence of multiple RNA-binding domains, either as copies of one module, combinations of different modules, or both.

A unique case of an RNA-binding protein with multiple RNA-binding domains is Rrp5, which contains 12 tandem S1 domains. Originally named for the *Escherichia coli* ribosomal protein S1 (Subramanian 1983), the S1 motif consists of approximately 70 amino acids that are folded into a five-stranded anti-parallel β -barrel that contains a short 3_{10} helix cap. S1 domains are part of the larger oligonucleotide/oligosaccharide binding (OB)-fold superfamily (Murzin 1993; Bycroft et al. 1997). Even though members of this superfamily have little, if any, sequence similarity, they fold into highly similar structures; many also bind their ligands on a conserved surface formed by the β -sheet (mostly β_2 and β_3) and the loops between β_1 and β_2 , as well as β_3 and α (Fig. 1C; Theobald et al. 2003).

Conserved from yeast to humans, Rrp5 is an essential protein required for ribosome assembly, where it functions as part of a large macromolecular machinery that is unique to eukaryotes. This machinery is composed of over 200 protein and RNA factors, which are required for facilitating the processing and folding of the rRNAs and the binding of the ribosomal proteins. Since three of the four rRNAs are transcribed as a single precursor, rRNA maturation must occur via a series of endo- and exonucleolytic steps, which

¹Present address: Department of Cancer Biology, The Scripps Research Institute, Jupiter, Florida 33458, USA.

Reprint requests to: Katrin Karbstein, Department of Cancer Biology, The Scripps Research Institute, 130 Scripps Way #2C1, Jupiter, FL 33458, USA; e-mail: kkarbst@scripps.edu; fax: (561) 228-2920.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.2458811>.

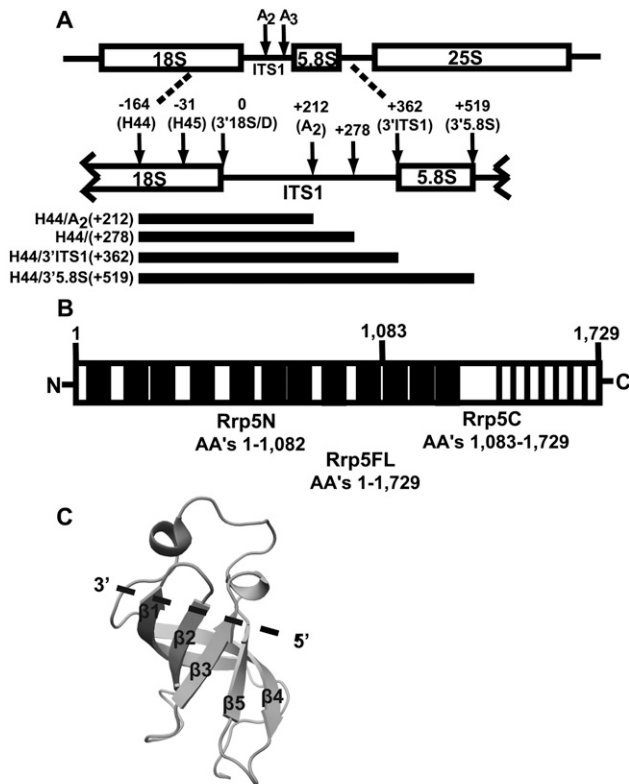


FIGURE 1. (A) Pre-rRNA fragments used in this study. The 18S, 5.8S, and 25S rRNAs are cotranscribed in a single transcript. Experiments described herein use rRNA fragments spanning the internal transcribed spacer 1 (ITS1), which includes cleavage sites A_2 and A_3 (Lamanna and Karbstein 2009). Specifically, rRNA mimics begin at helix 44 (H44) or helix 45 (H45) at the 5'-end and extend to either A_2 , (+278), 3'ITS1 or 3'5.8S at the 3'-end. (B) Rrp5 schematic and protein constructs. Rrp5 has 12 S1 RNA-binding domains (solid boxes) at its N terminus and seven tetratricopeptide (TPR) motifs (open boxes) at its C terminus. Full-length Rrp5 comprises amino acids (AA) 1–1729, the N-terminal fragment Rrp5N contains amino acids 1–1082, and the C-terminal fragment Rrp5C comprises amino acids 1083–1729. All proteins were expressed and purified from *E. coli*. (C) Cartoon model of the S1 domain of aIF2 α (PDB 1yz6). The five β strands characteristic of an S1 domain are labeled; the α -helix frequently found between strands 3 and 4 is also shown. The dashed line indicates the putative binding site of RNA and is shown using the conserved polarity of OB-fold proteins (N terminus binds the 3' end while the C terminus binds the 5' end) (Theobald et al. 2003).

have been well described. In contrast to almost all other ribosome assembly factors, Rrp5 is required for production of both the 40S and the 60S subunit (Venema and Tollervey 1996; Torchet et al. 1998; Eppens et al. 1999).

Previous work has indicated that Rrp5 can be physically and functionally separated between the ninth and 10th S1 motif. N-terminal or C-terminal truncation of Rrp5 between the ninth and 10th S1 motif is lethal. When provided in *trans* in Rrp5-depleted cells, however, the truncated proteins restore cell viability (Eppens et al. 1999). Furthermore, by analyzing the effects on rRNA processing when

mutations or deletions were introduced into the N- or C-terminal pieces, it was shown that the N-terminal portion was required for cleavage at the so-called site A_3 , while the C-terminal part was required for upstream cleavage at the so-called site A_2 (Fig. 1A; Torchet et al. 1998; Eppens et al. 1999). These cleavage sites are located ~ 70 nucleotides (nt) apart from each other within the internal transcribed spacer 1 (ITS1), the sequence between the 18S and 5.8S rRNAs (Fig. 1A), which is removed during maturation. This suggested that Rrp5 may have interactions within the ITS1, close to sites A_2 and A_3 .

Herein, we have mapped Rrp5's rRNA binding sites in a combination of RNA binding and DMS probing experiments that use a series of RNAs from the ITS1 and surrounding regions. The data pinpoint Rrp5's binding site to three regions within the ITS1, with most interactions occurring 3' to site A_2 . Furthermore, we have analyzed the contributions of two fragments of Rrp5 containing either the first nine (S1₁₋₉) or the last three (S1₁₀₋₁₂) S1 domains to RNA binding. Surprisingly, our data show that specificity for the RNA target arises almost exclusively from the last three RNA binding domains, while the first nine motifs contribute most of the RNA binding affinity. This finding also suggests that the specific interactions we have found in the DMS protection assay result from interactions with the last three S1 domains.

