ROK1ng the Ribosome Assembly Boat:
An Investigation of the DEAD-box Protein Rok1 and its Co-factor Rrp5

by

Crystal Lee Young

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2012

Doctoral Committee:

Associate Professor Katrin Karbstein, The Scripps Research Institute, Co-Chair
Professor Anna K. Mapp, Co-Chair
Professor Nils G. Walter
Associate Professor Bruce A. Palfey

Dedication

This dissertation is dedicated to all the ‘teachers’ in my life: to those that taught me basic life skills, love and compassion; to the many that exemplified dedication, persistence and patience; for those that taught me perspective and genuine priorities; to the teachers that both sparked my initial interest in science and helped me evolve into the independent biochemist I am today; to the mentors that encouraged me to take the initiative and gave me the opportunity to polish leadership skills; finally, to those that both gave me the opportunity to realize my passion for teaching and have been supportive as I work towards my career goal.

Thank you for making me who I am today.

I hope to one day be the role model that you all have been to me.

Acknowledgements

I am surrounded by an amazing group of family, friends and colleagues that I could never thank enough for all the support they have provided over these past five years (and in most cases, even longer). First and foremost, I credit my fiancé Matt with helping me to survive perhaps the most mentally, emotionally and physically exhausting adventure I have ever undertaken. When I lost faith in myself, you did not; you reminded me that I started this endeavor with a purpose, and that I am not one to give-up and settle. From the two-state to across-the-country distance, the postponed dinners, the missed and un-returned phone calls and the "I'll be home by 7" really meaning I will see you at 8:30, your patience has been never-ending. You have been there for both high school and college graduations, and now with the PhD, I promise to finally stop being a student! My parents have also provided infinite support in all I have done, and these past five years have been no exception. Mom – you joke that you have not been able to help me with school since eighth grade, but your constant reminder to "keep my chin up" has been more helpful than any answers to math problems, critiques of English essays or guidance with data interpretation. Your positive attitude helped me to maintain my optimism, and that advice is something I will certainly carry through life. Also, even though I will no longer be a student, you are still welcome to send care packages! Dad – you may not have much to say, but I know that my accomplishments are your accomplishments, and I feel most successful when I have made you proud. I have acquired your work ethic and persistence, and those qualities have contributed tremendously to all my successes. To my brother Scott and sister-in-law Maggi – even though we now live 1,093 miles apart, the weekly phone calls and times together over holidays make it seem like we never missed a beat. Thank you for both the good times in the past, and those to come! To Kolby and Kayci, my littlest inspirations in life. Kolby – I give you complete credit for my determination and perspective. Your willpower is unbelievable, and I hope to carry

all of your admirable qualities through life. Kayci – you are a mini-me in so many ways, and you are therefore my biggest motivation to always do my best. You are continually so helpful and compassionate, and you make me very proud! To Cindy and Eric – because two sets of parents are better than one. You both have been amazing, and I am so lucky to have you as a huge part of my life. No, I don't need any money. Yes, I have plenty of groceries. Do I have gas in my car? Well… most of the time! To Grandma and Papa – even though our lunch dates are now more infrequent, I always love seeing you and catching-up when I am home. Grandma, I especially appreciate our phone calls. You always make life sunnier, and the drive to and from lab seems much faster when you are on the other line! I also love when your chocolate chip cookies come in the mail, but I love even more when I can get them in person. Between the rest of the family I have known forever and that which I will soon marry into, I have more support than I can name individually. But thank you all for always thinking of me; asking about me; checking up on my progress; and inquiring about and listening attentively to descriptions of my research when we can finally visit. Because of all of you, my life has indeed been a lucky one.

Even though the moral support from my family has been wonderful, I credit the accomplishment of my PhD to the amazing group of people I work with every day. In particular, I acknowledge my advisor, Katrin Karbstein; you are the one who has made this whole experience possible. I thank you for taking a gamble on me five years ago – I came to your lab with minimal biochemistry knowledge and even less experience at the bench. You accepted me without a doubt and have since been unwavering in your guidance, patience and support. Your door is always open, and you are always selfless with your time. When I thought I set the bar high, you showed me it can go even higher. When I told you five years ago that I wanted to use the PhD to teach undergraduates, you did everything you could to help me achieve that goal. Your passion and dedication are inspiring, and I hope to be the role model to my students that you have been to me. While I think you initially took a gamble on me, hindsight tells me we took a gamble on each other: I never imagined I would spend part of my PhD in Florida. Even though I was hesitant about the move, I trusted your judgment and you did not disappoint. This

transition has been amazing for both my personal and professional development, and I thank you for making this PhD experience richer than I could ever have imagined. Although leaving your lab after calling it home for five years is going to be incredibly difficult, I am excited to see all the amazing things you will accomplish in the future.

In naming those that have made my PhD experience possible, I next recognize Beth Strunk. I have never known a day in the Karbstein Lab without Beth, and I feel honored to call her a lab mate, a role model and a best friend. Beth – I can say, without a doubt, that I have spent more time with you than anyone else in my life over the past five years, and I would not have it any other way. Graduate school is a roller coaster of the highs of the highs and the lows of the lows, and it has been great having you by my side. You are a beautiful person inside and out and there are so many qualities about you that I admire: your passion; your patience; your intellectual leadership; your willingness to drop anything (sometimes literally!) to help anyone; your take-a-step-back perspective; and, finally, your love for margaritas and dancing. No matter where our lives take us or how frequent our visits, I will always look forward to our next coffee date.

We have always had a great group dynamic in the Karbstein Lab, and now is no exception. Darryl – you have been part of us since our move to Florida, and I am sure you wonder every day how you ever survived without the Karbstein Lab girls. Although we will have to agree to disagree on the best NFL team, I still very much appreciate your expertise, brilliant ideas and, of course, humor (yes, I still think you are a talker of bs). To the rest of the Karbstein Lab – Bogusia; our newest post-docs, Juliette, Homa and Sohail (you are already proving to be great additions to the lab, and I am disappointed my time here is ending just as yours is beginning!); Caroline (thank you so much for keeping us organized!); and John – through the scientific conversations, practice talks and helpful suggestions you have all been so supportive. The best of luck to each of you in your future endeavors! Even though it has not been easy being the lone graduate student keeping all you post-docs in line (insert smiley face here), you have made the end of my time here memorable.

I also cannot forget the past members of the Karbstein Lab, especially Allison (without who my project would be non-existent; you are an inspiration in so many ways),

Melody (our move to Florida would have been nearly impossible without you; I miss your bubbly presence in lab, but I am hoping we still meet up once a year for a race, whether we are dressed as princesses or not!) and Annie (I love that you are just a phone call away, and I know I can always count on you for honest and great advice).

Finally, I would like to recognize my dissertation committee, Anna Mapp, Bruce Palfey and Nils Walter, for their support throughout my PhD endeavor. Although the move to Florida made our communication during the second half of my PhD more difficult and infrequent, I thank you for helping to make the best of the situation. Your willingness to have discussions via e-mail and critique data electronically are greatly appreciated, and I thank you for taking time out of your busy schedules to help me polish my scientific knowledge, my research perspective and, finally, my dissertation.

The aforementioned groups made completion of the PhD possible, but I credit the Chemistry Department at Washington & Jefferson College with giving me the desire to go to graduate school. The chance-of-a-lifetime teaching opportunity that the department gave me made me immediately realize my passion for teaching and my aspiration to achieve the PhD so I can eventually teach at the undergraduate level. That experience was incredible, and I am excited to officially begin my career in academia.

Last, but not least, I acknowledge my huge support system of friends. To my W&J girls: Alexis, Aimee, Jen, Jennie, Katie, Mary and Nikki. Even though the past six years have separated us all over the world and down different paths in life, I love that we can pick-up right where we left off when we reconnect. To my fellow Michigan "first years," especially Allison (and our weekly dinner dates), Jenna (my favorite 'Manager Special' roommate), Noah (and Erin – one of our favorite double-date couples) and Ted (my little bit of Pittsburgh in Ann Arbor): from Michigan football games to trivia nights to barbecues and walks to support JDRF, you made my first three years of graduate school amazing, and I miss you all. To my Michigan book club girls and my Florida book club girls: I love the monthly breaks to relax, catch-up and gossip over dinner (oh yeah, the books are always good, too!). Finally, to my fellow 'externals' – Candie, Thu and Heather: the move to Florida has both created and strengthened great friendships, and

you three have certainly contributed to making these last two years memorable. I look forward to all the great things you each will achieve!

Table of Contents

List of Figures

List of Tables

Abstract

Even though DEAD-box proteins are often referred to as RNA helicases, their described biochemical activities additionally include protein displacement from RNAs, ATP-dependent RNA binding and RNA annealing. *In vitro* analyses of DEAD-box proteins indicate that in all cases except one, they lack substrate specificity. However, since DEAD-box proteins have non-redundant functions and therefore high specificity *in vivo*, it is consequently believed that co-factors may increase the specificity of DEAD-box proteins by specifically binding individual RNA sequences. Surprisingly, until now, there is no experimental evidence for this simple hypothesis.

Here, I describe both *in vitro* and *in vivo* approaches to dissect the function of two assembly factors essential for 40S ribosome maturation in *Saccharomyces cerevisiae*. I show that Rrp5 is an RNA-binding protein that binds specifically to rRNA sequences within the intron-like segment between 18S and 5.8S rRNAs. Interestingly, this modular protein uses some of its RNA-binding motifs to interact with rRNA in a sequence-specific fashion, while others are used to provide very high affinity. Additionally, my data provide evidence for a direct interaction between Rrp5 and the DEAD-box protein Rok1. Results from assays developed to characterize Rok1 indicate that Rok1 is a unique DEAD-box protein: it preferentially binds double-stranded RNA over single-stranded RNA and has annealing but no unwinding activity. The presence of the C-terminus of Rrp5 greatly enhances this annealing activity in an RNA-sequence specific manner. This data therefore provides evidence that co-factors can enhance the sequence specificity of DEAD-box proteins. Furthermore, the preferentially annealed RNA duplex is part of an inhibitory duplex in the pre-rRNA that serves to regulate the final cleavage step in 18S rRNA maturation. These biochemical results suggest that Rok1 and Rrp5 promote formation of this inhibitory duplex during rDNA transcription; preliminary *in vivo*

structure probing experiments support this Rok1 requirement. Moreover, additional *in vivo* studies indicate that Rok1's ATPase activity is essential for Rok1 dissociation from the ribosome. These results suggest a model by which Rok1 anneals the inhibitory duplex and associates with the pre-ribosome; when the checkpoint for Rok1 removal occurs, Rok1 can then use its ATPase activity to dissociate from the pre-ribosome.

Chapter I

Introduction: Effects of co-factors on RNA helicases

Published in: Young, C.L., Karbstein, K. (2012) Analysis of Cofactor Effects on RNA Helicases in Methods in Enzymology. Elsevier. Vol 511. pp 213-37.

Why are co-factors needed?

Eukaryotic RNA helicases belong to either of the two helicase superfamilies (SF): SF1 or SF2 (Fairman-Williams et al., 2010; Jankowsky, 2011). All of these helicases share a structurally conserved helicase core that consists of two tandem RecA-like domains connected by a flexible linker. In most helicases, these domains contain sequence motifs that are conserved within the SF and are involved in ATP binding and hydrolysis, RNA binding and interdomain contacts. Since the helicase core is highly conserved between RNA helicases, specificity for particular cellular roles must arise outside of these conserved regions, and at least partially derive from the N- and C-terminal extensions that surround the two RecA-like domains. These additional sequences are unique to each helicase. However, the only RNA helicases known to have sequence specificity in the absence of a co-factor are bacterial orthologs of DbpA, an *Escherichia coli* DEAD-box protein that is involved in assembly of the large ribosomal subunit (Fuller-Pace et al., 1993; Iost and Dreyfus, 2006; Sharpe Elles et al., 2009). DbpA specifically recognizes hairpin 92 in 23S ribosomal RNA (Fuller-Pace et al., 1993). Yeast proteins such as Mss116 and Ded1 have a preference for structured RNA (Fairman et al., 2004; Mohr et al., 2008), but no specific target RNAs have been identified. It is therefore believed that by specifically binding individual RNA sequences, helicase co-factors may increase the specificity of RNA helicases, thereby contributing to

their unique and generally non-overlapping roles in various biological processes. Surprisingly, there is no strong experimental evidence for this simple hypothesis.

In addition to possibly increasing biological specificity for substrates, co-factors can also modulate enzymatic activity. RNA helicases such as Prp43 and Upf1 have very low intrinsic ATPase and helicase activities (Chakrabarti et al., 2011; Schutz et al., 2008). Their co-factors Pfa1 and Upf2, respectively, enhance ATPase and helicase activities, thereby contributing to the biological functions of these helicases (Table 1).

Examples in the literature indicate that co-factors can affect all steps of the RNA helicase kinetic cycle (Figure 1.1). For DEAD-box proteins (SF2), it has recently been shown that ATP hydrolysis regulates release of the single-stranded RNA product and not duplex unwinding (Liu et al., 2008). It is therefore not surprising that co-factors can also affect the rates of RNA release (Table 1). Furthermore, since phosphate-release, not ATP hydrolysis, is the irreversible step in the kinetic cycle of both helicases and other ATPases (Henn et al., 2008; Hilbert et al., 2011), co-factors also affect this step (Table 1). As reported for the helicase Dbp5 and its co-factors Nup159 and Gle1, different co-factors can have opposing effects on the same helicase (Alcazar-Roman et al., 2006; Montpetit et al., 2011; von Moeller et al., 2009; Weirich et al., 2006). *In vivo*, these co-factors therefore likely regulate the progression of Dbp5 through its catalytic cycle. In some cases, it has been shown that the same co-factor can have multiple effects

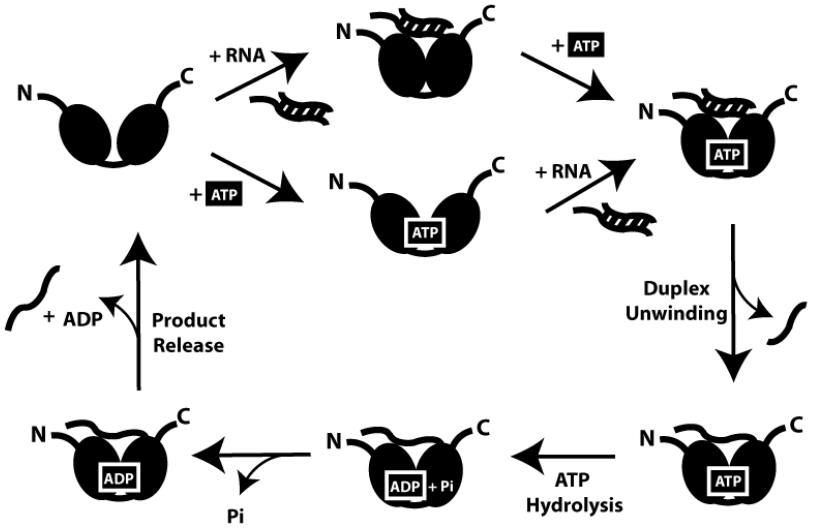


Figure 1.1. Kinetic cycle of RNA helicases. RNA duplex unwinding begins with independent association of the RNA helicase with both an RNA duplex and ATP. Crystal structures have shown that ATP binds at the base of the cleft of the two RecA-like domains while the RNA duplex is bent across the top (Sengoku et al., 2006). RNA helicases can then unwind the RNA duplex via local strand separation, an ATP-independent process (Liu et al., 2008). ATP is then hydrolyzed and dissociation of both the single-strand RNA and the ADP and P_i hydrolysis products completes the cycle. *In vitro* studies with RNA helicases have shown that co-factors can influence each step in the ATPase cycle (Table 1).

Table 1: *S. cerevisiae* RNA helicases and their co-factors

RNA Helicase	Class	Co-Factor	Means of ID	Co-factor Effects				References
				Substrate Binding	ATPase Activity	Helicase Activity	Product Release	
Prp43	DEAH	Ntr1/Ntr2	Genetic interaction[a] Yeast two-hybrid[b] Co-IP[b]; *in vitro* binding[a]	N/D**	No effect	Increase[a]	N/D	[a](Tanaka et al., 2007) [b](Tsai et al., 2005)
Prp43	DEAH	Pfa1	TAP-purification[c] *in vitro* binding[c]	ATP: No effect[c]	Increase[c]	Increase[c]	N/D	[c](Lebaron et al., 2009)
Prp2	DEAH	Spp2	Genetic interaction[d] Yeast two-hybrid[d,e] Reconstituted complex[f]	N/D	N/D	N/D	N/D	[d](Last et al., 1987) [e](Silverman et al., 2004) [f](Warkocki et al., 2009)
Upf1	SF1-like	Upf2	Genetic interaction[g] Co-IP[h;] Crystal structure[h]	N/D	Increase[h]	Increase[h]	N/D	[g](He et al., 1997) [h](Chakrabarti et al., 2011)
Fal1	DEAD	Sgd1	Genetic interaction[i] Co-IP[i]	N/D	N/D	N/D	N/D	[i](Alexandrov et al., 2011)
eIF4A	DEAD	eIF4G	Co-IP[j] *in vitro* binding[j,k] Crystal structure[k]	ATP and ADP: Decrease[l]	Increase[k,l]	N/D	P_i: Increase[l]	[j](Dominguez et al., 1999) [k](Schutz et al., 2008) [l](Hilbert et al., 2011)

RNA Helicase	Class	Co-Factor	Means of ID	Co-factor Effects				References
				Substrate Binding	ATPase Activity	Helicase Activity	Product Release	
eIF4A	DEAD	eIF4B	Genetic interaction[m] Reconstituted complex[n,o,p]	RNA[n] and ATP[o]: Increase	Increase[n,q,r]	Increase[p]	N/D	[m](Coppolecchia et al., 1993) [n](Abramson et al., 1988) [o](Bi et al., 2000) [p](Rozen et al., 1990) [q](Rogers et al., 2001) [r](Rogers et al., 1999)
Dbp8	DEAD	Esf2	Yeast two-hybrid[s] Co-IP[s]; *in vitro* binding[s]	N/D	Increase[s]	N/D	N/D	[s](Granneman et al., 2006)
Dbp5	DEAD	$Gle1_{InsP6}$*	Genetic interaction[t] Yeast two-hybrid[t,u] *in vitro* binding[t] Crystal structure[v]	RNA[u] and ATP[x]: Increase	Increase[u,w]	N/D	RNA: Increase[v]	[t](Strahm et al., 1999) [u](Weirich et al., 2006) [v](Montpetit et al., 2011) [w](Alcazar-Roman et al., 2006) [x](Noble et al., 2011)
Dbp5	DEAD	Nup159	Genetic interaction[y] Crystal structure[z]	RNA: Decrease[z,aa]	Decrease[z]	N/D	ADP: Increase[bb]	[y](Hodge et al., 1999) [z](Montpetit et al., 2011) [aa](von Moeller et al., 2009) [bb](Noble et al., 2011)
Ded1	DEAD	Gle1	Genetic interaction[cc] Co-IP[cc]; *in vitro* binding[cc]	N/D	Decrease[cc]	N/D	N/D	[cc](Bolger and Wente, 2011)
Brr2	Ski2-like	Prp8	Genetic interaction[dd] Yeast two-hybrid[dd] *in vitro* binding[ee,ff]	ssRNA: Decrease[ee] dsRNA: Increase[ff]	Decrease[ee] ***	Increase[ee,gg]	N/D	[dd](van Nues and Beggs, 2001) [ee](Maeder et al., 2009) [ff](Zhang et al., 2009) [gg](Pena et al., 2009)

*$Gle1_{InsP6}$: Gle1 has stronger effects on Dbp5's activities when it is bound to $InsP_6$, an endogenous small molecule (Alcazar-Roman et al., 2006; Weirich et al., 2006); **N/D indicates that this co-factor effect has not been determined *in vitro*; *** The effect on ATPase activity likely arises from weakened RNA binding[dd]

(Table 1). However, because the kinetic cycle in these examples is dominated by a single rate-limiting step, it is not clear whether many of these effects have *physiological* repercussions. Additional effects reported to arise from co-factors include an increase (Ballut et al., 2005; Weirich et al., 2006) or decrease (von Moeller et al., 2009) in RNA affinity and a decrease (Maeder et al., 2009) in ATPase activity (Table 1).

By modulating activity and potentially increasing specificity of RNA helicases, co-factors provide an additional level of regulation. Furthermore, by encoding the helicase core and the co-factor on two distinct polypeptides, modularity is achieved *in vivo*. As an example, the yeast helicase Prp43 functions in both pre-mRNA splicing and ribosome assembly, where it is associated with and regulated by Ntr1/Ntr2 and Pfa1, respectively (Lebaron et al., 2009; Martin et al., 2002; Tanaka et al., 2007; Tsai et al., 2005; Walbott et al., 2010). Therefore, in order to thoroughly understand their cellular functions, helicases must be studied together with the co-factors that modulate them.

How do co-factors modulate the enzymatic activities of RNA helicases?

While a mechanistic understanding of co-factor effects is still lacking for many helicases, recent biochemical and structural data have indicated a mechanism that is shared by at least two yeast RNA helicases: the DEAD-box protein Dbp5 and the SF1-like helicase Upf1 (Chakrabarti et al., 2011; Hodge et al., 2011; Noble et al., 2011). These helicase-co-factor complexes share structural similarity, but no detectable sequence homology. The regulatory N-terminal domain of Upf1, which inhibits its ATPase activity, is displaced upon binding of its co-factor Upf2; this rearrangement results in the stimulation of Upf1's ATPase activity (Chakrabarti et al., 2011). The inhibitory domain of Upf1 also extends the RNA binding site (Chakrabarti et al., 2011). As a result, it is expected to increase RNA binding affinity. The binding of the Upf1 co-factor is therefore predicted to weaken RNA binding by shortening the RNA-protein interface. Furthermore, it was shown that nucleotide binding moderately weakens RNA binding (Chakrabarti et al., 2011). In a similar mechanism, addition of the co-factor Gle1

to Dbp5 weakens RNA binding and is predicted to displace an N-terminal extension, which also has autoinhibitory effects on Dbp5's ATPase activity (Montpetit et al., 2011).

Even though the crystal structures of Dbp5 and Upf1 and their respective co-factors were tremendously insightful in predicting the mechanisms of co-factor activation, biochemical experiments were required to validate the hypothesized effects. Additionally, many complexes that contain combinations of helicase, co-factor, nucleotide and/or nucleic acid are highly unstable and therefore often poorly accessible with structural biology methods. In this case, biochemical approaches are perhaps the only available method for probing synergistic effects from multiple ligands (nucleotide, RNA, co-factor) on helicase function. For example, crystal structures also suggest that the effects from ADP and Gle1 on RNA binding should be synergistic, as both prevent RNA binding (Montpetit et al., 2011). Similarly, kinetic methods are the only methods to probe transition states, which are associated with the conformational changes that modulate the helicases conformational cycle. Before the effects of co-factors on RNA helicases can be studied, however, co-factors must first be identified. Below is a summary of the common approaches used to identify RNA helicase-specific co-factors.

Identification of co-factors

In many cases, the study of RNA helicases and their biological functions has been slowed by the lack of information about their co-factors (*e.g.*, in yeast, co-factors with known *in vitro* effects have been identified for only 7 of the 41 helicases). This has precluded a meaningful analysis, as the biologically active complex is undefined. Furthermore, and likely linked to this deficiency, RNA targets often remain unknown. Further dissection of the roles of yeast helicases will therefore almost certainly require a systematic approach to identify co-factors.

Yeast two-hybrid assays have many documented shortcomings (Bruckner et al., 2009), yet they are a strong starting point for the identification of helicase-co-factor interactions. Especially for the analysis of the roles of the many nuclear helicases involved in pre-mRNA splicing and ribosome assembly, this technique can be fruitful, as evidenced by a complete map of protein-protein interactions within pre-ribosomal

subcomplexes obtained with this technique (Champion et al., 2008; Charette and Baserga, 2010); most of the RNA helicase co-factors in Table 1 were also identified by yeast two-hybrid analyses. As an untested, yet potentially interesting, extension of this technique, three-hybrid assays using the helicase and/or the identified co-factor could be used to identify RNA binding sites.

Genetic interactions can also provide insight into potential co-factors. For example, a point mutation within a protein-protein interaction module of the RNA-binding protein Rrp5 is suppressed by overexpressing the helicase Rok1 (Torchet et al., 1998), indicating that this genetic interaction could reflect on a direct physical interaction. We have verified this hypothesis using recombinant proteins *in vitro* (see Chapter 3). Genetic interactions have also been substantiated *in vitro* for the helicases Prp43 and Upf1 and their respective co-factors Ntr1/Ntr2 and Upf2 (He et al., 1997; Tanaka et al., 2007). Similarly, a creative genetic strategy was recently described to isolate RNA targets for helicases (Proux et al., 2011). Iost and co-workers screened a library of RNA mutants for suppressors of growth defects observed in the absence of SrmB, an *Escherichia coli* DEAD-box protein involved in ribosome assembly (Charollais et al., 2003). The isolated mutants suggest a role for SrmB in resolving an inhibitory structure between 23S rRNA and 5S rRNA, although some data are not explained by this model (Proux et al., 2011).

In theory, co-immunoprecipitation assays are powerful tools to isolate protein-binding partners for any given protein; however, they have not always been fruitful for the study of RNA helicases. A likely reason is that a large majority of helicases are involved in either ribosome assembly or pre-mRNA splicing. In both cases, helicases associate with very large macromolecular complexes, which are also immunoprecipitated. As a result, dozens of associated proteins are often identified, but most are indirectly bound, and this methodology provides no means to distinguish between direct or indirect interactions.

Another method to identify helicase co-factors is cross-linking. Photoactivatable (e.g., diazirine analogs (Suchanek et al., 2005)) or chemical cross-linkers such as formaldehyde can be used to covalently link a helicase to any nearby protein, and

subsequent purification can be carried out under non-native conditions (*e.g.* using a His-tag on the helicase of interest*)* to retain co-factors that are bound to the helicase via covalent interactions. Similarly, RNA cross-linking has been recently used to map the sites of interactions for several ribosome assembly factors, including the helicase Prp43 (Bohnsack et al., 2009). Comparison of cross-links obtained from helicases to those of putative co-factors should reveal neighboring binding sites.

While these tools are of tremendous value for identifying potential co-factors, it is important to realize that they are merely a starting point. Interactions should be confirmed by combining the methods described above. For example, genetic interactions between a helicase and potential co-factor can be corroborated by yeast two-hybrid data and vice versa. Genetic experiments and yeast two-hybrid data can also be used in combination to map protein regions involved in interactions. In the case of Brr2 and Prp8, this approach has been used to demonstrate that both the N- and C-termini of Prp8 interact with Brr2 (van Nues and Beggs, 2001). The gold standard for a helicase-co-factor interaction is demonstration of direct protein-protein interactions using recombinant proteins and, ideally, the analysis of co-factor effects on the enzymatic activity of the RNA helicase.

Studying the effects of co-factors on RNA helicases

In order to gain a mechanistic understanding of how co-factors affect an RNA helicase and to delineate the function of the helicase-co-factor complex, *in vitro* studies should be conducted. This is especially important since helicases cycle through different steps. Thus, analyzing effects *in vivo* will only allow the analysis of one such step (generally the first). Effects on the entire catalytic cycle, including RNA and nucleotide binding and ATPase, helicase or annealing activities, should be analyzed. First, these characteristics for the RNA helicase alone can be analyzed; varying effects in the presence of co-factors can then be observed.

Chapter 3 focuses on the DEAD-box protein Rok1 and the enzymatic effects of its co-factor Rrp5, an RNA-binding protein. Before the effects of Rrp5 could be observed, however, its interactions with RNA, specifically the pre-ribosomal RNA (pre-rRNA), were analyzed; these results are presented in Chapter 2.

Chapter II

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

Published in: Young, C.L. Karbstein, K. (2011). The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA. *RNA.* **17**, 512-521.

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 *E. coli* ribosomal protein S1 (Subramanian, 1983), the S1 motif consists of ~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 (Bycroft et al., 1997; Murzin, 1993). 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 β2 and β3) and the loops between β1 and β2, as well as β3 and α ((Theobald et al., 2003); Figure 2.1C).

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 comprised 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 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 (Eppens et al., 1999; Torchet et al., 1998; Venema and Tollervey, 1996).

