UCLA
UCLA Electronic Theses and Dissertations

Title
The Control of Gene Expression by Nuclear RNA Degradation in Saccharomyces cerevisiae

Permalink
https://escholarship.org/uc/item/9769m7mz

Author
Roy, Kevin Richard Jones

Publication Date
2015

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA

Los Angeles

The Control of Gene Expression by Nuclear RNA Degradation in *Saccharomyces cerevisiae*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Molecular Biology

by

Kevin Richard Jones Roy

2015

© Copyright by

Kevin Richard Jones Roy

2015

ABSTRACT OF THE DISSERTATION

The Control of Gene Expression by Nuclear RNA Degradation in *Saccharomyces cerevisiae*

by

Kevin Richard Jones Roy

Doctor of Philosophy in Molecular Biology

University of California, Los Angeles, 2015

Professor Guillaume Chanfreau, Chair

Ribonucleases play critical roles in controlling the quantity and quality of gene expression through processing and degrading RNA. An important class of evolutionarily conserved ribonucleases is the RNase III family of enzymes, which are distinguished by their specificity for cleaving double-stranded RNA (dsRNA). RNase III enzymes perform diverse functions in RNA metabolism in all eukaryotes studied, yet numerous questions remain regarding their range of natural targets *in vivo*, how they achieve substrate specificity, and how their cleavage activity is regulated. The model eukaryote *Saccharomyces cerevisiae* harbors one RNase III homolog, Rnt1p, which is responsible for all known dsRNA cleavage activity in this organism. To better understand the substrate selectivity of Rnt1p, we examined how its double-stranded RNA binding domain (dsRBD) recognizes a non-canonical substrate containing an AAGU tetraloop

sequence differing from the NGNN consensus sequence. Surprisingly, we found that upon engaging the RNA, the dsRBD induces a structural change in the AAGU loop so that it closely adopts the structure of the NGNN loop. This suggested that the structures of isolated RNAs in solution are not necessarily predictive of substrate specificity. We next characterized how structural dynamics in the dsRBD mediate specific binding. We found that in order to bind substrate dsRNA with high affinity, the dsRBD must undergo a significant conformational change involving the first alpha helix and beta strand of the dsRBD. Next we implemented computational RNA secondary structure screens to scan the genome for potential Rnt1p targets. We identified a characteristic Rnt1p stem-loop in the *BDF2* mRNA, which is also subject to nuclear decay by the spliceosome through a first step splicing discard pathway. *Cis* acting mutations in *BDF2* blocking Rnt1p or spliceosome-mediated decay (SMD) conferred distinct phenotypes for each pathway, revealing that salt stress hyper-activates Rnt1p cleavage while spliceosome-mediated decay controls *BDF2* expression during DNA replication stress. To globally identify RNA targets of Rnt1p cleavage, we leveraged the fact that the 5′ product of Rnt1p cleavage is oligo-adenylated by Trf4/5-Air2/1-Mtr4 polyadenylation (TRAMP) complex prior to degradation by the nuclear exosome, a 3′-to-5′ exonuclease complex. We mapped TRAMP poly(A) tails genome-wide by high-throughput sequencing of 3′ ends of polyadenylated RNA in yeast cells lacking a nuclear exosome component. This revealed a global profile of destabilized 3′ ends arising from various nuclear RNA degradation mechanisms, including Rnt1p cleavage, transcription termination by the Nrd1p-Nab3p-Sen1p (NNS) pathway and roadblock transcription termination by Reb1p and TFIIIB DNA binding factors. While the NNS pathway was known to play a prominent role in limiting pervasive RNA polymerase II, we uncovered previously unappreciated roles for roadblocks and Rnt1p in controlling Pol II transcriptional output throughout the genome, revealing how cells use a multitude of nuclear mechanisms to regulate the levels of coding and cryptic transcripts.

The dissertation of Kevin Richard Jones Roy is approved.

Feng Guo

Elizabeta Nemeth

Guillaume Chanfreau, Committee Chair

University of California, Los Angeles

2015

To my friends and family for their unwavering support and encouragement

TABLE OF CONTENTS **PAGE**

ABSTRACT…… ii

COMMITTEE PAGE…….. iv

LIST OF FIGURES……..…. vii

ACKNOWLEDGEMENTS……..…. xii

VITA…….. xv

CHAPTER 1: Introduction to nuclear RNA processing and degradation pathways in *Saccharomyces cerevisiae*……..1

CHAPTER 2: The diverse functions of fungal RNase III enzymes in RNA metabolism [Article Reprint]…….. 27

CHAPTER 3: Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs [Article Reprint]……51

CHAPTER 4: Intrinsic dynamics of an extended hydrophobic core in the *S. cerevisiae* RNase III dsRBD contributes to recognition of specific RNA binding sites [Article Reprint]……..……..……..64

CHAPTER 5: Stress-induced nuclear RNA degradation pathways regulate yeast bromodomain factor 2 to promote cell survival [Article Reprint]...…….. 82

CHAPTER 6: Uncovering the dark side of poly(A) tails reveals genome-wide mechanisms of nuclear RNA processing and decay ...……...……...104

LIST OF FIGURES **PAGE**

Figure 1.1 – Quality control and regulated degradation of messenger RNA precursors…….. 14

Figure 1.2 – Diverse roles for the nuclear 5´-3´ exonuclease Rat1p in gene regulation……...15

Figure 1.3 – Diverse roles for the nuclear 3´-5´ exosome complex in RNA processing and quality control…..…….. 16

Figure 1.4 – Architecture of the nuclear RNA exosome……..……..17

Figure 1.5 – Coupling of the nuclear exosome to RNA processing pathways..…….. 18

Figure 2.1 – Architecture and evolutionary relationships of fungal RNase III representatives……...…… 30

Figure 2.2 – Cotranscriptional rRNA processing is coupled to Pol I termination…….. 33

Figure 2.3 – Small nuclear RNA processing by RNase III Rnt1p……..…….. 35

Figure 2.4 – Small nucleolar RNA processing by RNase III Rnt1p……..…… 38

Figure 2.5 – RNase III regulation of unspliced pre-mRNAs and lariat degradation……..…… 42

Figure 2.6 – Fail-safe transcription termination of RNA polymerase II by RNase III..…..…… 44

Figure 3.1 – snR47-Derived dsRNA Capped by an AAGU Tetraloop Binds to the Rnt1p dsRBD……..……..…… 54

Figure 3.2 – dsRNA Capped by an AAGU Tetraloop Is an Efficient Substrate for Rnt1p..…..… 54

Figure 3.3 – Solution Structure of the Rnt1p dsRBD/AAGU Complex..…..…… 55

Figure 3.4 – Correlation Plots between Experimental and Back-Calculated RDCs for the dsRBD/AAGU and dsRBD/AGAA Hairpin Complexes…… 57

Figure 3.5 – Comparison of the Structures of the AAGU and AGAA Tetraloops in the Free and dsRBD-Bound States..…..…….. 58

Figure 3.6 – Interactions between the Rnt1p dsRBD and RNA..…..…….. 59

Figure 3.7 – Comparison of the Structures of Free dsRBD and RNA-Bound dsRBD..... 60

Figure 4.1 – Solution structure of the Rnt1p dsRBD$_{366–453}$..... 67

Figure 4.2 – Comparison of the free RDC-refined dsRBD with the RDC-refined dsRBD/AGAA complex (PDB ID: 2LUP)..... 68

Figure 4.3 – The α1–α3 extended hydrophobic interface..... 69

Figure 4.4 – NMR spin-relaxation parameters. R_1, R_2, and heteronuclear NOE values for the free dsRBD..... 70

Figure 4.5 – Fast- and slow-timescale dynamics of the free dsRBD and the dsRBD/AGAA complex at 150 mM NaCl..... 71

Figure 4.6 – The effect of substrate length and salt concentration on chemical shift..... 72

Figure 4.7 – Fast- and slow-timescale dynamics of dsRBD/AGAA and dsRBD/AGAA22 complexes at 300 mM NaCl..... 73

Figure 4.8 – *In vivo* analysis of dsRBD α1–β1 loop mutants..... 75

Figure 4.9 – dsRBD dynamics are associated with concerted structural changes necessary for binding..... 77

Figure 5.1 – *BDF2* mRNA contains a canonical Rnt1p target stem-loop within its intron and is cleaved by Rnt1p *in vivo* and *in vitro*..... 86

Figure 5.2 – Mutation of the Rnt1p target stem-loop and 5′-splice site suppresses *bdf1Δ* growth defects in various environmental conditions..... 88

Figure 5.3 – Hyper-activation of Rnt1p cleavage of *BDF2* mRNA in osmotic stress..... 90

Figure 5.4 – Spliceosome-mediated decay of *BDF2* mRNA predominates over Rnt1p cleavage during DNA replication stress..... 92

Figure 5.5 – Environmental stress conditions control the expression of Bromodomain Factor 2 mRNA through modulating RNase III and spliceosome-mediated decay..... 93

Figure 6.1 – Genome-wide distribution of poly(A)+ 3´ ends in WT cells and cells lacking the nuclear exosome component Rrp6p..... 134

Figure 6.2 – The Reb1p DNA binding factor pauses Pol II genome-wide and destabilizes nascent transcripts through degradation by Rrp6p 135

Figure 6.3 – Reb1p-induced pausing of Pol II enhances termination efficiency of the Nrd1p-Nab3p-Sen1p pathway..... 136

Figure 6.4 – Reb1p-induced pausing of Pol II controls the selection of alternative cleavage and poly-adenylation sites..... 137

Figure 6.5 – Meta-gene analysis of Reb1 binding locations relative to transcription start sites and poly(A) sites for ORF transcripts…….....138

Figure 6.6 – The tRNA gene *tG(GCC)E* promotes optimal NNS-dependent termination of snR53 transcription…….....139

Figure 6.7 – Genome-wide roadblock by Pol III initiation factor TFIIIB restricts interference by cryptic Pol II transcription……...140

Figure 6.8 – Co-transcriptional cleavage of *SMF1* by Rnt1p protects cells against cadmium stress……...141

Figure 6.9 – Genome-wide co-transcriptional cleavage by Rnt1p controls expression of coding and non-coding transcripts in salt stress…….. 142

LIST OF TABLES **PAGE**

Table 3.1 – Structural Statistics of the Rnt1p dsRBD/AAGU Hairpin Complex……..56

Table 4.1 – Structural statistics of Rnt1p dsRBD…… 68

LIST OF SUPPLEMENTARY FIGURES **PAGE**

Figure 4.S1 – Northern blot analysis of genes identified as targets of spliceosome-mediated decay……98

Figure 4.S2 – RT-PCR analysis of full-length *BDF2* mRNA and spliced products generated by spliceosome-mediated decay in wild-type, the strain carrying the disruption of the Rnt1p stem loop (*bdf2 stem mutant*), and the deletion of nuclear exosome co-factor *RRP6* (*rrp6Δ*)……99

Figure 4.S3 – Probe walking to verify the identity of the *BDF2* mRNA degradation intermediates generated by spliceosome-mediated decay and Rnt1p…….....100

Figure 4.S4 – The nuclear exosome co-factor *RRP6*, the nonsense mediated decay (NMD) factor *UPF1*, and the pre-mRNA nuclear retention factors *MLP1/2* participate in the surveillance of *BDF2* transcripts…… 101

Figure 4.S5 – Wild-type and the specified *BDF2* mutants were shifted to 0.6 M NaCl for the time points indicated.…… 102

Figure 4.S6 – The wild-type strain and the specified *BDF2* mutants were grown on plates with the indicated media at 30°C for the indicated number of days……..…….. 103

LIST OF ABBREVIATIONS

S. cerevisiae	*Saccharomyces cerevisiae*
RNA	ribonucleic acid
mRNA	messenger RNA
pre-mRNA	precursor mRNA
poly(A)	poly-adenosine
Pol II	RNA polymerase II
TRAMP	Trf4/5p-Air2/1p-Mtr4p poly-adenylation complex
NNS	Nrd1p-Nab3p-Sen1p complex
DRS	Direct RNA Sequencing
RNase III	ribonuclease III
dsRNA	double-stranded RNA
dsRBD	double-stranded RNA binding domain
ncRNA	non-coding RNA
lncRNA	long non-coding RNA
CRAC	UV crosslinking and analysis of cDNAs

ACKNOWLEDGEMENTS

PERSONAL

First and foremost, I would like to thank my mentor, Dr. Guillaume Chanfreau, who steered the direction of this work and has fostered my development as a scientist in all facets. His constant support and belief in my abilities has allowed me to pursue this work fearlessly and encouraged me through uncertain times. From him, I have learned that with optimism and perseverance you can make your own good luck. I am truly fortunate to have trained in his lab and will always cherish the moments of discovery that came with patience and persistence. I am also thankful to my committee members for their support and advice on my thesis work.