RESULTS

To understand the interplay between Rrp5's S1 RNA-binding motifs in its interaction with pre-rRNA, we carried out an *in vitro* characterization with recombinant Rrp5 expressed in *E. coli*. Even though Rrp5 is a 193-kDa protein containing 12 tandem S1 RNA-binding motifs, prior *in vivo* work had shown that it can be effectively separated between the ninth and 10th S1 RNA-binding domains, with the two proteins providing for growth comparable to wild-type Rrp5 when provided in *trans* (Eppens et al. 1999). We prepared two truncations that are highly similar to these published fragments in that they contain the same number of S1 motifs; however, our truncation site was shifted by 48 amino acids in order to provide for more stable constructs, facilitating expression in bacteria. The N-terminal portion of Rrp5, called Rrp5N hereafter, comprises amino acids 1–1082 (S1₁₋₉), while Rrp5C, the C-terminal fragment, contains residues 1083–1729 (S1₁₀₋₁₂) (Fig. 1B). In agreement with prior *in vivo* studies, these two fragments fully complement growth in Rrp5-depleted cells when provided in *trans* (Fig. 2), indicating that they are fully functional *in vivo*. The observation that Rrp5C, when expressed alone, provides for more growth than does expression of Rrp5N alone can be explained by the previous finding that cleavage at the Rrp5N-dependent cleavage site A_3 is not essential (Henry et al. 1994). Both truncations, as well as full-length Rrp5 (Rrp5FL), were overexpressed and purified from *E. coli*.

from the observation that Rrp5C does not bind the TSL fragment at the highest concentration used (5 μM), and Rrp5N and Rrp5FL bind TSL more weakly, as their interactions can be competed off with both unlabeled H44/3'ITS1 and unlabeled polyU (Fig. 3B). However, it should be noted that Rrp5N's affinity for TSL is only twofold below its affinity for H44/3'ITS1 ($K_{1/2} = 0.1 \pm 0.01 \mu\text{M}^2$ and $0.2 \pm 0.1 \mu\text{M}^2$ for H44/3'ITS1 and TSL, respectively), while Rrp5FL's affinity is decreased at least 30-fold (data not shown). These data indicate that Rrp5N binding is relatively nonspecific (see Rrp5N Provides Affinity, Rrp5C Provides Specificity for pre-rRNA below); in contrast, binding of Rrp5C and Rrp5FL to the fragments containing ITS1 is specific.

Comparison of lane 1 with lanes 2, 6, and 10 in Figure 3A also indicates that upon binding of Rrp5C, RNA is shifted upward in the gel, while the same RNA is shifted into the well upon binding of Rrp5N or Rrp5FL. We believe that this observation simply reflects the large size of the RNA-protein complexes (560 kDa, 425 kDa, and 310 kDa for Rrp5FL, Rrp5N, and Rrp5C, respectively) and not nonspecific aggregation, as the complexes formed with Rrp5N and Rrp5FL are amenable to competition with unlabeled RNA (Fig. 3A,C).

Rrp5N provides affinity, Rrp5C provides specificity for pre-rRNA

Knowing that Rrp5N, Rrp5C, and Rrp5FL all have affinity for rRNA constructs including ITS1, we next wanted to quantitatively measure this interaction and determine which region(s) in the protein and rRNA are important for Rrp5 binding.

We first compared binding of Rrp5N, Rrp5C, and Rrp5FL to H44/3'ITS1. The data in Figure 4 show that Rrp5FL binds the strongest, followed by Rrp5N and then Rrp5C. For this RNA, binding to Rrp5FL is approximately fivefold stronger than is binding to Rrp5N and approximately 50-fold stronger than binding to Rrp5C, with Rrp5FL binding in the lower nanomolar range (Table 1). These data suggest that most of the binding energy arises from interactions made with the first nine S1 domains

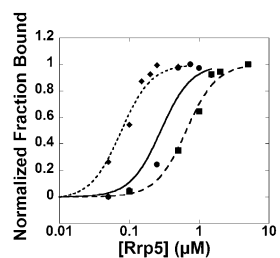


FIGURE 4. Affinities of Rrp5 constructs for H44/3'ITS1. Rrp5FL (◆) binds the H44/3'ITS1 rRNA fragment approximately fivefold stronger than Rrp5N (●) and approximately 50-fold stronger than Rrp5C (■).

TABLE 1. RNA binding affinities for Rrp5 constructs

rRNA construct	$K_{1/2}$ (μM^2)		
	Rrp5N	Rrp5C	Rrp5FL
H44/A₂(+212)	0.10 ± 0.01	14.5 ± 0.24	0.085 ± 0.01
H44/(+278)	0.10 ± 0.02	8.0 ± 0.90	0.047 ± 0.01
H44/3'ITS1(+362)	0.10 ± 0.01	0.93 ± 0.12	0.023 ± 0.001
H44/3'5.8S(+519)	0.08 ± 0.03	0.98 ± 0.03	0.009 ± 0.001

All RNA binding affinities were determined in gel-shift assays at 30°C in 10 mM MgCl₂, 100 mM KCl, and 50 mM HEPES (pH 7.5). Data are the averages of three or more experiments.

(Rrp5N), as deletion of S1₁₀₋₁₂ only has a fivefold effect on RNA binding. Since Rrp5N has nine S1 RNA-binding modules, this result was not unexpected; however, the fivefold higher affinity observed for Rrp5FL relative to Rrp5N does suggest that the last three S1 motifs (Rrp5C) also contribute to Rrp5's RNA binding activity.

To determine which portion(s) of ITS1 Rrp5 is recognizing and to better dissect the roles of the individual S1 modules in RNA binding, we compared Rrp5 binding to a number of rRNA constructs that start at H44(−164) and end at either A₂(+212), (+278) (6 nt upstream of A₃) or 3'ITS1(+362) at the 3' end (Fig. 1A).