Previous work has indicated that Rrp5 can be physically and functionally separated between the ninth and tenth S1 motif. N-terminal or C-terminal truncation of Rrp5 between the ninth and tenth 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 so-called site A_3, while the C-terminal part was required for upstream cleavage at so-called site A_2 (Figure 2.1A; (Eppens et al., 1999; Torchet et al., 1998)). These cleavage sites are located about 70 nucleotides apart from each other within the internal transcribed spacer 1 (ITS1), the sequence between 18S and 5.8S rRNAs (Figure 2.1A), which is removed during maturation. This suggested that Rrp5 may have interactions within 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 ITS1 and surrounding regions. The data pinpoint Rrp5's binding site to three regions within 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 ($S1_{1-9}$) or the last three ($S1_{10-12}$) 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

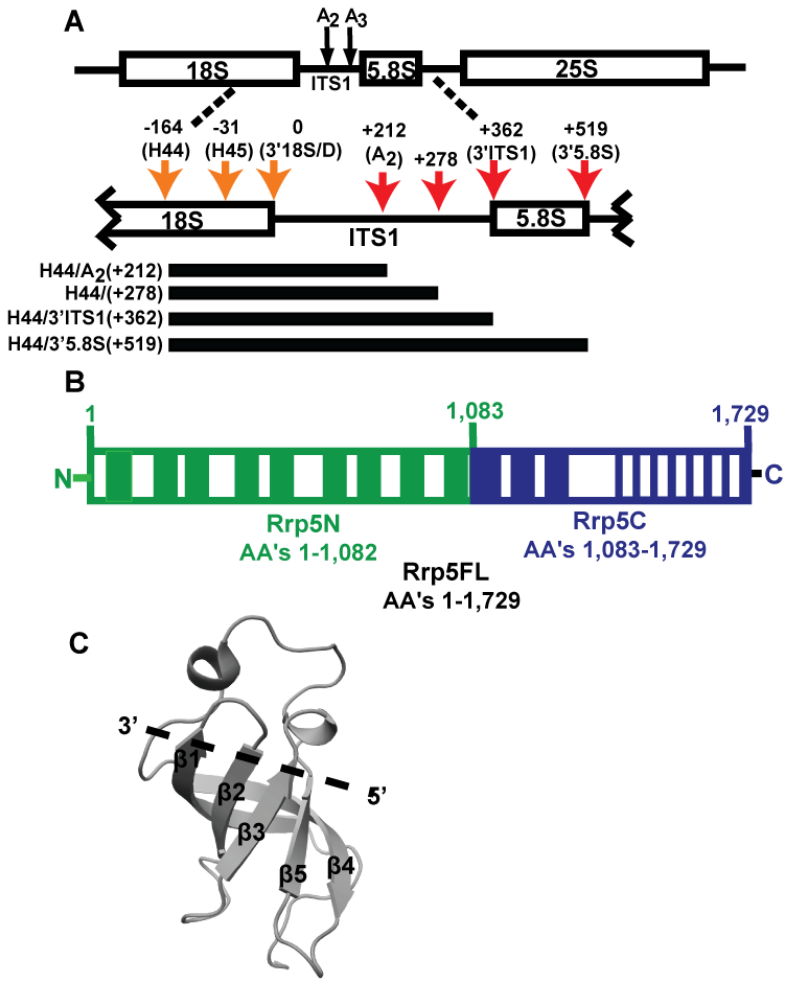


Figure 2.1. Pre-rRNA and Rrp5 constructs. (A) Pre-rRNA fragments used in this study. The 18S, 5.8S and 25S rRNAs are co-transcribed in a single transcript. Experiments described herein use rRNA fragments spanning the internal transcribed spacer 1 (ITS1), which includes cleavage sites A_2 and A_3 (Lamanna 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 - 1,729, the N-terminal fragment Rrp5N contains amino acids 1 - 1,082, and the C-terminal fragment Rrp5C comprises amino acids 1,083 - 1,729. All proteins were expressed and purified from *E. coli*. (C) S1 ribbon drawing modeled from domain 1 of aIF2α (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)).

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 tenth 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 1-1,082 ($S_{1\text{-}9}$), while Rrp5C, the C-terminal fragment, contains residues 1,083-1,729 ($S1_{10\text{-}12}$, Figure 2.1B). In agreement with prior *in vivo* studies, these two fragments fully complement growth in Rrp5-depleted cells when provided *in trans* (Figure 2.2), indicating that they are fully functional *in vivo*. The observation that Rrp5C, when expressed alone, provides for more growth than 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*.

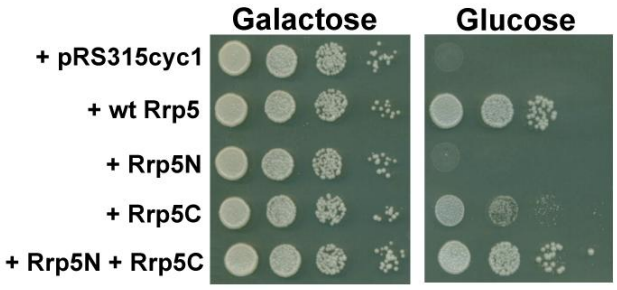


Figure 2.2. Rrp5N and Rrp5C constructs fully complement growth *in trans*. Ten-fold serial dilutions of gal:Rrp5-TAP strain transformed with Rrp5N, Rrp5C or Rrp5N and Rrp5C. Wild-type Rrp5 and pRS315cyc1 were used as positive and negative controls, respectively. Growth was compared on glucose- or galactose-containing plates.

Prior observations have shown that Rrp5 is required for cleavage at both sites A_2 and A_3 within ITS1 (Venema and Tollervey, 1996). Furthermore, *in vitro* binding experiments with immunopurified Rrp5 have provided evidence that Rrp5 binds U-rich sequences, akin to those found in ITS1 (de Boer et al., 2006). We therefore decided to use pre-rRNA fragments that include the 3'-minor domain of 18S rRNA in addition to partial or full ITS1 sequences (Figure 2.1A). In prior work, we have shown that these rRNA constructs fold into the tertiary structure expected based on the mature small subunit crystal structures (Lamanna and Karbstein, 2009). Additionally, the secondary structures we observe *in vitro* reflect those observed *in vivo* (Kaminishi et al., 2007;

Schuwirth et al., 2005). This agreement with *in vivo* structures has also been confirmed for the larger rRNA constructs used herein (see below), thus demonstrating that these small rRNA fragments are valid mimics of the predicted *in vivo* substrate of Rrp5.

Rrp5 specifically binds rRNAs containing ITS1

To study Rrp5's interaction with pre-rRNA, we developed a gel-shift assay. In this assay, recombinant Rrp5 (Rrp5C, Rrp5N or Rrp5FL) is incubated with pre-folded RNA. Protein-bound RNA is separated from free RNA using native PAGE (Figure 2.3), and the fraction of bound RNA is plotted as a function of Rrp5 concentration and fit to obtain $K_{1/2}$ values (see Materials and Methods).

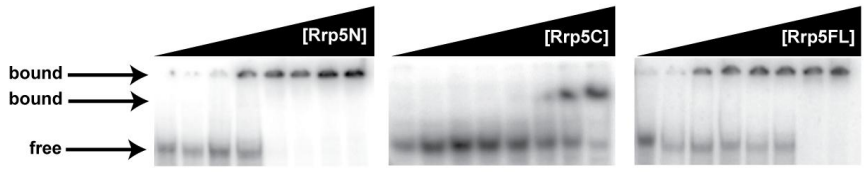


Figure 2.3. Representative gel-shift data using trace ^{32}P-labeled H44/3'ITS1 and excess Rrp5N, Rrp5C or Rrp5FL. H44/3'ITS1 was pre-folded and incubated with increasing amounts of each Rrp5 construct for 2 hours before being loaded on a 6% native acrylamide gel as described.

We first tested if the interactions between Rrp5 and the rRNA mimics are specific. Considering Rrp5's role in cleavage at sites A_2 and A_3, we used H44/3'ITS1, the pre-rRNA fragment that encompasses both of these cleavage sites, as a positive control. As a negative control, we tested binding to the thiostreptone loop (TSL), an RNA fragment from 25S rRNA (Karbstein et al., 2005). Native PAGE demonstrates that upon the addition of Rrp5N, Rrp5C or Rrp5FL, H44/3'ITS1 is indeed shifted upwards in the gel, indicating an interaction (Figure 2.4A, lanes 2, 6 and 10). These interactions with radiolabeled H44/3'ITS1 are competed off with excess unlabeled H44/3'ITS1 (lanes 3, 7 and 11), but not to the same extent with excess unlabeled polyU (lanes 4, 8 and 12), showing that the interaction between the Rrp5 constructs and H44/3'ITS1 is stronger than their interactions with polyU. In Figure 2.4A, lanes 3, 7, and 11 also show that after addition of unlabeled RNA, significant amounts of RNA remain in the wells. This is not because competition was ineffective, but is instead due to aggregation of the H44/3'-ITS1 RNA at concentrations above 1 μM RNA, as this is also observed when unlabeled RNA is added in the absence of protein (data not shown). To provide additional evidence that

our RNA constructs are specifically recognized by Rrp5, we carried out similar competition experiments using Rrp5N and Rrp5FL and labeled and unlabeled H44/A_2 RNA, for which the aggregation problem is ameliorated (Figure 2.4C). These data show that most of the label in the gel-shift is efficiently competed by addition of excess unlabeled H44/A_2 RNA, but only partial competition is observed with polyU at the same concentration. It was not possible to carry out this experiment with Rrp5C, as its weak affinity for H44/A_2 required us to use large amounts of Rrp5, and therefore excess amounts of RNA could not be achieved. Further evidence for specific binding of Rrp5C and Rrp5FL comes 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 (Figure 2.4B). However, it should be noted that Rrp5N's affinity for TSL is only two-fold below its affinity for H44/3'ITS1 ($K_{1/2}$ = 0.1±0.01 μM^2 and 0.2±0.1 μM^2 for H44/3'ITS1 and TSL, respectively), while Rrp5FL's affinity is decreased at least 30-fold, and could not be saturated at feasible Rrp5 concentrations (data not shown). These data indicate that Rrp5N binding is relatively non-specific (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.

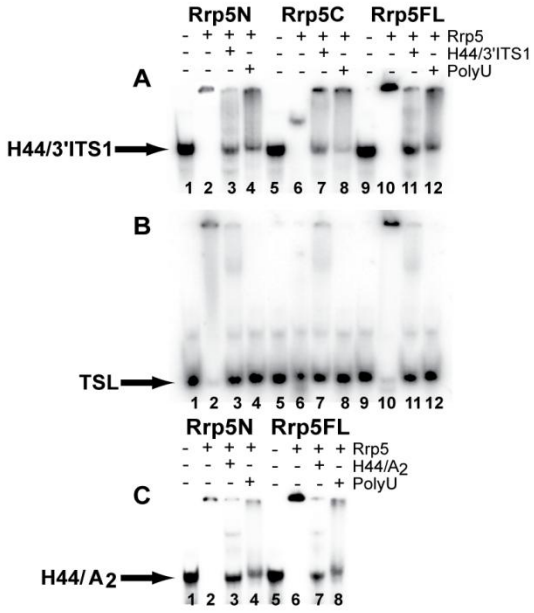


Figure 2.4. Rrp5C and Rrp5FL specifically bind H44/3'ITS1. Binding of Rrp5N, Rrp5C and Rrp5FL to H44/3'ITS1 (A), the 25S thiostreptone loop (TSL; (Karbstein et al., 2005); (B)) or H44/A_2 ((C); Rrp5N and Rrp5FL only) was observed via electrophoretic mobility shift assays. Excess unlabeled H44/3'ITS1 rRNA (5 μM with Rrp5N and Rrp5FL; 10 μM with Rrp5C), H44/A_2 rRNA (5 μM) and PolyU RNA (5 μM with Rrp5N and Rrp5FL; 10 μM with Rrp5C) were used in competition experiments. In (A) and (B), concentrations of 1 μM, 5 μM and 0.75 μM were used for Rrp5N, Rrp5C and Rrp5FL, respectively. In (C), both Rrp5N and Rrp5FL were used at 0.5 μM. After addition of unlabeled H44/3'-ITS1 to complexes containing labeled H44/3'-ITS1, the shift out of the well is not complete because H44/3'-ITS1 aggregates in a Rrp5-free state (data not shown). This aggregation is observed in both the presence and absence of Rrp5 at RNA concentrations above 3 μM (data not shown). However, as shown in (C), the addition of unlabeled H44/A_2 results in labeled H44/A_2 reappearing in the free form, indicating that the gel shift observed upon the addition of Rrp5N and Rrp5FL is not due to aggregation and instead is a native binding interaction.

Comparison of lane 1 with lanes 2, 6 and 10 in Figure 2.4A also indicates that upon binding of Rrp5C, RNA is shifted upwards 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 non-specific aggregation, as the complexes formed with Rrp5N and Rrp5FL are amenable to competition with unlabeled RNA (Figures 2.4A and 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 2.5 show that Rrp5FL binds the strongest, followed by Rrp5N and then Rrp5C. For this RNA, binding to Rrp5FL is approximately five-fold stronger than binding to Rrp5N and approximately 50-fold stronger than binding to Rrp5C, with Rrp5FL binding in the lower nanomolar range (Table 2). These data suggest that most of the binding energy arises from interactions made with the first nine S1 domains (Rrp5N), as deletion of $S1_{10\text{-}12}$ only has a five-fold effect on RNA binding. Since Rrp5N has nine S1 RNA-binding modules, this result was not unexpected; however, the five-fold 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.

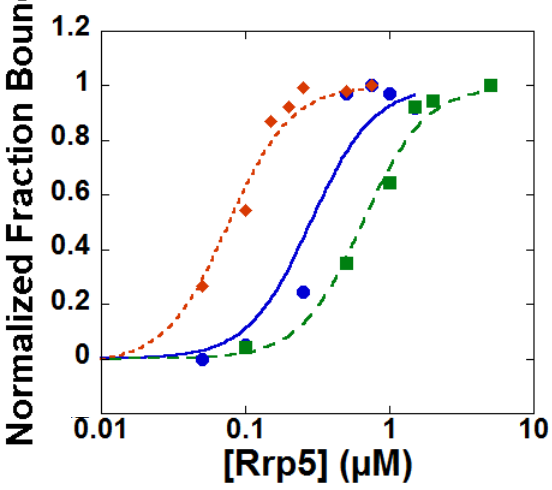


Figure 2.5. Affinities of Rrp5 constructs for H44/3'ITS1. Rrp5FL (♦) binds the H44/3'ITS1 rRNA fragment approximately five-fold stronger than Rrp5N (●), and approximately 50-fold stronger than Rrp5C (■).

Table 2: RNA binding affinities for Rrp5 constructs

rRNA Construct	$K_{1/2}$ (μM^2)		
	Rrp5N	**Rrp5C**	**Rrp5FL**
H44/A$_2$(+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 $MgCl_2$, 100 mM KCl and 50 mM HEPES, pH 7.5. Data are the averages of three or more experiments.

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 A_2(+212), (+278) (6 nucleotides upstream of A_3) or 3'ITS1(+362) at the 3' end (Figure 2.1A). 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 2). 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 (Figure 2.6A, Table 2). 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 either binding of 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 non-detectable. Together, these experiments indicate that

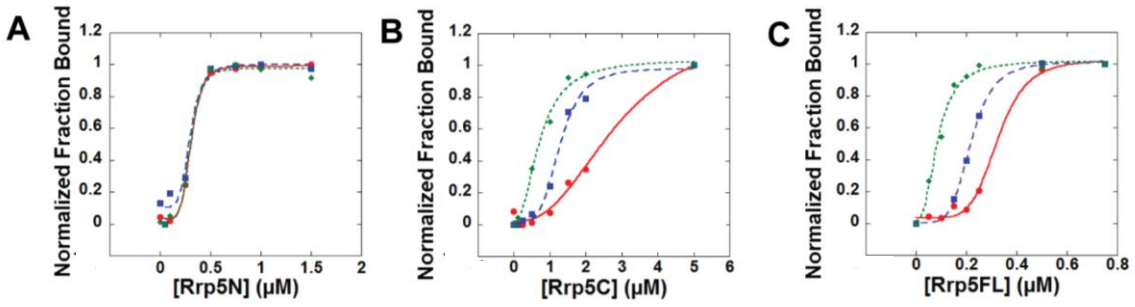


Figure 2.6. Binding of Rrp5N (A), Rrp5C (B) and Rrp5FL (C) to H44/A_2(●), H44/(+278)(■) and H44/3'ITS1(♦) rRNA mimics.

while Rrp5N binds to pre-rRNA fragments including regions of ITS1 relatively tightly, it does so relatively non-specifically. 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 Rrp5N binding (Table 2), 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 A_3 cleavage upon deletion of Rrp5N (Eppens et al., 1999; Torchet et al., 1998), 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 A_2, we determined Rrp5N's affinity for the A_2/3'-ITS1 RNA, which was also within less than three-fold of the H44/3'-ITS1 affinity ($K_{1/2}$ = 0.1±0.03 μM^2 and 0.27±0.03 μM^2 for H44/3'-ITS1 and A_2/3'-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 A_2 to (+278) to 3'ITS1 (Figure 2.6B, Table 2). 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 2 and data not shown). This result shows that Rrp5C makes specific interactions with ITS1, mainly 3' to site A_2.

Binding data for Rrp5FL combines the trends observed for Rrp5N and Rrp5C (Figure 2.6C, Table 2). While Rrp5FL binds even more strongly to rRNA than 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 2), indicating that Rrp5 does not make any interactions with 5.8S rRNA.

In summary, deleting $S1_{1-9}$ (Rrp5C) weakens RNA binding 50-fold, but does not affect binding specificity, while deletion of $S1_{10-12}$ (Rrp5N) only weakens RNA binding five-fold, 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 ITS1 that is 3'-to site A_2, 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 rRNA secondary structure can be obtained via DMS modification. In order to account for differences in the loading, or the efficiency of reverse transcription between different lanes in the gel, we also carried out line scans that were normalized to nearby residues (Figure 2.7B).

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 non-specific 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 higher than 1 μM, we used the shorter H45/3'5.8S rRNA fragment to eliminate the aggregation; binding experiments indicated that Rrp5FL has the same

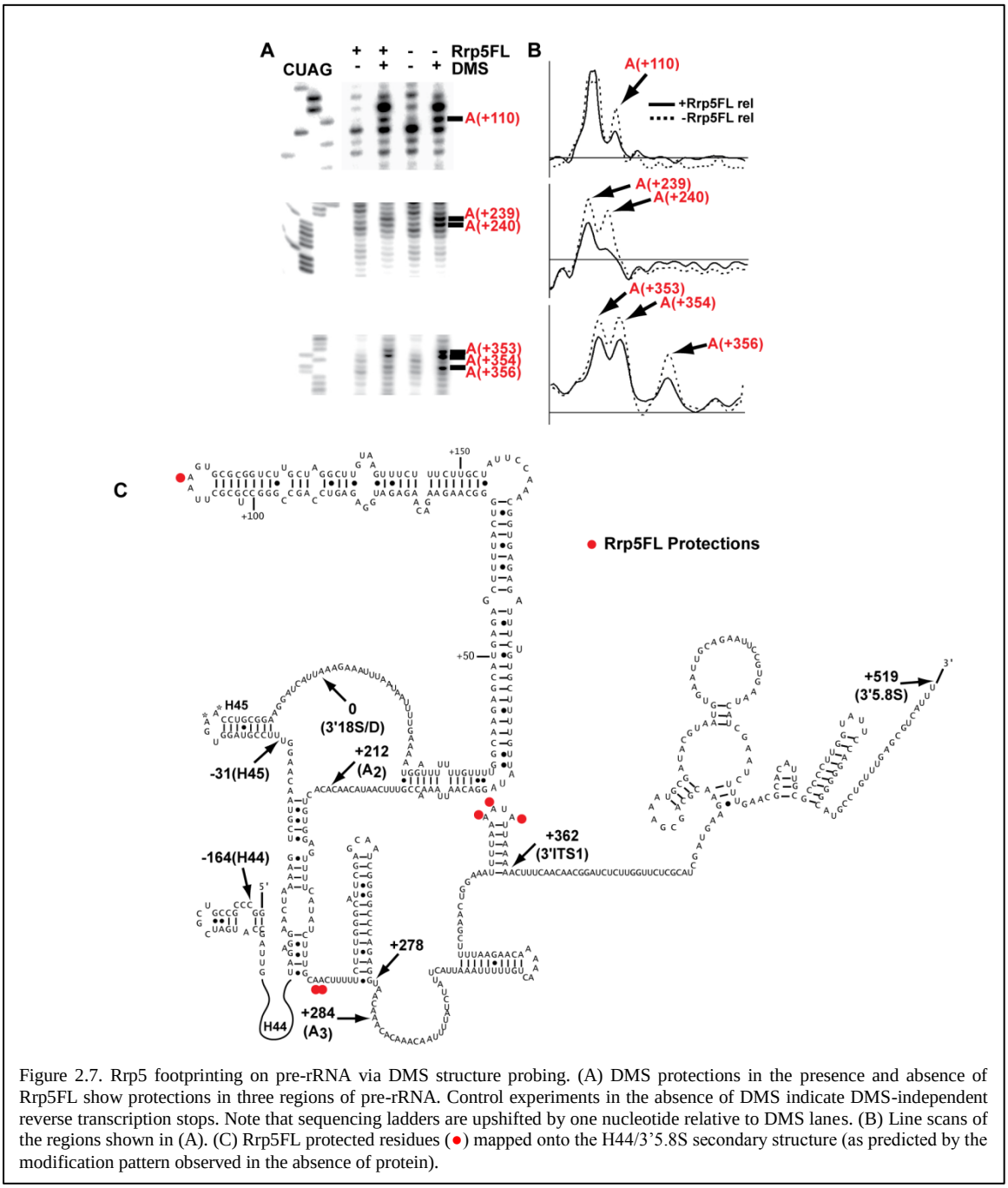


Figure 2.7. Rrp5 footprinting on pre-rRNA via DMS structure probing. (A) DMS protections in the presence and absence of Rrp5FL show protections in three regions of pre-rRNA. Control experiments in the absence of DMS indicate DMS-independent reverse transcription stops. Note that sequencing ladders are upshifted by one nucleotide 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).

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 secondary structure we previously reported, as well as the secondary structures expected for the 3'-minor domain of 18S rRNA and 5.8S rRNA (Figure 2.7C; (Lamanna and Karbstein, 2009, 2010; Wuyts et al., 2001)). Upon the addition of Rrp5FL, protected residues were identified in three areas (Figure 2.7). 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 of a helix formed between H44 and ITS1 ((Lamanna and Karbstein, 2010), +239/240); as well as protections close to the 3'-end of ITS1 (+353 region). These protections were quantified in line scans (Figure 2.7B) and are plotted onto the secondary structure in Figure 2.7C. The footprinting data fully support the binding data, which showed weak interactions with ITS1 prior to site A_2 and additional contributions both 5'-and 3'- to site A_3 (Table 2). 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 2)), 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 ITS1, with most of the binding energy arising from contacts occurring 3' to site A_2. 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 (Figure 2.1B), and do not contribute to RNA binding (de Boer et al., 2006). In surprising contrast, an N-terminal fragment (Rrp5N) containing the first nine S1 domains binds all pre-rRNA mimics strongly but non-specifically. 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 A_2 (in a stem loop where Nob1, the D-site endonuclease, also makes an interaction (Lamanna and Karbstein, 2009)), one 5' to site A_3, and one in a stem loop near the 3'-end of ITS1 (Figure 2.7).

These quantitative experiments confirm and extend prior qualitative experiments, largely undertaken via *in vivo* truncation analyses (Eppens et al., 1999; Torchet et al., 1998). Deletions or mutations in the first nine S1 domains (Rrp5N) affect cleavage at site A_3, while mutations or deletions in the last three S1 or the C-terminal TPR motifs (Rrp5C) inhibit cleavage at site A_2 (Eppens et al., 1999; Torchet et al., 1998); these results suggest that Rrp5 is interacting with 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 A_3, is essential ((Eppens et al., 1999) and Figure 2.2), while cleavage at site A_3 itself is not (Henry et al., 1994). While Rrp5C is sufficient to provide the necessary specificity for cleavage at site A_2, 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 *trans*) is required for strong binding to pre-rRNA, resulting in cleavage at site A_2. 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 and colleagues identified an Rrp5 interaction with rRNA 3' of site A_2 (de Boer et al., 2006); 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 A_3. 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 pre-ribosomes (Venema and Tollervey, 1996) and is required for both 40S and 60S assembly (Eppens et al., 1999; Torchet et al., 1998; Venema and Tollervey, 1996). Previous proteomic analysis of pre-ribosomal 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 A_2 and instead has the majority of its interactions in regions 3' to site A_2. Thus, it appears that after co-transcriptional cleavage at site A_2, Rrp5 releases the RNA 5' to site A_2 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).

ITS1 varies significantly in length (137 nucleotides in *C. albicans* vs. 1,095 nucleotides in humans) as well as sequence (Lamanna and Karbstein, 2010). The three regions of 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 (Figure 2.8). 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 ITS1. Interestingly, Rrp5's last five S1 domains are much less conserved than its first seven S1 domains

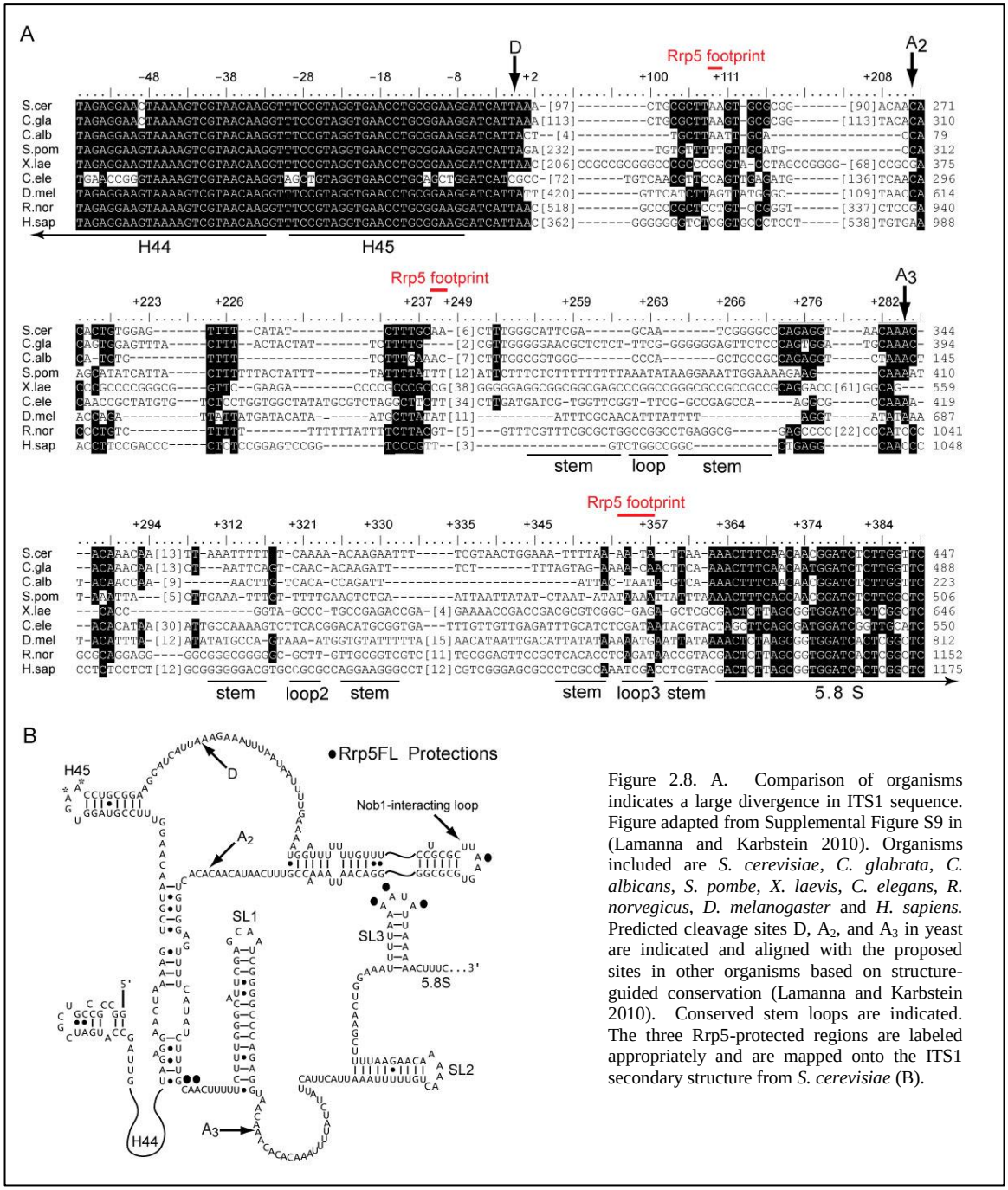
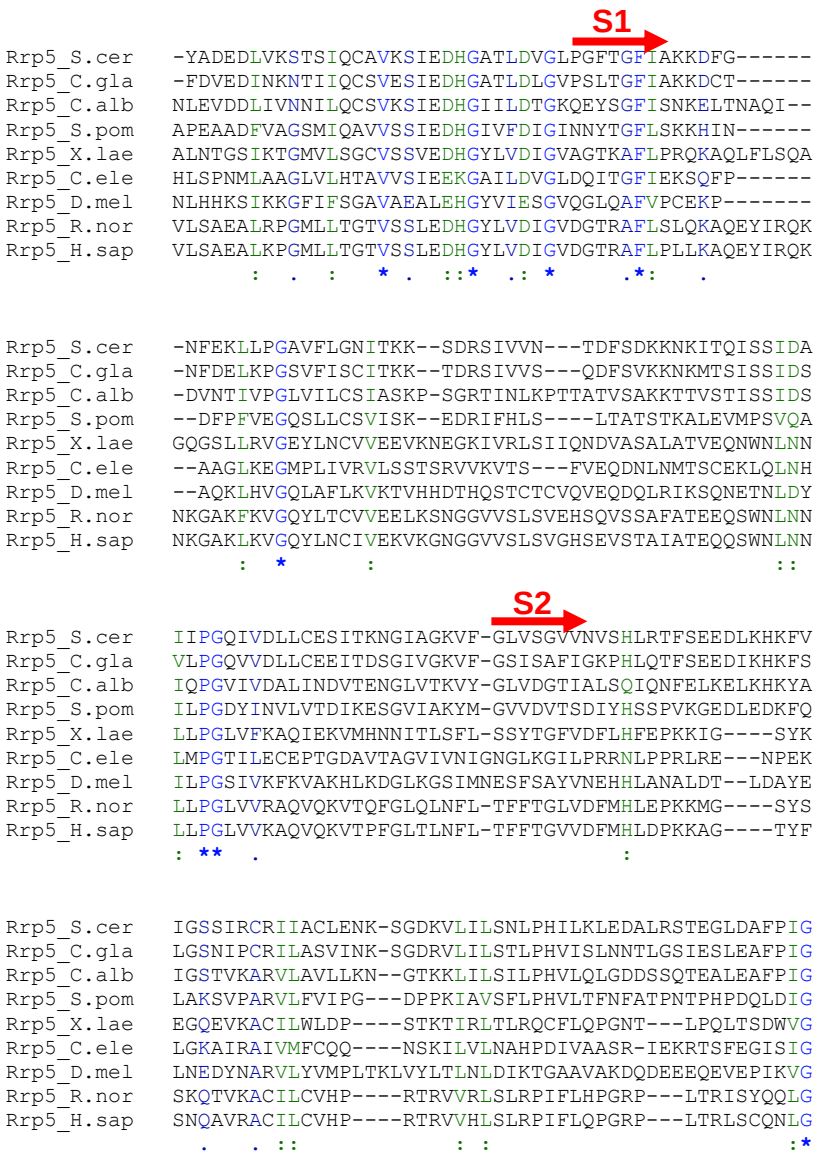
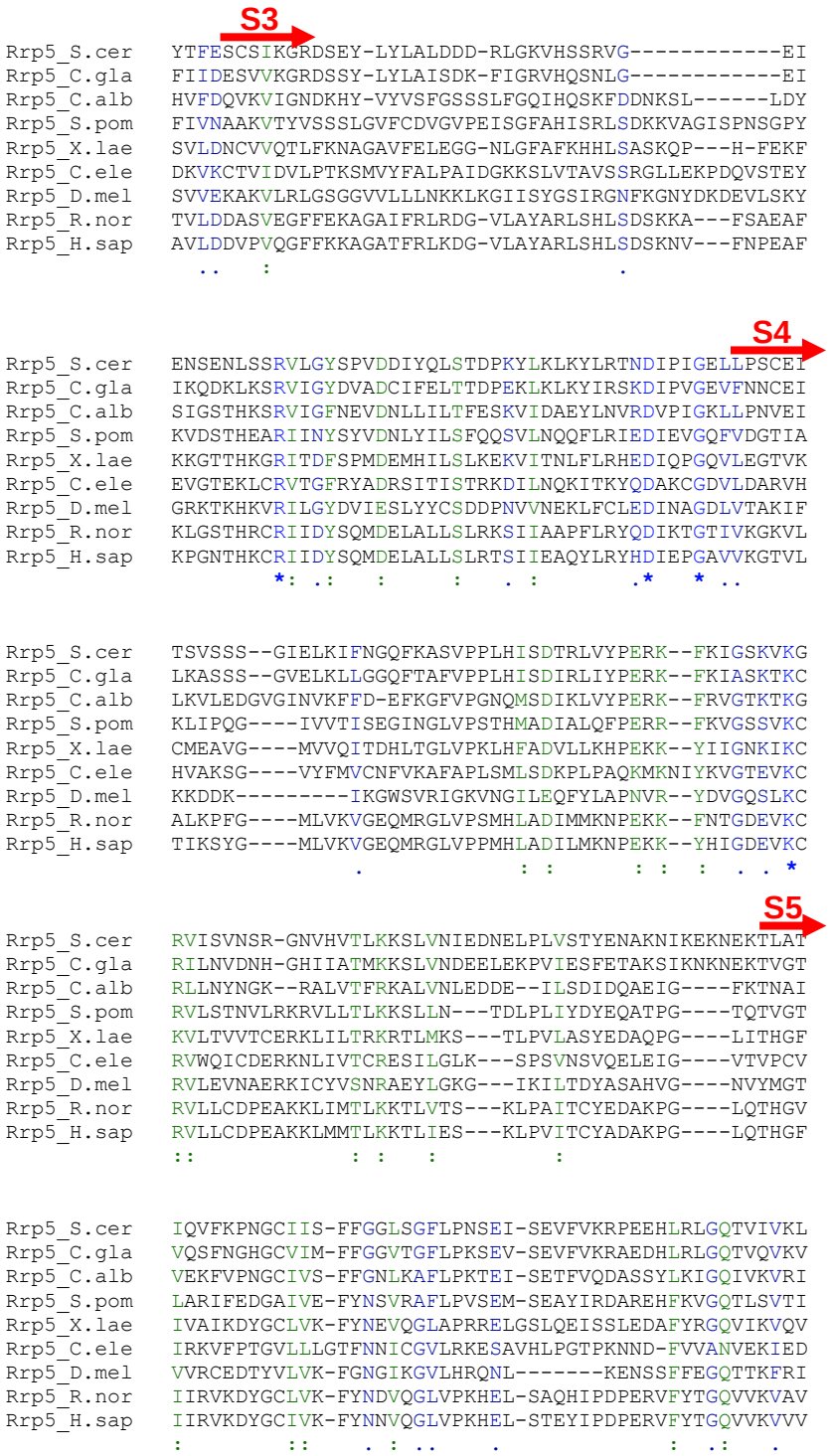