I would also like to acknowledge my colleagues in the lab, past and present. I am especially indebted to Shakir Sayani, who first mentored me in the lab and introduced me to experimental molecular biology. I am also grateful for past graduate students Tad Kawashima, Cynthia Lee, Camille Nery, Chrissie Lee, Michael Janis, and post-docs Fernando Gonzales-Zubiate, Isabelle Toesca, and Cesar Fernandez for their friendship and support. I have had the pleasure of working in an intellectually stimulating environment, thanks to fellow graduate students Domi Hodko, Jason Gabunilas and Charles Wang. I am especially fortunate to have collaborated with Jason, who was always willing to discuss project ideas and contributed to much of the work presented in Chapter 6.

This work has benefited from extensive technical assistance from numerous undergraduates, and from them I have learned a great deal about how to be an effective mentor and lead a team. I would like to acknowledge undergraduates Arena Park, Tiffany Nguyen, Anh Tran, Doug Zhang, Rebekah Liu, and Duy Ngo. Doug, Rebekah, and Duy gave considerable technical assistance for the work presented in Chapter 6. Our technician Abby Gillespie has also greatly contributed to Chapter 6.

Most importantly, I am fortunate to have had the constant love and support from my best friend, Chitrada Kaweeteerawat, who fostered my balance in work and in life. She provided me with a regular dose of rejuvenation and inspiration with many outdoor adventures and national park trips, and has made me a better person in all areas of my life. Equally important, I am grateful to my parents and brothers for providing constant emotional support and encouraging me to pursue my dreams.

PROFESSIONAL

As a graduate student, I received financial support from the Philip J. Whitcome Graduate Fellowship from the Molecular Biology Interdepartmental Ph.D. Program at UCLA. I also was supported by a Ruth L. Kirschstein National Research Service Award GM007185, and a UCLA Graduate Division Dissertation Year Fellowship. The research presented in this work was supported by funding from the National Institutes of Health Grant GM061518 to Guillaume Chanfreau.

Reprint of publications:

1. **Roy K**, Chanfreau G. The diverse functions of fungal RNase III enzymes in RNA metabolism. ***The Enzymes*.** Review. 2012. 31 part A: 123-145. doi: 10.1016/B978-0-12-404740-2.00010-0.

Chapter 2 is a reprint of this publication. I gratefully acknowledge Elsevier for allowing me to include a reprint of this chapter (license number 3499620051458).

2. Wang Z, Hartman E, **Roy K**, Chanfreau G, Feigon, J. Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. ***Structure.*** 2011 Jul 13:19(7): 999-1010. doi: 10.1016/j.str.2011.03.022.

Chapter 3 is a reprint of this publication. I performed the *in vivo* work presented in Figure 2D and 2E. I gratefully acknowledge Professor Juli Feigon for the opportunity to collaborate in this work and for granting permission to include this publication in my

thesis. I thank Elsevier for allowing me to include a reprint of this chapter (license number 3691730767502).

3. Hartman E, Wang Z, Zhang Q, **Roy K**, Chanfreau G, Feigon J. Intrinsic dynamics of an extended hydrophobic core in the *S. cerevisiae* RNase III dsRBD contributes to recognition of specific RNA binding sites. ***J Mol Biol.*** 2013 Feb 8;425(3): 546-62. doi: 10.1016/j.jmb.2012.11.025.

Chapter 4 is a reprint of this publication. I performed the *in vivo* work presented in Figure 8. I am grateful to Professor Juli Feigon for the opportunity to collaborate and for granting permission to include this publication in my thesis. I thank Elsevier for allowing me to include a reprint of this chapter (license number 3691730915282).

4. **Roy K** and Chanfreau G. Stress-induced nuclear RNA degradation pathways regulate yeast bromodomain factor 2 to promote cell survival. ***PLoS Genet.*** 2014 Sept 18; 10(9).

Chapter 5 is a reprint of this publication. The content of this work is licensed by the Creative Commons Attribution (CC BY).

Chapter 6 is a draft of a manuscript in preparation. This work would not have been possible without the following co-authors and their respective contributions: Jason Gabunilas, for yeast strain construction, culturing, Northern blot analysis, and assistance with preparation of DNA libraries for high-throughput RNA sequencing; Elisabeth Petfalski and Professor David Tollervey of the Wellcome Trust Centre for Cell Biology at the University of Edinburgh for performing the UV crosslinking and analysis of cDNAs (CRAC) on Rnt1p; Marco Morselli and Professor Matteo Pellegrini for preparing cDNA libraries for Illumina high-throughput RNA sequencing; Rebekah Liu, Duy Ngo, Douglas Zhang, and Abby Gillespie, for extensive technical assistance with yeast strain construction, culturing, and Northern blot analysis.

VITA

EDUCATION

2009……..…… B.S./M.S. in Biochemistry and Molecular Biology, **University of California, Los Angeles (UCLA)**, Los Angeles, CA

PUBLICATIONS

Roy K, Chanfreau G. The diverse functions of fungal RNase III enzymes in RNA metabolism. ***The Enzymes*.** Review. 2012. 31 part A: 123-145.

Wang Z, Hartman E, **Roy K**, Chanfreau G, Feigon, J. Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. ***Structure***. 2011 Jul 13:19(7): 999-1010.

Hartman E, Wang Z, Zhang Q, **Roy K**, Chanfreau G, Feigon J. Intrinsic dynamics of an extended hydrophobic core in the *S. cerevisiae* RNase III dsRBD contributes to recognition of specific RNA binding sites. ***J Mol Biol***. 2013 Feb 8;425(3): 546-62.

Al-Hadid Q, **Roy K**, Munroe W, Dzialo MC, Chanfreau GF, Clarke SG.Histidine methylation of yeast ribosomal protein Rpl3p is required for proper 60S subunit assembly. ***Mol Cell Biol.*** 2014 Aug 1; 34(15):2903-16.

Roy K and Chanfreau G. Stress-induced nuclear RNA degradation pathways regulate yeast bromodomain factor 2 to promote cell survival. ***PLoS Genet*.** 2014 Sept 18; 10(9).

Dzialo MC, Travaglini KJ, Shen S, **Roy K**, Chanfreau GF, Loo JA, Clarke SG. Translational roles of elongation factor 2 protein lysine methylation. ***J Biol Chem*.** 2014 Oct 31; 289(44):30511-30524.

Kaweeteerawat C, Chang CH, **Roy K**, Liu R, Li R, Toso D, Fischer H, Ivask A, Ji Z, Zink JI, Zhou ZH, Chanfreau GF, Telesca D, Cohen Y, Holden PA, Nel AE, Godwin HA. Cu nanoparticles have different impacts in Escherichia coli and Lactobacillus brevis than their microsized and ionic analogues. ***ACS Nano***. 2015 Jul 13. (Epub ahead of print)

ORAL PRESENTATIONS

RNA Society Annual Meeting, UW-Madison, May 2015. 'Exploring the dark side of poly(A) tails: Diverse mechanisms control 3´-end processing and transcription termination genome-wide'

MBI Annual Retreat, UCLA, April 2015. 'Extensive Control of Pervasive Transcription by RNase III'

Gene Regulation Group Monthly Seminar Series, UCLA, March 2015. 'Exploring the dark side of poly(A) tails reveals diverse mechanisms directing co-transcriptional RNA degradation and transcription termination genome-wide'

MBI Annual Retreat, UCLA, October 2010. 'The role of the mitochondrial degradosome and a putative RNAse III in mitochondrial RNA processing.'

Glenn T. Seaborg Symposium, UCLA, November 2014. (Poster) 'A Transcriptome-Wide Map of Nuclear Endonuclease Cleavage Sites'

David S. Sigman Symposium, UCLA, January 2013. (Poster) 'Uncovering New Functions for Yeast RNAse III'

Glenn T. Seaborg Symposium, UCLA, November 2012. (Poster) 'Global Identification of Yeast RNAse III Targets'

RNA Society Annual Meeting, U. of Michigan, May-June 2012. (Poster) 'Determinants and anti-determinants of double-stranded RNA substrate recognition by yeast RNAse III'

AWARDS

Jules Brenner Scholar's Achievement Fellowship	2014 – 2015
UCLA Graduate Division Dissertation Year Fellowship	2014 – 2015
Cellular and Molecular Biology NIH Training Grant	2010 – 2013
Philip J. Whitcome Graduate Fellowship	2009 – 2010

CHAPTER 1

Introduction to Nuclear RNA Processing and Degradation in *Saccharomyces cerevisiae*

RNA Maturation: Quality Control and Regulation

For optimal growth and survival, cells must precisely regulate the types and quantities of RNAs and proteins expressed from their genome. In eukaryotic cells, precursor messenger RNAs (pre-mRNAs) must undergo multiple processing steps in the nucleus prior to export of the mature mRNAs and translation in the cytoplasm. These steps include the addition of a protective 5´-5´ triphosphate-linked methyl-guanosine (5´ cap), the removal of introns through splicing, 3´-end cleavage and polyadenylation, and association with specific RNA-binding proteins to form export-competent messenger ribonucleoprotein particles (mRNPs) (Tutucci and Stutz, 2011). While transcriptional regulation dictates the initial levels of precursor RNA, the numerous steps required for RNA biogenesis provide multiple points of regulation that together determine what fraction of precursor RNAs are matured into functional RNAs (Fig. 1; Mata et al., 2005).

At each stage in RNA biogenesis, stalled processing intermediates or aberrantly processed RNAs may exit functional maturation pathways and be turned over by nuclear RNA degradation pathways (Das et al., 2003; Parker, 2012). Defects in 5´ cap addition, cap methylation, splicing, or 3´ end cleavage and polyadenylation all result in enhanced degradation of the aberrant transcripts by exoribonucleases (Jimeno-González et al., 2010; Bousquet-Antonelli et al., 2000; Jiao et al., 2010; Galy et al., 2004). Furthermore, co-transcriptional packaging of specific RNA binding proteins with nascent RNAs is required for efficient 3´-end cleavage and polyadenylation, localization of the mRNA with nuclear pores, and mRNA export (Tutucci and Stutz, 2011). Defects in mRNP formation result in nuclear retention and rapid destabilization of the transcript (Saguez et al., 2008). While many of these quality control mechanisms were initially identified by artificial mutation-induced defects, mounting evidence indicates that many of the steps in mRNP biogenesis are not constitutive but are dynamically regulated, such that a substantial fraction of precursors fail to pass through each stage (Fig. 1.1;

Bousquet-Antonelli et al., 2000; Gudipati et al., 2012b). An emerging concept is that RNA degradation machineries serve both an RNA surveillance function to remove defective molecules from the pool of cellular RNAs, and increase the cellular regulatory capacity to degrade a fraction of otherwise functional RNA processing intermediates or mature mRNAs (Houseley and Tollervey, 2009; Ghosh and Jacobson, 2010). Kinetic competition between maturation factors and degradation factors for precursor RNA substrate likely dictates the proportion of transcripts diverted to degradation pathways at each biogenesis step (Bousquet-Antonelli et al., 2000; Mühlemann and Jensen, 2012).