RNA-binding experiments with these labeled rRNA fragments and Rrp5N indicate that RNA binding is relatively strong, with $K_{1/2}$ values in the high nanomolar range (Table 1). Further comparison of the Rrp5N data also indicates that as the 3' end of the rRNA fragment is lengthened, the affinity of Rrp5N remains unchanged (Fig. 5A; Table 1). To rule out the possibility that the lack in apparent changes in the RNA affinity could be explained because the labeled RNA was already partially or fully saturating the RNA-binding affinity, we carried out control experiments in which the amount of labeled RNA was varied over a 10-fold range. If RNA binding was already saturated by the labeled RNA, we would expect the resulting apparent $K_{1/2}$ value to change with the RNA concentration. However, no such change was observed for binding of either Rrp5N or Rrp5FL (data not shown). Furthermore, the RNA-binding affinity observed for Rrp5FL is even stronger than Rrp5N and is responsive to changes in the RNA identity but not concentration; this indicates that binding to Rrp5N is not approaching a general threshold in the assay, such that stronger binding would be nondetectable. Together, these experiments indicate that while Rrp5N binds to pre-rRNA fragments including regions of ITS1 relatively tightly, it does so relatively nonspecifically. It is possible that this lack of specificity arises because we have not included the correct RNA-binding site for Rrp5N. However, both Rrp5FL and Rrp5C specifically bind the H44/3'-ITS1 RNA (see below), indicating that if there is a binding site for Rrp5N, it should be nearby. However, addition of the entire 5.8S sequence does not change

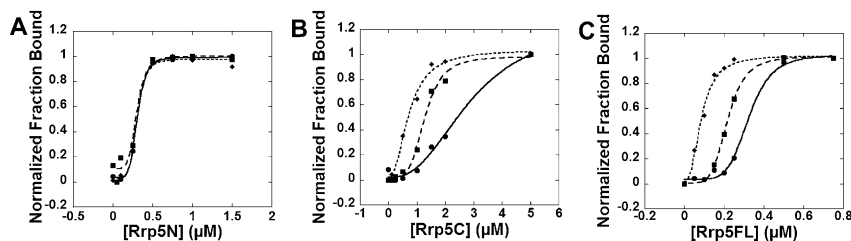


FIGURE 5. Binding of Rrp5N (A), Rrp5C (B), and Rrp5FL (C) to H44/A₂ (●), H44/(+278) (■), and H44/3'ITS1 (◆) rRNA mimics.

Rrp5N binding (Table 1), making it unlikely that Rrp5N binds further 3'. Because OB-fold domains generally have a polarity by which the N terminus binds the 3'-end and the C terminus binds the 5'-end (Murzin 1993; Theobald et al. 2003) and because of the effects on A₃ cleavage upon deletion of Rrp5N (Torchet et al. 1998; Eppens et al. 1999), we believe it is unlikely that Rrp5N binds 5' to RNAs tested herein. Consistent with this notion, fragments in which H44 is deleted bind with the same affinity as RNAs in which it is present (data not shown), indicating that H44 also does not constitute a binding site. Finally, to rule out that Rrp5N's binding site was contained 3' to site A₂, we determined Rrp5N's affinity for the A₂/3'-ITS1 RNA, which was also within less than threefold of the H44/3'-ITS1 affinity ($K_{1/2} = 0.1 \pm 0.03 \mu\text{M}^2$ and $0.27 \pm 0.03 \mu\text{M}^2$ for H44/3'-ITS1 and A₂/3'-ITS1, respectively). Together, these data indicate that Rrp5N binds RNAs tightly but nonspecifically.

In contrast, the affinity of Rrp5C increases as the length of the 3' end of the pre-rRNA fragment is increased from A₂ to (+278) to 3'ITS1 (Fig. 5B; Table 1). This increase in affinity does not simply reflect stronger binding due to a longer RNA, as addition of the 5.8S sequence or deletion of H44 have no effect on Rrp5C binding (Table 1; data not shown). This result shows that Rrp5C makes specific interactions with the ITS1, mainly 3' to site A₂.

Binding data for Rrp5FL combine the trends observed for Rrp5N and Rrp5C (Fig. 5C; Table 1). While Rrp5FL binds even more strongly to rRNA than does Rrp5N (reflecting the additional interactions provided by the C-terminal S1 domains), its binding does reflect the specificity observed with Rrp5C, as binding affinity increases as the 3'-end of ITS1 is lengthened. Addition of the 5.8S sequence at the 3' end has no effect on binding of any of the Rrp5 constructs (Table 1), indicating that Rrp5 does not make any interactions with 5.8S rRNA.

In summary, deleting S1₁₋₉ (Rrp5C) weakens RNA binding 50-fold but does not affect binding specificity, while deletion of S1₁₀₋₁₂ (Rrp5N) only weakens RNA binding fivefold but leads to a complete loss in RNA binding specificity. These data demonstrate a surprising division of labor between Rrp5's S1 RNA-binding modules, whereby the three last S1 domains provide the specificity and

the first nine S1 domains provide much of the affinity for Rrp5's interaction with ITS1.

DMS structure probing verifies that Rrp5 has interactions in ITS1

In order to validate the proposed interactions between Rrp5 and the region in the ITS1 that is 3'-to site A₂, we used dimethyl sulfate (DMS) structure probing to test the solvent accessibility of the

rRNA in the presence and absence of Rrp5. DMS methylates the N1 of adenosine and the N3 of cytosine unless they are protein protected or involved in base pairs or tertiary structures (Stern et al. 1988). Methylation at specific residues can then be detected by stops in reverse transcription and analyzed via sequencing gels. Comparing modification patterns of rRNA in the presence and absence of protein can lead to identification of residues protected from modification by the presence of protein, either directly or indirectly, via changes in the RNA structure. Additionally, because of its inability to methylate nucleotides involved in base pairs, information regarding the rRNA secondary structure can be obtained via DMS modification. In order to account for differences in the loading or in the efficiency of reverse transcription between different lanes in the gel, we also carried out line scans that were normalized to nearby residues (Fig. 6B).

Since RNA-binding analyses indicated that the affinities obtained with Rrp5FL reflect the trends observed for both Rrp5N and Rrp5C, we used the Rrp5FL construct for the DMS structure probing experiments. Probing experiments with Rrp5C were also attempted and resulted in changes in the DMS accessibility in the same general regions. However, due to its weak affinity for the pre-rRNA fragments, Rrp5C's transient interactions were difficult to map reproducibly to specific residues; instead, they varied between neighboring residues in different experiments. Because our RNA binding data indicated relatively nonspecific binding for Rrp5N, no DMS protection experiments were attempted for this fragment.

Since DMS probing requires higher concentrations of RNA in comparison to the binding experiments and since the H44/3'ITS1 construct tended to aggregate at concentrations >1 μM , we used the shorter H45/3'5.8S rRNA fragment to eliminate the aggregation; binding experiments indicated that Rrp5FL has the same affinity for both H44/3'5.8S and H45/3'5.8S (data not shown). Since the binding results indicated that Rrp5 has no interactions in 5.8S rRNA, this region was included as a negative control. Furthermore, the extended 3'-region provided a primer binding site for mapping of the 3'-end of ITS1.