Figure 2.8. A. Comparison of organisms indicates a large divergence in ITS1 sequence. Figure adapted from Supplemental Figure S9 in (Lamanna and Karbstein 2010). Organisms included are *S. cerevisiae*, *C. glabrata*, *C. albicans*, *S. pombe*, *X. laevis*, *C. elegans*, *R. norvegicus*, *D. melanogaster* and *H. sapiens.* Predicted cleavage sites D, A_2, and A_3 in yeast are indicated and aligned with the proposed sites in other organisms based on structure-guided conservation (Lamanna and Karbstein 2010). Conserved stem loops are indicated. The three Rrp5-protected regions are labeled appropriately and are mapped onto the ITS1 secondary structure from *S. cerevisiae* (B).

(Figure 2.9), 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 tenth S1 domain contains a large expansion of the loop between β2 and β3, which is part of the expected RNA-binding surface (Figures 2.1C and 2.9); interestingly, mammals also have the largest ITS1. It is tempting to speculate that the expansion in $S1_{10}$ in the evolutionary divergence of Rrp5 homologs arose to accommodate an expanded ITS1.




S3
Rrp5_S.cer YTFESCSIKGRDSEY-LYLALDDD-RLGKVHSSRVG----------EI
Rrp5_C.gla FIIDESVVKGRDSSY-LYLAISDK-FIGRVHQSNLG----------EI
Rrp5_C.alb HVFDQVKVIGNDKHY-VYVSFGSSSLFGQIHQSKFDDNKSL------LDY
Rrp5_S.pom FIVNAAKVTYVSSSLGVFCDVGVPEISGFAHISRLSDKKVAGISPNSGPY
Rrp5_X.lae SVLDNCVVQTLFKNAGAVFELEGG-NLGFAFKHHLSASKQP---H-FEKF
Rrp5_C.ele DKVKCTVIDVLPTKSMVYFALPAIDGKKSLVTAVSSRGLLEKPDQVSTEY
Rrp5_D.mel SVVEKAKVLRLGSGGVVLLLNKKLKGIISYGSIRGNFKGNYDKDEVLSKY
Rrp5_R.nor TVLDDASVEGFFEKAGAIFRLRDG-VLAYARLSHLSDSKKA---FSAEAF
Rrp5_H.sap AVLDDVPVQGFFKKAGATFRLKDG-VLAYARLSHLSDSKNV---FNPEAF
S4
Rrp5_S.cer ENSENLSSRVLGYSPVDDIYQLSTDPKYLKLKYLRTNDIPIGELLPSCEI
Rrp5_C.gla IKQDKLKSRVIGYDVADCIFELTTDPEKLKLKYIRSKDIPVGEVFNNCEI
Rrp5_C.alb SIGSTHKSRVIGFNEVDNLLILTFESKVIDAEYLNVRDVPIGKLLPNVEI
Rrp5_S.pom KVDSTHEARIINYSYVDNLYILSFQQSVLNQQFLRIEDIEVGQFVDGTIA
Rrp5_X.lae KKGTTHKGRITDFSPMDEMHILSLKEKVITNLFLRHEDIQPGQVLEGTVK
Rrp5_C.ele EVGTEKLCRVTGFRYADRSITISTRKDILNQKITKYQDAKCGDVLDARVH
Rrp5_D.mel GRKTKHKVRILGYDVIESLYYCSDDPNVVNEKLFCLEDINAGDLVTAKIF
Rrp5_R.nor KLGSTHRCRIIDYSQMDELALLSLRKSIIAAPFLRYQDIKTGTIVKGKVL
Rrp5_H.sap KPGNTHKCRIIDYSQMDELALLSLRTSIIEAQYLRYHDIEPGAVVKGTVL
Rrp5_S.cer TSVSSS--GIELKIFNGQFKASVPPLHISDTRLVYPERK--FKIGSKVKG
Rrp5_C.gla LKASSS--GVELKLLGGQFTAFVPPLHISDIRLIYPERK--FKIASKTKC
Rrp5_C.alb LKVLEDGVGINVKFFD-EFKGFVPGNQMSDIKLVYPERK--FRVGTKTKG
Rrp5_S.pom KLIPQG----IVVTISEGINGLVPSTHMADIALQFPERR--FKVGSSVKC
Rrp5_X.lae CMEAVG----MVVQITDHLTGLVPKLHFADVLLKHPEKK--YIIGNKIKC
Rrp5_C.ele HVAKSG----VYFMVCNFVKAFAPLSMLSDKPLPAQKMKNIYKVGTEVKC
Rrp5_D.mel KKDDK---------IKGWSVRIGKVNGILEQFYLAPNVR--YDVGQSLKC
Rrp5_R.nor ALKPFG----MLVKVGEQMRGLVPSMHLADIMMKNPEKK--FNTGDEVKC
Rrp5_H.sap TIKSYG----MLVKVGEQMRGLVPPMHLADILMKNPEKK--YHIGDEVKC
S5
Rrp5_S.cer RVISVNSR-GNVHVTLKKSLVNIEDNELPLVSTYENAKNIKEKNEKTLAT
Rrp5_C.gla RILNVDNH-GHIIATMKKSLVNDEELEKPVIESFETAKSIKNKNEKTVGT
Rrp5_C.alb RLLNYNGK--RALVTFRKALVNLEDDE--ILSDIDQAEIG----FKTNAI
Rrp5_S.pom RVLSTNVLRKRVLLTLKKSLLN---TDLPLIYDYEQATPG----TQTVGT
Rrp5_X.lae KVLTVVTCERKLILTRKRTLMKS---TLPVLASYEDAQPG----LITHGF
Rrp5_C.ele RVWQICDERKNLIVTCRESILGLK---SPSVNSVQELEIG----VTVPCV
Rrp5_D.mel RVLEVNAERKICYVSNRAEYLGKG---IKILTDYASAHVG----NVYMGT
Rrp5_R.nor RVLLCDPEAKKLIMTLKKTLVTS---KLPAITCYEDAKPG----LQTHGV
Rrp5_H.sap RVLLCDPEAKKLMMTLKKTLIES---KLPVITCYADAKPG----LQTHGF
Rrp5_S.cer IQVFKPNGCIIS-FFGGLSGFLPNSEI-SEVFVKRPEEHLRLGQTVIVKL
Rrp5_C.gla VQSFNGHGCVIM-FFGGVTGFLPKSEV-SEVFVKRAEDHLRLGQTVQVKV
Rrp5_C.alb VEKFVPNGCIVS-FFGNLKAFLPKTEI-SETFVQDASSYLKIGQIVKVRI
Rrp5_S.pom LARIFEDGAIVE-FYNSVRAFLPVSEM-SEAYIRDAREHFKVGQTLSVTI
Rrp5_X.lae IVAIKDYGCLVK-FYNEVQGLAPRRELGSLQEISSLEDAFYRGQVIKVQV
Rrp5_C.ele IRKVFPTGVLLLGTFNNICGVLRKESAVHLPGTPKNND-FVVANVEKIED
Rrp5_D.mel VVRCEDTYVLVK-FGNGIKGVLHRQNL-------KENSSFFEGQTTKFRI
Rrp5_R.nor IIRVKDYGCLVK-FYNDVQGLVPKHEL-SAQHIPDPERVFYTGQVVKVAV
Rrp5_H.sap IIRVKDYGCIVK-FYNNVQGLVPKHEL-STEYIPDPERVFYTGQVVKVVV


S6
S7
Rrp5_S.cer LDVDADRRRIIATCKVSNEQAAQQKDTIENIVPG-----RTIITVHVIEK
Rrp5_C.gla LEVDEERRRIIVTCKVSNEEAQQQKSIIESLKIG-----QSIIETVVVEK
Rrp5_C.alb LDVNKEQKRLVVTLKQSSELSNAQKNEISKLVSG-----KSIVKTVVVEK
Rrp5_S.pom VSCDPENRKMRVGCREQSWDAKRLER-FENIKAG------SVLSGIVLQK
Rrp5_X.lae LECNPQTQQLLLSFRITEEGHTEQEQRFLKKIKAVKLDVGKLVDIRVLSK
Rrp5_C.ele NRVVFVLRDVNPSAAVANGQNSEKKLVQRPNPADGGISIGKIYKGSLCAK
Rrp5_D.mel LTRNKDQITLTLPEDKFQLGEICPVEITNALDAG-----LEVKITFAAED
Rrp5_R.nor LNCEPSKERMLLSFKLLSGSEPKDESVKNSKKKGSTVNTGQLVDVKVLEK
Rrp5_H.sap LNCEPSKERMLLSFKLSSDPEPKKEPAGHSQKKGKAINIGQLVDVKVLEK
Rrp5_S.cer TKDSVIVEIPDVGLRGVIY--VGHLSDSRIEQNRAQLKKLRIGTELTGLV
Rrp5_C.gla TKDSVIVEIPDVGIRGVIY--VGHISDERIEQCRAEIKKIRIGSKLTGLV
Rrp5_C.alb TKDSVLVELEGSHLRGVIY--DGQLSDGNYEQNRALAKRLAIGETLEVLI
Rrp5_S.pom TEDSVIVDLGDK-VTGVIT--LGQLCDGDLNKCSKVMNKLRASTKLAEVL
Rrp5_X.lae TDKGVNVLVLPEESPAFIP--KMHLSDHVSN-CELLWHTLQEGDDIPGAM
Rrp5_C.ele AGEKANVTFTGEGKKEVYASVDDHLLSDLLDAPIGLTKRLLIENKEDIST
Rrp5_D.mel DEEDEDGNPKLEEFVGLIP--LRLLSDHLELLHAQMRVHPAGSYTDAACI
Rrp5_R.nor TKNGLEVAILPHNIPAFLP--TPHLSDHVAN-GPLLHHWLQTGDTLHRVL
Rrp5_H.sap TKDGLEVAVLPHNIRAFLP--TSHLSDHVAN-GPLLHHWLQAGDILHRVL
Rrp5_S.cer ID-KDTRTRVFNMSLKSSLIKDAKKETLPLTYDDVKDLNKDVPMHAYIKS
Rrp5_C.gla ID-KDSRTQIFNLSLKKSLMKDAQNKRLPTTFAEITKFKKTDPLHGYVKS
Rrp5_C.alb LD-KDLKARTVIATAKKSLIDASKSKSFPVEFSDIAVN---DVVRGYVKS
Rrp5_S.pom VLRKDTSKKLISLSLKKSLVEAAKENRMPINITDLKEG---IKYFGFVRN
Rrp5_X.lae CL--SSLKGHSILTKKSLLISSVEKGSCVKVISEVQTG---MHLTGFVKS
Rrp5_C.ele IVPMGKMAAINRACVKRSVASFVKGMKLPKKIDELKVG---KVIVGIVGQ
Rrp5_D.mel MQ------NIFSLRDVPYFSGQLTKDWQSVQVGDIIRS----YVKHATDQ
Rrp5_R.nor CL--SQSEKHILLCRKPALVSTVEGGQDPKSLSEIQPG---MLLIGFVKS
Rrp5_H.sap CL--SQSEGRVLLCRKPALVSTVEGGQDPKNFSEIHPG---MLLIGFVKS
Rrp5_S.cer ISD-KGLFVAFNGKFIGLVLPSYAVDSRDIDISKAFYINQSVTVYLLRTD
Rrp5_C.gla ISS-TGVFVAFTGKFVGLVLPSYAVESRQVDIEKAFYSNQSVTAYLLRTD
Rrp5_C.alb VTS-LGLFVTFTGRLTGLILAKYVTKNANEDLSKKFHKYQSINCRVLSVD
Rrp5_S.pom ATT-FGVFVEFCDGLVALVPKAYISEEYVPVPSAVYKPQQSVTCVCLSVE
Rrp5_X.lae IMP-YGVFVEFPYGLFGLVPMSEISDKFVTNIRDHFVEGNTVVATVIKMD
Rrp5_C.ele VITNVGVFVELVGGSG--LVGKVLERKNAKNTSELLEVGQVIVGTIESID
Rrp5_D.mel VVD-LMVCVRNYNKPVKVHVKMLRLN-AVKNAPVELVPEQLLWVKVLSKE
Rrp5_R.nor IKD-YGVFVQFPSGLSGLSPKTIMSDKFVTTPSEHFVEGQTVVAKVTNVD
Rrp5_H.sap IKD-YGVFIQFPSGLSGLAPKAIMSDKFVTSTSDHFVEGQTVAAKVTNVD
Rrp5_S.cer DKNQKFLLSLKAPKVKEE----KKKVESNIEDPVDSSIKSWDDLSIGSIV
Rrp5_C.gla DDNERFLLTLKAPKVEKA----AETISA--ENIIDTSIKSVKDIKLGKIL
Rrp5_C.alb KENERFLLTLNN---SSS----FTKDDEQLIKPIDATKKIVADYQPGETT
Rrp5_S.pom LSQEKAFMSFKPLAQKQEKAVEFMESKYDIDNPVDETIKKTYDYVAGKIT
Rrp5_X.lae EEKKRFLLTLKMSECAPD---DYSIEGLFLLSQCFSELQLMKGLLARKGD
Rrp5_C.ele TTKKSFLIDP-CTDLETGELMLKKYALPLLESIVEEVKWLAEQSNYPIPG
Rrp5_D.mel VETKTLTVSAKLTDVWSG--------------------------------
Rrp5_R.nor ESKQRMLLSLRLSDCSLG---DSASTSFLLLCQCLEELQGIRSLMSNQ-D
Rrp5_H.sap EEKQRMLLSLRLSDCGLG---DLAITSLLLLNQCLEELQGVRSLMSNR-D


S8
Rrp5_S.cer KAKIKSVKKNQLNVILAANLHGRVDIAEVFDTYEEIT-------DKKQPL
Rrp5_C.gla DAKIKGVKKNQLNIILADNVHGRVDISEVFDNYNDIK-------DKKHPL
Rrp5_C.alb KVVIKAVKGTQLNVQLADNLQGRVDITQCFKSIKDIK-------NLHQPL
Rrp5_S.pom WAVVTSAKASQLNVDLAANVHGRVDVSEVFDNFGEIV-------DPNKPL
Rrp5_X.lae PEDELSIYTLIAGQKLTLVVENAEENGPVQFSAGSISGAQTVSATQYHIG
Rrp5_C.ele SKVNGKVTKELDDLTLVEFEHNGKKIAGILPKMVDDK--------KTTEK
Rrp5_D.mel -DLSDTAKLVEGYLNEVAQIKAGLEEASAPIS----------
Rrp5_R.nor SVLIQTLADMTPGMVLDAMVQEVLENGSVVFGGGPVP-DLILRASRYHRA
Rrp5_H.sap SVLIQTLAEMTPGMFLDLVVQEVLEDGSVVFSGGPVP-DLVLKASRYHRA
Rrp5_S.cer S-NYKKDDVIKVKIIGNHDVKSHKFLPITHKIS--KASVLELSMKPSELK
Rrp5_C.gla S-HYKANDKVRVKIIGHHDLKSHKSLPITHNFV--KGTVFELTMKPSQLK
Rrp5_C.alb ASSFQKGETLDVKVIGIHDAKNHTFLPITHRKSN-RTTILELSLNNPQIN
Rrp5_S.pom K-RFHKGDKIRVRVLGIHDSRNHKFLPISHRVS--PKQFLELSVRPSILN
Rrp5_X.lae DKALVAGQKVKAVVLHVDMLTLHVHVSLNQTLLKKKQNVPKMNSSHSADV
Rrp5_C.ele TKKKKATKSFIVVDINHSTNEVVLATPSTEQSRIVAIRRDYLCAVSPDGL
Rrp5_D.mel --KYSVGEKINVVFKGIDATTNDWVYTVEGN----------GKVSALLL
Rrp5_R.nor GQELEPGQKKKVVVLHVDALKLEVHVSLHQDLVNRKARKLRKNSRHQGIV
Rrp5_H.sap GQEVESGQKKKVVILNVDLLKLEVHVSLHQDLVNRKARKLRKGSEHQAIV
S9
Rrp5_S.cer SKEVHTKSLEEINIGQELTGFVNNSSGNHLWLTISPVLKARISLLDLADN
Rrp5_C.gla SKDVKELDIRDITVGDEIIAFVNNYQNSTLWLTVTPTIKAKLSIFDLSEE
Rrp5_C.alb NKLLTDLKLADFKQGDEVLAFVNNITNGMVWVSISPTIKGRIPMLDLTED
Rrp5_S.pom -MEPFSMKEPQFKKGDEVTGFVNNVSKECVWVSLTPSVNGRIPILDLTTD
Rrp5_X.lae QHVAEEFAVVSLADSAHLIAVPVSSHLNDTFRFESEKLKVGETISVILKT
Rrp5_C.ele IYLPTRLHPNHLPTSDSKVKLHTVIDLNDQKSVGEGVFVATRGGGDVDDV
Rrp5_D.mel SSLVGTAKAPEMGSKHEAVILWIEYSSDVLLISNKKLDIAHISPSGELST
Rrp5_R.nor QHLEESFAVASLVETGHLVAFSLISHLNDTFHFDSEKLRVGQGVCLTLKT
Rrp5_H.sap QHLEKSFAIASLVETGHLAAFSLTSHLNDTFRFDSEKLQVGQGVSLTLKT
Rrp5_S.cer DSNFSENIESVFPLGSALQVKVASIDREHGFVNAIG----------
Rrp5_C.gla TLMNIKNVEDDFPLGSVLKVNVTGKDQNKSILQVTQ----------
Rrp5_C.alb GS-IFQDIDNKLPIGNAINVKVKQVDLQHQILVLTA----------
Rrp5_S.pom -VKELNSLQKHFFLGKAIKCYVVNAEDSITLSAIGP----------
Rrp5_X.lae TTVNEHGLLLAVQNKAASKTSKNLGRTTQSA----------
Rrp5_C.ele ARTTLINKEFVGDEKKKKQTAEGSDGAVSTSTKSVKN----------
Rrp5_D.mel N----------
Rrp5_R.nor TEPGVTGLILAVEGPASKRIRMPAQRDSETVDDKGEEEEEEEEED-----
Rrp5_H.sap TEPGVTGLLLAVEGPAAKRTMRPTQKDSETVD------EDEEVDP-----
S10
Rrp5_S.cer -KSHVDINMSTIKVGDELPGRVLKIAEKYVLLDLGNKVTGISFITDALND
Rrp5_C.gla -RNGKINSIDDLKVGGHTIGKIVKVTPKYLLIELENKITGISTALEALND
Rrp5_C.alb -RKNFIEKFEDVRQDQTYPARIIKIKPNHVLVELGNNVIASSFVTDALND
Rrp5_S.pom -LQG----FENLTPGSRLVGKVTNVNEAGAILQLPGHMSGRVSRIDMFDD
Rrp5_X.lae -MRVRGAISHGLKIGDLVTGTVKSIKPTQVTVSINDNVFGFIHVSQIMDE
Rrp5_C.ele -----FGVYSGVVIGTANLEENRKRNSLFADIRLPGDNIGRLHVSELPPS
Rrp5_D.mel ----------LIGKAGMKAKVLLKLESVAVCSLKKGTNPLVICPIRLHP
Rrp5_R.nor -LTVRSKKSHSLAIGDKVTGTIKSVKATHAVVTLDDGFIGCIHVSRILDD
Rrp5_H.sap ALTVGTIKKHTLSIGDMVTGTVKSIKPTHVVVTLEDGIIGCIHASHILDD

```
Rrp5_S.cer   FSLTLKEAFEDKINNVIPTTVLS--------------------------
Rrp5_C.gla   FTEPLDQVFAGKENEFVKAKITS--------------------------
Rrp5_C.alb   YSDKLDHVFN--VNDYVSAKVLD--------------------------
Rrp5_S.pom   YDILPETKFTR--NNLVGVCVLS--------------------------
Rrp5_X.lae   TPQGCFPTSKLNPKQEVTCRVIGGREVKTHRYLPITHPDFIHSVPELSLL
Rrp5_C.ele   LLKAENPLDEFVKRNVRKAVIVRIIG----------------------
Rrp5_D.mel   NDIEN-------------------------------------------
Rrp5_R.nor   VPEGVSPTTTLKAGKKVTARVIGGRDVKTSKFLPISHPRFVLTILELSVR
Rrp5_H.sap   VPEGTSPTTKLKVGKTVTARVIGGRDMKTFKYLPISHPRFVRTIPELSVR

Rrp5_S.cer   ------------------------------------------------
Rrp5_C.gla   ------------------------------------------------
Rrp5_C.alb   ------------------------------------------------
Rrp5_S.pom   ------------------------------------------------
Rrp5_X.lae   P--------ELINTDNVPKPRALKTFNPGDKVTCYVNKFNTETKYLEVEI
Rrp5_C.ele   ----------------------------------------------FIKSS
Rrp5_D.mel   ------------------------------------------------
Rrp5_R.nor   PSELKGD-YRALNTHSVSPMEKIRQYQAGQTVTCFLKKYNVMKKWLEVDI
Rrp5_H.sap   PSEL-EDGHTALNTHSVSPMEKIKQYQAGQTVTCFLKKYNVVKKWLEVEI

Rrp5_S.cer   -------------------VDEQNKKIELSLRPATA-----------KTR
Rrp5_C.gla   -------------------VIADEKKVQLQLAEDES-----------TVQ
Rrp5_C.alb   -------------------IDSTNKRISVSLRTDKAG----------SDK
Rrp5_S.pom   -------------------VDVPNRKVALSARNSRTQSQPV----EIKDK
Rrp5_X.lae   TPEIRGRIELLLLSQTPKNLKRPEKLFKNGQALSATVVGPD-AVHKHLCL
Rrp5_C.ele   KGPRIAELTMIPSKIQAGKVRASNLSYKSNYKIGDVVKCFG-----ANTM
Rrp5_D.mel   ------------------------------------------------
Rrp5_R.nor   EPDIRGRIPLLLTSLSFKVLKHPDKKFQIGQAIKATVVGPD-IPRAFLCL
Rrp5_H.sap   APDIRGRIPLLLTSLSFKVLKHPDKKFRVGQALRATVVGPD-SSKTLLCL

                                  S11
Rrp5_S.cer   SIKSHEDLKQGEIVDG-IVKNVNDKGIFVYLSRK------------VEAF
Rrp5_C.gla   KITSHSDLKVGEVVNG-LVKTVTDKGLFVFLGKS------------VEAF
Rrp5_C.alb   VINSISDLTRGQVIKG-FVKNISNNGVYVSLGRS-------------IYAL
Rrp5_S.pom   EINSVDDLKIGDICRG-FVCNVANQGLFVTIGHN------------LIAR
Rrp5_X.lae   SLTGIHSLEEGAVTVACVTKVVKGSGLTLSLPFG------------KTGN
Rrp5_C.ele   TEKQELKVEVNPVWTGTISRENLSDDLKVTAADGGIVDLALKKGEMREAK
Rrp5_D.mel   --SGSAELRQGDFCNIAFIHDKLHIAVPETVWRL------------WRGV
Rrp5_R.nor   SLIGPYKLEEGDVAMGRVLKVLPNKGLTVSFPFG------------RIGK
Rrp5_H.sap   SLTGPHKLEEGEVAMGRVVKVTPNEGLTVSFPFG------------KIGT
                    :  . .    .        :                    .

Rrp5_S.cer   VPVSKLSDSYLKEWKKFYKPMQYVLGKVVTCD----EDS-RISLTLRESE
Rrp5_C.gla   VPVSKLSDSYLKEWKKFYKPMQPVVGKIVSCN----EDD-RILLTLRETE
Rrp5_C.alb   VRVSDLSDSYLKDWQNFFKPNQPVIGKIVNCK----QEG-RILMTLKESE
Rrp5_S.pom   VKIGELFDTFIKDWKPHFHVNQLVKGSIVGID----NDSKRIEMSLKQSK
Rrp5_X.lae   ANMFHLCDKYAEASLEKFTPGKFVRCAILSNSK-IVKVSLRQSRVNQQAQ
Rrp5_C.ele   ITAVDRKNMSLKLTLDTAEEAFEIGASVTGRVFFVSKTYIRLKLSSGQQA
Rrp5_D.mel   KRTAGTEVAPVKAKKAKVEESPKQKKTTIEETS-----AQKKVKAKAKVE
Rrp5_R.nor   VSVFHLSDSYSEEPLTDFCPQKIVRCYILSTAHRMLALSLRSSRTNKETK
Rrp5_H.sap   VSIFHMSDSYSETPLEDFVPQKVVRCYILSTADNVLTLSLRSSRTNPETK
                       :                                 :     :
```

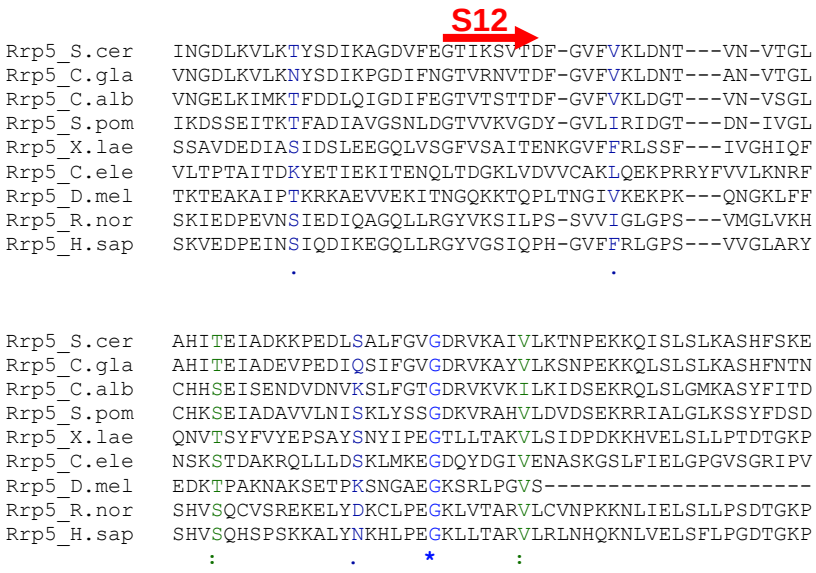


Figure 2.9. Sequence alignment of *S. cerevisiae* Rrp5 and its homologs in *C. glabrata*, *C. albicans*, *S. pombe*, *X. laevis*, *C. elegans*, *D. melanogaster*, *R. norvegicus* and *H. sapiens*. Universally conserved, strongly conserved and weakly conserved residues are labeled with (*), (:) and (.), respectively. S1 motifs in Rrp5 were identified via sequence alignments with the S1 domains of the *E. coli* S1 protein (data not shown) and are indicated.