In addition to turnover of precursor and mature RNAs, a significant role for RNA degradation systems appears to be the removal of pervasive Pol II transcripts, which are defined as poorly characterized ncRNAs that arise from promiscuous Pol II activity throughout the genome (Tisseur et al., 2011; Jensen et al., 2013). For the majority of these pervasive ncRNAs, the extent to which these transcripts have cellular functions, and whether such functions are mediated merely *in cis* by the act of their transcription or *in trans* by the transcript per se, has yet to be fully unraveled (Clark et al., 2011; Guil and Esteller, 2012; Tudek et al., 2015). The primary RNA degradation pathways operating in the nucleus and their associated ribonucleases are introduced below.

Nuclear 5´ to 3´ RNA Degradation Pathways

The predominant mode of nuclear and cytoplasmic RNA turnover in eukaryotes involves exonucleases, which are distinguished by their directionality (5´ to 3´ or 3´ to 5´). In eukaryotes, a single enzyme conducts the vast majority of nuclear 5´ to 3´ degradation. The gene encoding this enzyme was initially isolated through multiple independent genetic screens (Amberg et al., 1992; Di Segni et al., 1993; Kenna et al., 1993). The first involved screening temperature sensitive yeast strains for defects in mRNA export, and led to the identification of ribonucleic

acid trafficking 1 (*RAT1*), recognized to have homology to the previously characterized exoribonuclease 1 (*XRN1*), a highly processive enzyme favoring RNA substrates with 5′ mono-phosphates (Stevens, 1978; Stevens and Maupin, 1987). Rat1p is an essential gene in budding yeast, and has diverse functions in stable non-coding RNA (ncRNA) maturation, control of RNA polymerase II (Pol II) elongation and termination, and turnover of aberrantly processed ncRNAs or cryptic ncRNAs (Krzyszton et al., 2012). A major function of Rat1p is to trim 5′ extensions from ncRNA precursors involved in ribosome biogenesis, including the 5.8S and 25S ribosomal RNAs and small nucleolar RNAs (Fig. 1.2A; Henry et al., 1994; Qu et al., 1999; Lee et al., 2003; Henras et al., 2015). Rat1p exonuclease digestion stops at the mature 5′ ends when it encounters stable RNA secondary structures or 5′ ends protected by RNA-binding proteins (Granneman et al., 2011; Petfalski et al., 1998). Rat1p, along with the major cytoplasmic 5′ to 3′ exonuclease Xrn1p, also degrades endonuclease cleavage products generated by the yeast RNase III homolog Rnt1p, which targets various stem-loop structures present in pre-mRNA introns and open reading frames (Danin-Kreiselman et al., 2003; Ge et al., 2005).

Another critical cellular function for Rat1p is to promote transcription termination of RNA polymerase II (Pol II). After cleavage and polyadenylation of the nascent RNA, the downstream 3′ RNA fragment is recognized and degraded by Rat1p (Kim et al., 2004; Luo et al., 2006). When the Rat1p-RNA degradation complex catches up to the Pol II elongation complex, the Pol II-RNA-DNA ternary complex is destabilized to promote Pol II release from the template, in a manner that is enhanced by the Rat1p-binding partner Rai1p and the Sen1p helicase (Hage et al., 2008; Kawauchi et al., 2008). In addition, co-transcriptional cleavage of pre-rRNA by Rnt1p elicits termination in a Rat1p-dependent manner, in an analogous mechanism to Pol II termination (Fig. 1.2B; Hage et al., 2008). This was shown to provide a fail-safe mechanism for Pol II termination at various protein-coding genes in cases of defective cleavage and poly-adenylation, whereby Rnt1p cleavage at specific stem-loop structures downstream of poly(A)

sites generates an entry site for Rat1p-dependent degradation and termination (Ghazal et al., 2009; Rondón et al., 2009). Importantly, this was shown to prevent read-through of Pol II into downstream transcriptional units (Ghazal et al., 2009; Rondón et al., 2009). Another quality control function for Rat1p, along with its phosphohydrolase binding partner Rai1p, is to degrade nascent RNAs that fail to undergo co-transcriptional 5´ cap addition or cap methylation, resulting in the degradation of the nascent transcript and early termination of transcription (Fig. 1.2C; Jimeno-González et al., 2010).

As Rat1p recognizes 5´-phosphate containing substrates independent of the polymerase or cleavage mechanism that generated them, it serves a broad role in the quality control of nuclear-localized transcripts. Indeed, Rat1p has been shown to rapidly degrade mis-folded or hypo-modified tRNAs in the nucleus in a manner that directly correlates with the accessibility of these 5´ ends to purified phosphatase or exonuclease *in vitro* (Fig. 1.2D; Dewe et al., 2012). Another recently characterized function for Rat1p is in the degradation of Pol II transcripts arising from telomeric and subtelomeric regions (Fig. 1.2E). The sense transcripts from these regions have been termed telomeric repeat-containing RNA (TERRA), and their antisense transcripts ARRET. These transcripts differ significantly from canonical Rat1p substrates, as they harbor canonical 5´ caps which normally protect against Rat1p activity (Porro et al., 2010; Feuerhahn et al., 2010). It is currently unknown what targets Rat1p to specifically degrade these transcripts. Overall, it is clear that Rat1p exhibits diverse roles in maturation, quality control, and degradation pathways (Krzyszton et al., 2012).

Nuclear 3´ to 5´ RNA Degradation by the Exosome

The vast majority of 3´ to 5´ degradation in both the nucleus and cytoplasm is conducted by the RNA exosome, an exonuclease complex conserved throughout eukaryotes and archaea. Similarly to the 5´ to 3´ exonuclease Rat1p, the nuclear localized RNA-exosome has diverse

functions in ncRNA maturation, including 3′ end formation of 5.8S rRNA, small nuclear and small nucleolar RNAs (sn(o)RNAs) (Fig. 1.3A; Allmang et al., 1999a). The nuclear exosome also mediates the turnover of hypomodified tRNA, cryptic unstable transcripts (CUTs), and pre-mRNAs at various stages of mRNP biogenesis (Fig. 1.3B, C; Kadaba et al., 2004; Kadaba et al., 2006; Schneider and Tollervey, 2013).

Structural and Catalytic Components of the Exosome

The first indications of the existence of the exosome complex came through genetics by screening temperature-sensitive yeast strains for defects in the 3′ processing of 5.8S rRNA, resulting in the identification of the *RRP4* subunit (Mitchell et al., 1996). Subsequent glycerol gradient and tandem mass spectrometry analyses of immunoprecipitates of protein A-tagged Rrp4p revealed that Rrp4p exists in a multi-protein complex with the RNase PH homologs Rrp41p, Rrp42p, Rrp43p, as well as a bacterial RNase R-related protein Rrp44p/Dis3p (Mitchell et al., 1997). Sequence-homology based searches for RNase PH-like sequences suggested additional yeast homologs, and led to the identification of Rrp40p, Rrp45p, Rrp46p, Mtr3p, and Csl4p as additional components of the exosome, confirmed by tandem mass spectrometry of affinity purified complexes obtained with protein A-tagged Rrp4p (Allmang et al., 1999b). Yet another component, Rrp6p, related to the bacterial RNase D, was found exclusively in the nuclear exosome (Allmang et al., 1999b).

Structural characterization of the eukaryotic exosome revealed a ring consisting of three different heterodimers Rrp41p-Rrp45p, Rrp46p-Rrp43p, and Mtr3-Rrp42p (Fig. 1.4A), a structural arrangement conserved in the archaeal exosome (Liu et al., 2006; Lorentzen et al., 2005). In eukaryotes, the ring requires three proteins (Rrp4p, Csl4p, and Rrp40p) to form a lid on one end for stability, and together these nine proteins make up the “core” exosome (Fig. 1.4B; Liu et al., 2006). The three lid proteins contain the K-homology (KH) and ribosomal

protein S1 (S1-type) domains known to bind RNA. However, the KH domain likely serves only a role in stabilizing the core structure as the conserved RNA binding motif is buried in the S1-KH interface (Januszyk and Lima, 2014). On the other hand, the S1 domain contains an extended loop with conserved basic residues that is well positioned for binding or guiding RNA through the central channel of the ring (Januszyk and Lima, 2010).

While structurally similar to the archaeal exosome and bacterial polynucleotide phosphorylase (PNPase), the eukaryotic exosome employs only hydrolytic cleavage and is devoid of any phosphorolytic activity, with the exception of some plants (Chekanova et al., 2002). Rrp44p/Dis3p and the nuclear specific-component Rrp6p are responsible for all catalytic activity of the yeast exosome complex *in vivo* and *in vitro* (Fig 1.4C; Dziembowski et al., 2006; Liu et al., 2006). Indeed, further characterization of the archaeal complex revealed that catalytic residues required for activity of the archaeal core subunits are not conserved in the respective eukaryotic homologues (Dziembowski et al., 2006; Lorentzen et al., 2005). Hints to the function of the catalytically inactive core came from biochemical and structural analyses showing that the ring harbors an 8-10 angstrom channel just large enough to accommodate single-stranded RNA but not double-stranded RNA (Liu et al., 2006). The active nuclease Rrp44p binds the bottom of the core ring, where it was proposed to receive the 3´ end of single stranded RNA threaded from the top of the core through the channel of the ring (Fig 1.4C; Bonneau et al., 2009). The feeding of unstructured single-stranded RNA through the channel to the Rrp44p exonuclease domain was proposed to contribute to the highly processive degradation observed *in vivo* (Bonneau et al., 2009). Rrp44p/Dis3p also contains an endonuclease PIN domain, which requires manganese for activity *in vitro* (Lebreton et al., 2008), and may help to release RNA substrates that are stalled during Rrp44p exonuclease digestion (Schneider et al., 2012).

The other active exoribonuclease of the nuclear exosome, Rrp6p, is a member of the death effector domain containing DNA-binding protein (DEDD) family of 3´ to 5´ exonucleases

(Midtgaard et al., 2006). The gene encoding Rrp6p was initially identified through a genetic screen for suppressors of the poly(A) polymerase mutant *pap1-1*, and Rrp6p was shown to participate in the 3′ trimming of 5.8S rRNA precursor (Briggs et al., 1998). Rrp6p consists of an N-terminal domain (NTD), an exonuclease domain, and a helicase and RNase D C-terminal (HRDC) domain (Phillips and Butler, 2003). The NTD harbors a polycystin 2 N-terminus (PC2NT) domain, which creates a binding site for Rrp47p, a co-factor of the exosome that promotes Rrp6p exonuclease activity and allows Rrp6p to bind various RNA secondary structures (Stead et al., 2007). Rrp6p binds to the lid of the exosome at the top of the channel where it is implicated in the initial threading of RNA into the central channel (Fig 1.4C; Makino et al., 2015). The current model is that Rrp44p degrades the RNA completely or until the lid of the exosome encounters a steric blockage, as is the case for 3′ trimming of stable ncRNAs such as the 5.8S rRNA (Makino et al., 2015). Subsequently, multiple subunits in the exosome are thought to undergo conformational changes to release the 3′ end of the RNA back through the channel where it is processed further by Rrp6p in a distributive mode of degradation (Makino et al., 2013; Makino et al., 2015).