DMS probing of the folded H45/3'5.8S rRNA fragment alone resulted in a modification pattern consistent with the

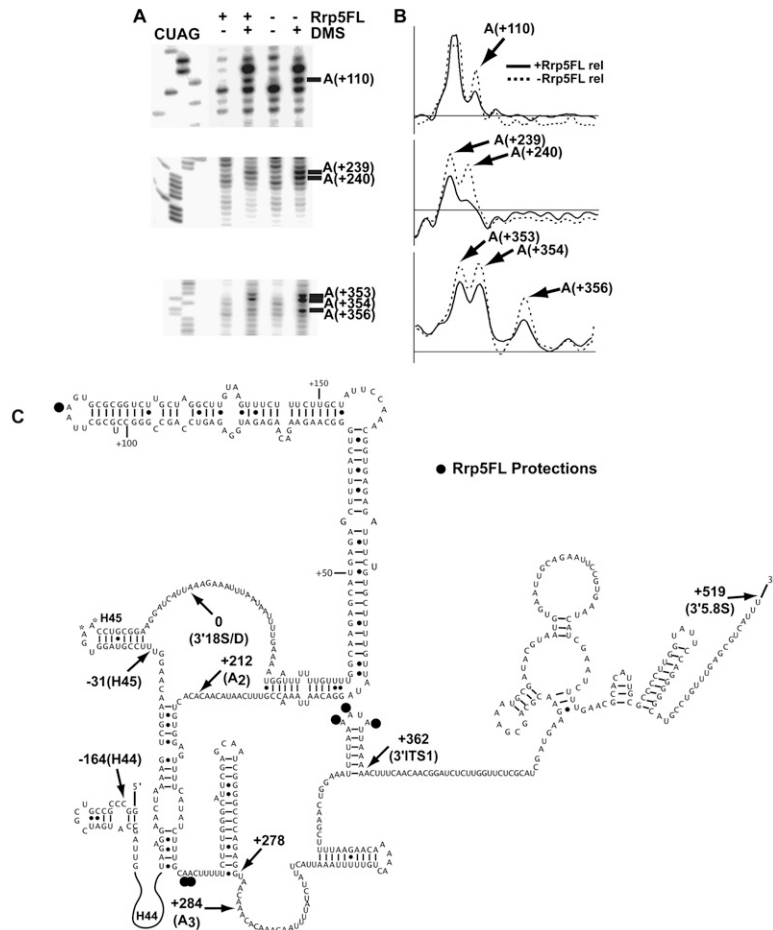


FIGURE 6. Rrp5 footprinting on pre-rRNA via DMS structure probing. (A) DMS protections in the presence and absence of Rrp5FL show protections in three regions of pre-rRNA. Control experiments in the absence of DMS indicate DMS-independent reverse transcription stops. Note that sequencing ladders are upshifted by 1 nt relative to DMS lanes. (B) Line scans of the regions shown in A. (C) Rrp5FL protected residues (●) mapped onto the H44/3'5.8S secondary structure (as predicted by the modification pattern observed in the absence of protein).

secondary structure we previously reported, as well as the secondary structures expected for the 3'-minor domain of 18S rRNA and 5.8S rRNA (Fig. 6C; Wuyts et al. 2001; Lamanna and Karbstein 2009, 2011). Upon the addition of Rrp5FL, protected residues were identified in three areas (Fig. 6). While these residues were not completely protected, significant differences between pre-rRNA with and without protein were nevertheless observed in four different experiments obtained with two independent samples, indicating reproducibility. We observe a protection within the distal stem-loop of ITS1 (+110; at the same position where we have previously observed a Nob1-dependent footprint) (Lamanna and Karbstein 2009), immediately downstream from a helix formed between H44 and ITS1 (+239/240) (Lamanna and Karbstein 2011), as well as protections close to the 3'-end of ITS1 (+353 region). These protections were quantified in line scans (Fig. 6B) and are plotted onto the secondary structure in Figure 6C. The footprinting data fully support the binding data,

which showed weak interactions with the ITS1 prior to site A₂ and additional contributions both 5'- and 3'- to site A₃ (Table 1). Finally, in further support of our binding data (which indicates that the addition of 5.8S rRNA has no additional contribution to Rrp5 binding) (Table 1), no DMS protections were observed in 5.8S rRNA.

DISCUSSION

The binding site for Rrp5

Rrp5 is unique among ribosome assembly factors, as it is required for the production of both large and small ribosomal subunits. To better understand the role of Rrp5 in ribosome assembly, we have characterized Rrp5's pre-rRNA binding site. Comparison of Rrp5's affinity to a number of different pre-rRNA analogs indicates that Rrp5 makes contacts over much of the length of the ITS1, with most of the binding energy arising from contacts occurring 3' to site A₂. The same trend in pre-rRNA binding affinity observed for full-length protein is also seen with a C-terminal fragment (Rrp5C) containing the last three S1 domains, as well as the seven tetratricopeptide (TPR) motifs, which are involved in protein-protein interactions (Fig. 1B) and do not contribute to RNA binding (de Boer et al. 2006). In surprising contrast, an N-terminal frag-

ment (Rrp5N) containing the first nine S1 domains binds all pre-rRNA mimics strongly but nonspecifically. This finding indicates that the majority of specific interactions in ITS1 are made by these last three S1 RNA-binding domains. This quantitative functional analysis is corroborated by DMS protection experiments, which show protections in three regions: one between sites D and A₂ (in a stem-loop where Nob1, the D-site endonuclease, also makes an interaction) (Lamanna and Karbstein 2009), one 5' to site A₃, and one in a stem-loop near the 3'-end of ITS1 (Fig. 6).

These quantitative experiments confirm and extend prior qualitative experiments, largely undertaken via *in vivo* truncation analyses (Torchet et al. 1998; Eppens et al. 1999). Deletions or mutations in the first nine S1 domains (Rrp5N) affect cleavage at site A₃, while mutations or deletions in the last three S1 or the C-terminal TPR motifs (Rrp5C) inhibit cleavage at site A₂ (Torchet et al. 1998; Eppens et al. 1999); these results suggest that Rrp5 is interacting with the pre-rRNA encompassing these regions.