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 ((Lamanna and Karbstein, 2010); see Figure 6C). This interaction prevents the formation of the decoding site and inhibits premature Nob1-dependent cleavage at site D. Cleavage at site A_2 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 A_2 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 of the inhibitory duplex (Figure 2.7). As described above, the observation that the binding affinities for Rrp5C parallel those for full length Rrp5 indicate 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, A_2 cleavage can be bypassed for the production of 18S rRNA (Torchet and Hermann-Le Denmat, 2000; Vos et al., 2004b). Thus, the three C-

terminal S1 domains are necessary to enforce the requirement for A_2 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-A_2 cleavage structure. This would explain the requirement of Rrp5 for cleavage at site A_2 (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). Most recently, Matsumoto and co-workers reported the structure of a bacterial virulence factor, CvfB, which contains three S1 domains and a winged helix (WH) domain (Matsumoto et al., 2010). 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 between multiple S1 domains in Rrp5, we find that the RNA binding affinities for Rrp5N and Rrp5C are non-additive: Rrp5FL binds RNA more strongly than either Rrp5N or Rrp5C alone, but not as strong as expected if these fragments acted entirely independently (Table 2), 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 non-specifically, 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 β2 and β3) provide rather non-specific interaction surfaces, while the loops (often those between the second and third β-sheets, and the third beta 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 are likely interacting via 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 A_3, while Rrp5C is required for cleavage at the more 5' site A_2 (Figure 2.1A). 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 $S1_{10\text{-}12}$, whereby $S1_{10}$ would bind the stem loop near 5.8S rRNA, and $S1_{12}$ could bind between sites D and A_2.

Comparison to other RNA binding proteins with multiple RBDs

Most defined RNA-binding domains recognize only between four and eight nucleotides at a time, providing for little specificity and affinity (Lunde et al., 2007). Thus, many proteins contain multiple RNA-binding domains, either as different modules or multiple copies of the same module[1] (Lunde et al., 2007). General considerations suggest that such combinations should allow for a large expansion of both the specificity

[1] Ribosomal 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.

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 (Diaz-Moreno et al., 2010; Garcia-Mayoral et al., 2007). Furthermore, the increase is completely additive in the case of KH3 and KH4, which are linked by a flexible linker, while the increase appears sub-additive in the case of KH2 and KH3 (Garcia-Mayoral et al., 2008; Garcia-Mayoral et al., 2007). 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 ten-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 finger IIIA have shown that distinct fingers provide for most of the interactions with RNA relative to DNA (Clemens et al., 1993; Liao et al., 1992). 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).

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, $S1_{10\text{-}12}$ (contained in Rrp5C) provide all of the specificity in the recognition of pre-rRNA, while $S1_{1\text{-}9}$ (contained in Rrp5N) provide 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 five-fold loss in RNA binding affinity.

Materials and Methods

Cloning of Rrp5 Constructs. All three proteins were amplified from genomic DNA. Rrp5C (amino acids 1,083 – 1,729) was cloned between the SfoI and BamHI sites of pSV272; Rrp5N (amino acids 1 – 1,082) and Rrp5FL (amino acids 1 – 1,729) were cloned between the NcoI and BamHI sites of a pET23-derivative containing a TEV-cleavable His_6-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 OD_{600} of ~0.6 before inducing with 1 mM IPTG at 30°C for 5 hours. Cells were resuspended in lysis buffer (containing 0.1 mM PMSF and 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 His_6-MBP tag. Rrp5C was further purified over a MonoS column in a linear gradient from 30 mM to 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 $M^{-1}cm^{-1}$.

Rrp5N-containing fractions were dialyzed for 5 hours in 150 mM KCl, 50 mM Tris, pH 7.8, 1 mM TCEP and 1 mM DTT. Rrp5FL-containing fractions were dialyzed for 3 hours 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 mM to 660 mM KCl; Rrp5FL was separated

over a 12 column volume linear gradient from 100 mM to 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 $M^{-1}cm^{-1}$ and 111,100 $M^{-1}cm^{-1}$ 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 ^{32}P-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 Mg^{2+} 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, pre-folded rRNA and Rrp5 were incubated together at 30°C for 2 hours before being loaded on a 6% acrylamide/THEM (Tris, HEPES, EDTA, pH 7.5, $MgCl_2$) gel (Karbstein et al., 2002) at 4°C for 3 hours. Rrp5-bound and unbound fractions were quantified using phosphoimager software, and data was fit to Eq. 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 Rrp5's lower pI value of 4.9 (compared to 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'ITS1-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 Eq. 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:

$$K_{1/2} = \frac{[\mathrm{Rrp5}]^2 \times [\mathrm{rRNA}]}{[\mathrm{Rrp5_2 \cdot rRNA}]} \qquad \text{Eq. 1}$$

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

$$\mathrm{fraction_{bound}} = \frac{[\mathrm{Rrp5_2 \cdot rRNA}]}{[\mathrm{Rrp5_2 \cdot rRNA}] + [\mathrm{rRNA}]} \qquad \text{Eq. 2}$$

Solving Eq. 2 for [$Rrp5_2$•rRNA] and substituting this into Eq. 1 gives Eq. 3:

$$\mathrm{fraction_{bound}} = \frac{\mathrm{fraction_{bound,max}}[\mathrm{Rrp5}]^2}{[\mathrm{Rrp5}]^2 + K_{1/2}} \qquad \text{Eq. 3}$$

Competition Experiments. rRNA was folded as described above. For binding to H44/3'ITS1 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/A_2, 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). 1 μM of unlabeled H45/3'5.8S was folded and incubated in the presence or absence of 6 μM Rrp5FL, exposed to either 1.5% (H45/3'5.8S only) or 2.5% (H45/3'5.8S and Rrp5FL) DMS and then incubated for 3 min. Reactions were quenched with β-mercaptoethanol and RNA was phenol extracted. Reverse transcription and sequencing gels were done as described (Lamanna and Karbstein, 2009).

Chapter III

Rrp5 enhances the specificity and annealing activity of the DEAD-box protein Rok1

Introduction

DEAD-box proteins are RNA-dependent ATPases that are involved in all aspects of RNA metabolism: translation initiation, pre-mRNA splicing, mRNA decay and export and ribosome biogenesis. In these processes, their functions include RNA duplex unwinding, RNA-protein complex remodeling, RNA duplex annealing and ATP-dependent RNA binding (Jankowsky, 2011; Jankowsky and Fairman, 2007; Jarmoskaite and Russell, 2011; Linder, 2006).

Despite their non-redundant and therefore substrate-specific functions *in vivo*, to date, only DbpA, a DEAD-box protein in involved in bacterial ribosome assembly (Fuller-Pace et al., 1993; Iost and Dreyfus, 2006; Sharpe Elles et al., 2009), has been shown to have sequence specific ATPase activity (Fuller-Pace et al., 1993; Tsu and Uhlenbeck, 1998), which arises from unique sequences in the C-terminus (Kossen et al., 2002; Wang et al., 2006). Such sequence specificity has not been demonstrated for any other helicase and is consistent with the highly conserved helicase core, which includes an RNA-binding site that primarily contacts the sugar-phosphate backbone (Schutz et al., 2010; Sengoku et al., 2006). It has therefore been postulated that RNA-binding co-factors confer specificity for certain RNA sequences. Co-factors have been identified for a number of RNA helicases, for which they regulate the catalytic activity. Nevertheless, co-factor effects on specificity have not been demonstrated.

The three best studied DEAD-box proteins, Ded1, eIF4A and Mss116 display RNA-dependent ATPase and RNA duplex unwinding activities that are not sequence

specific (Halls et al., 2007; Iost et al., 1999; Ozes et al., 2011; Schutz et al., 2008; Schutz et al., 2010). This is consistent with the expectation that translation, a process regulated by Ded1 and eIF4A, is independent of RNA sequence. These DEAD-box proteins may alternatively recognize particular RNA structures (Fairman et al., 2004) in order to carry out their essential *in vivo* roles. Furthermore, Gle1, a binding partner of Ded1, inhibits its activity and therefore renders this system unsuitable for testing co-factor effects on substrate specificity (Bolger and Wente, 2011). In contrast, DEAD-box proteins involved in cellular processes such as ribosome biogenesis and pre-mRNA splicing are likely to encounter specific substrates. As a result, these enzymes may be more appropriate targets for understanding how DEAD-box proteins achieve specificity. However, co-factors have not been identified for Mss116, a pre-mRNA splicing factor, and therefore co-factor effects on specificity have not been determined. Consequently, until now, no experimental evidence has supported the hypothesis that co-factors enhance DEAD-box protein specificity.

Here, we present a biochemical characterization of the DEAD-box protein Rok1 and its co-factor Rrp5, both of which are required for nucleolar 40S ribosome maturation (Venema et al., 1997; Venema and Tollervey, 1996). *In vitro* pull-down experiments verify that Rrp5 directly binds Rok1. We recently reported that the binding site for Rrp5 is located near a pre-rRNA duplex that undergoes a conformational change to regulate cleavage steps during ribosome assembly (Lamanna and Karbstein, 2011; Young and Karbstein, 2011). Using RNA strands that mimic the duplex near the Rrp5 binding site, we show that Rok1 binds the double-stranded duplex ~30-fold stronger than either of the single-stranded substrates. We also demonstrate that Rok1 lacks RNA unwinding activity but stabilizes duplexes and catalyzes the reverse reaction, strand annealing. While Rok1 alone binds and anneals the test duplex with moderate specificity relative to other duplexes, the addition of Rrp5 enhances Rok1 annealing of the wild-type duplex ~15-fold while essentially not affecting control duplexes. Rrp5 therefore enhances Rok1's specificity for its target. This finding expands the role of 'helicase' co-factors from regulators to specificity factors.

Results

Rok1 directly binds Rrp5

In vitro studies have shown that co-factors can regulate the activities of RNA helicases ((Young and Karbstein, 2012) and references therein). For the DEAD-box protein Rok1, previous data implicate the RNA-binding protein Rrp5 as a likely co-factor: Rok1 overexpression suppresses the temperature-sensitive phenotype of a mutation in the first tetratricopeptide (TPR) motif of Rrp5 (Torchet et al., 1998). Furthermore, Rok1 requires the presence of Rrp5 in order to associate with the pre-rRNA during ribosome assembly (Vos et al., 2004a). To verify that Rrp5 binds Rok1, as expected from a co-factor, we first carried out an *in vitro* pull-down assay. Rrp5 is a 192 kDa protein with a large number of repeating motifs: 12 S1 RNA-binding domains towards its N-terminus, seven TPR motifs (known to be involved in protein-protein interactions (Lamb et al., 1995)) towards its C-terminus. Since the N- and C-termini of Rrp5 function independently (Eppens et al., 1999; Torchet et al., 1998; Young and Karbstein, 2011) and the C-terminus contains the TPR motifs, we used the C-terminal fragment of Rrp5. This fragment, together with the N-terminal piece, fully complements the Rrp5 deletion (Young and Karbstein, 2011). Additionally, this truncation binds specifically but weakly to rRNA mimics containing the Rok1 binding site (Young and Karbstein, 2011) and is also more stable and therefore easier to prepare in larger quantities in comparison to the full-length protein. We therefore used the C-terminus of Rrp5, containing the last three S1 RNA-binding motifs and all seven TPR motifs (Young and Karbstein, 2011), to test for a direct interaction with Rok1

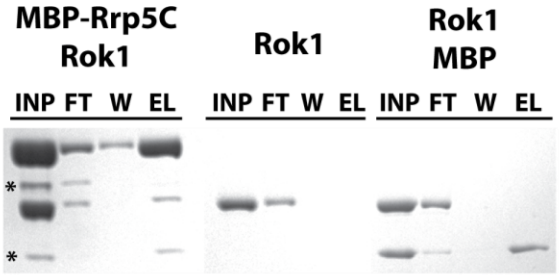


Figure 3.1. Rok1 binds (MBP-)Rrp5C. Rok1 binds an MBP-tagged C-terminal fragment of Rrp5 that contains the TPR motifs and three S1 RNA-binding domains (Eppens et al., 1999; Torchet et al., 1998; Young and Karbstein, 2011). Control experiments indicate that Rok1 does not stick to the amylose resin or the MBP tag alone. (*) indicates contaminating MBP and Rrp5C from the (MBP)-Rrp5C protein preparation.

in *in vitro* pull-down assays. Recombinant Rok1 elutes from amylose beads in the presence of (MBP)-Rrp5C, but not in the presence of MBP alone (Figure 3.1), indicating that Rok1 directly binds Rrp5C.

Rok1 specifically binds the pre-A_2 duplex

Before testing the effects of the co-factor Rrp5 on Rok1 activity, we first carried out an *in vitro* biochemical characterization of Rok1 alone.

Consistent with previous Northern analysis of rRNA processing phenotypes during ribosome maturation (Eppens et al., 2002; Venema and Tollervey, 1996), we have recently shown that Rrp5 binds within ITS1, the intron-like sequence between 18S and 5.8S sequences in rRNA precursors (Young and Karbstein, 2011). Within this sequence, Rrp5 binds adjacent to a duplex that is known to be required for ordering cleavage at sites A_2 (separates the pre-40S rRNA from pre-rRNAs destined for the 60S subunit) and D (produces the mature 3'-end of 18S rRNA) (Figure 3.2; (Lamanna and Karbstein, 2011;

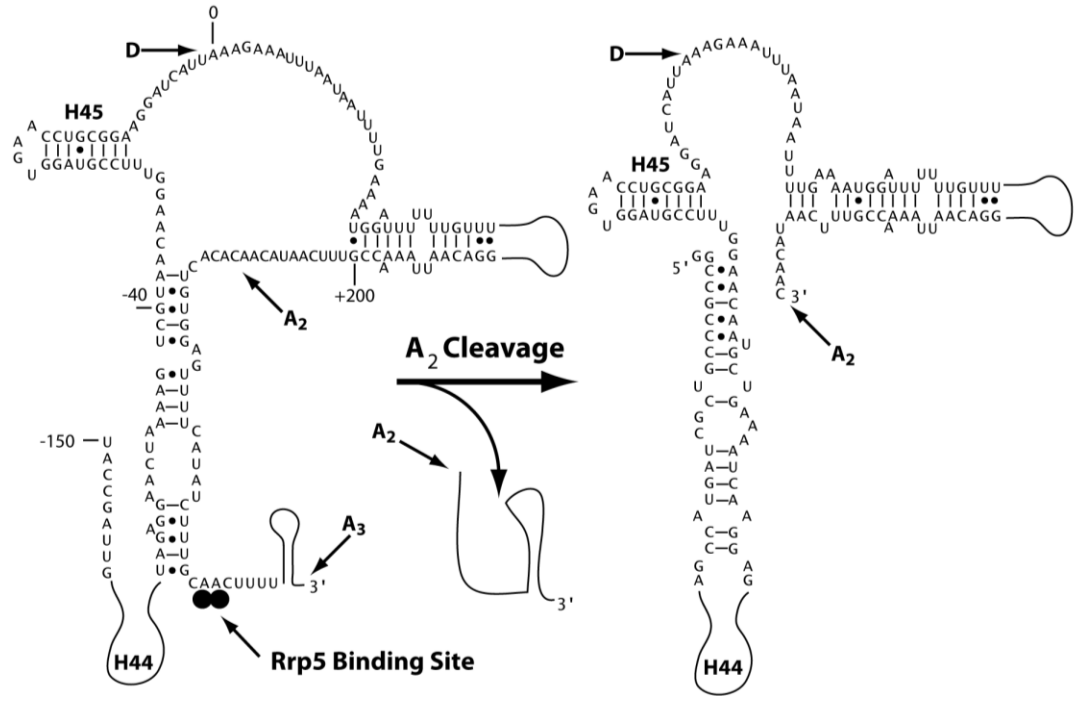


Figure 3.2. An essential conformational change regulates cleavage at sites A_2 and D (Lamanna and Karbstein, 2011). Prior to cleavage at A_2, sequences in helix 44 (H44) form an inhibitory duplex with sequences downstream of the A_2 cleavage site. After cleavage at A_2, this duplex dissociates and is replaced with the decoding site H44. This latter conformation is preferred for cleavage at site D to produce the mature 3' end of 18S rRNA. Our recent characterization of Rrp5 interactions with the pre-rRNA identified two Rrp5 binding sites (●) at the base of the inhibitory duplex that orders these cleavage steps (Young and Karbstein, 2011).

Young and Karbstein, 2011)). Early in assembly, this duplex is formed between one strand of helix 44 (H44; the decoding site helix in mature 18S rRNA) and sequences downstream of the A_2 cleavage site. Later in assembly, this duplex dissociates to be replaced by the mature decoding site H44. Since its secondary structure was rearranged during assembly, and because of both its proximity to the Rrp5 binding site and our results that validated Rrp5 as an interacting partner of Rok1, we postulated that Rok1 is involved in changes in the conformation of this helix. We therefore designed our RNA substrate to encompass this pre-A_2 inhibitory duplex. RNA oligonucleotides were designed to mimic either the strand in H44 that is involved in the duplex interaction (H44) or the strand immediately downstream of the A_2 cleavage site (DownA_2); to increase the thermodynamic stability of the duplex, four alternating G-C base pairs were added to the base of the duplex (Figure 3.3A).

We used gel-shift analysis with ^{32}P-labeled RNA substrates to first test for Rok1 binding to the wild-type pre-A_2 duplex versus each of the single-stranded RNAs, H44 or DownA_2. As shown in Figure 3.3B, Rok1 binds the duplex ~20-fold more tightly than

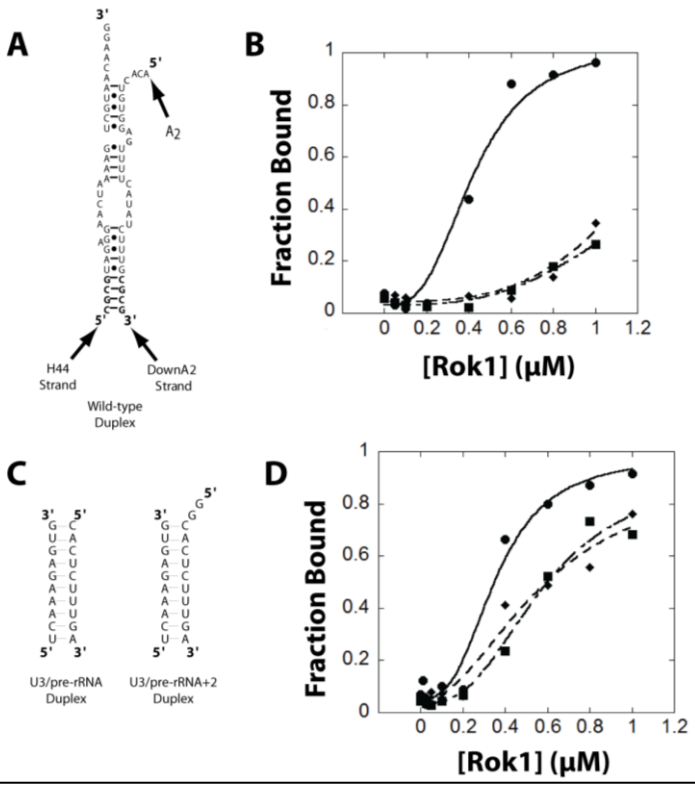


Figure 3.3. Rok1 specifically binds the pre-A_2 duplex. A. Wild-type duplex is based upon Rok1's predicted *in vivo* substrate. Four alternating G-C base pairs (bold) were added to the base of the duplex to improve its thermodynamic stability. B. Rok1 binds the wild-type duplex (●; 0.5 nM ^{32}P-DownA_2, 20 nM H44) with a higher affinity than either the DownA2 (■) or H44 (♦) single-stranded RNAs (0.5 nM ^{32}P-DownA_2 or ^{32}P-H44). Fitting to Eq. 6 gives $K_{1/2}$ values of 0.17 μM, 3.6 μM and 3.4 μM for the wild-type duplex, DownA_2 and H44 strands, respectively. C. Control duplexes include regions from U3 snoRNA that form a duplex with sequences within pre-18S rRNA. D. Rok1 has moderately increased affinity for the wild-type duplex over both control duplexes (0.5 nM radiolabeled U3, 5 nM pre-rRNA or 15 nM pre-rRNA+2). Fitting to Eq. 3 gives $K_{1/2}$ values of 0.16 μM, 0.46 μM and 0.35 μM for the wild-type, U3/pre-rRNA and U3/pre-rRNA+2 duplexes, respectively.

either of the single-stranded RNAs ($K_{1/2}$ is 0.17 µM for the duplex, 3.6 µM and 3.4 µM for the DownA_2 and H44 strands, respectively). These results are surprising as they differ from those observed with other well-studied DEAD-box proteins such as eIF4A, Dbp5 and Ded1. All of these indicate a strong preference for single-stranded RNA over double-stranded RNA (Banroques et al., 2008; Lorsch and Herschlag, 1998; Weirich et al., 2006). Interestingly, this preference for single-stranded regions is thought to stimulate strand separation by DEAD-box proteins ((Tijerina et al., 2006; Yang and Jankowsky, 2006); reviewed in (Jankowsky, 2011; Jarmoskaite and Russell, 2011)).

To confirm that the pre-A_2 duplex studied here is indeed a specific target for Rok1, we next compared Rok1's affinity for control duplexes that contain regions from U3 snoRNA that base pair with sequences within pre-18S rRNA (Figure 3.3C); these are 10 base pair duplexes with either a blunt end (U3/pre-rRNA) or a two nucleotide extension (U3/pre-rRNA+2). Figure 3.3D indicates that Rok1 has moderate preference for the wild-type duplex compared to other control duplexes, with $K_{1/2}$ values of 0.16 µM, 0.46 µM and 0.35 µM for the wild-type, U3/pre-rRNA and U3/pre-rRNA+2 duplexes, respectively. While a 2-3-fold preference for any given substrate is not impressive, it should be pointed out that such specificity has only been observed for one other DEAD-box protein, DbpA, which interacts with hairpin 92 of 23S rRNA 1000-fold more specifically than other control RNAs (Tsu and Uhlenbeck, 1998). Lack of strong sequence specificity is also consistent with structures of DEAD-box proteins bound to RNA, as interactions are primarily made with the sugar-phosphate backbone (Schutz et al., 2010; Sengoku et al., 2006).

Rok1 has RNA-independent ATPase activity

We next analyzed Rok1's ATPase activity under single-turnover conditions using trace ^{32}P-γ-ATP and TLC analysis to monitor the production of $^{32}P_i$. Unlike most DEAD-box proteins, Rok1 has RNA and nucleotide-independent ATPase activity ((Linder et al., 2001); Figure 3.4A). These findings are in agreement with those observed for Rok1 in prior *in vitro* ATPase experiments (Garcia and Uhlenbeck, 2008; Oh and Kim, 1999). The Walker A mutant K172A completely abolishes this activity, verifying that the

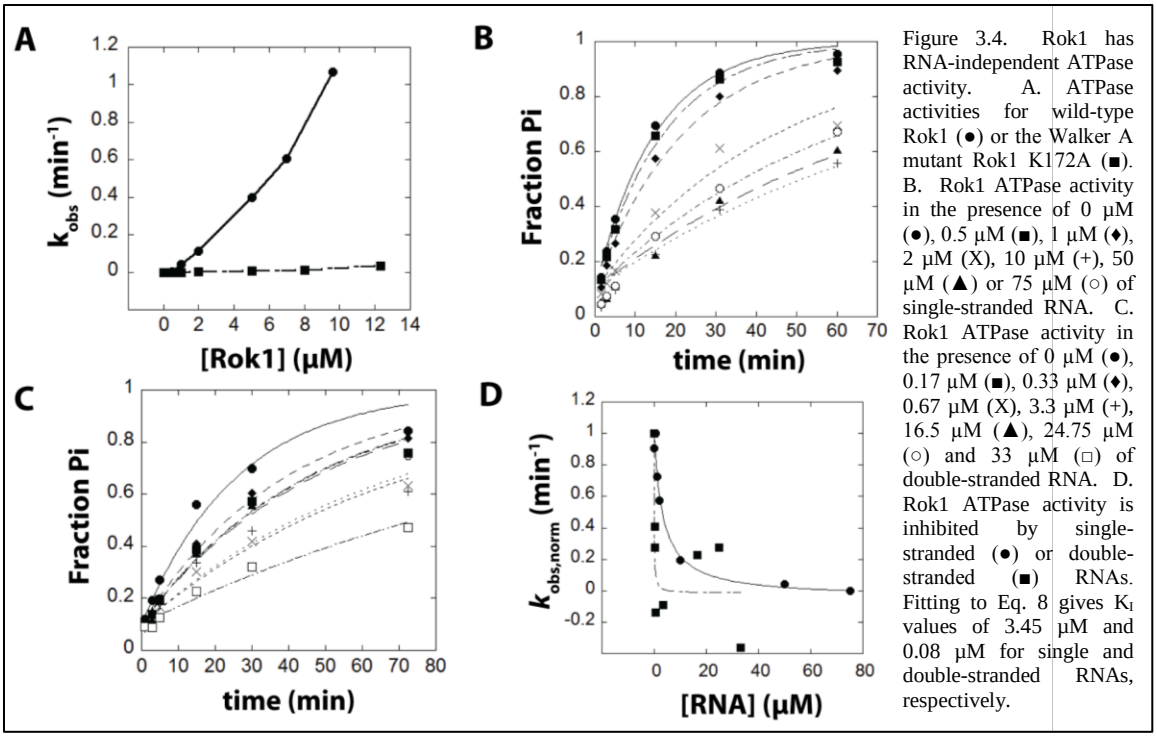


Figure 3.4. Rok1 has RNA-independent ATPase activity. A. ATPase activities for wild-type Rok1 (●) or the Walker A mutant Rok1 K172A (■). B. Rok1 ATPase activity in the presence of 0 μM (●), 0.5 μM (■), 1 μM (♦), 2 μM (X), 10 μM (+), 50 μM (▲) or 75 μM (○) of single-stranded RNA. C. Rok1 ATPase activity in the presence of 0 μM (●), 0.17 μM (■), 0.33 μM (♦), 0.67 μM (X), 3.3 μM (+), 16.5 μM (▲), 24.75 μM (○) and 33 μM (□) of double-stranded RNA. D. Rok1 ATPase activity is inhibited by single-stranded (●) or double-stranded (■) RNAs. Fitting to Eq. 8 gives K_I values of 3.45 μM and 0.08 μM for single and double-stranded RNAs, respectively.

observed ATPase activity is indeed coming from wild-type Rok1 and not a contaminating ATPase (Figure 3.4A). Furthermore, RNA inhibits Rok1 ATPase activity in a concentration dependent manner (Figures 3.4B and 3.4C, respectively). In their study of DEAD-box proteins, Garcia and Uhlenbeck observed similar atypical results: unlike the other DEAD-box proteins they were analyzing, Rok1 had little ATPase stimulation by RNAs that ranged from 10-25 nucleotides in length (Garcia and Uhlenbeck, 2008).

As shown in Figure 3.4D, double-stranded RNA binds and inhibits Rok1 ATPase activity ~40-fold more tightly than single-stranded RNA (K_I values of 0.08 μM and 3.45 μM, respectively). These results are quantitatively consistent with those observed in the gel-shift binding experiments and further demonstrate Rok1's preference for double-stranded RNA.

Rok1 stabilizes the pre-A_2 duplex

Since Rok1 preferentially binds double-stranded RNA, Rok1 binding to RNA must similarly stabilize duplex formation. To test this prediction, we measured duplex formation at increasing H44 concentrations in the absence and presence of 400 nM Rok1 using gel-shift analysis (Figure 3.5; $K_{1/2}$ values of 65.6 nM and 2.4 nM for the absence and presence of Rok1, respectively). These results indicate that the presence of Rok1 indeed stabilizes the pre-A_2 duplex ~30-fold. Again, this enhancement agrees quantitatively with the 20-40-fold stronger binding of Rok1 to double-stranded RNA observed in the gel-shift and ATPase experiments and therefore demonstrates the robustness of our system.