Enhancement of Exosome Activity by the TRAMP Complex

The prokaryotic equivalent of the exosome, the bacterial degradosome, contains the ribonucleases PNPase and RNase E, and the helicase RhlB (Py et al., 1996). Additionally, the prokaryotic poly(A) polymerase PAPI physically interacts with the degradosome, and poly-adenylates the 3′ ends of mature mRNAs and degradation intermediates (Raynal and Carpousis, 1999). The single-stranded poly(A) tail allows the degradosome to stably anchor onto RNAs with 3′ terminal secondary structures or bound proteins, which otherwise would be refractory to degradosome binding (Carpousis, 2007). The helicase activity of the degradosome subsequently unwinds the 3′ secondary structures and removes bound proteins to initiate exonucleolytic degradation (Carpousis, 2007). The destabilizing function of the poly(A) tail in

prokaryotes stood in stark contrast to the well-characterized role of the eukaryotic poly(A) tail in promoting mRNA stability and translation in eukaryotes (Dreyfus and Régnier, 2002). However, it was subsequently uncovered that an ancient role of the poly(A) tail was likely conserved at least partially in eukaryotic nuclei, as cells deficient in exosome activity were found to accumulate poly(A) tails on incompletely processed precursors to stable ncRNAs (van Hoof et al., 2000; Kadaba et al., 2006).

These exosome-sensitive poly(A) tails were subsequently found to be added by a poly(A) polymerase complex distinct from that of the cleavage and polyadenylation machinery (LaCava et al., 2005; Vanácová et al., 2005). This complex was found to contain a non-canonical poly(A) polymerase (either Trf4p or Trf5p), a zinc-knuckle RNA binding protein (either Air1p or Air2p), and the essential RNA helicase Mtr4p, and was termed the TRAMP complex (Trf4/5p-Air1/2p-Mtr4p poly-adenylation complex) (Fig. 1.5A; LaCava et al., 2005; Vanácová et al., 2005). These studies demonstrated that in contrast to the processive canonical poly(A) polymerase Pap1p, the TRAMP poly(A) polymerase added adenosines in a distributive fashion. Additional work demonstrated that TRAMP specifically enhanced the Rrp6p activity of the exosome, with no effect on Rrp44p *in vitro* (Callahan and Butler, 2010). This explained why polyadenylated ncRNA processing intermediates specifically accumulated in the absence of Rrp6p (van Hoof et al., 2000). The Mtr4p component of the TRAMP complex interacts with the nuclear exosome component Rrp6p and its cofactor Rrp47p, suggesting a mechanism for coupling TRAMP oligo-adenylation to rapid degradation by the nuclear exosome (Fig. 1.4A; Schuch et al., 2014).

A major function for TRAMP and the nuclear exosome is in the degradation of cryptic pervasive transcripts (Wyers et al., 2005). These transcripts were found to arise from genomic regions previously thought to be transcriptionally silent, as well as from widespread bidirectional promoters (Davis and Ares, 2006; Xu et al., 2009). In budding yeast, the TRAMP-exosome

complex is recruited co-transcriptionally by the RNA binding proteins Nrd1p and Nab3p (Vasiljeva and Buratowski, 2006). Nrd1p and Nab3 associate with the Sen1p helicase to form the NNS complex, which promotes termination of transcription in a Rat1p-independent manner (Vasiljeva and Buratowski, 2006). Nrd1p specifically binds UGUA and GUAG, while Nab3p specifically binds UCUU and CUUG (Carroll et al., 2004; Carroll et al., 2007). A sufficient enrichment of these motifs within a region of nascent RNA likely controls the selectivity of the NNS complex for termination of Pol II at specific genomic regions (Schulz et al., 2013; Webb et al., 2014). The 3´ ends generated by NNS are not protected by poly(A) binding proteins, and are freely susceptible for TRAMP adenylation and exosome degradation (Lykke-Andersen et al., 2009). Detailed mechanistic work showed that the Nrd1p component of the NNS pathway interacts with both Pol II and Trf4p in a mutually exclusive manner, suggesting that subsequent to transcription termination and release of the nascent transcript from Pol II, Nrd1p recruits the TRAMP complex via Trf4p interaction to oligo-adenylate the bound transcript (Tudek et al., 2014).

In addition to contributing to degrading cryptic transcripts, the NNS pathway also operates on a substantial fraction of protein coding genes, through attenuation of transcription or by competing with cleavage and polyadenylation (Gudipati et al., 2012a; Webb et al., 2014). The activity of this pathway appears to be regulated by growth conditions, with nutrient depletion leading to a redistribution of Nrd1p and Nab3p onto RNAs not bound in rich medium (Darby et al., 2012). It has been proposed that when transcription of consensus motif-containing mRNA and snoRNA precursors decreases, Nrd1p and Nab3p are more available to bind transcripts with binding motifs deviating from the consensus sequences (Darby et al., 2012).

While there are mechanisms for direct recruitment of TRAMP to specific transcripts, the emerging view is that TRAMP has no inherent specificity for particular RNA sequences, and will oligo-adenylate any RNA 3´ end to which it can access *in vivo* (Belostotsky, 2009; Lykke-

Andersen et al., 2009). In this capacity, the TRAMP-exosome pathway constantly surveys the nuclear RNA population for RNAs that present single-stranded 3´ ends, which can arise from mis-folding of normally stable ncRNAs or from mRNAs that fail to associate with protective RNA-binding proteins (Fig. 1.5B). The addition of the oligo-adenosine tail to these RNAs provides an anchor for the Mtr4p helicase component of TRAMP to stably bind the RNA substrate, and proceed to unwind the 3´ end of the RNA and deliver ssRNA to the nuclear exosome (Jia et al., 2012). Accordingly, the TRAMP complex has been found to be required for a variety of nuclear exosome-mediated RNA surveillance and processing pathways, including the degradation of hypo-modified tRNAs, final maturation steps of rRNAs, snRNAs, and snoRNAs, and in processing and degrading cleavage products generated by Rnt1p (Fig. 1.5C; Kadaba et al., 2006; Allmang et al., 1999a; Egecioglu et al., 2006). In addition, the Reb1p DNA binding protein has recently been shown to block the progression of Pol II at various genomic loci and promote targeting of the nascent transcript to TRAMP and the nuclear exosome for degradation (Fig. 1.5C) (Colin et al., 2014).

Originally believed to function solely on transcripts after their release from Pol II, recent findings have demonstrated that the exosome can directly impact transcription. In *Schizosaccharomyces pombe,* the exosome functions in transcription termination by engaging the nascent 3´ ends of transcripts exposed by backtracked Pol II complexes (Lemay et al., 2014). In B-cells, the exosome is required for efficient class-switch recombination and somatic hypermutation, by recruiting the mutator protein activation-induced cytidine deaminase (AID) to genes expressing exosome-targeted anti-sense ncRNAs (Pefanis et al., 2014). These functions indicate that the nuclear exosome can impact DNA as well as RNA metabolism, and has acquired roles unique to specific cell types. Therefore, uncovering targets of the nuclear exosome is a high priority to understand how nuclear RNA degradation impacts the quality and quantity of gene expression.

Overview and Significance of the Thesis

Various nuclear RNA processing mechanisms are coupled to degradation by the nuclear exosome (Fig. 1.5C). While the NNS pathway has been well characterized to control gene expression through coupling Pol II termination with TRAMP-exosome degradation, it is unknown the extent to which the yeast RNase III homolog Rnt1p or other nuclear RNA processing mechanisms directly control gene expression. Cleavage by the nuclear endoribonuclease Rnt1p promotes the maturation of stable ncRNAs through cleavage of precursors, as well as degrades normal and aberrant forms of various mRNAs (Chanfreau et al., 1998; Egecioglu et al., 2006; Rondón et al., 2009; Ghazal et al., 2009). In chapter 2, established roles of the yeast nuclear RNase III homolog Rnt1p and its fungal orthologues in RNA metabolism are reviewed in detail. Rnt1p targets double-stranded RNA stems with single-stranded tetraloops of the sequence NGNN or AAGU. However, a full understanding of its binding and cleavage specificity is still lacking. In chapter 3, we present structural work showing how Rnt1p specifically recognizes a non-canonical substrate, providing key insights into its flexible rules of substrate selection. In chapter 4, we demonstrate how structural dynamics mediated by a hydrophobic hinge in the double-stranded RNA binding domain contribute to tight binding of Rnt1p to its substrates *in vitro* and *in vivo.*

Next we uncover physiological roles for Rnt1p cleavage in controlling gene expression. The full range of Rnt1p cleavage targets *in vivo* and its direct impact on the transcriptome, as well as whether this enzyme is regulated by environmental conditions, is not fully understood. In Chapter 5, we identify *BDF2* mRNA as a new mRNA target for Rnt1p, and show that Rnt1p cleavage activity on *BDF2* is significantly activated during salt stress. Interestingly, *BDF2* is also targeted by the spliceosome in another specialized decay pathway termed spliceosome-mediated decay (SMD) (Volanakis et al., 2013). We demonstrate that SMD plays a minor role

during salt stress but is important for *BDF2* repression during DNA replication stress, a condition that down-regulates Rnt1p cleavage of *BDF2*.

In Chapter 6, we present a genome-wide analysis of RNA decay pathways coupled to the nuclear exosome, and discover a global role for Rnt1p in controlling pervasive transcription and gene expression during salt stress. We also uncover widespread roles for the TFIIIB transcription initiation factor and Reb1p DNA binding protein in defining transcriptional boundaries through coupling the pausing of Pol II to RNA 3´-end processing. We show that TFIIIB blocks Pol II elongation and enhances NNS termination on snoRNAs and cryptic transcripts. We also demonstrate that Reb1p does not simply act as a roadblock for Pol II progression, but cooperates with the NNS pathway to terminate transcription of snoRNAs. We show that Reb1p also controls the selection of alternative poly(A) sites by the cleavage and poly-adenylation machinery by pausing the elongation of Pol II and enhancing the usage of upstream poly(A) sites. Taken together, these results demonstrate that rather than relying on a single pathway, cells depend on a combination of diverse mechanisms to promote Pol II termination and control pervasive transcription.