Furthermore, these *in vivo* studies indicated that deletion of one to three consecutive S1 motifs spanning the first nine S1 motifs of Rrp5 was not lethal. Our finding that Rrp5N provides affinity but no specificity for the pre-rRNA explains this observation: The roles of the first nine S1 motifs are redundant, and as long as at least six S1 domains are present, Rrp5 has enough affinity for the pre-rRNA. Furthermore, our results also explain the seemingly paradoxical observation that Rrp5N, required only for cleavage at site A₃, is essential (Fig. 2; Eppens et al. 1999), while cleavage at site A₃ itself is not (Henry et al. 1994). While Rrp5C is sufficient to provide the necessary specificity for cleavage at site A₂, it binds only weakly to pre-rRNA, such that pre-rRNA will not have Rrp5C bound most of the time. The presence of Rrp5N (in *cis* or in *trans*) is required for strong binding to pre-rRNA, resulting in cleavage at site A₂. Apparently these two protein fragments can associate, or cooperatively bind rRNA, as evidenced by their complementation of the Rrp5 deletion.

Our findings are also consistent with *in vitro* binding experiments using Rrp5 immunopurified from yeast, which indicated that Rrp5 preferentially binds U-rich single-stranded sequences; we have delineated Rrp5's binding site in regions of ITS1, a U-rich sequence (de Boer et al. 2006). Furthermore, one of the three Rrp5-interacting regions that is protected from DMS modification is flanked on both sides by stretches of uridine residues.

In their pioneering *in vitro* work with Rrp5, de Boer et al. (2006) identified an Rrp5 interaction with rRNA 3' of site A₂; our DMS probing experiments also map an Rrp5 binding site to this region. Differing from these prior results, but instead consistent with the *in vivo* experiments described above, we also uncovered an Rrp5 binding site 3' of site A₃. We attribute these differences to the nature of Rrp5 in the binding conditions: since our Rrp5 constructs are not immobilized on resin, perhaps additional RNA-binding residues are exposed and therefore further interactions in ITS1 are observed.

During *in vivo* rRNA processing, Rrp5 binds relatively early to assembling preribosomes (Venema and Tollervey 1996) and is required for both 40S and 60S assembly (Venema and Tollervey 1996; Torchet et al. 1998; Eppens et al. 1999). Previous proteomic analysis of preribosomal complexes indicates that Rrp5 remains bound to pre-60S ribosomes upon separation of the pre-40S and pre-60S rRNAs (Lebreton et al. 2008). Our data explain this observation, as Rrp5 binds only weakly to pre-rRNA analogs ending at site A₂ and instead has the majority of its interactions in regions 3' to site A₂. Thus, it appears that after cotranscriptional cleavage at site A₂, Rrp5 releases the RNA 5' to site A₂ either spontaneously or, perhaps, via interaction with another ribosome assembly factor. Intriguingly, the DEAD-box helicase Rok1 interacts genetically with the TPR motifs in Rrp5 (Torchet et al. 1998).

The ITS1 is very divergent in length (137 nt in *Candida albicans* vs. 1095 nt in humans) as well as sequence (Lamanna and Karbstein 2011). The three regions of the

ITS1 protected by Rrp5 *in vitro* are all present in even the shortest ITS1 sequence, yet the ITS1 sequence is not conserved in these areas (Supplemental Fig. S2). Thus, given the lack of ITS1 sequence conservation, one would expect that in order to maintain sequence-specific interactions, Rrp5 must evolve in conjunction with the ITS1. Interestingly, Rrp5's last five S1 domains are much less conserved than its first seven S1 domains (Supplemental Fig. S3), consistent with the observation that RNA-binding specificity arises from the last three S1 domains. Furthermore, it should be noted that in mammalian Rrp5, the 10th S1 domain contains a large expansion of the loop between β 2 and β 3, which is part of the expected RNA-binding surface (Fig. 1C; Supplemental Fig. S2); interestingly, mammals also have the largest ITS1. It is tempting to speculate that the expansion in S1₁₀ in the evolutionary divergence of Rrp5 homologs arose to accommodate an expanded ITS1.

The role of Rrp5 in ribosome assembly

We have recently shown that early in 40S assembly, sequences in ITS1 form base pairs with the decoding site strand at the top of helix 44 (see Fig. 6C; Lamanna and Karbstein 2011). This interaction prevents the formation of the decoding site and inhibits premature Nob1-dependent cleavage at site D. Cleavage at site A₂ allows removal of the inhibitory nucleotides and subsequent formation of the mature base pairing in the top of helix 44. These data provide a molecular rationale for the observed requirement of cleavage at site A₂ for subsequent cleavage at site D. The RNA-binding and DMS probing data herein place Rrp5 directly adjacent to this switch region, with two of the protections directly downstream from the inhibitory duplex (Fig. 6). As described above, the observation that the binding affinities for Rrp5C parallel those for full-length Rrp5 indicates that the last three S1 domains are involved in these interactions. Studies of the rRNA processing phenotype of an Rrp5 mutant in which the last three S1 motifs are deleted show that, in this mutant, A₂ cleavage can be bypassed for the production of 18S rRNA (Torchet and Hermann-Le Denmat 2000; Vos et al. 2004). Thus, the three C-terminal S1 domains are necessary to enforce the requirement for A₂ cleavage and bind directly adjacent to the helix that produces this requirement. Together, these data indicate that Rrp5 could play an essential role in stabilizing the early pre-A₂ cleavage structure. This would explain the requirement of Rrp5 for cleavage at site A₂ (Venema and Tollervey 1996). It should be noted that a similar, although less specific, proposal was made previously (de Boer et al. 2006).

Implications for the structure of the Rrp5•RNA complex

The structures of a few S1 domain-containing proteins and several other OB-fold proteins (some bound to RNA) have

been solved (Theobald et al. 2003 and references therein). Recently, Matsumoto et al. (2010) reported the structure of a bacterial virulence factor, CvfB, which contains three S1 domains and a winged helix (WH) domain. This structure represents the first complete structure of a protein with multiple S1 domains and so far provides the best model for considerations of Rrp5's structure. The CvfB structure shows a roughly L-shaped molecule, where the first two S1 domains (S1A and S1B) form one leg, and the third S1 (S1C) and WH domain form the second leg. An extended potential RNA-binding surface was observed on the S1C/WH leg, and binding experiments indicate that these two domains alone bind polyU RNA with the same affinity as the full-length protein. Furthermore, another extended region of positive charge was found on the S1A/S1B surface, but the significance of this region for RNA binding and/or protein function remains unclear. The CvfB structure demonstrates substantial interdomain interactions between all three S1 domains, indicating that these are not free to act on their own. Consistent with similar extensive interactions among multiple S1 domains in Rrp5, we find that the RNA binding affinities for Rrp5N and Rrp5C are nonadditive: Rrp5FL binds RNA more strongly than either Rrp5N or Rrp5C alone but not as strongly as expected if these fragments acted entirely independently (Table 1), as binding affinities observed for Rrp5FL are not equivalent to the product of the Rrp5N and Rrp5C affinities observed for the same rRNA mimic.