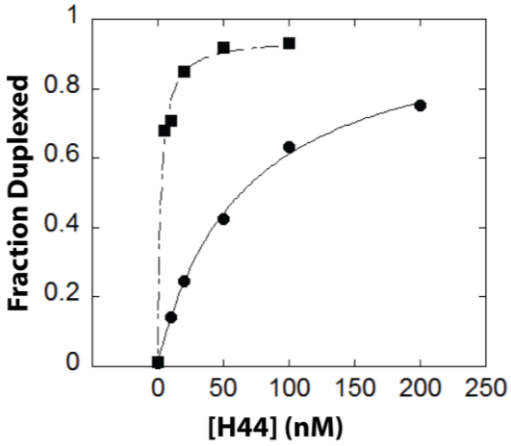


Figure 3.5. Stability of the wild-type duplex in the presence and absence of Rok1. The fraction of duplex formed at 0.5 nM ^{32}P-DownA_2 and varying H44 concentrations were quantified from gel-shift experiments. Observed $K_{1/2}$ values are 65.6 nM and 2.4 nM in the absence (●) and presence (■) of 400 nM Rok1, respectively.

Rok1 promotes annealing but not unwinding

DEAD-box proteins are classically described as RNA helicases as some family members have the ability to unwind a short RNA duplex. Nevertheless, only a few DEAD-box proteins have demonstrated unwinding activity (Bizebard et al., 2004; Halls et al., 2007; Iost et al., 1999; Kikuma et al., 2004; Lawson et al., 1989; Weirich et al., 2006). In contrast, some have the ability to promote duplex formation, dissociate RNA-protein complexes or function as ATP-dependent RNA-binding proteins (Jankowsky and Bowers, 2006; Jankowsky and Fairman, 2007; Le Hir and Andersen, 2008). Two DEAD-box proteins, Ded1 and Mss116, both of which can dissociate duplexes, are also super-annealers: they enhance the rate of duplex formation to that of diffusion (Halls et al., 2007; Yang and Jankowsky, 2005). Considering Rok1's unique preference for double-stranded RNA over single-stranded RNA and its ability to stabilize RNA duplexes, we tested Rok1's annealing activity. Gel-shift analysis indicates that Rok1

fully anneals the pre-A_2 wild-type duplex within three minutes (Figure 3.6A), and that this ATP-independent annealing activity is Rok1 concentration dependent (Figure 3.6B).

We next tested Rok1's ability to unwind the wild-type duplex in the presence of ATP/$MgCl_2$. Unlike other characterized DEAD-box proteins, Rok1 has no unwinding activity in either the absence or presence of Rrp5 (Figure 3.6C). This lack of unwinding activity was also observed with a U3/pre-rRNA duplex with a 27 nucleotide single-stranded overhang, even after 10 minutes (data not shown). Rok1's low affinity for single-stranded RNA supports this observation: since Rok1 preferentially binds double-stranded RNA, it promotes duplex formation and does not have the ability to unwind RNA duplexes.

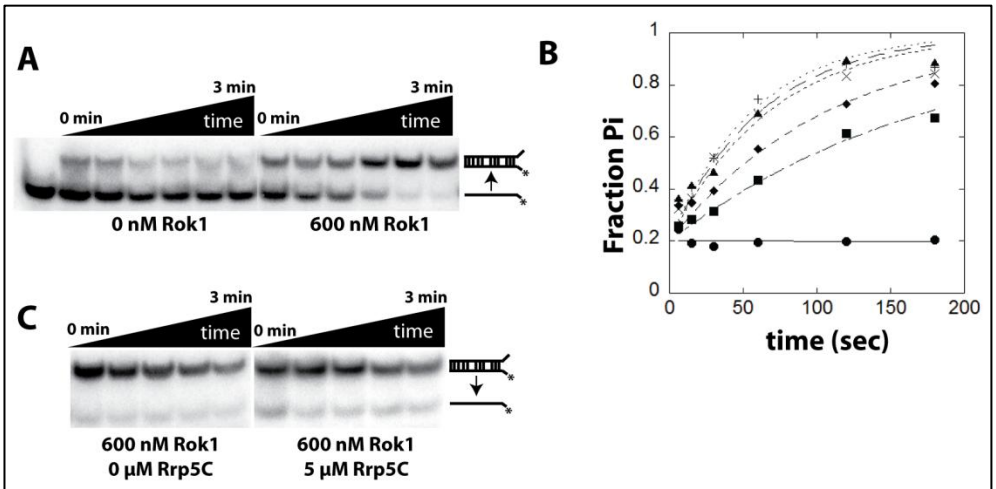


Figure 3.6. Rok1 promotes annealing but not unwinding. A. Gel-shift analysis under protein-denaturing conditions indicates that the presence of 600 nM Rok1 supports complete duplex formation within 3 minutes. To ensure ~20% initial duplex formation in the absence of Rok1, 0.5 nM ^{32}P-DownA_2 and 20 nM H44 were used. B. Annealing activity is Rok1 concentration dependent: 0 nM (●), 50 nM (■), 100 nM (♦), 200 nM (X), 400 nM (+) and 600 nM (▲) Rok1. C. Gel-shift analysis indicates that in the presence of 2 mM ATP/$MgCl_2$ and either 600 nM Rok1 or 600 nM Rok1 and 5 µM Rrp5C, unwinding of the wild-type duplex does not occur. RNA concentrations: 0.5 nM ^{32}P-DownA_2 and 40 nM H44.

Our data so far indicate that Rok1 has the ability to stabilize RNA duplexes thermodynamically. Furthermore, the data also show that Rok1 can promote duplex formation. To test if Rok1 can catalyze duplex formation without stably binding to duplexes, we observed Rok1 annealing and binding of the pre-A_2 duplex under gel-shift conditions that are native with respect to both RNA duplexes and protein. Figure 3.7 shows that at concentrations as low as 10 nM, Rok1 catalyzes duplex formation without stably binding to the double-stranded RNA; in contrast, stable Rok1 binding requires high Rok1 concentrations (≥ 400 nM Rok1). These results indicate that Rok1 is actively

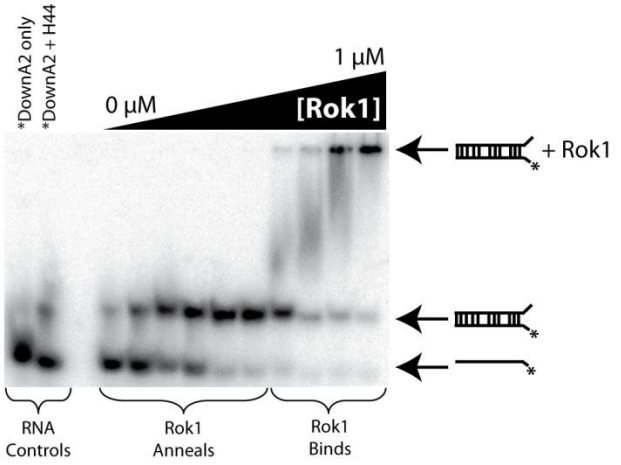


Figure 3.7. Rok1 catalyzes pre-A_2 duplex formation. Native gel analysis of Rok1 annealing and binding activities indicates that Rok1 fully anneals the wild-type duplex at concentrations below those required for stable binding. Rok1 concentrations are 0, 10, 25, 50, 100, 200, 400, 600, 800 and 1000 nM. RNA controls include ^{32}P-DownA_2 alone (lane 1) and unduplexed ^{32}P-DownA_2 and H44 mixture prior to 10-fold dilution into reactions containing Rok1 (lane 2). Final RNA concentrations are 0.5 nM ^{32}P-DownA_2 and 20 nM H44.

catalyzing pre-A_2 duplex formation in a manner that could allow for turnover. Interestingly, Rok1 does so in a transient manner that is nucleotide-independent.

Rrp5 enhances Rok1 annealing and specificity

DEAD-box proteins have specific non-redundant *in vivo* roles and are therefore essential (Jarmoskaite and Russell, 2011). However, because of their substantial sequence and structural similarity, and since DEAD-box proteins recognize primarily the backbone of bound RNAs (Schutz et al., 2010), it is not clear how DEAD-box proteins recognize their specific substrates. *In vitro* studies with DEAD-box proteins and co-factors have already shown that co-factors can modulate ATPase, duplex unwinding and RNA and nucleotide binding activities ((Young and Karbstein, 2012) and references therein). Similarly, co-factors could provide substrate specificity. However, there has been no evidence of a co-factor regulating DEAD-box protein activities in a sequence-specific fashion.

As shown above, Rrp5 is an RNA-binding protein that binds directly to Rok1 as expected for a DEAD-box protein co-factor. We therefore decided to use our system to test the hypothesis that co-factors enhance the specificity of DEAD-box proteins. To determine if Rok1 alone has specificity for the test duplex, we compared the annealing rate for the wild-type duplex to those for two control duplexes. As shown in Figure 3.8A (solid symbols), Rok1 anneals the wild-type duplex ~5-fold faster than the reverse

complement duplex (k_{obs} values of 0.02 sec^{-1} and 0.0039 sec^{-1}, respectively). Similarly, the U3/pre-rRNA duplex (including a 27 nucleotide extension at the 3' end of the pre-rRNA strand) is annealed ~20-fold more slowly (k_{obs} = 0.001 sec^{-1}; data not shown). These results indicate that Rok1 annealing activity is moderately sequence specific.

Next, we tested the effect of Rrp5 on the annealing activity of Rok1. As in the binding experiments above, we used the Rrp5C construct containing the TPR motifs. Comparison of duplex formation over time in the presence of Rrp5C indicates that Rrp5C enhances the annealing activity of Rok1 at least 15-fold with the wild-type duplex (Figures 3.8A (open symbols) and 3.8B; k_{obs} values of 0.30 sec^{-1} and 0.02 sec^{-1} for the presence and absence of Rrp5, respectively).[2] In contrast, Rrp5 enhances the annealing rate of the reverse complement duplex only ~2-fold (Figures 3.8A (open symbols) and 3.8B; k_{obs} values of 0.0082 sec^{-1} and 0.0039 sec^{-1} for the presence and absence of Rrp5C, respectively). The presence of Rrp5 provided no annealing enhancement for the U3/pre-rRNA duplex (Figure 3.8B; k_{obs} values of 0.0003 sec^{-1} and 0.001 sec^{-1} for the presence and absence of Rrp5, respectively). The presence of Rrp5

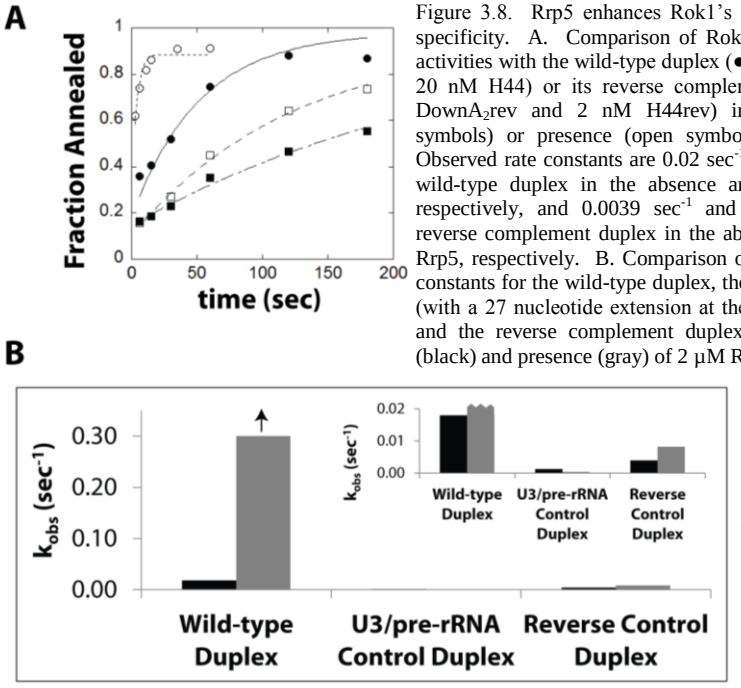


Figure 3.8. Rrp5 enhances Rok1's annealing activity and specificity. A. Comparison of Rok1 (400 nM) annealing activities with the wild-type duplex (●; 0.5 nM ^{32}P-DownA_2, 20 nM H44) or its reverse complement (■; 0.5 nM ^{32}P-DownA_2rev and 2 nM H44rev) in the absence (filled symbols) or presence (open symbols) of 2 μM Rrp5C. Observed rate constants are 0.02 sec^{-1} and 0.30 sec^{-1} for the wild-type duplex in the absence and presence of Rrp5, respectively, and 0.0039 sec^{-1} and 0.0082 sec^{-1} for the reverse complement duplex in the absence and presence of Rrp5, respectively. B. Comparison of Rok1 annealing rate constants for the wild-type duplex, the U3/pre-rRNA duplex (with a 27 nucleotide extension at the 3'-end of pre-rRNA) and the reverse complement duplex in both the absence (black) and presence (gray) of 2 μM Rrp5C. The arrow above the bar representing the rate constant for annealing of the wild-type duplex in the presence of Rrp5 indicates that this observed rate constant is an underestimate, as the reaction was too fast to be accurately measured (see panel A).

[2] Since annealing of the wild-type duplex in the presence of Rrp5C was too fast to be measured accurately (Figure 3.8A), we believe that the observed rate constant is an underestimate (indicated by arrow in Figure 3.8B).

therefore increases Rok1's specificity such that in the presence of Rrp5, Rok1's annealing of the wild-type duplex is ~40-fold faster than annealing of the reverse complement control duplex. For the first time, we present here results that exemplify a co-factor modulating DEAD-box protein activity sequence specifically.

Discussion

Expanding the role of RNA helicase co-factors

Of the 41 RNA helicases in *Saccharomyces cerevisiae*, eight are known to be regulated by co-factors. Activities such as nucleotide and RNA binding, ATP hydrolysis, RNA duplex unwinding and phosphate and RNA release can all be modulated by the presence of a co-factor ((Young and Karbstein, 2012) and references therein). In all of these cases, however, activity is regulated in a sequence-independent manner. Since DEAD-box proteins, the most abundant type of RNA helicases in *S. cerevisiae*, have both high sequence and structure conservation yet have non-redundant and therefore essential functions *in vivo*, how do DEAD-box proteins acquire specificity?

Substrate specificity could arise from the N- and C-terminal extensions that surround the conserved helicase core; however, the only demonstrated case in which an RNA helicase has intrinsic specificity is the *E. coli* DEAD-box protein DbpA. DbpA has ATPase and helicase activities that are specifically activated by helix 92 in 23S rRNA, which is recognized by unique sequences in DbpA's C-terminus (Fuller-Pace et al., 1993; Kossen et al., 2002; Tsu and Uhlenbeck, 1998; Wang et al., 2006). It is therefore postulated that DEAD-box proteins require additional means for conferring substrate specificity. Here, we provided insight into this outstanding question by showing that not only can the presence of a co-factor enhance the activity of a DEAD-box protein, but it can do so in a sequence-dependent way.

Prior data suggests that the RNA-binding protein Rrp5 is a co-factor of the DEAD-box protein Rok1 (Torchet et al., 1998; Vos et al., 2004a). Here we show that Rrp5, specifically the C-terminal fragment containing the TPR protein-protein interaction motifs, directly binds Rok1. Additionally, we show that Rok1 actively stabilizes and

catalyzes duplex formation; this annealing activity is moderately specific for the pre-A_2 wild-type duplex. Importantly, in the presence of Rrp5, annealing activity of this duplex is enhanced 15-fold while that of control duplexes is essentially unaffected.

Reconsidering the roles of DEAD-box proteins

Rok1 differs from 'classical' DEAD-box proteins in that it has both RNA-independent ATPase activity and lacks duplex unwinding activity. It is believed that an increased affinity for single-stranded RNA drives the unwinding activities of classical helicases. Consistent with the lack of unwinding activity, Rok1 preferentially binds double-stranded over single-stranded RNA. This affinity for double-stranded RNA also likely underlies Rok1's ability to promote duplex annealing. These results with Rok1 further validate roles for DEAD-box proteins outside of strictly helicase activities.

DEAD-box proteins are generally known as RNA-dependent ATPases that bind and release RNA in a regulated manner during the ATPase cycle (Jarmoskaite and Russell, 2011). Because they comprise the largest family of superfamily 2 (SF2) helicases, these proteins have also been described as DEAD-box helicases. However, only five of the 25 DEAD-box proteins in *S. cerevisiae* have exemplified *in vitro* duplex unwinding activity (Halls et al., 2007; Iost et al., 1999; Kikuma et al., 2004; Lawson et al., 1989; Weirich et al., 2006). In addition, DEAD-box proteins are known to have activities such as RNA-protein complex remodeling, RNA duplex annealing and ATP-dependent RNA binding (Jankowsky, 2011; Jankowsky and Fairman, 2007; Jarmoskaite and Russell, 2011; Linder, 2006).

Even though DEAD-box proteins are known to have a variety of activities, *in vitro* observations do not immediately translate to intracellular functions. A more comprehensive understanding of DEAD-box protein function includes validating biochemical characterizations *in vivo*; determining the relevance of Rok1 annealing activity and its role in 40S pre-rRNA processing will be essential for elucidating Rok1's in ribosome assembly.

Co-factors confer specificity in other biological processes

The DEAD-box proteins eIF4A and Ded1 have RNA-dependent ATPase and RNA duplex unwinding activities that are not sequence specific (Iost et al., 1999; Ozes et al., 2011). Since they are regulators of translation, a process that is independent of RNA sequence, this lack of substrate specificity is not surprising. Furthermore, these results are consistent with the observation that DEAD-box proteins have a highly conserved helicase core, which includes an RNA-binding site that primarily contacts the sugar-phosphate backbone (Schutz et al., 2010; Sengoku et al., 2006). For cellular processes such as ribosome assembly and pre-mRNA splicing, however, DEAD-box protein specificity for particular RNA sequences is likely required. As shown in this work with two 40S ribosome assembly factors, the RNA-binding protein Rrp5 can increase the activity of the DEAD-box protein Rok1 in an RNA sequence-dependent fashion; this confirms the hypothesis that co-factors can enhance the specificity of DEAD-box proteins (Jankowsky, 2011). While shown here for the first time for this large class of RNA-dependent ATPases that regulate all aspects of RNA metabolism, this concept of co-factors conferring substrate specificity is not unique to DEAD-box proteins and the cellular processes in which they are involved.

During transcription regulation, a highly specific process, the synergistic recruitment of co-factors by transcription factors is one of the ways by which transcription factor binding sites are differentiated (Carey, 1998; Merika and Thanos, 2001). For example, in *Saccharomyces cerevisiae*, the Met4 transcriptional co-factor regulates genes that control sulfur metabolism (Kuras et al., 1996). Met4 and another co-factor, Met28, however, only have DNA-binding activity when they are part of a trimeric complex with the DNA-binding transcription factor Cbf1 (Siggers et al., 2011). In the absence of the Met4-Met28 complex, Cbf1 specifically binds promoter regions containing an E-box sequence (CACGTG); when Cbf1 is complexed with Met4 and Met28, though, its binding becomes even more specific in that it primarily recognizes E-box sequences that are two base pairs away from a highly conserved RYAAT (Pur-Pyr-Ade-Ade-Thy) motif (Siggers et al., 2011). This increased specificity thereby allows a certain sub-class of the Met4 regulon genes to be transcriptionally activated.

A similar co-factor-dependent specificity enhancement is observed during the ubiquitylation and intracellular trafficking of membrane proteins. Rsp5, an essential yeast HECT ligase, is responsible for transferring ubiquitin to membrane proteins and consequently sorting them to the internal membranes of multivesicular bodies (MVBs) (Galan et al., 1996; Hein et al., 1995; Hicke, 1999). Since Rsp5 is essential for the endocytosis of most plasma membrane proteins (Morvan et al., 2004), it requires specificity co-factors to target certain protein substrates. For example, interaction of Rsp5 with the co-factor Bsd2 results in efficient ubiquitylation and sorting of the yeast vacuolar enzyme carboxypeptidase S (Cps1) and the polyphosphatase Phm5 into the MVB pathway (Hettema et al., 2004). Alternatively, in the absence of the co-factors endosomal adaptor of Rsp5 (Ear1) and its homolog Ssh4, the ubiquitylation and MVB sorting of Rsp5's target cargoes Sit1 and Fur4, respectively, are greatly decreased (Leon et al., 2008).

Exosome degradation of RNAs is another example of a biological process with conferred specificity. Consisting of a hexameric complex with a trimeric cap, the archaeal exosome has both 3'-5' exoribonuclease and heteropolymeric RNA-tail processing activities (Evguenieva-Hackenberg et al., 2008; Portnoy et al., 2005; Walter et al., 2006). The trimeric cap may contain the RNA-binding subunits Rrp4, Csl4 or a mixture of both (Buttner et al., 2005). The composition of this cap dictates exosome substrate specificity: while the Rrp4-exosome strongly prefers and degrades RNAs with A-rich tails, the Csl4-exosome targets both A-rich and A-poor RNAs equally (Roppelt et al., 2010). Additionally, the Rrp4-binding cap allows the exosome to interact more efficiently with longer RNA substrates (*i.e.* 90 nucleotides) in comparison to the Csl4-exosome (Roppelt et al., 2010).

As exemplified in these biological processes, co-factor-dependent substrate recognition is imperative for cellular function. In systems that have minimal intrinsic specificity, however, substrate recognition cannot be thoroughly characterized until co-factors have first been identified. In the case of RNA helicases in *S. cerevisiae*, yeast two-hybrid analyses and large scale affinity purifications in combination with mass spectrometry approaches have identified many potential interacting factors; in most

cases, however, direct interactions have yet to be shown and therefore bonafide co-factors have not been discovered for a majority of the RNA helicases. As more RNA helicase co-factors are identified and effects on activity and substrate specificity can be analyzed, it is likely that more examples of co-factors enhancing RNA helicase specificity will arise.

Materials and Methods

Cloning of Rok1 and Rrp5C. The Rok1 and Rrp5C ORFs were amplified from *Saccharomyces cerevisiae* genomic DNA. Rrp5C was cloned into pSV272 as in (Young and Karbstein, 2011). Rok1 was cloned between the SfoI and HindIII sites of pSV272. Rok1 K172A was derived from this plasmid using site-directed mutagenesis. All primers are listed in Table 3**.**

Rrp5C and Rok1 Expression and Purification. Rrp5C was purified as described in (Young and Karbstein, 2011). The purification of MBP-tagged Rrp5C differs in that it was dialyzed overnight in 50 mM NaCl, 25 mM HEPES, pH 7.6, 1 mM TCEP and 1 mM DTT (in the absence of TEV protease); all remaining purification steps were identical. MBP-Rrp5C concentration was determined by absorbance at 280 nm using a calculated extinction coefficient of 128,160 $M^{-1}cm^{-1}$. For both wild-type and Rok1 K172A overexpression, *Escherichia coli* Rosetta cells containing the Rok1-encoding pSV272 plasmid were grown at 37°C in LB Miller media (supplemented with 25 µg/mL and 34 µg/mL of Kanamaycin and Chloramphenicol, respectively) to an OD_{600} ~0.6 before inducing with 1 mM IPTG in the presence of 2% ethanol at 18°C for ~18 hours. Cells were resuspended in lysis buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 10 mM imidazole, 0.1 mM PMSF and 5 mM benzamidine) and sonicated. After pelleting, the soluble fraction was purified over Ni-NTA resin (Qiagen) according to the manufacturer's protocol. Rok1-containing elution fractions were pooled and further purified using a 35% 3.9 M ammonium sulfate cut to eliminate contaminants followed by a 50% cut to precipitate Rok1. Rok1 was resuspended in elution buffer (50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl and 250 mM imidazole) and dialyzed overnight in 200 mM NaCl, 50 mM Tris, pH 7.8, 2 mM EDTA, 1 mM TCEP and 1 mM DTT; TEV

protease was also added to remove the His_6-MBP tag. Rok1 was further purified over a MonoS column in a linear gradient from 440 mM to 720 mM NaCl over 12 column volumes in the presence of 2 mM EDTA. Rok1-containing fractions were then purified over a Superdex200 column into 150 mM KCl, 50 mM HEPES, pH 7.6, 1 mM DTT and 1 mM TCEP. Rok1 was stored in 15% glycerol at -80°C. Protein concentration was determined by absorbance at 280 nm using a calculated extinction coefficient of 23,980 $M^{-1}cm^{-1}$.

RNA transcription, purification and labeling. RNAs smaller than 35 nucleotides were ordered from IDT and subsequently purified over a 15% native acrylamide gel in TBE buffer (50 mM Tris, 40 mM Boric Acid and 0.5 mM EDTA). RNA was excised from the gel, passively eluted overnight in TE buffer (10 mM Tris, pH 8.0 and 1 mM EDTA) and concentration was determined by absorbance at 260 nm. Longer RNAs were *in vitro* transcribed using T7 RNA polymerase and a primer containing the T7 promoter in combination with a primer encompassing the T7 promoter sequence followed by the RNA specific sequence (T7 Promoter and T7-RNA, respectively; Table 3). Primers were mixed and denatured at 95°C for 1 min before incubating at room temperature for at least 10 min to allow primers to fully anneal. *In vitro* transcription was conducted as previously described (Karbstein et al., 2005; Lamanna and Karbstein, 2009) using 300 nM of primer template, and the RNA was purified over a 10% TBE/native acrylamide gel. RNA was passively eluted overnight in TE buffer or precipitated and resuspended in water. Concentration was determined by absorbance at 260 nm.

RNA was 5'-end-labeled using T4 polynucleotide kinase (New England BioLabs) and equimolar amounts of ^{32}P-γ-ATP. *In vitro* transcribed RNAs were first de-phosphorylated with Antarctic Phosphatase (New England BioLabs). RNAs were then gel purified using a TBE/native acrylamide gel, excised from the gel and passively eluted overnight in water.

Rok1 and Rrp5C Interaction. After a 15 minute pre-incubation on ice, either 5 µM Rok1; 5 µM Rok1 and 2.5 µM (MBP)-Rrp5C; or 5 µM Rok1 and 2.5 µM MBP were mixed in the presence of binding buffer (100 mM KCl, 30 mM HEPES, pH 7.6) and 15 µL amylose resin (New England BioLabs) and incubated on a rotating platform at 4°C for 30

minutes. Resin was washed four times with 55 µL binding buffer before being eluted with 30 µL binding buffer supplemented with 50 mM maltose.

Rok1 RNA Binding Experiments. For Rok1 binding to single-stranded RNA, 5 nM of ^{32}P-labeled RNA was used; for binding to double-stranded RNA, 5 nM ^{32}P-labeled RNA was pre-incubated at 95°C for 1 minute with enough unlabeled RNA to give full duplex formation (see Figure legends for exact concentrations) in the presence of 10 mM MOPS, pH 6.5, 1 mM EDTA and 50 mM KCl. RNAs were slowly cooled to 35°C and then incubated on ice for 10 minutes. RNA was diluted 10-fold into pre-incubated (20°C) reaction mixtures containing Rok1 in the presence of 40 mM Tris, pH 8.3, 50 mM KCl, 0.5 mM $MgCl_2$, and 2 mM DTT. Reactions were incubated at 20°C for 7 minutes before being separated on a 10% native acrylamide/THEM (Tris, HEPES, EDTA, pH 7.5, $MgCl_2$) gel (Karbstein et al., 2002). Gels were exposed to a phosphor screen, quantified using phosphorimager software and analyzed as described in (Young and Karbstein, 2011). Briefly, data was analyzed as follows:

$$\text{fraction}_{\text{bound}} = \frac{[\text{protein}_n \cdot \text{RNA}]}{[\text{protein}_n \cdot \text{RNA}] + [\text{RNA}]} \qquad \text{Eq. 4}$$

where n = the number of protein molecules bound.

The binding affinity, $K_{1/2}$, is defined by the binding equilibrium shown in Eq. 5:

$$K_{1/2} = \frac{[\text{protein}]_f^n \times [\text{RNA}]}{[\text{protein}_n \cdot \text{RNA}]} \qquad \text{Eq. 5}$$

where $[\text{protein}]_f$ = concentration of free protein.

Assuming that $[\text{protein}]_f \sim$ [protein] added (which is true since trace RNA is present), solving Eq. 5 for [protein$_n$•RNA] and substituting this into Eq. 4 gives Eq. 6:

$$\text{fraction}_{\text{bound}} = \frac{\text{fraction}_{\text{bound,max}}[\text{protein}]^n}{[\text{protein}]^n + K_{1/2}} \qquad \text{Eq. 6}$$

Rok1 ATPase Experiments. Rok1 was incubated at 20°C for 5 minutes in the presence of 40 mM Tris, pH 8.3, 50 mM KCl, 2 mM DTT and 10 mM $MgCl_2$. After the addition of gel-purified ^{32}P-γ-ATP to initiate the ATPase reactions, reaction mixtures were quenched

at various time points with an equal volume of 0.75 M KH_2PO_4, pH 3.3. Quenched reaction mixtures were spotted onto a PEI nitrocellulose TLC plate and eluted in 1 M LiCl, 300 mM NaH_2PO_4, pH 3.8. TLC plates were exposed to a phosphor screen and the amount of ^{32}P-γ-ATP vs. free ^{32}P-P_i was determined using phosphorimager software. Fraction P_i versus time was plotted and data was fit to the following single exponential equation to determine the k_{obs} for ATP hydrolysis:

$$\text{fraction } P_i = \text{fraction}_{\text{reacted}}^{max} - \text{fraction}_{\text{reacted}}^{max} \bullet exp(-k_{obs} \bullet t) \quad \text{Eq. 7}$$

To measure the effects of RNA on Rok1 ATPase activity, reactions were prepared as described above, except 1 μM Rok1 was incubated in the presence of increasing concentrations of double-stranded or single-stranded RNA. To determine K_I, k_{obs} versus RNA concentration was plotted and fit to the following equation:

$$k_{obs} = \frac{k_{obs}^{max}}{1 + \frac{[\text{nucleotide}]}{K_I}} \quad \text{Eq. 8}$$

Rok1 Annealing and Unwinding Experiments. Annealing and unwinding experiments were modified from (Jankowsky and Putnam, 2009). Briefly, annealing reactions were initiated by the 10-fold dilution of RNA (5 nM ^{32}P-labeled and enough unlabeled to give ~20% initial duplex formation – see Figure legends for exact concentrations) in 100 mM MOPS, pH 6.5, 10 mM EDTA and 500 mM KCl) to pre-incubated (20°C for 5 minutes) reaction tubes containing Rok1 in the presence of 40 mM Tris, pH 8.3, 50 mM KCl and 2 mM DTT. Unwinding reactions were initiated by the dilution of ATP/$MgCl_2$ to a final concentration of 2 mM into pre-incubated (20°C for 5 minutes) reaction tubes containing 0.5 nM ^{32}P-labeled and optimized concentration of unlabeled RNA, 10 mM MOPS, 40 mM Tris, pH 8.3, 50 mM KCl, 1 mM EDTA and 2 mM DTT. Reactions were quenched at various time points into an equal volume of 50 mM EDTA, 1% SDS and 0.1% bromophenol blue in 20% glycerol and iced. Reactions were separated on a 15% native acrylamide/THEM (Tris, HEPES, EDTA, pH 7.5, $MgCl_2$) gel (Karbstein et al., 2002). Gels were exposed to a phosphor screen, quantified using phosphorimager software and data were fit to Eq. 4 to determine k_{obs}.