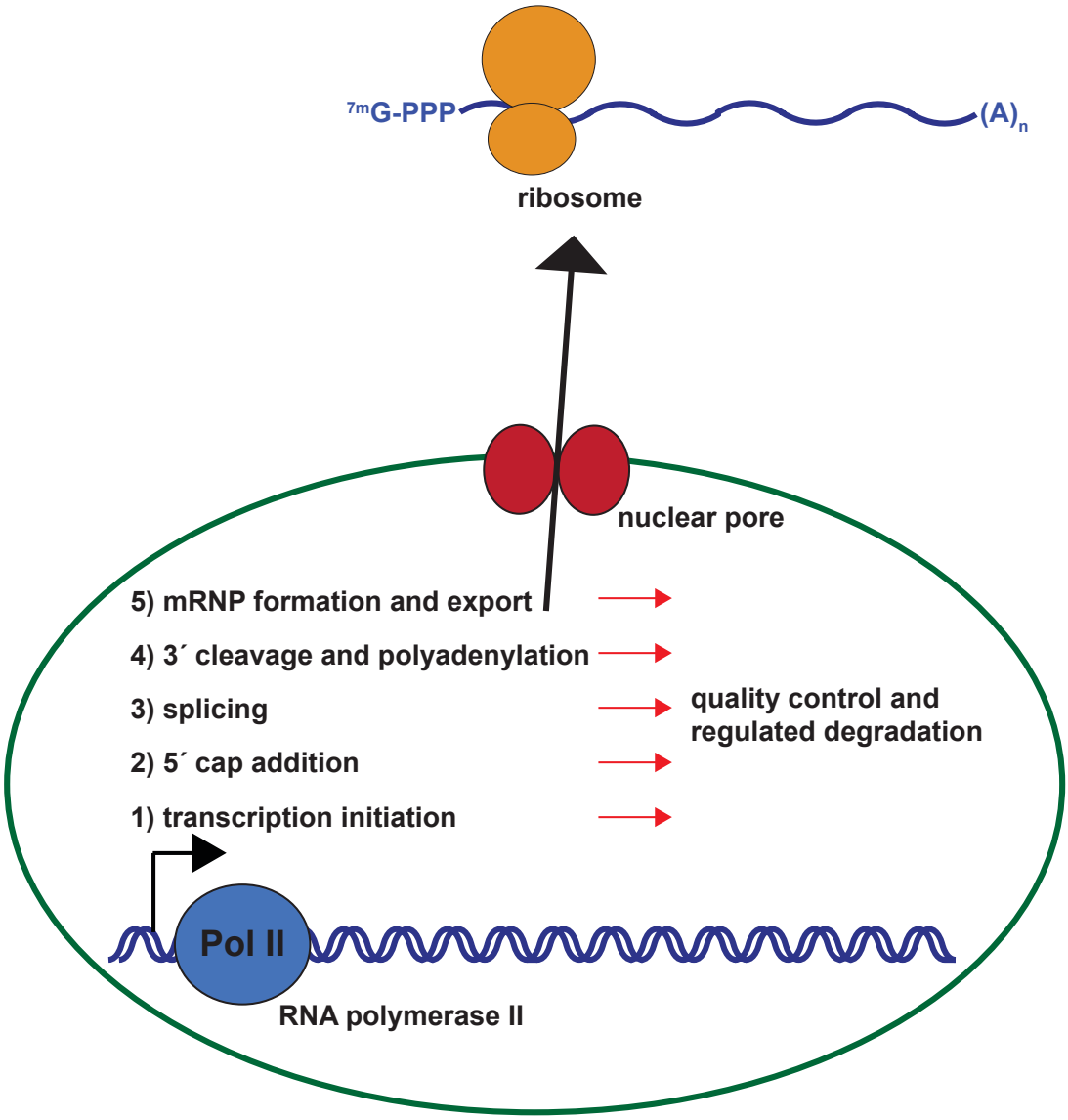


Figure 1.1. Quality control and regulated degradation of messenger RNA precursors. Messenger RNA (mRNA) biogenesis involves extensive RNA processing, as well as the association of the mRNA with specific proteins to form messenger ribonucleoproteins (mRNPs). Defects in these processes or regulated inhibition of these processes can result in a lack of nuclear export and subsequent degradation by nuclear ribonucleases.

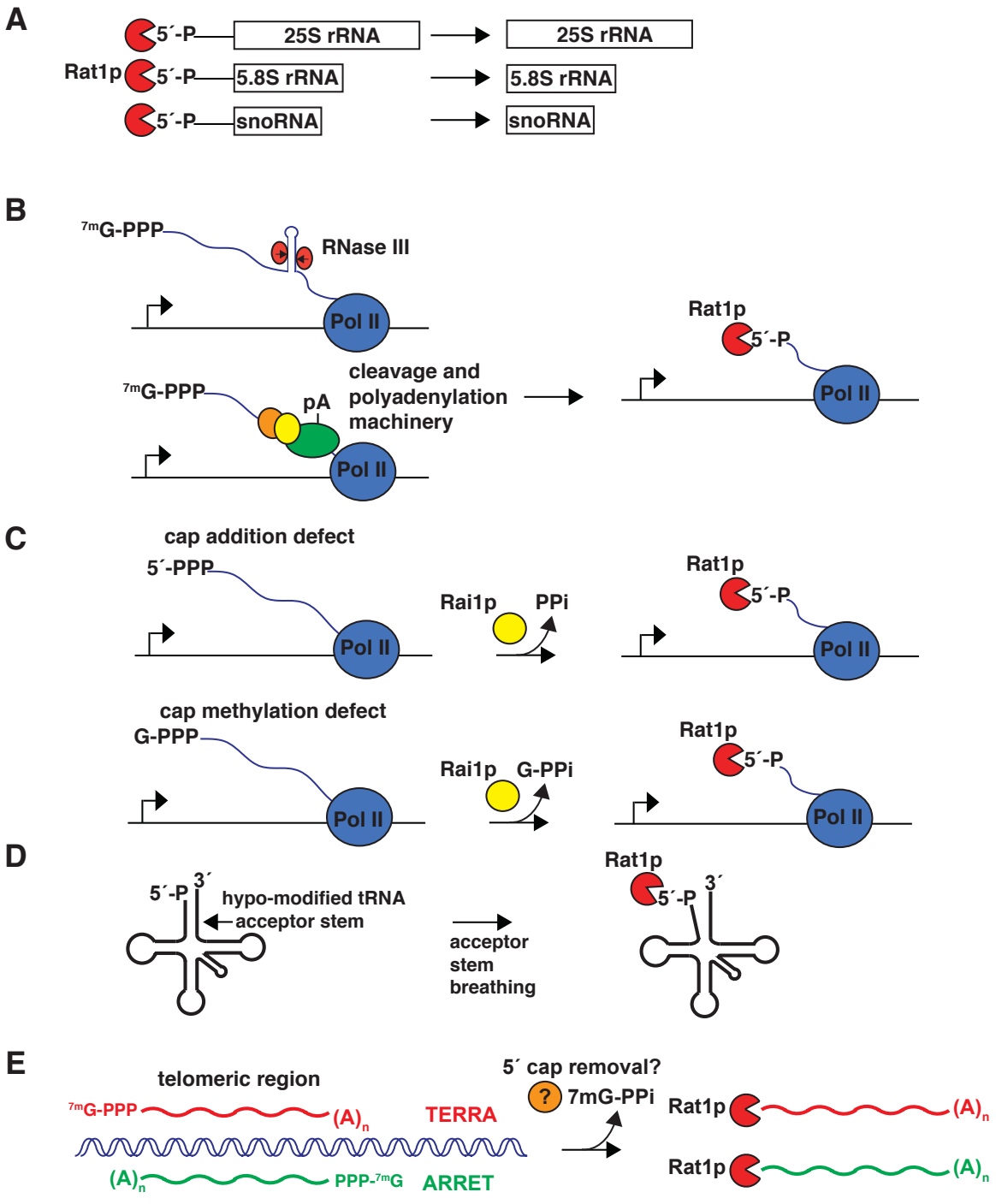


Figure 1.2. Diverse roles for the nuclear 5´-3´ exonuclease Rat1p in gene regulation. Rat1p mediates 5´ trimming of noncoding RNA precursors **(A)**, transcription termination (B), quality control of pre-mRNA capping and tRNA modifications (C,D) and regulation of telomeric and sub-telomeric repeat containing RNAs (sense = TERRA; antisense = ARRET) (E).

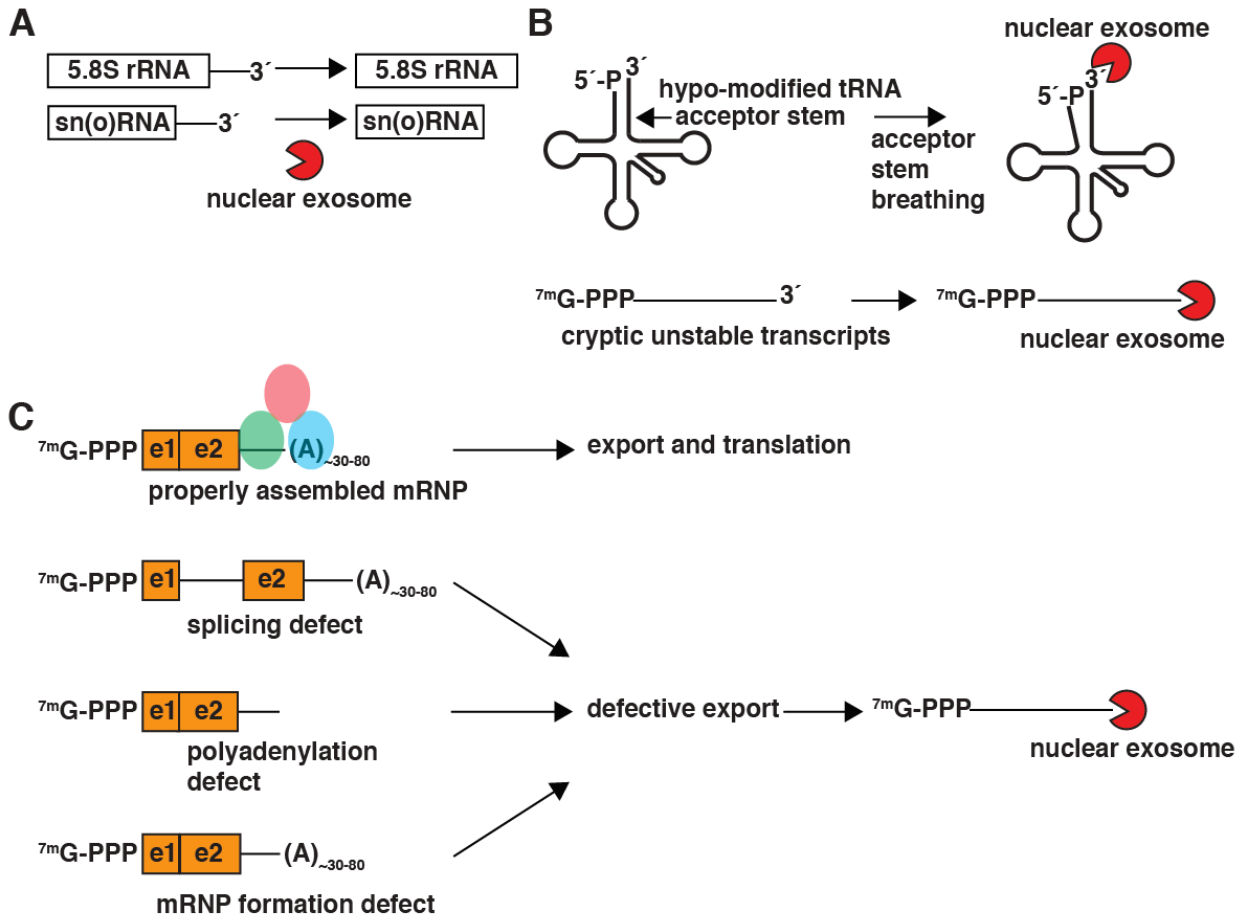


Figure 1.3. Diverse roles for the nuclear 3´-5´ exosome complex in RNA processing and quality control. The nuclear exosome mediates the 3 trimming of noncoding RNA precursors **(A)**, quality control of tRNA modifications and degradation of cryptic unstable transcripts (CUTs) **(B)**, quality control of pre-mRNA capping and tRNA modifications **(C, D)** and regulation of telomeric and sub-telomeric repeat containing RNAs (sense = TERRA; antisense = ARRET) **(E)**.