The finding that Rrp5's first nine S1 domains bind the pre-rRNA relatively nonspecifically, while the last three S1 domains bind RNA specifically, demonstrates that S1 domains can provide surfaces for specific as well as non-specific RNA interactions. Based on structures of other OB-fold proteins with RNA bound, it seems likely that interactions with the β -sheet (mostly β_2 and β_3) provide rather nonspecific interaction surfaces, while the loops (often those between the second and third β -sheets, and the third β and the α -helix) have the ability to form sequence specific contacts (Theobald et al. 2003 and references therein). Thus, we suggest that Rrp5's first nine S1 motifs interact with rRNA largely via their β -sheets, while the last three motifs likely form additional interactions with the loops.

OB-fold proteins bind RNA with a conserved polarity, where the N terminus of the protein is close to the 3'-end and the C terminus of the protein is close to the 5'-end (Murzin 1993; Theobald et al. 2003). This polarity is consistent with the early *in vivo* data showing that Rrp5N is required for cleavage at site A₃, while Rrp5C is required for cleavage at the more 5' site A₂ (Fig. 1A). Herein, we have refined the placement of Rrp5, and our data indicate that Rrp5C interacts specifically with three regions of ITS1. Within these three interactions, we can assume a similar polarity for S1₁₀₋₁₂, whereby S1₁₀ would bind the stem-loop near 5.8S rRNA, and S1₁₂ could bind between sites D and A₂.

Comparison to other RNA-binding proteins with multiple RBDs

Most defined RNA-binding domains recognize only between 4 and 8 nt at a time, providing for little specificity and affinity (Lunde et al. 2007). Thus, many proteins contain multiple RNA-binding domains, as either different modules or multiple copies of the same module² (Lunde et al. 2007). General considerations suggest that such combinations should allow for a large expansion of both the specificity and affinity of RNA recognition. Increases in affinity are generally due to large interaction surfaces and/or loss of entropy when one domain is already immobilized on the RNA. Increased specificity can be obtained from larger interaction surfaces or by strategic positioning of multiple RNA-binding domains in a fixed orientation (Lunde et al. 2007).

These first principle considerations have been confirmed by studies with several proteins. For example, KSRP, a protein involved in splicing and degradation of specific target mRNAs, has four KH domains (Gherzi et al. 2004). While each one of the domains alone recognizes RNA with weak, although not identical, affinity, combining any two increases RNA-binding strength (Garcia-Mayoral et al. 2007; Diaz-Moreno et al. 2010). Furthermore, the increase is completely additive in the case of KH3 and KH4, which are linked by a flexible linker, while the increase appears subadditive in the case of KH2 and KH3 (Garcia-Mayoral et al. 2007, 2008). Finally, disrupting the interaction between KH2/3 via a point mutation increases the RNA affinity, albeit weakly (Diaz-Moreno et al. 2010). These findings are consistent with the notion that the interaction between KH2 and KH3 positions their substrate binding surfaces for specific interactions, although such experiments have not been performed. Similarly, studies with the Nova-1 and Nova-2 proteins, also regulators of splicing, have shown that the last KH domain, KH3, provides most of the affinity, with addition of the first two KH domains only providing about a 10-fold increase in RNA binding affinity (Buckanovich and Darnell 1997). Furthermore, it appears qualitatively that all three KH domains contribute equally toward specificity (Musunuru and Darnell 2004). Finally, studies of the nine Zn-finger domains in transcription factor IIIA have shown that distinct fingers provide for most of the interactions with RNA relative to DNA (Liao et al. 1992; Clemens et al. 1993). Nevertheless, Zn-fingers 1–3 provide most of the DNA binding affinity and bind with the same specificity as the full-length protein (Liao et al. 1992).

²Ribosomal proteins form a notable exception as they often do not contain defined RNA-binding modules and are generally small but interact with extended stretches of RNA. As a result, ribosomal proteins are highly insoluble outside of the ribosome. Perhaps an advantage of the recurring, well-defined RDBs with small RNA binding surfaces is that such proteins remain soluble.

In contrast to these well-studied examples, the RNA-binding data with Rrp5 presented herein indicate a division of labor that is unique. In this example, S1₁₀₋₁₂ (contained in Rrp5C) provides all of the specificity in the recognition of pre-rRNA, while S1₁₋₉ (contained in Rrp5N) provides most of the affinity. As a result, deleting Rrp5N results in a 50-fold loss in RNA binding affinity, while deleting Rrp5C only leads to a fivefold loss in RNA binding affinity.

MATERIALS AND METHODS

Cloning of Rrp5 constructs

All three proteins were amplified from genomic DNA. Rrp5C (amino acids 1083–1729) was cloned between the SfoI and BamHI sites of pSV272; Rrp5N (amino acids 1–1082) and Rrp5FL (amino acids 1–1729) were cloned between the NcoI and BamHI sites of a pET23-derivative containing a TEV-cleavable His₆-tag.

Rrp5 expression and purification

Rosetta cells transformed with a plasmid encoding either Rrp5C, Rrp5N, or Rrp5FL were grown at 37°C in LB Miller media (supplemented with the appropriate antibiotic) to an OD₆₀₀ of ~0.6 before inducing with 1 mM IPTG for 5 h at 30°C. Cells were resuspended in lysis buffer (containing 0.1 mM PMSF, 5 mM benzamidine) and sonicated. After pelleting, the soluble fraction was purified over Ni-NTA (Qiagen) resin according to the manufacturer's protocol. Subsequent steps differ from construct to construct and are described individually below.

Rrp5C-containing elution fractions were dialyzed overnight in 70 mM NaCl, 25 mM HEPES (pH 7.6), 1 mM TCEP, and 1 mM DTT; TEV protease was also added to remove the His₆-MBP tag. Rrp5C was further purified over a MonoS column in a linear gradient from 30–420 mM NaCl over 12 column volumes. Rrp5C-containing fractions were then purified over a Superdex200 column into 100 mM KCl, 50 mM HEPES (pH 7.6), 1 mM TCEP, and 1 mM DTT. Rrp5C was stored in 15% glycerol at –80°C. Protein concentration was determined by absorbance at 280 nm using a calculated extinction coefficient of 60,900 M⁻¹cm⁻¹.