To analyze Rok1 annealing and binding under native conditions, reactions were prepared as described above for annealing conditions. After incubation at 20°C for 10 minutes, reactions were immediately separated on a 10% native acrylamide/THEM (Tris, HEPES, EDTA, pH 7.5, $MgCl_2$) gel (Karbstein et al., 2002). Gels were exposed to a phosphor screen.

To measure duplex stability in the presence and absence of Rok1, annealing reactions were prepared as normal, except 0.5 nM of radiolabeled $DownA_2$ was incubated with varying H44 concentrations in the presence of 400 nM Rok1. Fraction RNA duplexed versus H44 concentration was plotted and data was fit to Eq. 9 to determine $K_{1/2}$:

$$\text{fraction}_{\text{duplexed}} = \frac{\text{fraction}_{\text{duplexed,max}}[\text{RNA}]^n}{[\text{RNA}]^n + K_{1/2}} \qquad \text{Eq. 9}$$

Table 3: Oligonucleotides used in Chapter III

Name	Sequence
Rok1-SfoI	5’-GATCGAGGCGCCATGGATATTTTTAGAGTATTAACTAG-3’
Rok1-HindIII	5’-TCAGACAAGCTTTTATTTCGAGAAATGTTTTTTTGAAAG-3’
Rok1-K172A	5’-CCACAGGGTCTGGTGCGACGTTAGCATTC-3’
T7 Promoter	5’-GCTCGGTACCCGGGGATCTAATACGACTCACTATAGG-3’
T7+RNA	5’-CGCGCAAAGATATGAAAACT CCACAGTGTGTTGTATTGAAACGGTTTTAACCCTATAGTGAGTCGTATTAGATC-3’

Chapter IV

Rok1 annealing and ATPase activities regulate its pre-ribosome association

Introduction

During eukaryotic ribosome maturation, the four ribosomal RNAs (rRNAs) must be processed, folded and assembled with the 78 ribosomal proteins. Since three of the four rRNAs, the 18S, 5.8S and 25S rRNAs, are co-transcribed in a single transcript, these rRNAs must undergo processing that involves well-ordered cleavage steps to release the mature rRNAs (Strunk and Karbstein, 2009; Venema and Tollervey, 1999). rRNA maturation, folding and binding of ribosomal proteins are both temporally and spatially controlled by a molecular machinery of ~200 transiently associated protein and RNA factors. These assembly factors include RNA-binding proteins; kinases; endonucleases; ATPases and GTPases; and DEAD-box proteins (Strunk and Karbstein, 2009).

Fifteen DEAD-box proteins are involved in eukaryotic ribosome assembly; most of them are essential and therefore have non-redundant functions (Jarmoskaite and Russell, 2011). In almost all cases, however, even though it is known whether they are required for 40S or 60S assembly, their specific roles in rRNA processing are mostly unknown. An exception to this is Dbp4, which is required for release of the U14 small nucleolar RNA (snoRNA) from pre-ribosomes (Kos and Tollervey, 2005; Liang et al., 1997). Similarly, it has been suggested that Rok1 releases snR30 from pre-40S ribosomes, as snR30 accumulates 5-fold in the absence of Rok1 (Bohnsack et al., 2008). A mutation in the Walker A motif gives a 2-fold accumulation of snR30. However, because Rok1 is required for early events in 40S maturation, this effect could be direct or

indirect. Interestingly, depletion of Rps3 has a similar 3-fold effect, perhaps arguing that an indirect effect is more likely.

Of the fifteen DEAD-box proteins involved in ribosome assembly, seven have been analyzed *in vitro*. Their activities include duplex unwinding, ATP hydrolysis and RNA binding (Garcia and Uhlenbeck, 2008; Granneman et al., 2006; Kikuma et al., 2004; Rocak et al., 2005). Nevertheless, even though some instances have linked specific residues required for *in vitro* activity to general effects on cell growth, these biochemical observations have not been extended to specific *in vivo* functions. Knowing that the DEAD-box protein Rok1 has annealing activity that is enhanced by the RNA-binding protein Rrp5 in a sequence-specific manner (Chapter 3), we decided to test the significance of this annealing activity and further elucidate Rok1's function *in vivo*.

Here, we present preliminary insights into the role for Rok1 in 40S ribosome maturation *in vivo*. Preliminary *in vivo* structure probing data are consistent with a role for Rok1 in formation of a duplex between one strand of H44 and a sequence downstream of cleavage site A_2. Later in assembly, this duplex is required for temporal control of the 3'-end formation of 18S rRNA (Lamanna and Karbstein, 2011). Surprisingly, we also find that the K172A mutant of Rok1, which is unable to hydrolyze ATP, can promote annealing *in vitro* but displays a strong growth phenotype *in vivo*. Sucrose gradient centrifugation demonstrates that ATP hydrolysis is required for release of Rok1 from pre-ribosomes, as the K172A mutant accumulates on later assembly intermediates. Together, these data demonstrate for the first time that a DEAD-box protein functions *in vivo* to promote duplex formation and the switch to ATPase activity is then used to promote dissociation of Rok1 from the pre-ribosome.

Results

Rok1 is required for cleavage at A_1

Results in Chapter 3 indicate that Rok1 has RNA duplex annealing but not unwinding activity. Since biochemical observations do not directly indicate cellular relevance, we decided to analyze the role of Rok1 and test the significance of its annealing activity *in vivo*. Using a *S. cerevisiae* strain in which endogenous Rok1 is under the control of a galactose-inducible promoter and plasmids expressing wild-type Rok1 or the Walker A motif mutant K172A under the constitutive TEF1 promoter (Figure 4.1A; (Mumberg et al., 1995)), we first tested cell viability in the presence and absence of Rok1. In agreement with prior results (Oh and Kim, 1999; Song et al., 1995), plasmid-encoded wild-type Rok1 is required for cell growth on glucose but not galactose-containing medium. The Walker A mutant K172A produces a severe slow growth phenotype, but is viable (Figure 4.1B).

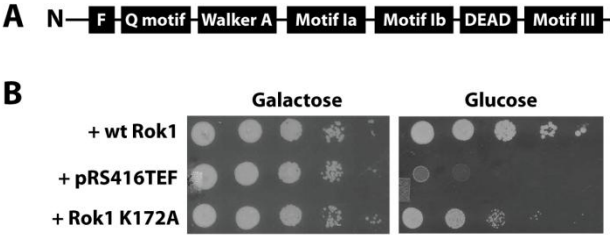


Figure 4.1. Rok1 and its ATPase activity are required for optimal cell growth. A. Rok1 contains the 11 highly conserved motifs found in DEAD-box proteins. B. Ten-fold serial dilutions of Gal1::Rok1 transformed with either wild-type Rok1, pRS416TEF vector or Rok1 K172A were compared on galactose- and glucose-containing plates.

Since Rok1 is known to be involved in 40S ribosome biogenesis (Venema et al., 1997), we next analyzed pre-rRNA processing phenotypes in the presence and absence of Rok1 using our Gal1::Rok1 strain. Consistent with previous results (Venema et al., 1997), wild-type Rok1 is required for 18S but not 25S rRNA production (Figure 4.2A). Additionally, in the absence of Rok1, the 20S precursor to 18S rRNA is depleted while the 35S (primary transcript that contains 18S, 5.8S and 25S rRNAs) and 22S/23S rRNAs accumulate. This accumulation of both 35S and 22S/23S rRNAs suggests that cleavages at sites A_1 and A_2 are inhibited (see rRNA processing schematic in Figure 4.2A). Similar to the absence of Rok1, yeast cells containing Rok1 K172A have decreased levels of 18S rRNA and its precursor 20S while levels of earlier pre-rRNA species are increased. To confirm the Northern analysis, we analyzed initial cleavage events using primer

extension. Results in Figure 4.2B indicate that while cleavage at A_0 is unaffected by the absence of Rok1, cleavage at A_1 is greatly decreased in the absence of Rok1 relative to its presence. These results demonstrate that Rok1 is required for rRNA processing events early in the 40S assembly pathway; this is in agreement with nucleolar localization of Rok1 ((Venema et al., 1997) and data not shown), as early cleavage events (A_0, A_1 and A_2; see Figure 4.2A) occur co-transcriptionally in the nucleolus (Kos and Tollervey, 2010)

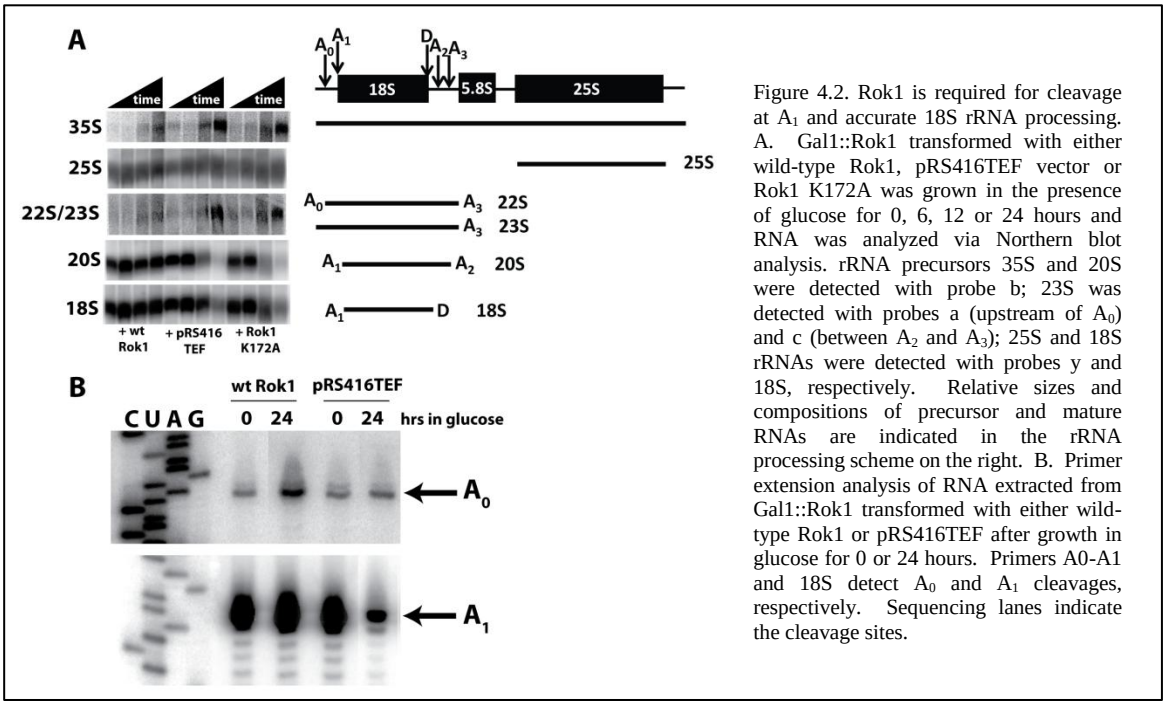


Figure 4.2. Rok1 is required for cleavage at A_1 and accurate 18S rRNA processing. A. Gal1::Rok1 transformed with either wild-type Rok1, pRS416TEF vector or Rok1 K172A was grown in the presence of glucose for 0, 6, 12 or 24 hours and RNA was analyzed via Northern blot analysis. rRNA precursors 35S and 20S were detected with probe b; 23S was detected with probes a (upstream of A_0) and c (between A_2 and A_3); 25S and 18S rRNAs were detected with probes y and 18S, respectively. Relative sizes and compositions of precursor and mature RNAs are indicated in the rRNA processing scheme on the right. B. Primer extension analysis of RNA extracted from Gal1::Rok1 transformed with either wild-type Rok1 or pRS416TEF after growth in glucose for 0 or 24 hours. Primers A0-A1 and 18S detect A_0 and A_1 cleavages, respectively. Sequencing lanes indicate the cleavage sites.

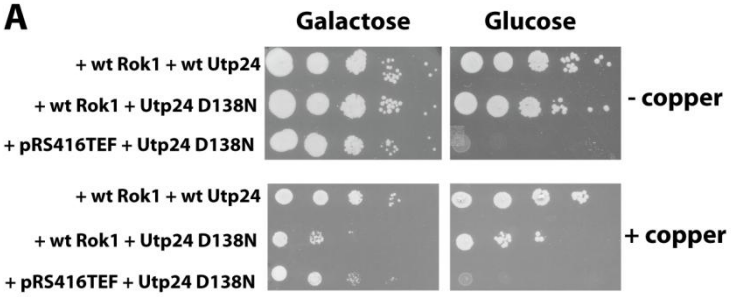
Absence of Rok1 results in different rRNA conformations near an essential switch region

Recent work form our lab uncovered an essential conformational change that is required for ordering cleavage at sites A_2 and D during 18S rRNA maturation (Lamanna and Karbstein, 2011). Prior to cleavage at site A_2, rRNA from one strand of helix 44 (H44) forms an inhibitory duplex with sequences downstream of A_2 (see Figure 3.2). After cleavage at site A_2, this duplex is replaced by the mature decoding site H44. The RNA-binding protein Rrp5, which is also required for 40S maturation, binds residues just downstream of this inhibitory duplex (Young and Karbstein, 2011). Results in Chapter 3 indicate that Rrp5 directly binds Rok1 (thereby placing it in the vicinity of this switch region). Furthermore, Rrp5 enhances Rok1 annealing of this duplex in a sequence

specific manner. Based on these biochemistry data, we hypothesized that Rok1 and Rrp5 promote formation of the inhibitory duplex during rDNA transcription. This could explain how this inhibitory duplex, which is formed between two strands ~300 nucleotides away from each other, form a local duplex. Since *in vivo* results thus far support Rok1's involvement in early pre-rRNA processing, we decided to test for changes in the rRNA near the essential switch region in both the presence and absence of Rok1.

Northern blot and primer extension results in Figure 4.2 suggest that 22S precursors accumulate in the absence of Rok1 due to decreased cleavage at A_1 while 20S rRNA accumulates in the presence of wild-type Rok1. In order to accurately analyze changes in the rRNA structure near the switch region due to the presence and absence of Rok1 and not other changes associated with these different 40S assembly intermediates, we first needed to ensure that we were accumulating and comparing the same rRNA species (i.e. 22S rRNA) in both the presence and absence of Rok1. Since A_1 cleavage is inhibited in the absence of Rok1, we decided to also inhibit A_1 cleavage in the presence of Rok1; this would ensure that we were accumulating the same rRNA precursors in both strains. Utp24 is suggested to be the endonuclease required for A_1 cleavage (Bleichert et al., 2006). The D138N mutant in its active site blocks 40S assembly prior to cleavage at site A_1 and provides for a dominant negative growth defect (Bleichert et al., 2006). We therefore constructed a copper-inducible plasmid encoding Utp24 D138N and transformed it into the Gal1::Rok1 strain.

We first confirmed that the copper-inducible presence of Utp24 D138N provides the expected dominant negative phenotype. As shown in Figure 4.3A, when strains with Utp24 D138N are grown on copper-containing plates, a slow growth phenotype is observed in both the presence and absence of Rok1 on galactose and glucose-containing plates; this verifies the dominant negative effect of Utp24 D138N. Demonstrating that the effect arises from the active site mutation, in the presence of wild-type Utp24, growth is unaffected on the copper-containing plates. Copper-free galactose plates verify that an equal number of cells were plated while copper-free glucose plates verify the Rok1 phenotype observed in Figure 4.1.



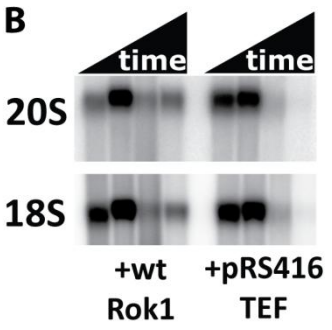


Figure 4.3. Utp24 D138N is dominant and stalls pre-rRNA processing. A. Ten-fold serial dilutions of Gal1::Rok1 + wild-type Utp24 (pRS413cup1) + wild-type Rok1 (pRS416TEF) or Gal1::Rok1 + Utp24 D138N (pRS413cup1) + wild-type Rok1 (pRS416TEF)/empty pRS416TEF were compared on galactose- and glucose-containing plates in both the presence and absence of 100 µM $CuSO_4 \bullet 5H_20$ at 30°C. B. Northern analysis of Gal1::Rok1 + Utp24 D138N (pRS413cup1) + wild-type Rok1 (pRS416TEF) or empty pRS416TEF grown in synthetic glucose (-His) media and 100 µM $CuSO_4 \bullet 5H_20$ at 30°C for 0, 10, 25 or 30 hours. 20S and 18S rRNA were detected with probes b and 18S, respectively.

Next we tested effects on pre-rRNA processing, particularly the accumulation of 18S rRNA, arising from expression of Utp24 D138N. As shown in Figure 4.3B, as expected from the dominant negative effect on rRNA processing (Bleichert et al., 2006), 18S and 20S rRNA levels are decreased in both the presence and absence of Rok1 after 25 hours of growth in glucose.

Since the cell growth and northern results suggested that the strains containing Utp24 D138N accumulate 22S rRNA both in the presence and absence of Rok1, we next conducted *in vivo* dimethyl sulfate (DMS) structure probing experiments to test the rRNA conformations near the switch region. DMS methylates adenosine and cytosine residues unless they are protein protected or involved in base pair interactions; RNA modifications can be detected by sequencing gel analysis as they lead to reverse transcription stops (Wells et al., 2000). Preliminary results identify one region of difference between the presence and absence of Rok1 (Figure 4.4A): the two adenosine residues at the very 3' end of the single-stranded region that encompasses cleavage site D (Figure 4.4B). These adenosines are exposed and consequently methylated in the presence of wild-type Rok1, while they are protected in the absence of Rok1. Mapping these protections onto the secondary structure in Figure 4.4B indicates that in the presence of Rok1, the adenosine residues are exposed as they would be in the pre-A_2 structure. In the absence of Rok1, they are protected; interestingly, in the post-A_2 structure, the base pairing interactions of these adenosines would inhibit them from being methylated.

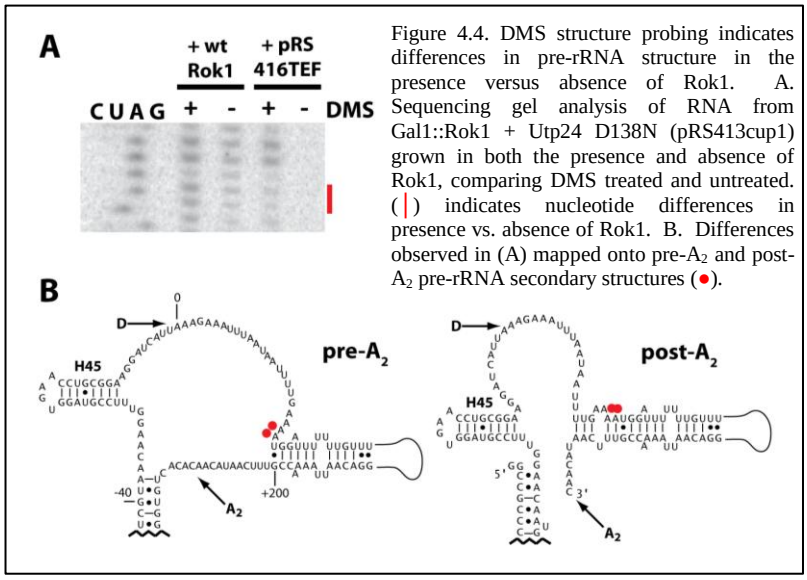


Figure 4.4. DMS structure probing indicates differences in pre-rRNA structure in the presence versus absence of Rok1. A. Sequencing gel analysis of RNA from Gal1::Rok1 + Utp24 D138N (pRS413cup1) grown in both the presence and absence of Rok1, comparing DMS treated and untreated. (|) indicates nucleotide differences in presence vs. absence of Rok1. B. Differences observed in (A) mapped onto pre-A_2 and post-A_2 pre-rRNA secondary structures (●).

Current aims include probing other regions of the rRNA, especially those that contain the inhibitory duplex. Preliminary investigations, however, have been complicated due to the remnants of 18S rRNA in these samples, even after 30 hours of growth in glucose (see Figure 4.3B). Therefore, when we sequence rRNA regions present in mature 18S rRNA, the results are tainted by sequences arising from 18S rRNA. To circumvent this problem, we have included a nuclear and cytoplasmic fractionation step after the *in vivo* DMS probing; this has allowed us to successfully isolate precursor RNAs from the mature 18S rRNA. Structure probing of these RNAs is currently in progress.

An additional concern with the current experimental set-up is the effect of 30 hours of growth in copper medium on yeast metabolism and therefore ribosome assembly. Because of the dominant negative effect on growth, there is selective pressure against the Utp24 D138N plasmid in the presence of copper. Strains must therefore be grown in minimal media to maintain the plasmid; however, this results in slower cell growth than normal. Consequently, to ensure that Rok1 is successfully depleted, we must grow these cells for 30 hours. Unfortunately, we have evidence that under these conditions, recombination between the endogenous wild-type Utp24 locus and the plasmid-encoded Utp24 can occur, which leads to poor reproducibility. To reduce these additional variables in this experimental set-up, we are currently creating a new yeast strain in which both endogenous Rok1 and Utp24 are under the control of galactose-inducible promoters. Rok1 and either wild-type Utp24 or Utp24 D138N can then be provided on constitutively active plasmids.

Rok1 preferentially binds ADP and has ATP-inhibited annealing activity

Results above indicate that Rok1 is required for accurate processing of early rRNA precursors to the 40S ribosome; preliminary *in vivo* structure probing results also suggest that a pre-A_2-like structure is favored in the presence of Rok1. These results support our current model: Rrp5 enhances the annealing activity of Rok1 to promote formation of the inhibitory duplex that orders cleavage at sites A_2 and D early in rRNA processing. If Rok1 anneals and then binds the duplex to stable it (as suggested by the biochemical data in Chapter 3), what triggers Rok1 release so that the conformational change can occur?

DEAD-box proteins are known as RNA-dependent ATPases that bind and release RNA in a regulated manner during the ATPase cycle (Jarmoskaite and Russell, 2011). Considering this, we decided to better characterize Rok1's ATPase activity with regard to our current model.

First, we tested Rok1's affinity for ADP versus ATP. As shown in Figure 4.5A, Rok1 binds ADP ~750-fold more tightly than the non-hydrolyzable ATP analog AMPPNP (K_I values of 0.26 µM and 196 µM, respectively). Interestingly, intracellular concentrations of ATP are only ~2-3-fold higher than ADP (Theobald et al., 1997), suggesting that at the steady state, Rok1 is bound to ADP *in vivo*. Preference for ADP over ATP is common amongst other DEAD-box proteins (Karow et al., 2007; Lorsch and Herschlag, 1998; Theobald et al., 1997), although the

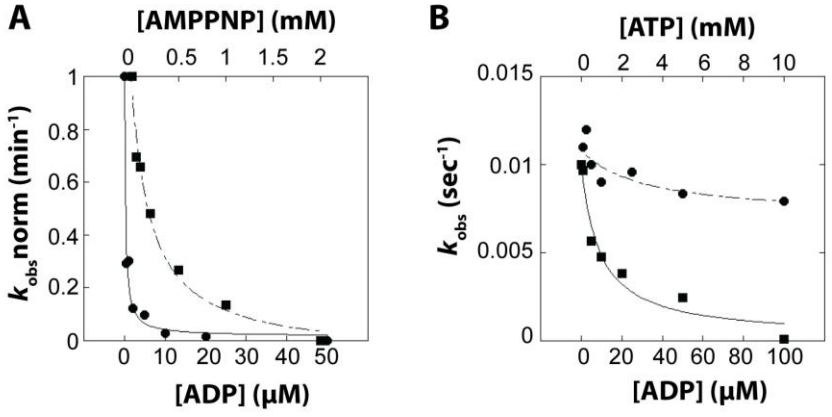


Figure 4.5. Rok1 has a higher affinity for ADP than ATP, but ATP inhibits Rok1 annealing activity. A. Normalized results from ATPase inhibition assays in the presence of AMPPNP (■) and ADP (●) (K_I values are 0.26 µM and 196 µM for ADP and AMPPNP, respectively). B. Comparison of Rok1 annealing rates in the presence of ATP (■) and ADP (●) (K_I values are 30 µM and 882 µM for ADP and ATP, respectively).

magnitude is much larger than previously reported for other DEAD-box proteins. In some cases, it has been shown that co-factors modulate the ADP/ATP preference ((Young and Karbstein, 2012) and references therein).

According to our model, Rok1's annealing activity is required for formation of the inhibitory duplex and subsequent association with the pre-ribosome. This model together with the finding that Rok1 is primarily ADP-bound, predicts that ADP should not affect Rok1 annealing activity. Annealing assays in Chapter 3 demonstrated that Rok1 annealing is an ATP-independent process; we therefore decided to also analyze the effects of ATP and ADP on Rok1 annealing activity. Figure 4.5B shows that indeed in the presence of ADP, Rok1 annealing activity is minimally affected. In the presence of ATP, however, Rok1 annealing activity is fully inhibited. This suggests that at the steady state, Rok1 is ADP-loaded and promotes annealing of the pre-A_2 duplex.

Rok1 ATPase-deficient mutant has annealing activity

If Rok1 is ADP-bound and ATP binding is not required and even inhibits annealing activity, why does the Walker A K172A mutant have slow growth (Figure 4.1B)? The simplest possibility is that this mutant is impaired in its ability to promote duplex annealing. We therefore tested this *in vitro*.

We recombinantly expressed and purified this mutant and tested its *in vitro* ATPase and annealing activities. Comparison of wild-type and Rok1 K172A activities indicated that although this Walker A mutant is deficient in ATPase activity (Figure 4.6A), it can anneal an RNA duplex as efficiently as wild-type Rok1 (Figure 4.6B). This is consistent with the observation that ATP is not required and inhibiting

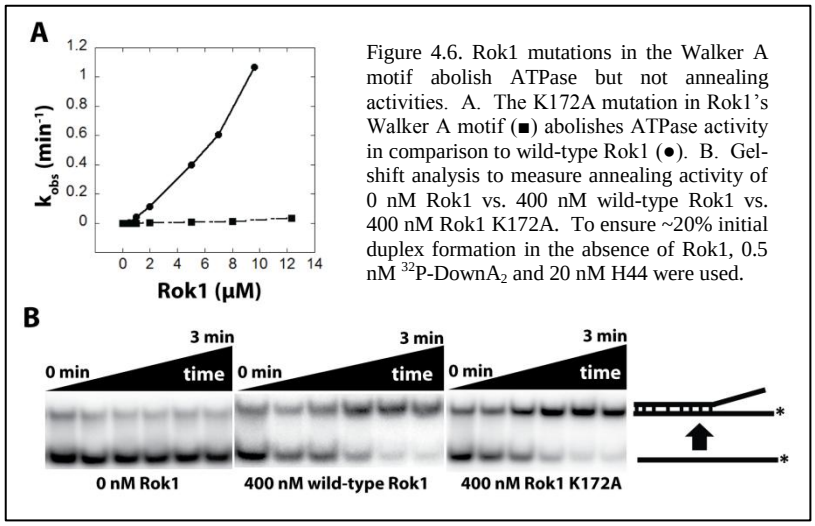


Figure 4.6. Rok1 mutations in the Walker A motif abolish ATPase but not annealing activities. A. The K172A mutation in Rok1's Walker A motif (■) abolishes ATPase activity in comparison to wild-type Rok1 (●). B. Gel-shift analysis to measure annealing activity of 0 nM Rok1 vs. 400 nM wild-type Rok1 vs. 400 nM Rok1 K172A. To ensure ~20% initial duplex formation in the absence of Rok1, 0.5 nM ^{32}P-DownA_2 and 20 nM H44 were used.

of Rok1 annealing activity. As suggested by the Rok1 K172A slow growth phenotype, the finding that this ATPase-deficient mutant still has annealing activity suggests that ATPase activity is required for some other aspect of Rok1 function.