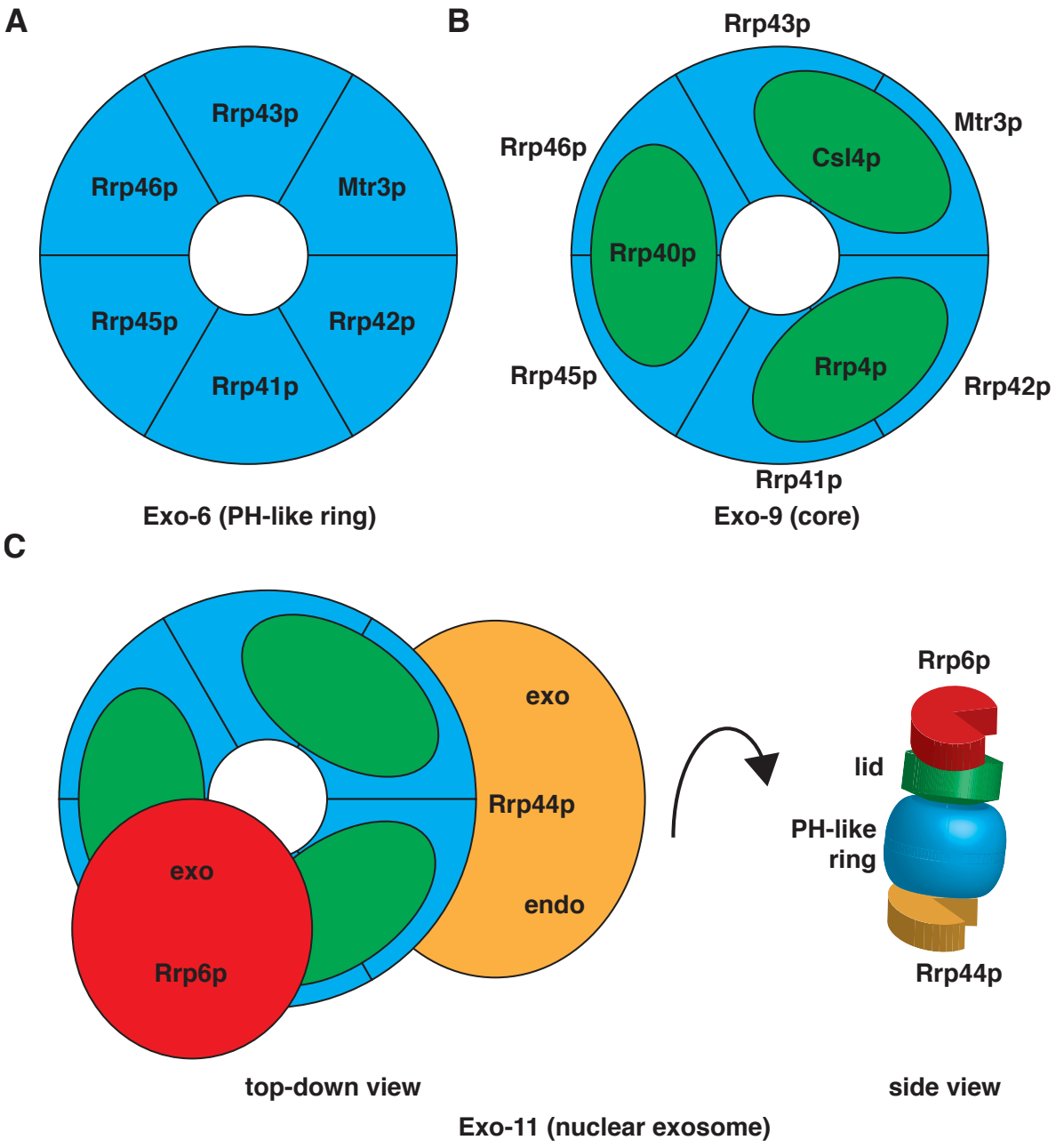


Figure 1.4. Architecture of the nuclear RNA exosome. The exosome core ring contains six catalytically inactive RNase PH-like proteins **(A)**. Three proteins containing K-homology / S1 RNA binding domains form a lid on top of the ring **(B)**. The Rrp6p 3′-5′ exonuclease associates with the lid, while the essential Rrp44p 3′-5′ exonuclease binds to the bottom of the ring. Rrp44p also contains an endonuclease domain **(C)**.

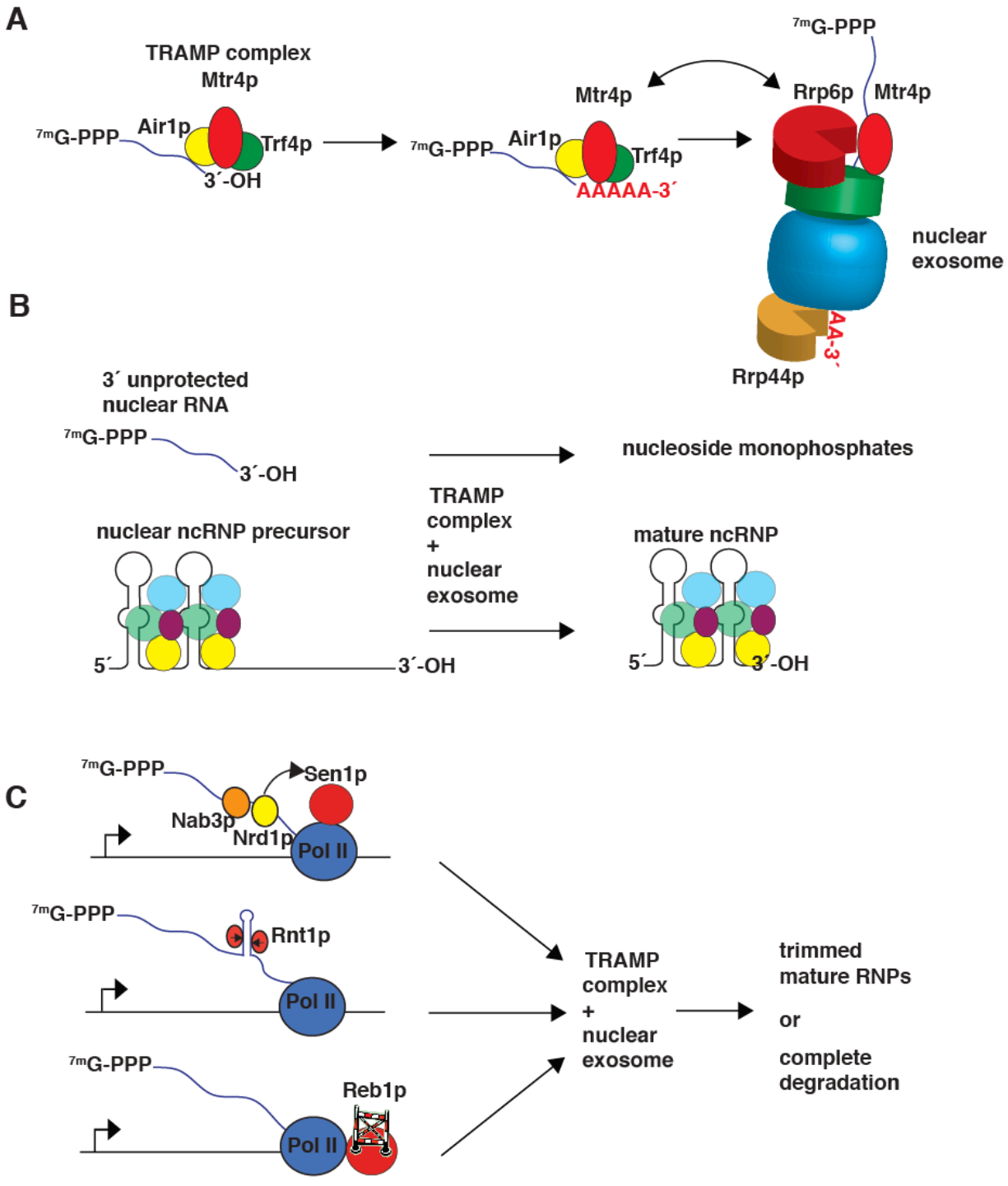


Figure 1.5. Coupling of the nuclear exosome to RNA processing pathways. The Trf4/5p-Air1/2p-Mtr4p poly-adenylation complex (TRAMP complex) oligo-adenylates RNA 3´ ends, providing an unstructured tail for Mtr4p to bind the RNA. An interaction between Rrp6p and

Mtr4p is proposed to recruit Mtr4p and its bound RNA to the lid of the exosome. Mtr4p helicase activity likely promotes the threading of the 3´ end of the RNA through the exosome core ring to reach the Rrp44p exonuclease active site. **(A)**. The combination of TRAMP and exosome activities constitutes a dual surveillance/processing function in the nucleus, where RNAs with unprotected 3´ ends are degraded completely, or trimmed until a stable structure blocks exosome 3´-5´ progression **(B)**. Various termination mechanisms, distinct from the canonical cleavage and poly-adenylation machinery, generate unprotected 3´ ends that are subsequently targeted by TRAMP / exosome activity for trimming or degradation **(C)**.

References

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999a). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. Embo J. *18*, 5399–5410.

Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999b). The yeast exosome and human PM-Scl are related complexes of 3“ --> 5” exonucleases. Genes Dev. *13*, 2148–2158.

Amberg, D.C., Goldstein, A.L., and Cole, C.N. (1992). Isolation and characterization of RAT1: an essential gene of Saccharomyces cerevisiae required for the efficient nucleocytoplasmic trafficking of mRNA. Genes Dev. *6*, 1173–1189.

Belostotsky, D. (2009). Exosome complex and pervasive transcription in eukaryotic genomes. Curr. Opin. Cell Biol. *21*, 352–358.

Bonneau, F., Basquin, J., Ebert, J., Lorentzen, E., and Conti, E. (2009). The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. Cell *139*, 547–559.

Bousquet-Antonelli, C., Presutti, C., and Tollervey, D. (2000). Identification of a regulated pathway for nuclear pre-mRNA turnover. Cell *102*, 765–775.

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

Callahan, K.P., and Butler, J.S. (2010). TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. J. Biol. Chem. *285*, 3540–3547.

Carpousis, A.J. (2007). The RNA degradosome of Escherichia coli: an mRNA-degrading machine assembled on RNase E. Annu. Rev. Microbiol. *61*, 71–87.

Carroll, K.L., Ghirlando, R., Ames, J.M., and Corden, J.L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. Rna *13*, 361–373.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D., and Corden, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol. Cell. Biol. *24*, 6241–6252.

Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. Embo J. *17*, 3726–3737.

Chekanova, J.A., Dutko, J.A., Mian, I.S., and Belostotsky, D.A. (2002). Arabidopsis thaliana exosome subunit AtRrp4p is a hydrolytic 3“-->5” exonuclease containing S1 and KH RNA-binding domains. Nucleic Acids Res. *30*, 695–700.

Clark, M.B., Amaral, P.P., Schlesinger, F.J., Dinger, M.E., Taft, R.J., Rinn, J.L., Ponting, C.P., Stadler, P.F., Morris, K.V., Morillon, A., et al. (2011). The reality of pervasive transcription. PLoS Biol. *9*, e1000625–discussione1001102.

Colin, J., Candelli, T., Porrua, O., Boulay, J., Zhu, C., Lacroute, F., Steinmetz, L.M., and Libri, D. (2014). Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol. Cell *56*, 667–680.

Danin-Kreiselman, M., Lee, C.Y., and Chanfreau, G. (2003). RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol. Cell *11*, 1279–1289.