Rrp5N-containing fractions were dialyzed for 5 h in 150 mM KCl, 50 mM Tris (pH 7.8), 1 mM TCEP, and 1 mM DTT. Rrp5FL-containing fractions were dialyzed for 3 h in 100 mM KCl, 50 mM Tris (pH 7.8), 1 mM TCEP, and 1 mM DTT. Rrp5N and Rrp5FL were further purified over a MonoQ column: Rrp5N was separated over a 12 column volume linear gradient from 150–660 mM KCl; Rrp5FL was separated over a 12 column volume linear gradient from 100–640 mM KCl. Rrp5N- or Rrp5FL-containing fractions were then purified over a Superdex200 column into 200 mM KCl, 50 mM HEPES (pH 7.6), 1 mM TCEP, and 1 mM DTT. Rrp5N and Rrp5FL were stored in 15% glycerol at –80°C. Protein concentration was determined by absorbance at 280 nm using calculated extinction coefficients of 45,200 M⁻¹cm⁻¹ and 111,100 M⁻¹cm⁻¹ for Rrp5N and Rrp5FL, respectively.

RNA cloning, transcription, and purification

rDNA fragments were amplified from genomic DNA and cloned into pUC19 as previously described (Lamanna and Karbstein

2009). BtsI was used to linearize the plasmid, and ³²P-labeled and unlabeled rRNA were transcribed and purified as described previously (Karbstein et al. 2005; Lamanna and Karbstein 2009).

Rrp5 and rRNA binding experiments

rRNA was folded in the presence of 10 mM Mg²⁺ as described (Karbstein et al. 2005). Binding reactions of Rrp5 to folded rRNA were carried out as reported previously for Nob1 (Lamanna and Karbstein 2009). Briefly, prefolded rRNA and Rrp5 were incubated together for 2 h at 30°C before being loaded on a 6% acrylamide/THEM (Tris, HEPES, EDTA at pH 7.5, MgCl₂) gel (Karbstein et al. 2002) for 3 h at 4°C. Rrp5-bound and unbound fractions were quantified using phosphorimager software, and data were fit to equation 3 using Kaleidagraph (Synergy Software). Control experiments in which the amount of labeled RNA was varied indicate that the RNA concentration remains substantially below the *K*_{1/2} value (data not shown), such that the binding equilibrium can be simplified as shown in the derivations below.

Well-shifts were obtained for Rrp5N and Rrp5FL, but not Rrp5C, for which RNA–protein complexes migrated into the gel. This could be a result of Rrp5C's lower molecular weight (68 kDa for Rrp5C vs. 125 kDa for Rrp5N and 193 kDa for Rrp5FL), or alternatively, the higher mobility may be due to Rrp5C's lower pI value of 4.9 (compared with 7.1 and 6.1 for Rrp5N and Rrp5FL, respectively). Since binding reactions are done at a pH of 7.5 and then separated on a native gel (pH 7.5), it is possible that the charged state of Rrp5C contributes to its higher mobility. Varying the pH of the native gel between 6.1 and 8.7 had no effect on either the mobility or the affinity of the H44/3'TTS1–protein complex for all three Rrp5 constructs (data not shown); due to concerns of protein and RNA stability, we did not further increase or decrease the pH.

Purification of each Rrp5 construct over the Superdex200 size exclusion column indicates that Rrp5C, Rrp5N, and Rrp5FL all elute as dimers. Correspondingly, the binding isotherms are indicative of binding of at least two subunits. This is considered in the binding equilibrium in Equation 1. Note that *K*_{1/2} is used herein instead of the perhaps more familiar *K*_d, as it is an overall value describing two binding events by two Rrp5 molecules:

$$K_{1/2} = \frac{[\text{Rrp5}]^2 \times [\text{rRNA}]}{[\text{Rrp5}_2 \cdot \text{rRNA}]} \quad (1)$$

To quantify the fraction of Rrp5 bound to rRNA, Equation 2 is used:

$$\text{fraction}_{\text{bound}} = \frac{[\text{Rrp5}_2 \cdot \text{rRNA}]}{[\text{Rrp5}_2 \cdot \text{rRNA}] + [\text{rRNA}]} \quad (2)$$

Solving Equation 2 for [Rrp5₂•rRNA] and substituting this into Equation 1 gives Equation 3:

$$\text{fraction}_{\text{bound}} = \frac{\text{fraction}_{\text{bound, max}}[\text{Rrp5}]^2}{[\text{Rrp5}]^2 + K_{1/2}} \quad (3)$$

Competition experiments

rRNA was folded as described above. For binding to H44/3'TTS1 and TSL, protein–rRNA complexes containing either 5 μM Rrp5C,

1 μM Rrp5N, or 0.75 μM Rrp5FL and trace ^{32}P -labeled rRNA were formed as described above in the presence or absence of 5 μM unlabeled RNA. For binding to H44/A₂, Rrp5N and Rrp5FL were used at a concentration of 0.5 μM .

DMS probing experiments

DMS probing was performed as described (Doherty et al. 1999; Lamanna and Karbstein 2009). We folded and incubated 1 μM of unlabeled H45/3'5.8S in the presence or absence of 6 μM Rrp5FL, which was then exposed to either 1.5% (H45/3'5.8S only) or 2.5% (H45/3'5.8S and Rrp5FL) DMS and incubated for 3 min. Reactions were quenched with β -mercaptoethanol and RNA was phenol extracted. Reverse transcription and sequencing gels were performed as described (Lamanna and Karbstein 2009).

SUPPLEMENTAL MATERIAL

Supplemental material is available for this article.

ACKNOWLEDGMENTS

C.L.Y. was partially supported by an Excellence in Research Award from the Department of Chemistry at the University of Michigan. This work was supported by NIH grant R01-GM086451.

Received September 8, 2010; accepted December 9, 2010.