Rok1 K172A is stalled on 60S-like ribosomes

Since ATP binding and hydrolysis are not required for Rok1 annealing activity but strongly enhance growth, ATP binding must be regulated at a different step. We therefore tested if Rok1 ATP binding and hydrolysis are required for release of Rok1 from pre-ribosomes. If Rok1 ATPase activity is required for Rok1 dissociation from the pre-ribosome, then we would predict that in strains with the ATPase-deficient mutant Rok1 K172A, Rok1 should be associated with later assembly intermediates than strains containing wild-type Rok1. To test this, we analyzed sedimentation of wild-type and Rok1 K172A via sucrose density gradients and Western blot. Figure 4.7 shows the expected polysome profiles for a 40S assembly factor: in both the absence of Rok1 and presence of Rok1 K172A, there is a decrease in free 40S ribosomes, which consequently causes a decrease in the levels of 80S ribosomes. Since there are fewer 40S particles to make 80S ribosomes, there is also an apparent increase in free 60S ribosomes. Western analysis indicates that wild-type Rok1 sediments in 40S, 60S and 80S sized fractions as expected for an early 40S assembly factor; a portion of Rok1 is also not associated with any ribosomes (Fraction 1; Figure 4.7). On the contrary, Rok1 K172A sediments only with 60S-like ribosomes (Fraction 7); notably, Rok1 is no longer observed in the free

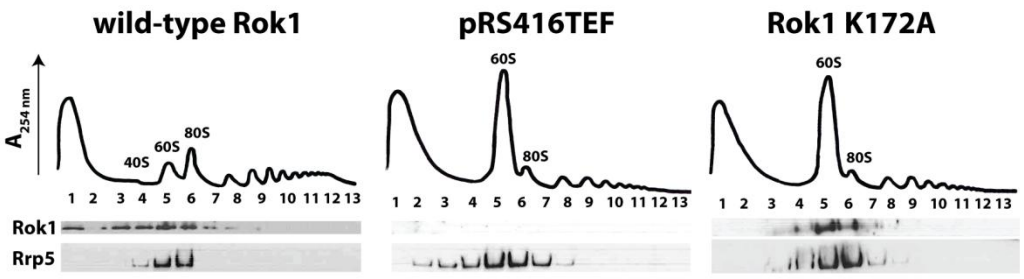


Figure 4.7. Rok1 K172A is stalled on 60S-like ribosomes. Gal1::Rok1 strains with either wild-type Rok1, pRS416TEF vector or Rok1 K172A were grown in glucose-containing medium for 18 hours and analyzed via sucrose gradient analysis. Absorbance at 254 nm was used to detect 40S, 60S and 80S fractions. Western blots analysis with polyclonal antibodies Rok1 and Rrp5 were used to detect Rok1 and Rrp5 sedimentation in these fractions.

fraction (Fraction 1; Figure 4.7). As expected, no Rok1 is observed in the presence of pRS416TEF only (Figure 4.7).

The depletion of the free Rok1 fraction and the shift into 60S-like ribosomes demonstrate that ATP hydrolysis by Rok1 is required for its release from 40S precursors. The sedimentation of a 40S assembly factor with 60S-like ribosomes can be explained by the unique positioning of the Rok1-Rrp5 complex: both proteins are bound to the rRNA segment 3' to site A_2 and Rrp5 binds even 3' to site A_3 and is known to partition with 60S precursors after A_2/A_3 cleavage (Nissan et al., 2002; Young and Karbstein, 2011); this is consistent with its sedimentation here (Figure 4.7). We are currently testing if ATP binding to Rok1 leads to release of Rrp5.

Discussion

Rok1's function in vivo

In this study, we present both *in vivo* and *in vitro* data that support a model by which Rok1's annealing and ATPase activities regulate pre-ribosome formation and Rok1 association, respectively. Our preliminary structure probing experiments suggest that Rok1 is required for accurate formation of the pre-A_2 duplex. This activity is in agreement with Rok1's robust duplex annealing and lack of duplex unwinding *in vitro* and provides the first example of a DEAD-box protein having *in vivo* relevant annealing activity. Since Rrp5 specifically enhances Rok1 annealing of the pre-A_2 duplex, we propose that Rok1 and Rrp5 promote formation of this inhibitory duplex during rDNA transcription. This finding explains how the pre-A_2 duplex forms from two RNA strands that are ~300 nucleotides apart instead of simply extending H44. Rok1's nucleotide affinities suggest that Rok1 is ADP-bound in the steady state *in vivo*; *in vitro* analyses indicate that ADP-Rok1 actively catalyzes duplex formation. Nevertheless, the observation that the ATPase deficient Rok1 K172A mutant has a severe slow growth phenotype *in vivo* suggests that ATPase activity and/or ATP binding is essential for cellular function. Sucrose density analysis indicates that Rok1's ATPase activity is required for release from pre-ribosomes, as Rok1 K172A remains bound to later

precursors. In further support of this model, Rok1 K172A retains its ability to anneal RNA duplexes and ATP inhibits wild-type Rok1 annealing activity. These data therefore suggest that when Rok1 dissociation is appropriate, presumably after cleavage at A_2, Rok1 changes its nucleotide preference to ATP (likely via an external stimulus), which consequently inhibits annealing activity and activates ATPase activity for subsequent pre-ribosome dissociation. It therefore seems that Rok1 is an "inverted" DEAD-box protein: in addition to preferentially binding double-stranded over single-stranded RNA and having RNA-independent ATPase activity, it also requires ATP-dependent release from complexes.

What causes Rok1's nucleotide affinity to switch from ADP to ATP?

DEAD-box proteins generally bind ADP one order of magnitude more tightly than ATP (Karow et al., 2007; Lorsch and Herschlag, 1998; Talavera and De La Cruz, 2005). Since they are classically known as RNA-dependent ATPases, DEAD-box proteins must change their affinity for ATP in order to promote this enzymatic activity. *In vitro* studies with DEAD-box proteins have shown that binding to co-factors is one means by which this affinity can be reversed. For example, the affinity of eIF4A, a DEAD-box protein involved in translation initiation, for ATP is increased ~7-fold in the presence of its binding partner eIF4B, an RNA-binding protein (Bi et al., 2000). At the same time, eIF4B decreases eIF4A's affinity for ADP ~2-fold, thereby contributing to an overall preference of eIF4A for ATP (Bi et al., 2000). In another example, the DEAD-box protein Dbp5, which is required for nuclear mRNA export, preferentially binds ATP over ADP when bound to the nuclear-pore protein co-factor Gle1 and its co-activator InsP6, an endogenous small molecule (Noble et al., 2011; Weirich et al., 2006).

Rok1's prolonged binding to later ribosomal precursors upon ATPase inactivation indicates that this ATP hydrolysis activity is required for Rok1 dissociation from pre-ribosomes. Since Rok1 is preferentially bound to ADP at the steady state, what causes its reversal in affinity from ADP to ATP? Similar to eIF4A and Dbp5, one possibility is that Rok1 interacts with a co-factor that increases its affinity for ATP, decreases its affinity for ADP, or both. We are currently testing if Rrp5 plays this role. An additional

candidate is XrnI, a 5'-3' exoribonuclease that functions in turnover of many RNAs, including the A_2-A_3 fragment (Hiley et al., 2005). Rok1 was originally identified as a high-copy-number suppressor of the XrnI null mutation (Kim et al., 2004). Perhaps A_2 cleavage allows access of XrnI to the 5'-end, leading to an interaction with Rok1 and consequently enhanced ATP binding to Rok1.

Another possible Rok1 co-factor is Rcl1, the endonuclease that cleavages at A_2 (Horn et al., 2011). Since Rok1 likely binds and stabilizes the pre-A_2 duplex until the pre-rRNA conformational change occurs (after cleavage at A_2), perhaps the enzyme that also controls A_2 cleavage then triggers Rok1 for release. Consistent with this proposal, we detect weak interactions between these proteins in pull-downs (data not shown).

Extending in vitro activities to cellular functions

Biochemical experiments have shown that DEAD-box proteins have a variety of activities including RNA duplex unwinding, RNA-protein complex remodeling, RNA duplex annealing and ATP-dependent RNA binding (Jankowsky, 2011; Jankowsky and Fairman, 2007; Jarmoskaite and Russell, 2011; Linder, 2006). In addition to identifying activities of particular DEAD-box proteins, these studies were also devoted to understanding how these RNA-dependent ATPases bind and release RNA in a regulated manner during the ATPase cycle. Further biochemical analyses have shown that co-factors can modulate nucleotide and RNA binding, ATP hydrolysis, RNA duplex unwinding and phosphate and RNA release ((Young and Karbstein, 2012) and references therein), all of which are crucial steps in the DEAD-box protein ATPase cycle. Finally, as demonstrated in biochemical approaches in Chapter 3, co-factors can also increase substrate specificity of a DEAD-box protein. Especially for DEAD-box proteins that function in cellular processes such as pre-rRNA splicing and ribosome assembly, exquisite substrate specificity is likely essential.

Even though all of these elegant studies have led to a better understanding of the general enzymatic activities of DEAD-box proteins, it is impossible to predict whether these *in vitro* activities translate to cellular function without extending the studies *in vivo*. While the current frontier involves developing complete models for the functions of

DEAD-box proteins *in vivo*, progress has been a challenge due to the lack of information on RNA targets or RNA binding sites (Jankowsky, 2011).

One of the most comprehensive functional models for a DEAD-box protein thus far is that of eIF4AIII, a component of the exon junction complex (EJC) in higher eukaryotes (Tange et al., 2004). The EJC is a multi-protein complex that is deposited on the mRNA ~25 nucleotides upstream of the exon-exon junction during splicing (Le Hir et al., 2000a; Le Hir et al., 2000b). This association with the mRNA is important for downstream events such as nonsense-mediated mRNA decay, mRNA localization, mRNA export from the nucleus and the enhancing effect of splicing on translation (Le Hir et al., 2001; Lykke-Andersen et al., 2001; Nott et al., 2004; Palacios et al., 2004). Interestingly, eIF4AIII can unwind duplexes *in vitro* (Li et al., 1999); in the cell, however, its role is to ensure stable association of the EJC core with RNA (Ballut et al., 2005). As part of the complex, the binding partners MAGOH and Y14 inhibit eIF4AIII ATPase activity; in their absence, ATP hydrolysis and subsequent dissociation of the complex from RNA occur (Ballut et al., 2005). Therefore, while *in vitro* experiments suggest that eIF4AIII's primary role is to dissociate an RNA duplex, further studies in the context of components from the EJC and comparative analyses *in vivo* indicate that the principal role of eIF4AIII is to actually stabilize the EJC core-mRNA interaction.

The eIF4AIII studies have indicated that its primary role *in vivo* is contradictory to its observed activities *in vitro*; on the contrary, our analyses of Rok1 have validated its *in vitro* annealing activity *in vivo*. As exemplified by these two cases, extending *in vitro* activities to cellular functions is absolutely essential for completely understanding the roles of DEAD-box proteins and their involvements in RNA metabolism. Most generally, this Rok1 characterization provides an important reminder that thinking of DEAD-box proteins as helicases is simplistic and can be incorrect.

Do other DEAD-box proteins have similar annealing roles?

In ribosome assembly, snoRNAs modify pre-rRNA (specifically methylation and pseudouridylation), inhibit premature formation of rRNA secondary structures and aid in endonucleolytic cleavage events (Karbstein, 2011) by interacting directly with the pre-

rRNA. This requires that snoRNAs bind their target sequences with high specificity. Furthermore, these interacting regions are typically long duplexes that are created from strands that are distant in primary sequence and often competing with local secondary structure. Considering these challenges in addition to the significant number of DEAD-box proteins that are known to have indirect and/or genetic interactions with snoRNAs, it is possible that future studies will implicate a subset of DEAD-box proteins as having annealing roles that are required for snoRNA specificity and function.

Materials and Methods

Rok1 overexpression and purification. As in Chapter 3 (see Materials and Methods), the Rok1 ORF was amplified and cloned into pSV272 and the Rok1 K172A mutant was derived via site-directed mutagenesis. Recombinant wild-type and Rok1 K172A were subsequently expressed in *E. coli* and purified to homogeneity (Chapter 3, Materials and Methods).

RNA labeling. RNAs were labeled as described (Chapter 3, Materials and Methods).

Rok1 ATPase experiments. ATPase assays were carried out as in Chapter 3 (Materials and Methods). To measure Rok1's affinities for ADP vs. ATP, ATPase assays were set-up as described, except 1 µM Rok1 was pre-incubated in the presence of increasing concentrations of either ADP or the non-hydrolyzable ATP analog AMPPNP. For concentrations of AMPPNP greater than 1 mM, an equimolar amount of $MgCl_2$ was added. After pre-incubating at 20°C for 5 minutes, ATPase reactions were initiated by the addition of purified ^{32}P-γ-ATP and conducted as normal (Chapter 3, Materials and Methods). Data was quantified and fit to Eq. 8 (Chapter 3) to determine the K_I, and consequently the $K_{1/2}$, for Rok1's affinities for ADP and AMPPNP. Control experiments using carrying concentrations of Rok1 were carried out to ensure Rok1 was subsaturating with respect to nucleotide concentration. Inhibition experiments with 0.2, 0.5 and 1 µM Rok1 give K_I values of 0.21, 0.23 and 0.26 µM, respectively.

Rok1 annealing experiments. Annealing experiments were conducted as described in Chapter 3 (Materials and Methods). To determine the effects of ATP and ADP on

Rok1's annealing activity, 400 nM Rok1 (wild-type or K172A mutant) was pre-incubated at 20°C for 5 minutes in the presence of increasing concentrations of either ADP/$MgCl_2$ or AMPPNP/$MgCl_2$. Annealing reactions were initiated by the addition of ^{32}P-DownA_2 and unlabeled H44 and conducted as normal (Chapter 3, Materials and Methods). Data was quantified and fit to Eq. 8 (Chapter 3) to determine the K_I for ATP and ADP.

Yeast strains. The galactose-inducible Rok1 strain was created by PCR-based recombination (Longtine et al., 1998) into the *Saccharomyces cerevisiae* strain BY4741 strain (Gal1:Rok1-For and Gal1:Rok1-Rev primers; Table 4). Correct integration was verified by PCR and Western blot analysis. Rok1 was cloned between the XhoI and XbaI sites of pRS416TEF. Rok1 K172A was derived from this plasmid using site-directed mutagenesis. The Utp24 ORF was amplified from *Saccharomyces cerevisiae* genomic DNA and cloned between the SfoI and SalI sites of the copper-inducible plasmid pRS413cup1 (Labbe and Thiele, 1999). Utp24 D138N was derived from this plasmid using site-directed mutagenesis. All primers are listed in Table 4.

Northern blot and primer extension analysis. RNA was isolated from 10-20 OD units of yeast cells via hot phenol/chloroform extraction and concentration was determined via absorbance at 260 nm. For northern blot analysis, 5 µg of RNA was separated over a 1.25% agarose/6.7% formaldehyde gel in 200 mM MOPS, 80 mM $NaC_2H_3O_2$ and 10 mM EDTA, pH 7.0 and passively transferred to an Amersham Hybond-N membrane (GE Healthcare). ^{32}P-labeled primers were used to probe for both precursor (probe b for ITS1 between D and A2, probe a for upstream of cleavage site A_0; probe c for between cleavage sites A_2 and A_3) and mature rRNAs (18S (probe 18S) and 25S (probe y); sequences in Table 4).

For primer extension analysis, gel-purified ^{32}P-labeled primers (A0-A1 to detect A0; 18S to detect A1; sequences in Table 4) were mixed with 2 µg of RNA (additionally purified using Qiagen RNeasy Mini Kit) in 50 mM Tris, pH 8.3, 68 mM KCl and 10 mM DTT, denatured at 95°C for 1 minute and slowly cooled to room temperature to allow complete annealing. Primer extension reactions were done with 20 U of SuperScript III Reverse Transcriptase (Invitrogen) in the presence of 9 mM $MgCl_2$ and 1.3 mM dNTPs. Sequencing lanes were obtained by dideoxy sequencing of rDNA using Sequenase 2.0

(USB). Reactions were separated on an 8% sequencing gel and exposed to a phosphor screen.

DMS structure probing and sequencing gel analysis. Gal::Rok1, Utp24 D138N (cup1) strains with either wild-type Rok1 (pRS416TEF) or pRS416TEF vector were grown in synthetic dextrose (-His) medium and 100 μM $CuSO_4 \bullet 5H_20$ at 30°C for 30 hours. Dimethyl sulfate (DMS) probing was done as in (Lamanna and Karbstein, 2011; Wells et al., 2000). Briefly, cells were treated with 0.5% DMS or ethanol (control sample) and shaken at 30°C for 2 minutes. Reactions were quenched with β-mercaptoethanol and isoamyl alchohol and RNA was extracted via hot phenol/chloroform extraction. Reverse transcription reactions and sequencing lane preparations were carried out as described above using the mid1 primer (see *Northern blot and primer extension analysis*). Primer sequence is in Table 4.

Sucrose density gradients and Western blot analysis. Sucrose density gradients were carried out as described in (Strunk et al., 2011). Briefly, Gal1::Rok1 strains with either wild-type Rok1 (pRS416TEF), Rok1 K172A (pRS416TEF) or pRS416TEF only were grown to mid-log phase in galactose before being switched to glucose. After growth at 30°C for 18 hours (OD_{600} ~0.5), cells were harvested in the presence of 0.1 mg/mL cycloheximide to stall translating ribosomes on mRNA. Cells were resuspended in the presence of protease and RNase inhibitors, frozen with liquid nitrogen and lysed via mechanical grounding using a mortar and pestle. Lysate was cleared using centrifugation and separated over 11 mL 10-50% sucrose gradients. Gradients were scanned by UV absorbance and fractionated using a fraction collector and UV detector from Brandel, Inc. 10 μL of each fraction was separated on an 8% SDS-PAGE gel and transferred to membranes for western detection. Rok1 and Rrp5 were detected with polyclonal antibodies generated against the full-length recombinant proteins in rabbits at Josman, LLC.

Table 4: Oligonucleotides used in Chapter IV

Name	Sequence
Gal1::Rok1-For	5’-CTGCAAAAAGGTACATACGATAGTAGAAGTTATCGAACAGGAATTCGAGCTCGTTAAAC-3’
Gal1::Rok1-Rev	5’-CTTCACGGAAGCTCCTCTAGTTAATACTCTAAAAATATCCATCATTTTGAGATCCGGGTTTT-3’
Rok1-XhoI	5’-TCAGACCTCGAGTTATTTCGAGAAATGTTTTTTTGAAAG-3’
Rok1-XbaI	5’-GATCGATCTAGAATGGATATTTTTAGAGTATTAACTAG-3’
Utp24-SfoI	5’-GATCGAGGCGCCATGGGTAAAGCTAAGAAACAAG-3’
Utp24-SalI	5’-GATCGAGTCGACTTAAAAGACATCTGGCAATTTTTC-3’
Utp24 D138N	5’-CAAGGGTACGTACGCGAATGACTGTTTAGTGC-3’
probe b	5’-GCTCTCATGCTCTTGCC-3’
probe 18S	5’-CATGGCTTAATCTTTGAGAC-3’
probe y	5’-GCCCGTTCCCTTGGCTGT-3’
probe a	5’-CGCTGCTCACCAATGG-3’
probe c	5’-ATGAAAACTCCACAGTG-3’
A0-A1	5’-CCAGATAACTATCTTAAAAG-3’
mid1	5’-GCTCTCATGCTCTTGCC-3’

Chapter V

Closing remarks and future directions

How do DEAD-box proteins achieve RNA substrate specificity?

With more than 300 identified members, RNA helicases are the largest class of enzymes involved in RNA metabolism. Their varying roles include translation initiation, pre-mRNA splicing, mRNA decay and export and ribosome biogenesis. The simple eukaryote *Saccharomyces cerevisiae* has 41 RNA helicases; of these, 25 are DEAD-box proteins, the largest class of superfamily 2 (SF2) RNA helicases, generally known as RNA-dependent ATPases. These proteins bind and release RNA in a regulated manner during the ATPase cycle. Since most DEAD-box proteins are essential, they therefore have non-redundant and consequently substrate specific functions *in vivo*. However, even though DEAD-box proteins were first identified as a family in the late 1980s (Linder et al., 1989), until now, only unconfirmed hypotheses suggested how DEAD-box proteins achieve RNA substrate specificity.

As shown in Chapter 3, our data indicates that the presence of a protein co-factor can increase DEAD-box protein specificity. In the presence of a C-terminal fragment (which contains all motifs involved in protein-protein interactions) of the RNA-binding protein Rrp5, the RNA duplex annealing rate of the DEAD-box protein Rok1 is increased ~15-fold for the pre-A_2 duplex (a duplex that serves as a switch region to regulate 18S rRNA maturation) in comparison to a ~2-fold enhancement of the annealing rate of the reverse complement control duplex. Therefore, not only can co-factors modulate DEAD-box protein activities such as RNA duplex annealing, but they can also increase activity in an RNA sequence specific fashion.

Identifying the RNA substrates of DEAD-box proteins

An outstanding problem in the field is that specific RNA binding sites or RNA target substrates are unknown for most DEAD-box proteins. To circumvent this problem, we began by first mapping the pre-rRNA binding sites of Rrp5, a known co-factor of the DEAD-box protein Rok1. Using structure probing experiments, we identified that Rrp5 has interactions directly at the base of the pre-A_2 duplex. Considering the known *in vitro* activities of other DEAD-box proteins, we therefore pursued Rok1's interactions with this duplex in comparison to control duplexes. Via this approach, we were able to test Rok1's specificity using the predicted *in vivo* substrate. To determine DEAD-box protein substrates using the aforementioned method, co-factors must first be identified. Combining approaches such as yeast two-hybrid, genetic interactions, co-immunoprecipitation and cross-linking can be useful for identifying potential co-factors of DEAD-box proteins; ideally, this interaction should be verified *in vitro* using recombinant proteins. Next, using observed *in vivo* phenotypes (i.e. specific defects in ribosome assembly or mRNA splicing) upon co-factor mutation/deletion, potential target RNAs can be narrowed down. Additionally, RNA-protein cross-linking can be used to identify specific RNA targets (Bohnsack et al., 2009; Granneman et al., 2009). Finally, co-immunoprecipitation of a co-factor and its bound complexes may be useful for identifying RNA substrates. Follow-up structure probing experiments with recombinant protein co-factors and *in vitro* transcribed RNA can confirm specific binding interactions. Once RNA substrates are verified, RNA specific interactions and activities of DEAD-box proteins can be investigated.

Rok1: a unique DEAD-box protein

Our results indicate that Rok1 differs from 'classical' DEAD-box proteins in that it has both RNA-independent ATPase activity and lacks duplex unwinding activity. To provide further insight into these distinctive observations, a structural analysis of Rok1 and its interactions with single-stranded versus double-stranded RNA in the presence or absence of nucleotide would be tremendously useful for comparing and contrasting Rok1 to other well studied DEAD-box proteins with solved structures. Is there something

intrinsic to Rok1 that makes its RecA-like domains less flexible, therefore eliminating the requirement for RNA in order to have robust ATPase activity? As predicted from its 20-fold preference for double-stranded RNA versus single-stranded RNA, does Rok1 make significantly more contacts with double-stranded RNA in comparison to single-stranded RNA? Does Rok1 interact with double-stranded RNA in such a way that there is no kink in the RNA backbone, therefore explaining why local stand separation does not occur?

Rok1 structures solved in the presence of Rrp5C would also be very insightful. Does Rrp5C enhance Rok1 RNA binding and therefore annealing activity by proximally stabilizing the otherwise floppy RecA-like domains with respect to each other? Or does Rrp5C also enhance RNA duplex annealing activity by binding and increasing the local concentration of RNA? How does the Rrp5C-Rok1 structure in the presence of the pre-A_2 duplex compare to that of the reverse complement control duplex? Using the aforementioned structures, functional approaches can then be carried out to verify essential Rok1 residues. For example, perhaps Rok1 residues involved in forming additional contacts with double-stranded RNA can be mutated. Does the resulting construct now have similar affinities for single-stranded and double-stranded RNA? If Rok1 residues that form specific contacts with Rrp5C are mutated, is the enhancement of Rok1 annealing activity eliminated?

Even though DEAD-box protein crystal structures can be infinitely insightful, structures often cannot be solved due to RNA/protein stability, inability to achieve optimal crystallization conditions, etc. In the case of a few questions addressed above, biochemical approaches can answer the question. For example, to test the effects of Rrp5C on Rok1 binding, the Rrp5C construct can be shortened to exclude the three S1 RNA-binding motifs. After verifying that this Rrp5C variant has lost RNA binding activity yet retained the ability to bind Rok1, the affinity of Rok1 in the presence versus absence of the shortened Rrp5C can be compared to see if only the Rok1-Rrp5C interaction is required for enhancing Rok1 binding and therefore annealing activity.

Rok1 truncations can also be biochemically analyzed. Perhaps the two most interesting truncations to start with are the elimination of the N- or C-termini domains. How does this affect RNA binding, RNA duplex annealing activity or interactions with

Rrp5C? Does Rok1 still retain its RNA-independent ATPase activity? Collectively, these biochemical approaches will be quite useful for further characterizing Rok1 and comparing and contrasting it to other well-studied DEAD-box proteins.

Extending in vitro observations to intracellular function

Finally, this work provides an example of the importance of corroborating *in vitro* observations with intracellular functions. Even though other DEAD-box proteins have been shown to have annealing activity, no additional experimentation was conducted to indicate the significance of this activity *in vivo*. Here, we provide preliminary structure probing results that suggest that Rok1 annealing activity is required for the accurate formation of the pre-A_2 duplex *in vivo*. While probing other regions of the pre-rRNA is required to confidently draw this conclusion, these results serve as the first example of an attempt to extend the significance of this annealing activity *in vivo*.

The combination of *in vitro* and *in vivo* observations is also essential for thoroughly understanding the intracellular role and regulation of a DEAD-box protein. Prior to this work, it was known that Rok1 was required for accurate 40S maturation (Venema et al., 1997), but its exact role had yet to be determined. Data in the literature had suggested that Rok1 may be required for removal of snR30 from the pre-ribosome, but whether this was a direct effect or simply reflected assembly defects upstream of snR30 dissociation was unknown (Bohnsack et al., 2008). Furthermore, this latter conclusion lacks an explanation of Rok1's interaction with Rrp5, which is known to be required for events on the opposite side of the pre-ribosome. In addition to providing insight into Rok1's requirement for an interaction with Rrp5, we also provided data that suggested a model for the roles of Rok1's activities in modulating its association with the pre-ribosome. Combining our *in vitro* and *in vivo* results, we propose the following model: ADP-bound Rok1 uses its annealing activity (enhanced by the presence of Rrp5) to anneal the pre-A_2 duplex and consequently associate with the pre-ribosome. When the conformational change (to order the essential A_2- and D-site cleavage events) is ready to occur, Rok1 then changes its preference from ADP to ATP (via an external stimulus) and

uses its ATPase activity to dissociate from the pre-ribosome. ADP-Rok1 can then be recycled for further pre-40S maturation events.

Even though this model nicely incorporates both the *in vivo* and *in vitro* work presented here, there are a few additional aspects that can be tested. First, we have only tested the effects of ADP or ATP on the annealing activity of Rok1 alone. Our model predicts that Rok1 is bound to ADP *in vivo*. If this is true, and if Rrp5 is required for enhancing Rok1 annealing activity intracellularly, then Rrp5 should also enhance the annealing activity of Rok1 in the presence of ADP. This can easily be tested using the gel-shift assays described prior.

We could provide additional insight to our model by testing the Rok1-Rrp5C interaction in the presence of ADP versus ATP. If Rrp5 can enhance the annealing activity of Rok1 in the presence of ADP, then we would predict that Rrp5C also binds ADP-Rok1. This can be tested using an *in vitro* pull-down assay. Results in the absence of nucleotide versus ADP or AMPPNP (non-hydrolyzable ATP analog) can be compared. Perhaps the binding of Rok1 to AMPPNP weakens its association with Rrp5 and contributes to Rok1 dissociation from the pre-ribosome; these pull-down results, in addition to those obtained in the aforementioned experiments, will be insightful in elucidating such details in our proposed model.

References

Abramson, R.D., Dever, T.E., Merrick, W.C. (1988). Biochemical evidence supporting a mechanism for cap-independent and internal initiation of eukaryotic mRNA. *J Biol Chem.* **263**, 6016-6019.

Alcazar-Roman, A.R., Tran, E.J., Guo, S., Wente, S.R. (2006). Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export. *Nat Cell Biol.* **8**, 711-716.

Alexandrov, A., Colognori, D., Steitz, J.A. (2011). Human eIF4AIII interacts with an eIF4G-like partner, NOM1, revealing an evolutionarily conserved function outside the exon junction complex. *Genes Dev.* **25**, 1078-1090.

Ballut, L., Marchadier, B., Baguet, A., Tomasetto, C., Seraphin, B., Le Hir, H. (2005). The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nat Struct Mol Biol.* **12**, 861-869.

Banroques, J., Cordin, O., Doere, M., Linder, P., Tanner, N.K. (2008). A conserved phenylalanine of motif IV in superfamily 2 helicases is required for cooperative, ATP-dependent binding of RNA substrates in DEAD-box proteins. *Mol Cell Biol.* **28**, 3359-3371.

Bi, X., Ren, J., Goss, D.J. (2000). Wheat germ translation initiation factor eIF4B affects eIF4A and eIFiso4F helicase activity by increasing the ATP binding affinity of eIF4A. *Biochemistry.* **39**, 5758-5765.

Bizebard, T., Ferlenghi, I., Iost, I., Dreyfus, M. (2004). Studies on three E. coli DEAD-box helicases point to an unwinding mechanism different from that of model DNA helicases. *Biochemistry.* **43**, 7857-7866.

Bleichert, F., Granneman, S., Osheim, Y.N., Beyer, A.L., Baserga, S.J. (2006). The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc Natl Acad Sci U S A.* **103**, 9464-9469.

Bohnsack, M.T., Kos, M., Tollervey, D. (2008). Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. *EMBO Rep.* **9**, 1230-1236.