Darby, M.M., Serebreni, L., Pan, X., Boeke, J.D., and Corden, J.L. (2012). The Saccharomyces cerevisiae Nrd1-Nab3 transcription termination pathway acts in opposition to Ras signaling and mediates response to nutrient depletion. Mol. Cell. Biol. *32*, 1762–1775.

Das, B., Butler, J.S., and Sherman, F. (2003). Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. Mol. Cell. Biol. *23*, 5502–5515.

Davis, C.A., and Ares, M.J. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.a. *103*, 3262–3267.

Dewe, J.M., Whipple, J.M., Chernyakov, I., Jaramillo, L.N., and Phizicky, E.M. (2012). The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. Rna *18*, 1886–1896.

Di Segni, G., McConaughy, B.L., Shapiro, R.A., Aldrich, T.L., and Hall, B.D. (1993). TAP1, a yeast gene that activates the expression of a tRNA gene with a defective internal promoter. Mol. Cell. Biol. *13*, 3424–3433.

Dreyfus, M., and Régnier, P. (2002). The poly(A) tail of mRNAs: bodyguard in eukaryotes, scavenger in bacteria. Cell *111*, 611–613.

Dziembowski, A., Lorentzen, E., Conti, E., and Séraphin, B. (2006). A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat. Struct. Mol. Biol. *14*, 15–22.

Egecioglu, D.E., Henras, A.K., and Chanfreau, G.F. (2006). Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. Rna *12*, 26–32.

Feuerhahn, S., Iglesias, N., Panza, A., Porro, A., and Lingner, J. (2010). TERRA biogenesis, turnover and implications for function. FEBS Lett. *584*, 3812–3818.

Galy, V., Gadal, O., Fromont-Racine, M., Romano, A., Jacquier, A., and Nehrbass, U. (2004). Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell *116*, 63–73.

Ge, D., Lamontagne, B., and Elela, S.A. (2005). RNase III-mediated silencing of a glucose-dependent repressor in yeast. Curr Biol *15*, 140–145.

Ghazal, G., Gagnon, J., Jacques, P.-E., Landry, J.-R., Robert, F., and Elela, S.A. (2009). Yeast RNase III triggers polyadenylation-independent transcription termination. Mol. Cell *36*, 99–109.

Ghosh, S., and Jacobson, A. (2010). RNA decay modulates gene expression and controls its fidelity. Wiley Interdisciplinary Reviews: RNA *1*, 351–361.

Granneman, S., Petfalski, E., and Tollervey, D. (2011). A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. Embo J. *30*, 4006–4019.

Gudipati, R.K., Neil, H., Feuerbach, F., Malabat, C., and Jacquier, A. (2012a). The yeast RPL9B gene is regulated by modulation between two modes of transcription termination. Embo J. *31*, 2427–2437.

Gudipati, R.K., Xu, Z., Lebreton, A., Séraphin, B., Steinmetz, L.M., Jacquier, A., and Libri, D. (2012b). Extensive degradation of RNA precursors by the exosome in wild-type cells. Mol. Cell *48*, 409–421.

Guil, S., and Esteller, M. (2012). Cis-acting noncoding RNAs: friends and foes. Nat. Struct. Mol. Biol. *19*, 1068–1075.

Hage, El, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5' exonuclease Rat1 in yeast. Genes Dev. *22*, 1069–1081.

Henras, A.K., Plisson Chastang, C., O'Donohue, M.F., Chakraborty, A., and Gleizes, P.E. (2015). An overview of pre-ribosomal RNA processing in eukaryotes. Wiley Interdisciplinary Reviews: RNA *6*, 225–242.

Henry, Y., Wood, H., Morrissey, J.P., Petfalski, E., Kearsey, S., and 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.

Houseley, J., and Tollervey, D. (2009). The Many Pathways of RNA Degradation. Cell *136*, 763–776.

Januszyk, K., and Lima, C.D. (2010). Structural components and architectures of RNA exosomes. Adv. Exp. Med. Biol. *702*, 9–28.

Januszyk, K., and Lima, C.D. (2014). The eukaryotic RNA exosome. Current Opinion in Structural Biology *24*, 132–140.

Jensen, T.H., Jacquier, A., and Libri, D. (2013). Dealing with pervasive transcription. Mol. Cell *52*, 473–484.

Jia, H., Wang, X., Anderson, J.T., and Jankowsky, E. (2012). RNA unwinding by the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex. Proc. Natl. Acad. Sci. U.S.a. *109*, 7292–7297.

Jiao, X., Xiang, S., Oh, C., Martin, C.E., Tong, L., and Kiledjian, M. (2010). Identification of a quality-control mechanism for mRNA 5′-end capping. Nature *467*, 608–611.

Jimeno-González, S., Haaning, L.L., Malagon, F., and Jensen, T.H. (2010). The Yeast 5′-3′ Exonuclease Rat1p Functions during Transcription Elongation by RNA Polymerase II. Mol. Cell *37*, 580–587.

Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G., and Anderson, J. (2004). Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. Genes

Dev. *18*, 1227–1240.

Kadaba, S., Wang, X., and Anderson, J.T. (2006). Nuclear RNA surveillance in Saccharomyces cerevisiae: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. Rna *12*, 508–521.

Kawauchi, J., Mischo, H., Braglia, P., Rondon, A., and Proudfoot, N.J. (2008). Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev. *22*, 1082–1092.

Kenna, M., Stevens, A., McCammon, M., and Douglas, M.G. (1993). An essential yeast gene with homology to the exonuclease-encoding XRN1/KEM1 gene also encodes a protein with exoribonuclease activity. Mol. Cell. Biol. *13*, 341–350.

Kim, M., Krogan, N.J., Vasiljeva, L., Rando, O.J., Nedea, E., Greenblatt, J.F., and Buratowski, S. (2004). The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. Nature *432*, 517–522.

Krzyszton, M., Zakrzewska-Placzek, M., Koper, M., and Kufel, J. (2012). Rat1 and Xrn2: The Diverse Functions of the Nuclear Rat1/Xrn2 Exonuclease. In Eukaryotic RNases and Their Partners in RNA Degradation and Biogenesis, Part A, (Elsevier), pp. 131–163.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713–724.

Lebreton, A., Tomecki, R., Dziembowski, A., and Séraphin, B. (2008). Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature *456*, 993–996.

Lee, C.Y., Lee, A., and Chanfreau, G. (2003). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. Rna *9*, 1362–1370.

Lemay, J.-F., Larochelle, M., Marguerat, S., Atkinson, S., Bähler, J., and Bachand, F. (2014). The RNA exosome promotes transcription termination of backtracked RNA polymerase II. Nat. Struct. Mol. Biol. *21*, 919–926.

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

Lorentzen, E., Walter, P., Fribourg, S., Evguenieva-Hackenberg, E., Klug, G., and Conti, E. (2005). The archaeal exosome core is a hexameric ring structure with three catalytic subunits. Nat. Struct. Mol. Biol. *12*, 575–581.

Luo, W., Johnson, A.W., and Bentley, D.L. (2006). The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. Genes Dev. *20*, 954–965.

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

Makino, D.L., Baumgärtner, M., and Conti, E. (2013). Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. Nature *495*, 70–75.

Makino, D.L., Schuch, B., Stegmann, E., Baumgärtner, M., Basquin, C., and Conti, E. (2015). RNA degradation paths in a 12-subunit nuclear exosome complex. Nature *524*, 54–58.

Mata, J., Marguerat, S., and Bähler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. Trends Biochem. Sci. *30*, 506–514.

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

Mitchell, P., Petfalski, E., and Tollervey, D. (1996). The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. Genes Dev. *10*, 502–513.

Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., and Tollervey, D. (1997). The Exosome: A Conserved Eukaryotic RNA Processing Complex Containing Multiple 3′→5′ Exoribonucleases. Cell *91*, 457–466.

Mühlemann, O., and Jensen, T.H. (2012). mRNP quality control goes regulatory. Trends Genet. *28*, 70–77.

Parker, R. (2012). RNA degradation in Saccharomyces cerevisae. Genetics *191*, 671–702.

Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N., and Basu, U. (2014). Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature *514*, 389–393.

Petfalski, E., Dandekar, T., Henry, Y., and Tollervey, D. (1998). Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. Mol. Cell. Biol. *18*, 1181–1189.

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

Porro, A., Feuerhahn, S., Reichenbach, P., and Lingner, J. (2010). Molecular dissection of telomeric repeat-containing RNA biogenesis unveils the presence of distinct and multiple regulatory pathways. Mol. Cell. Biol. *30*, 4808–4817.

Py, B., Higgins, C.F., Krisch, H.M., and Carpousis, A.J. (1996). A DEAD-box RNA helicase in the Escherichia coli RNA degradosome. Nature *381*, 169–172.

Qu, L.H., Henras, A., Lu, Y.J., Zhou, H., Zhou, W.X., Zhu, Y.Q., Zhao, J., Henry, Y., Caizergues-Ferrer, M., and Bachellerie, J.P. (1999). Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. Mol. Cell. Biol. *19*, 1144–1158.

Raynal, L.C., and Carpousis, A.J. (1999). Poly(A) polymerase I of Escherichia coli: characterization of the catalytic domain, an RNA binding site and regions for the interaction with

proteins involved in mRNA degradation. Mol. Microbiol. *32*, 765–775.

Rondón, A.G., Mischo, H.E., Kawauchi, J., and Proudfoot, N.J. (2009). Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol. Cell *36*, 88–98.

Saguez, C., Schmid, M., Olesen, J.R., Ghazy, M.A.E.-H., Qu, X., Poulsen, M.B., Nasser, T., Moore, C., and Jensen, T.H. (2008). Nuclear mRNA Surveillance in THO/sub2 Mutants Is Triggered by Inefficient Polyadenylation. Mol. Cell *31*, 91–103.

Schneider, C., and Tollervey, D. (2013). Threading the barrel of the RNA exosome. Trends Biochem. Sci. *38*, 485–493.

Schneider, C., Kudla, G., Wlotzka, W., Tuck, A., and Tollervey, D. (2012). Transcriptome-wide analysis of exosome targets. Mol. Cell *48*, 422–433.

Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P., and Conti, E. (2014). The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. Embo J. *33*, 2829–2846.

Schulz, D., Schwalb, B., Kiesel, A., Baejen, C., Torkler, P., Gagneur, J., Soeding, J., and Cramer, P. (2013). Transcriptome surveillance by selective termination of noncoding RNA synthesis. Cell *155*, 1075–1087.

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

Stevens, A. (1978). An exoribonuclease from Saccharomyces cerevisiae: effect of modifications of 5“ end groups on the hydrolysis of substrates to 5” mononucleotides. Biochem. Biophys. Res. Commun. *81*, 656–661.

Stevens, A., and Maupin, M.K. (1987). A 5“----3” exoribonuclease of Saccharomyces cerevisiae: size and novel substrate specificity. Arch. Biochem. Biophys. *252*, 339–347.

Tisseur, M., Kwapisz, M., and Morillon, A. (2011). Pervasive transcription - Lessons from yeast. Biochimie *93*, 1889–1896.

Tudek, A., Candelli, T., and Libri, D. (2015). Non-coding transcription by RNA polymerase II in yeast: Hasard or nécessité? Biochimie.

Tudek, A., Porrua, O., Kabzinski, T., Lidschreiber, M., Kubicek, K., Fortova, A., Lacroute, F., Vanácová, S., Cramer, P., Stefl, R., et al. (2014). Molecular basis for coordinating transcription termination with noncoding RNA degradation. Mol. Cell *55*, 467–481.