REFERENCES

- Buckanovich RJ, Darnell RB. 1997. The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol Cell Biol* **17**: 3194–3201.
- Bycroft M, Hubbard TJ, Proctor M, Freund SM, Murzin AG. 1997. The solution structure of the S1 RNA binding domain: A member of an ancient nucleic acid-binding fold. *Cell* **88**: 235–242.
- Clemens KR, Wolf V, McBryant SJ, Zhang P, Liao X, Wright PE, Gottesfeld JM. 1993. Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science* **260**: 530–533.
- de Boer P, Vos HR, Faber AW, Vos JC, Raue HA. 2006. Rrp5p, a *trans*-acting factor in yeast ribosome biogenesis, is an RNA-binding protein with a pronounced preference for U-rich sequences. *RNA* **12**: 263–271.
- Diaz-Moreno I, Hollingworth D, Kelly G, Martin S, Garcia-Mayoral M, Briata P, Gherzi R, Ramos A. 2010. Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. *Nucleic Acids Res* **38**: 5193–5205.
- Doherty EA, Herschlag D, Doudna JA. 1999. Assembly of an exceptionally stable RNA tertiary interface in a group I ribozyme. *Biochemistry* **38**: 2982–2990.
- Eppens NA, Rensen S, Granneman S, Raue HA, Venema J. 1999. The roles of Rrp5p in the synthesis of yeast 18S and 5.8S rRNA can be functionally and physically separated. *RNA* **5**: 779–793.
- Garcia-Mayoral MF, Hollingworth D, Masino L, Diaz-Moreno I, Kelly G, Gherzi R, Chou CF, Chen CY, Ramos A. 2007. The structure of the C-terminal KH domains of KSRP reveals a noncanonical motif important for mRNA degradation. *Structure* **15**: 485–498.
- Garcia-Mayoral MF, Diaz-Moreno I, Hollingworth D, Ramos A. 2008. The sequence selectivity of KSRP explains its flexibility in the recognition of the RNA targets. *Nucleic Acids Res* **36**: 5290–5296.
- Gherzi R, Lee KY, Briata P, Wegmuller D, Moroni C, Karin M, Chen CY. 2004. A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol Cell* **14**: 571–583.
- Henry Y, Wood H, Morrissey JP, Petfalski E, Kearsey S, Tollervey D. 1994. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO J* **13**: 2452–2463.
- Kaminishi T, Wilson DN, Takemoto C, Harms JM, Kawazoe M, Schluenzen F, Hanawa-Suetsugu K, Shirouzu M, Fucini P, Yokoyama S. 2007. A snapshot of the 30S ribosomal subunit capturing mRNA via the Shine-Dalgarno interaction. *Structure* **15**: 289–297.
- Karbstein K, Carroll KS, Herschlag D. 2002. Probing the Tetrahymena group I ribozyme reaction in both directions. *Biochemistry* **41**: 11171–11183.
- Karbstein K, Jonas S, Doudna JA. 2005. An essential GTPase promotes assembly of preribosomal RNA processing complexes. *Mol Cell* **20**: 633–643.
- Lamanna AC, Karbstein K. 2009. Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proc Natl Acad Sci* **106**: 14259–14264.
- Lamanna AC, Karbstein K. 2011. A conformational switch regulates pre-18S cleavage. *J Mol Biol* **405**: 3–17.
- Lebreton A, Roussele JC, Lenormand P, Namane A, Jacquier A, Fromont-Racine M, Saveanu C. 2008. 60S ribosomal subunit assembly dynamics defined by semi-quantitative mass spectrometry of purified complexes. *Nucleic Acids Res* **36**: 4988–4999.
- Liao XB, Clemens KR, Tennant L, Wright PE, Gottesfeld JM. 1992. Specific interaction of the first three zinc fingers of TFIIIA with the internal control region of the *Xenopus* 5 S RNA gene. *J Mol Biol* **223**: 857–871.
- Lunde BM, Moore C, Varani G. 2007. RNA-binding proteins: Modular design for efficient function. *Nat Rev Mol Cell Biol* **8**: 479–490.
- Matsumoto Y, Xu Q, Miyazaki S, Kaito C, Farr CL, Axelrod HL, Chiu HJ, Klock HE, Knuth MW, Miller MD, et al 2010. Structure of a virulence regulatory factor CvfB reveals a novel winged helix RNA binding module. *Structure* **18**: 537–547.
- Murzin AG. 1993. OB(oligonucleotide/oligosaccharide binding)-fold: Common structural and functional solution for non-homologous sequences. *EMBO J* **12**: 861–867.
- Musunuru K, Darnell RB. 2004. Determination and augmentation of RNA sequence specificity of the Nova K-homology domains. *Nucleic Acids Res* **32**: 4852–4861.
- Schuwirth BS, Borovinskaya MA, Hau CW, Zhang W, Vila-Sanjurjo A, Holton JM, Cate JH. 2005. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**: 827–834.
- Stern S, Moazed D, Noller HF. 1988. Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol* **164**: 481–489.
- Subramanian AR. 1983. Structure and functions of ribosomal protein S1. *Prog Nucleic Acid Res Mol Biol* **28**: 101–142.
- Theobald DL, Mitton-Fry RM, Wuttke DS. 2003. Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**: 115–133.
- Torchet C, Hermann-Le Denmat S. 2000. Bypassing the rRNA processing endonucleolytic cleavage at site A2 in *Saccharomyces cerevisiae*. *RNA* **6**: 1498–1508.
- Torchet C, Jacq C, Hermann-Le Denmat S. 1998. Two mutant forms of the S1/TPR-containing protein Rrp5p affect the 18S rRNA synthesis in *Saccharomyces cerevisiae*. *RNA* **4**: 1636–1652.
- Venema J, Tollervey D. 1996. RRP5 is required for formation of both 18S and 5.8S rRNA in yeast. *EMBO J* **15**: 5701–5714.
- Vos HR, Faber AW, de Gier MD, Vos JC, Raue HA. 2004. Deletion of the three distal S1 motifs of *Saccharomyces cerevisiae* Rrp5p abolishes pre-rRNA processing at site A(2) without reducing the production of functional 40S subunits. *Eukaryot Cell* **3**: 1504–1512.
- Wuyts J, De Rijk P, Van de Peer Y, Winkelmans T, De Wachter R. 2001. The European Large Subunit Ribosomal RNA Database. *Nucleic Acids Res* **29**: 175–177.



RNA

A PUBLICATION OF THE RNA SOCIETY

The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA

Crystal L. Young and Katrin Karbstein

RNA 2011 17: 512-521 originally published online January 13, 2011
Access the most recent version at doi:[10.1261/ma.2458811](https://doi.org/10.1261/ma.2458811)

Supplemental Material <http://rnajournal.cshlp.org/content/suppl/2010/12/30/rna.2458811.DC1>

References This article cites 31 articles, 9 of which can be accessed free at:
<http://rnajournal.cshlp.org/content/17/3/512.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *RNA* go to:
<http://rnajournal.cshlp.org/subscriptions>