Bohnsack, M.T., Martin, R., Granneman, S., Ruprecht, M., Schleiff, E., Tollervey, D. (2009). Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol Cell.* **36**, 583-592.

Bolger, T.A., Wente, S.R. (2011). Gle1 is a multifunctional DEAD-box protein regulator that modulates Ded1 in translation initiation. *J Biol Chem.* **286**, 39750-39759.

Bruckner, A., Polge, C., Lentze, N., Auerbach, D., Schlattner, U. (2009). Yeast two-hybrid, a powerful tool for systems biology. *Int J Mol Sci.* **10**, 2763-2788.

Buckanovich, R.J., Darnell, R.B. (1997). The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol Cell Biol.* **17**, 3194-3201.

Buttner, K., Wenig, K., Hopfner, K.P. (2005). Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol Cell.* **20**, 461-471.

Bycroft, M., Hubbard, T.J., Proctor, M., Freund, S.M., Murzin, A.G. (1997). The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell.* **88**, 235-242.

Carey, M. (1998). The enhanceosome and transcriptional synergy. *Cell.* **92**, 5-8.

Chakrabarti, S., Jayachandran, U., Bonneau, F., Fiorini, F., Basquin, C., Domcke, S., Le Hir, H., Conti, E. (2011). Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol Cell.* **41**, 693-703.

Champion, E.A., Lane, B.H., Jackrel, M.E., Regan, L., Baserga, S.J. (2008). A direct interaction between the Utp6 half-a-tetratricopeptide repeat domain and a specific peptide in Utp21 is essential for efficient pre-rRNA processing. *Mol Cell Biol.* **28**, 6547-6556.

Charette, J.M., Baserga, S.J. (2010). The DEAD-box RNA helicase-like Utp25 is an SSU processome component. *RNA.* **16**, 2156-2169.

Charollais, J., Pflieger, D., Vinh, J., Dreyfus, M., Iost, I. (2003). The DEAD-box RNA helicase SrmB is involved in the assembly of 50S ribosomal subunits in Escherichia coli. *Mol Microbiol.* **48**, 1253-1265.

Clemens, K.R., Wolf, V., McBryant, S.J., Zhang, P., Liao, X., Wright, P.E., Gottesfeld, J.M. (1993). Molecular basis for specific recognition of both RNA and DNA by a zinc finger protein. *Science.* **260**, 530-533.

Coppolecchia, R., Buser, P., Stotz, A., Linder, P. (1993). A new yeast translation initiation factor suppresses a mutation in the eIF-4A RNA helicase. *EMBO J.* **12**, 4005-4011.

de Boer, P., Vos, H.R., Faber, A.W., Vos, J.C., Raue, H.A. (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, E.A., Herschlag, D., Doudna, J.A. (1999). Assembly of an exceptionally stable RNA tertiary interface in a group I ribozyme. *Biochemistry.* **38**, 2982-2990.

Dominguez, D., Altmann, M., Benz, J., Baumann, U., Trachsel, H. (1999). Interaction of translation initiation factor eIF4G with eIF4A in the yeast Saccharomyces cerevisiae. *J Biol Chem.* **274**, 26720-26726.

Eppens, N.A., Faber, A.W., Rondaij, M., Jahangir, R.S., van Hemert, S., Vos, J.C., Venema, J., Raue, H.A. (2002). Deletions in the S1 domain of Rrp5p cause processing at a novel site in ITS1 of yeast pre-rRNA that depends on Rex4p. *Nucleic Acids Res.* **30**, 4222-4231.

Eppens, N.A., Rensen, S., Granneman, S., Raue, H.A., 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.

Evguenieva-Hackenberg, E., Roppelt, V., Finsterseifer, P., Klug, G. (2008). Rrp4 and Csl4 are needed for efficient degradation but not for polyadenylation of synthetic and natural RNA by the archaeal exosome. *Biochemistry.* **47**, 13158-13168.

Fairman-Williams, M.E., Guenther, U.P., Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol.* **20**, 313-324.

Fairman, M.E., Maroney, P.A., Wang, W., Bowers, H.A., Gollnick, P., Nilsen, T.W., Jankowsky, E. (2004). Protein displacement by DExH/D "RNA helicases" without duplex unwinding. *Science.* **304**, 730-734.

Fuller-Pace, F.V., Nicol, S.M., Reid, A.D., Lane, D.P. (1993). DbpA: a DEAD box protein specifically activated by 23s rRNA. *EMBO J.* **12**, 3619-3626.

Galan, J.M., Moreau, V., Andre, B., Volland, C., Haguenauer-Tsapis, R. (1996). Ubiquitination mediated by the Npi1p/Rsp5p ubiquitin-protein ligase is required for endocytosis of the yeast uracil permease. *J Biol Chem.* **271**, 10946-10952.

Garcia-Mayoral, M.F., 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.

Garcia-Mayoral, M.F., Hollingworth, D., Masino, L., Diaz-Moreno, I., Kelly, G., Gherzi, R., Chou, C.F., Chen, C.Y., 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, I., Uhlenbeck, O.C. (2008). Differential RNA-dependent ATPase activities of four rRNA processing yeast DEAD-box proteins. *Biochemistry.* **47**, 12562-12573.

Gherzi, R., Lee, K.Y., Briata, P., Wegmuller, D., Moroni, C., Karin, M., Chen, C.Y. (2004). A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Mol Cell.* **14**, 571-583.

Granneman, S., Kudla, G., Petfalski, E., Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A.* **106**, 9613-9618.

Granneman, S., Lin, C., Champion, E.A., Nandineni, M.R., Zorca, C., Baserga, S.J. (2006). The nucleolar protein Esf2 interacts directly with the DExD/H box RNA helicase, Dbp8, to stimulate ATP hydrolysis. *Nucleic Acids Res.* **34**, 3189-3199.

Halls, C., Mohr, S., Del Campo, M., Yang, Q., Jankowsky, E., Lambowitz, A.M. (2007). Involvement of DEAD-box proteins in group I and group II intron splicing. Biochemical characterization of Mss116p, ATP hydrolysis-dependent and -independent mechanisms, and general RNA chaperone activity. *J Mol Biol.* **365**, 835-855.

He, F., Brown, A.H., Jacobson, A. (1997). Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Mol Cell Biol.* **17**, 1580-1594.

Hein, C., Springael, J.Y., Volland, C., Haguenauer-Tsapis, R., Andre, B. (1995). NPl1, an essential yeast gene involved in induced degradation of Gap1 and Fur4 permeases, encodes the Rsp5 ubiquitin-protein ligase. *Mol Microbiol.* **18**, 77-87.

Henn, A., Cao, W., Hackney, D.D., De La Cruz, E.M. (2008). The ATPase cycle mechanism of the DEAD-box rRNA helicase, DbpA. *J Mol Biol.* **377**, 193-205.

Henry, Y., Wood, H., Morrissey, J.P., 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.

Hettema, E.H., Valdez-Taubas, J., Pelham, H.R. (2004). Bsd2 binds the ubiquitin ligase Rsp5 and mediates the ubiquitination of transmembrane proteins. *EMBO J.* **23**, 1279-1288.

Hicke, L. (1999). Gettin' down with ubiquitin: turning off cell-surface receptors, transporters and channels. *Trends Cell Biol.* **9**, 107-112.

Hilbert, M., Kebbel, F., Gubaev, A., Klostermeier, D. (2011). eIF4G stimulates the activity of the DEAD box protein eIF4A by a conformational guidance mechanism. *Nucleic Acids Res.* **39**, 2260-2270.

Hiley, S.L., Babak, T., Hughes, T.R. (2005). Global analysis of yeast RNA processing identifies new targets of RNase III and uncovers a link between tRNA 5' end processing and tRNA splicing. *Nucleic Acids Res.* **33**, 3048-3056.

Hodge, C.A., Colot, H.V., Stafford, P., Cole, C.N. (1999). Rat8p/Dbp5p is a shuttling transport factor that interacts with Rat7p/Nup159p and Gle1p and suppresses the mRNA export defect of xpo1-1 cells. *EMBO J.* **18**, 5778-5788.

Hodge, C.A., Tran, E.J., Noble, K.N., Alcazar-Roman, A.R., Ben-Yishay, R., Scarcelli, J.J., Folkmann, A.W., Shav-Tal, Y., Wente, S.R., Cole, C.N. (2011). The Dbp5 cycle at the nuclear pore complex during mRNA export I: dbp5 mutants with defects in RNA binding and ATP hydrolysis define key steps for Nup159 and Gle1. *Genes Dev.* **25**, 1052-1064.

Horn, D.M., Mason, S.L., Karbstein, K. (2011). Rcl1 protein, a novel nuclease for 18 S ribosomal RNA production. *J Biol Chem.* **286**, 34082-34087.

Iost, I., Dreyfus, M. (2006). DEAD-box RNA helicases in Escherichia coli. *Nucleic Acids Res.* **34**, 4189-4197.

Iost, I., Dreyfus, M., Linder, P. (1999). Ded1p, a DEAD-box protein required for translation initiation in Saccharomyces cerevisiae, is an RNA helicase. *J Biol Chem.* **274**, 17677-17683.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. *Trends Biochem Sci.* **36**, 19-29.

Jankowsky, E., Bowers, H. (2006). Remodeling of ribonucleoprotein complexes with DExH/D RNA helicases. *Nucleic Acids Res.* **34**, 4181-4188.

Jankowsky, E., Fairman, M.E. (2007). RNA helicases--one fold for many functions. *Curr Opin Struct Biol.* **17**, 316-324.

Jankowsky, E., Putnam, A. (2009). *Helicases*. Springer-Verlag New York, LLC.

Jarmoskaite, I., Russell, R. (2011). DEAD-box proteins as RNA helicases and chaperones. *Wiley Interdiscip Rev RNA.* **2**, 135-152.

Kaminishi, T., Wilson, D.N., Takemoto, C., Harms, J.M., 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. (2011). Inside the 40S ribosome assembly machinery. *Curr Opin Chem Biol.* **15**, 657-663.

Karbstein, K., Carroll, K.S., Herschlag, D. (2002). Probing the Tetrahymena group I ribozyme reaction in both directions. *Biochemistry.* **41**, 11171-11183.

Karbstein, K., Jonas, S., Doudna, J.A. (2005). An essential GTPase promotes assembly of preribosomal RNA processing complexes. *Mol Cell.* **20**, 633-643.

Karow, A.R., Theissen, B., Klostermeier, D. (2007). Authentic interdomain communication in an RNA helicase reconstituted by expressed protein ligation of two helicase domains. *FEBS J.* **274**, 463-473.

Kikuma, T., Ohtsu, M., Utsugi, T., Koga, S., Okuhara, K., Eki, T., Fujimori, F., Murakami, Y. (2004). Dbp9p, a member of the DEAD box protein family, exhibits DNA helicase activity. *J Biol Chem.* **279**, 20692-20698.

Kim, J., Jeon, S., Yang, Y.S. (2004). Posttranscriptional regulation of the karyogamy gene by Kem1p/Xrn1p exoribonuclease and Rok1p RNA helicase of Saccharomyces cerevisiae. *Biochem Biophys Res Commun.* **321**, 1032-1039.

Kos, M., Tollervey, D. (2005). The Putative RNA Helicase Dbp4p Is Required for Release of the U14 snoRNA from Preribosomes in Saccharomyces cerevisiae. *Mol Cell.* **20**, 53-64.

Kos, M., Tollervey, D. (2010). Yeast pre-rRNA processing and modification occur cotranscriptionally. *Mol Cell.* **37**, 809-820.

Kossen, K., Karginov, F.V., Uhlenbeck, O.C. (2002). The carboxy-terminal domain of the DExDH protein YxiN is sufficient to confer specificity for 23S rRNA. *J Mol Biol.* **324**, 625-636.

Kuras, L., Cherest, H., Surdin-Kerjan, Y., Thomas, D. (1996). A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism. *EMBO J.* **15**, 2519-2529.

Labbe, S., Thiele, D.J. (1999). Copper ion inducible and repressible promoter systems in yeast. *Methods Enzymol.* **306**, 145-153.

Lamanna, A.C., 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 U S A.* **106**, 14259-14264.

Lamanna, A.C., Karbstein, K. (2010). A Conformational Switch Regulates Pre-18S Cleavage. *J Mol Biol.* **in the press**.

Lamanna, A.C., Karbstein, K. (2011). An RNA conformational switch regulates pre-18S rRNA cleavage. *J Mol Biol.* **405**, 3-17.

Lamb, J.R., Tugendreich, S., Hieter, P. (1995). Tetratrico peptide repeat interactions: to TPR or not to TPR? *Trends Biochem Sci.* **20**, 257-259.

Last, R.L., Maddock, J.R., Woolford, J.L., Jr. (1987). Evidence for related functions of the RNA genes of Saccharomyces cerevisiae. *Genetics.* **117**, 619-631.

Lawson, T.G., Lee, K.A., Maimone, M.M., Abramson, R.D., Dever, T.E., Merrick, W.C., Thach, R.E. (1989). Dissociation of double-stranded polynucleotide helical structures by eukaryotic initiation factors, as revealed by a novel assay. *Biochemistry.* **28**, 4729-4734.

Le Hir, H., Andersen, G.R. (2008). Structural insights into the exon junction complex. *Curr Opin Struct Biol.* **18**, 112-119.

Le Hir, H., Gatfield, D., Izaurralde, E., Moore, M.J. (2001). The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J.* **20**, 4987-4997.

Le Hir, H., Izaurralde, E., Maquat, L.E., Moore, M.J. (2000a). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J.* **19**, 6860-6869.

Le Hir, H., Moore, M.J., Maquat, L.E. (2000b). Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* **14**, 1098-1108.

Lebaron, S., Papin, C., Capeyrou, R., Chen, Y.L., Froment, C., Monsarrat, B., Caizergues-Ferrer, M., Grigoriev, M., Henry, Y. (2009). The ATPase and helicase activities of Prp43p are stimulated by the G-patch protein Pfa1p during yeast ribosome biogenesis. *EMBO J.* **28**, 3808-3819.

Lebreton, A., Rousselle, J.C., 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.

Leon, S., Erpapazoglou, Z., Haguenauer-Tsapis, R. (2008). Ear1p and Ssh4p are new adaptors of the ubiquitin ligase Rsp5p for cargo ubiquitylation and sorting at multivesicular bodies. *Mol Biol Cell.* **19**, 2379-2388.

Li, Q., Imataka, H., Morino, S., Rogers, G.W., Jr., Richter-Cook, N.J., Merrick, W.C., Sonenberg, N. (1999). Eukaryotic translation initiation factor 4AIII (eIF4AIII) is functionally distinct from eIF4AI and eIF4AII. *Mol Cell Biol.* **19**, 7336-7346.

Liang, W.Q., Clark, J.A., Fournier, M.J. (1997). The rRNA-processing function of the yeast U14 small nucleolar RNA can be rescued by a conserved RNA helicase-like protein. *Mol Cell Biol.* **17**, 4124-4132.

Liao, X.B., Clemens, K.R., Tennant, L., Wright, P.E., Gottesfeld, J.M. (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.

Linder, P. (2006). Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Res.* **34**, 4168-4180.

Linder, P., Lasko, P.F., Ashburner, M., Leroy, P., Nielsen, P.J., Nishi, K., Schnier, J., Slonimski, P.P. (1989). Birth of the D-E-A-D box. *Nature.* **337**, 121-122.

Linder, P., Tanner, N.K., Banroques, J. (2001). From RNA helicases to RNPases. *Trends Biochem Sci.* **26**, 339-341.

Liu, F., Putnam, A., Jankowsky, E. (2008). ATP hydrolysis is required for DEAD-box protein recycling but not for duplex unwinding. *Proc Natl Acad Sci U S A.* **105**, 20209-20214.

Longtine, M.S., McKenzie, A., 3rd, Demarini, D.J., Shah, N.G., Wach, A., Brachat, A., Philippsen, P., Pringle, J.R. (1998). Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast.* **14**, 953-961.

Lorsch, J.R., Herschlag, D. (1998). The DEAD box protein eIF4A. 1. A minimal kinetic and thermodynamic framework reveals coupled binding of RNA and nucleotide. *Biochemistry.* **37**, 2180-2193.

Lunde, B.M., Moore, C., Varani, G. (2007). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol.* **8**, 479-490.

Lykke-Andersen, J., Shu, M.D., Steitz, J.A. (2001). Communication of the position of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1. *Science.* **293**, 1836-1839.

Maeder, C., Kutach, A.K., Guthrie, C. (2009). ATP-dependent unwinding of U4/U6 snRNAs by the Brr2 helicase requires the C terminus of Prp8. *Nat Struct Mol Biol.* **16**, 42-48.

Martin, A., Schneider, S., Schwer, B. (2002). Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem.* **277**, 17743-17750.

Matsumoto, Y., Xu, Q., Miyazaki, S., Kaito, C., Farr, C.L., Axelrod, H.L., Chiu, H.J., Klock, H.E., Knuth, M.W., Miller, M.D., Elsliger, M.A., Deacon, A.M., Godzik, A., Lesley, S.A., Sekimizu, K., Wilson, I.A. (2010). Structure of a virulence regulatory factor CvfB reveals a novel winged helix RNA binding module. *Structure.* **18**, 537-547.

Merika, M., Thanos, D. (2001). Enhanceosomes. *Curr Opin Genet Dev.* **11**, 205-208.

Mohr, G., Del Campo, M., Mohr, S., Yang, Q., Jia, H., Jankowsky, E., Lambowitz, A.M. (2008). Function of the C-terminal domain of the DEAD-box protein Mss116p analyzed in vivo and in vitro. *J Mol Biol.* **375**, 1344-1364.

Montpetit, B., Thomsen, N.D., Helmke, K.J., Seeliger, M.A., Berger, J.M., Weis, K. (2011). A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature.* **472**, 238-242.

Morvan, J., Froissard, M., Haguenauer-Tsapis, R., Urban-Grimal, D. (2004). The ubiquitin ligase Rsp5p is required for modification and sorting of membrane proteins into multivesicular bodies. *Traffic.* **5**, 383-392.

Mumberg, D., Muller, R., Funk, M. (1995). Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene.* **156**, 119-122.

Murzin, A.G. (1993). OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* **12**, 861-867.

Musunuru, K., Darnell, R.B. (2004). Determination and augmentation of RNA sequence specificity of the Nova K-homology domains. *Nucleic Acids Res.* **32**, 4852-4861.

Nissan, T.A., Bassler, J., Petfalski, E., Tollervey, D., Hurt, E. (2002). 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO J.* **21**, 5539-5547.

Noble, K.N., Tran, E.J., Alcazar-Roman, A.R., Hodge, C.A., Cole, C.N., Wente, S.R. (2011). The Dbp5 cycle at the nuclear pore complex during mRNA export II: nucleotide cycling and mRNP remodeling by Dbp5 are controlled by Nup159 and Gle1. *Genes Dev.* **25**, 1065-1077.

Nott, A., Le Hir, H., Moore, M.J. (2004). Splicing enhances translation in mammalian cells: an additional function of the exon junction complex. *Genes Dev.* **18**, 210-222.

Oh, J.Y., Kim, J. (1999). ATP hydrolysis activity of the DEAD box protein Rok1p is required for in vivo ROK1 function. *Nucleic Acids Res.* **27**, 2753-2759.

Ozes, A.R., Feoktistova, K., Avanzino, B.C., Fraser, C.S. (2011). Duplex Unwinding and ATPase Activities of the DEAD-Box Helicase eIF4A Are Coupled by eIF4G and eIF4B. *J Mol Biol.* **412**, 674-687.

Palacios, I.M., Gatfield, D., St Johnston, D., Izaurralde, E. (2004). An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature.* **427**, 753-757.

Pena, V., Jovin, S.M., Fabrizio, P., Orlowski, J., Bujnicki, J.M., Luhrmann, R., Wahl, M.C. (2009). Common design principles in the spliceosomal RNA helicase Brr2 and in the Hel308 DNA helicase. *Mol Cell.* **35**, 454-466.

Portnoy, V., Evguenieva-Hackenberg, E., Klein, F., Walter, P., Lorentzen, E., Klug, G., Schuster, G. (2005). RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polynucleotidylates RNA in Sulfolobus. *EMBO Rep.* **6**, 1188-1193.

Proux, F., Dreyfus, M., Iost, I. (2011). Identification of the sites of action of SrmB, a DEAD-box RNA helicase involved in Escherichia coli ribosome assembly. *Mol Microbiol.*

Rocak, S., Emery, B., Tanner, N.K., Linder, P. (2005). Characterization of the ATPase and unwinding activities of the yeast DEAD-box protein Has1p and the analysis of the roles of the conserved motifs. *Nucleic Acids Res.* **33**, 999-1009.

Rogers, G.W., Jr., Richter, N.J., Lima, W.F., Merrick, W.C. (2001). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem.* **276**, 30914-30922.

Rogers, G.W., Jr., Richter, N.J., Merrick, W.C. (1999). Biochemical and kinetic characterization of the RNA helicase activity of eukaryotic initiation factor 4A. *J Biol Chem.* **274**, 12236-12244.

Roppelt, V., Klug, G., Evguenieva-Hackenberg, E. (2010). The evolutionarily conserved subunits Rrp4 and Csl4 confer different substrate specificities to the archaeal exosome. *FEBS Lett.* **584**, 2931-2936.

Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C., Sonenberg, N. (1990). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol.* **10**, 1134-1144.

Schutz, P., Bumann, M., Oberholzer, A.E., Bieniossek, C., Trachsel, H., Altmann, M., Baumann, U. (2008). Crystal structure of the yeast eIF4A-eIF4G complex: an RNA-helicase controlled by protein-protein interactions. *Proc Natl Acad Sci U S A.* **105**, 9564-9569.

Schutz, P., Karlberg, T., van den Berg, S., Collins, R., Lehtio, L., Hogbom, M., Holmberg-Schiavone, L., Tempel, W., Park, H.W., Hammarstrom, M., Moche, M., Thorsell, A.G., Schuler, H. (2010). Comparative structural analysis of human DEAD-box RNA helicases. *PLoS One.* **5**.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., Cate, J.H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. *Science.* **310**, 827-834.

Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S., Yokoyama, S. (2006). Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. *Cell.* **125**, 287-300.

Sharpe Elles, L.M., Sykes, M.T., Williamson, J.R., Uhlenbeck, O.C. (2009). A dominant negative mutant of the E. coli RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.* **37**, 6503-6514.

Siggers, T., Duyzend, M.H., Reddy, J., Khan, S., Bulyk, M.L. (2011). Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. *Mol Syst Biol.* **7**, 555.

Silverman, E.J., Maeda, A., Wei, J., Smith, P., Beggs, J.D., Lin, R.J. (2004). Interaction between a G-patch protein and a spliceosomal DEXD/H-box ATPase that is critical for splicing. *Mol Cell Biol.* **24**, 10101-10110.

Song, Y., Kim, S., Kim, J. (1995). ROK1, a high-copy-number plasmid suppressor of kem1, encodes a putative ATP-dependent RNA helicase in Saccharomyces cerevisiae. *Gene.* **166**, 151-154.

Stern, S., Moazed, D., Noller, H.F. (1988). Structural analysis of RNA using chemical and enzymatic probing monitored by primer extension. *Methods Enzymol.* **164**, 481-489.

Strahm, Y., Fahrenkrog, B., Zenklusen, D., Rychner, E., Kantor, J., Rosbach, M., Stutz, F. (1999). The RNA export factor Gle1p is located on the cytoplasmic fibrils of the NPC and physically interacts with the FG-nucleoporin Rip1p, the DEAD-box protein Rat8p/Dbp5p and a new protein Ymr 255p. *EMBO J.* **18**, 5761-5777.

Strunk, B.S., Karbstein, K. (2009). Powering through ribosome assembly. *RNA.* **15**, 2083-2104.

Strunk, B.S., Loucks, C.R., Su, M., Vashisth, H., Cheng, S., Schilling, J., Brooks, C.L., 3rd, Karbstein, K., Skiniotis, G. (2011). Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science.* **333**, 1449-1453.

Subramanian, A.R. (1983). Structure and functions of ribosomal protein S1. *Prog Nucleic Acid Res Mol Biol.* **28**, 101-142.

Suchanek, M., Radzikowska, A., Thiele, C. (2005). Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat Methods.* **2**, 261-267.

Talavera, M.A., De La Cruz, E.M. (2005). Equilibrium and kinetic analysis of nucleotide binding to the DEAD-box RNA helicase DbpA. *Biochemistry.* **44**, 959-970.

Tanaka, N., Aronova, A., Schwer, B. (2007). Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev.* **21**, 2312-2325.

Tange, T.O., Nott, A., Moore, M.J. (2004). The ever-increasing complexities of the exon junction complex. *Curr Opin Cell Biol.* **16**, 279-284.

Theobald, D.L., Mitton-Fry, R.M., Wuttke, D.S. (2003). Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct.* **32**, 115-133.

Theobald, U., Mailinger, W., Baltes, M., Rizzi, M., Reuss, M. (1997). In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae : I. Experimental observations. *Biotechnol Bioeng.* **55**, 305-316.

Tijerina, P., Bhaskaran, H., Russell, R. (2006). Nonspecific binding to structured RNA and preferential unwinding of an exposed helix by the CYT-19 protein, a DEAD-box RNA chaperone. *Proc Natl Acad Sci U S A.* **103**, 16698-16703.

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.

Tsai, R.T., Fu, R.H., Yeh, F.L., Tseng, C.K., Lin, Y.C., Huang, Y.H., Cheng, S.C. (2005). Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev.* **19**, 2991-3003.

Tsu, C.A., Uhlenbeck, O.C. (1998). Kinetic analysis of the RNA-dependent adenosinetriphosphatase activity of DbpA, an Escherichia coli DEAD protein specific for 23S ribosomal RNA. *Biochemistry.* **37**, 16989-16996.

van Nues, R.W., Beggs, J.D. (2001). Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. *Genetics.* **157**, 1451-1467.

Venema, J., Bousquet-Antonelli, C., Gelugne, J.P., Caizergues-Ferrer, M., Tollervey, D. (1997). Rok1p is a putative RNA helicase required for rRNA processing. *Mol Cell Biol.* **17**, 3398-3407.

Venema, J., Tollervey, D. (1996). RRP5 is required for formation of both 18S and 5.8S rRNA in yeast. *EMBO J.* **15**, 5701-5714.

Venema, J., Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. *Annu Rev Genet.* **33**, 261-311.

von Moeller, H., Basquin, C., Conti, E. (2009). The mRNA export protein DBP5 binds RNA and the cytoplasmic nucleoporin NUP214 in a mutually exclusive manner. *Nat Struct Mol Biol.* **16**, 247-254.

Vos, H.R., Bax, R., Faber, A.W., Vos, J.C., Raue, H.A. (2004a). U3 snoRNP and Rrp5p associate independently with Saccharomyces cerevisiae 35S pre-rRNA, but Rrp5p is essential for association of Rok1p. *Nucleic Acids Res.* **32**, 5827-5833.

Vos, H.R., Faber, A.W., de Gier, M.D., Vos, J.C., Raue, H.A. (2004b). 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.

Walbott, H., Mouffok, S., Capeyrou, R., Lebaron, S., Humbert, O., van Tilbeurgh, H., Henry, Y., Leulliot, N. (2010). Prp43p contains a processive helicase structural architecture with a specific regulatory domain. *EMBO J.* **29**, 2194-2204.

Walter, P., Klein, F., Lorentzen, E., Ilchmann, A., Klug, G., Evguenieva-Hackenberg, E. (2006). Characterization of native and reconstituted exosome complexes from the hyperthermophilic archaeon Sulfolobus solfataricus. *Mol Microbiol.* **62**, 1076-1089.

Wang, S., Hu, Y., Overgaard, M.T., Karginov, F.V., Uhlenbeck, O.C., McKay, D.B. (2006). The domain of the Bacillus subtilis DEAD-box helicase YxiN that is responsible for specific binding of 23S rRNA has an RNA recognition motif fold. *RNA.* **12**, 959-967.

Warkocki, Z., Odenwalder, P., Schmitzova, J., Platzmann, F., Stark, H., Urlaub, H., Ficner, R., Fabrizio, P., Luhrmann, R. (2009). Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. *Nat Struct Mol Biol.* **16**, 1237-1243.

Weirich, C.S., Erzberger, J.P., Flick, J.S., Berger, J.M., Thorner, J., Weis, K. (2006). Activation of the DExD/H-box protein Dbp5 by the nuclear-pore protein Gle1 and its coactivator InsP6 is required for mRNA export. *Nat Cell Biol.* **8**, 668-676.

Wells, S.E., Hughes, J.M., Igel, A.H., Ares, M., Jr. (2000). Use of dimethyl sulfate to probe RNA structure in vivo. *Methods Enzymol.* **318**, 479-493.

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.

Yang, Q., Jankowsky, E. (2005). ATP- and ADP-dependent modulation of RNA unwinding and strand annealing activities by the DEAD-box protein DED1. *Biochemistry.* **44**, 13591-13601.

Yang, Q., Jankowsky, E. (2006). The DEAD-box protein Ded1 unwinds RNA duplexes by a mode distinct from translocating helicases. *Nat Struct Mol Biol.* **13**, 981-986.

Young, C.L., Karbstein, K. (2011). The roles of S1 RNA-binding domains in Rrp5's interactions with pre-rRNA. *RNA.* **17**, 512-521.

Young, C.L., Karbstein, K. (2012). *RNA Helicases*. Elsevier.

Zhang, L., Xu, T., Maeder, C., Bud, L.O., Shanks, J., Nix, J., Guthrie, C., Pleiss, J.A., Zhao, R. (2009). Structural evidence for consecutive Hel308-like modules in the spliceosomal ATPase Brr2. *Nat Struct Mol Biol.* **16**, 731-739.