Tutucci, E., and Stutz, F. (2011). Keeping mRNPs in check during assembly and nuclear export. Nat. Rev. Mol. Cell Biol. *12*, 377–384.

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

Vanácová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. *3*, e189.

Vasiljeva, L., and Buratowski, S. (2006). Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Mol. Cell *21*, 239–248.

Volanakis, A., Passoni, M., Hector, R.D., Shah, S., Kilchert, C., Granneman, S., and Vasiljeva, L. (2013). Spliceosome-mediated decay (SMD) regulates expression of nonintronic genes in budding yeast. Genes Dev. *27*, 2025–2038.

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, R8.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., et al. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725–737.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

CHAPTER 2

The Diverse Functions of Fungal RNase III Enzymes in RNA Metabolism

The Diverse Functions of Fungal RNase III Enzymes in RNA Metabolism

Kevin Roy, Guillaume F. Chanfreau[1]
Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, California, USA
[1]Corresponding author: e-mail address: guillom@chem.ucla.edu

Contents

1. Introduction
2. Phylogenetic Distribution and Conservation of RNase III Enzymes in Fungi
3. Ribosomal RNA Processing and RNA Polymerase I Transcriptional Termination
4. Small Nuclear RNAs Processing
5. Functions in Small Nucleolar RNAs Processing
   - 5.1 Processing of independently transcribed snoRNA precursors
   - 5.2 Cleavage of polycistronic transcription units
   - 5.3 Processing of intron-encoded snoRNAs
6. (Pre)-mRNA Surveillance, Degradation, and Regulation
7. RNA Polymerase II Termination
8. Conclusions and Perspectives

Acknowledgements
References

Abstract

Enzymes from the ribonuclease III family bind and cleave double-stranded RNA to initiate RNA processing and degradation of a large number of transcripts in bacteria and eukaryotes. This chapter focuses on the description of the diverse functions of fungal RNase III members in the processing and degradation of cellular RNAs, with a particular emphasis on the well-characterized representative in *Saccharomyces cerevisiae*, Rnt1p. RNase III enzymes fulfill important functions in the processing of the precursors of various stable noncoding RNAs such as ribosomal RNAs and small nuclear and nucleolar RNAs. In addition, they cleave and promote the degradation of specific mRNAs or improperly processed forms of certain mRNAs. The cleavage of these mRNAs serves both surveillance and regulatory functions. Finally, recent advances have shown that RNase III enzymes are involved in mediating fail-safe transcription termination by RNA polymerase II (Pol II), by cleaving intergenic stem-loop structures present downstream from Pol II transcription units. Many of these processing functions appear to be conserved in fungal species close to the *Saccharomyces* genus, and even in more distant eukaryotic species.

1. INTRODUCTION

RNase III enzymes were discovered over 45 years ago as enzymatic activities specific for cleaving double-stranded RNA in prokaryotes [1,2]. Representatives of this class of enzymes were characterized only more recently in eukaryotic cells. This chapter focuses on fungal representatives of this class of enzymes, with a special emphasis on describing their diverse functions in RNA metabolism. We mostly exclude from this discussion RNase III from the Dicer and Drosha families and their involvement in the RNA interference and microRNA biogenesis pathways, as these topics will be extensively treated in chapters from the next volume of this series. Instead, we focus on the function of RNase III enzymes in the processing of noncoding RNAs involved in splicing and ribosomal RNA processing and in the cleavage of specific mRNAs and read-through transcripts. Many of these enzymes are not essential in fungi, facilitating the steady-state analysis of potential RNA substrates of these enzymes. However, the fact that strains carrying deletions of the gene encoding these enzymes exhibit substantial growth defects reflects their functional importance in RNA metabolism. In the next sections, we will review the phylogenetic distribution and conservation of RNase III enzymes in fungi and describe the various RNA processing and degradation pathways that are known to involve cleavage by these enzymes.

2. PHYLOGENETIC DISTRIBUTION AND CONSERVATION OF RNase III ENZYMES IN FUNGI

RNase III enzymes are present in all kingdom and species except in *Archaea*, where the bulge-helix-bulge nucleases cleave endogenous dsRNA substrates [3]. RNase III proteins are identified by a characteristic catalytic domain exhibiting the signature motif NERLEFLGD (Fig. 10.1A and B). In addition, nearly all contain a C-terminal dsRNA-binding domain (dsRBD; Fig. 10.1A) with the conserved αββββα fold, with the one known exception of *Giardia* Dicer [5]. In addition to the catalytic and dsRBD domains, eukaryotic RNases III can contain additional accessory domains important for substrate specificity and proper positioning on the substrate (Fig. 10.1A; [6]). Based on domain architecture and evolutionary relationships, RNase III proteins can be divided into four classes [7]: class I, bacterial RNase III; class II, Drosha; class IIIa, canonical Dicer; and class IIIb, fungal

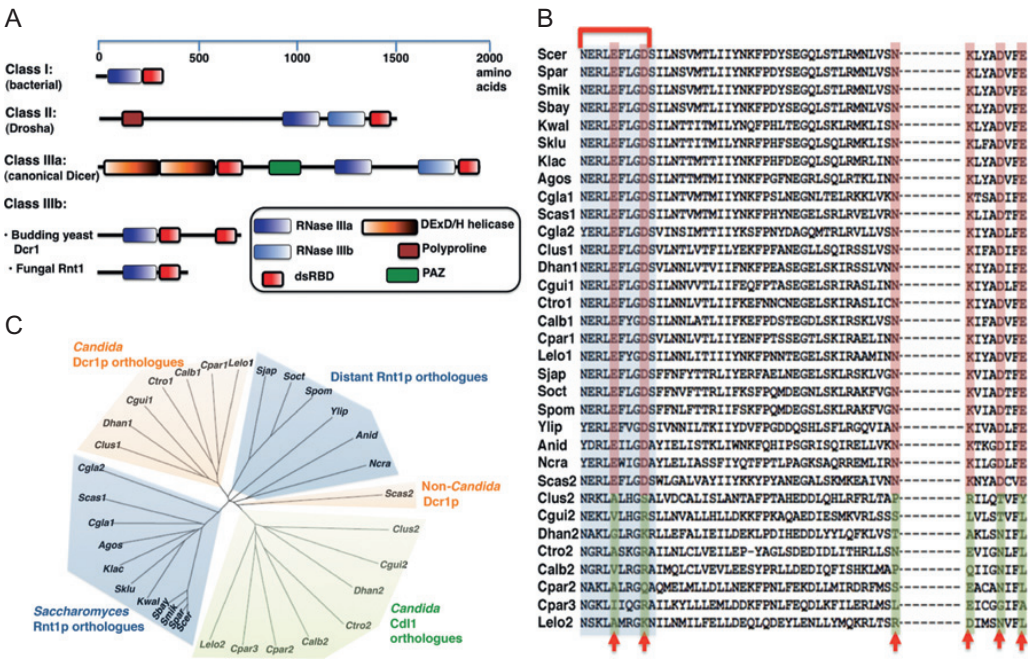


Figure 10.1 Architecture and evolutionary relationships of fungal RNase III representatives. (A) Classification scheme for RNase III enzymes, which range in size from ~200 (class I) to ~2000 amino acids (class IIIa). Class I includes bacterial RNase III (homologues of *E. coli rnc*). Class II contains Drosha, which is involved in pre-rRNA processing and in the conversion of primary microRNA (pri-miRNA) transcripts to precursor microRNA (pre-miRNA) hairpins. Class IIIa contains the canonical Dicer enzymes involved in processing pre-miRNA to mature miRNA. Class IIIb includes the fungal RNase III enzymes, which are further divided into two types depending on the presence (Dcr1) or absence (Rnt1) of a second dsRBD at the C-terminus. (C) Alignment of the class IIIb RNase III domains around the highly conserved signature motif NERLEFLGDS indicated by the bracket. The six catalytic residues found in all active eukaryotic RNase III domains are highlighted in red; for *S. cerevisiae* Rnt1, these correspond to E241, D245, N278, K313, D317, and E320 (class I prokaryotic RNase III contains only four catalytic residues, lacking residues corresponding to N278 and K313). In the *Candia dicer*-like (Cdl1/2) class of enzymes, these residues have been altered to become catalytically inactive. The function of these catalytically "dead" RNase III orthologues remains unknown. The multiple sequence alignment was performed using sequences from the Fungal Orthogroups Repository (http://www.broadinstitute.org/regev/orthogroups/) with the exception that *Candida parapsilosis* orthologues Cpar1/2/3 were identified by BLAST. ClustalX 2.1 program produced the alignment with default settings. Species abbreviations: *Scer*, *Saccharomyces cerevisiae*; *Spar*, *Saccharomyces paradoxus*; *Smik*, *Saccharomyces mikatae*; *Sbay*, *Saccharomyces bayanus*; *Kwal*, *Kluyveromyces waltii*; *Sklu*, *Saccharomyces kluyveri*; *Klac*, *Kluyveromyces lactis*; *Agos*, *Ashbya gossypii*; *Cgla*, *Candida glabrata*; *Scas*, *Saccharomyces castellii*; *Clus*, *Clavispora lusitaniae*; *Dhan*, *Debaryomyces hansenii*; *Cgui*, *Candida guilliermondi*; *Ctro*, *Candida tropicalis*; *Calb*, *Candida albicans*; *Cpar*, *Candida parapsilosis*; *Lelo*, *Lodderomyces elongisporus*; *Sjap*, *Schizosaccharomyces japonicus*; *Soct*, *Schizosaccharomyces octosporus*; *Spom*, *Schizosaccharomyces pombe*; *Ylip*, *Yarrowia lipolytica*; *Anid*, *Aspergillus nidulans*; *Ncra*, *Neurospora crassa*. (C) Phylogenetic tree of class IIIb fungal RNase III enzymes using the multiple sequence alignment from (B) visualized in the DrawTree 3.66 program on the Phylogeny.fr server [4].

RNases III, including the noncanonical Dicers (Dcr1) and the orthologues of *Saccharomyces cerevisiae* Rnt1p (Fig. 10.1A). In addition to lacking helicase and PAZ domains, noncanonical Dicers direct dsRNA cleavage internally, while canonical Dicers require dsRNA termini with a 3′ OH overhang [7]. These mechanistic and sequence differences highlight a case of convergent evolution where RNase III enzymes of different origins evolved independently to take on the function of Dicer in the RNA interference pathway.

Modern-day budding yeasts outside of the *Candida* clade exhibit two different types of class IIIb enzymes, the first typified by *S. cerevisiae* Rnt1p and the second by *Saccharomyces castelli* noncanonical Dicer (Dcr1), which contains a second C-terminal dsRBD (Fig. 10.1A). The *Candida* lineage presents an exceptional case, as it contains two types of RNase III similar in domain architecture to Dcr1, but not syntenic with either *DCR1* or *RNT1* [8]. Class IIIb orthologues are found in all fungi, indicating that they evolved from an ancient fungal ancestor (Fig. 10.1B). The noncanonical Dicers are exclusive to the budding yeast clade, whereas class IIIa canonical dicers are exclusive to nonbudding yeasts and higher eukaryotes. The sequence similarities of noncanonical Dicers to Rnt1 suggest that they evolved from the *RNT1* gene duplication and acquisition of a second dsRBD in an early budding yeast ancestor, perhaps with concomitant loss of the canonical Dicer [8]. Interestingly, the *Candida* lineage appears to have lost Rnt1p, while members of the *Saccharomyces* lineage, including *S. cerevisiae*, have lost Dcr1 and thus RNAi as well. In *Candida albicans*, a single noncanonical Dicer has evolved to perform both RNAi- and Rnt1-like processing functions, highlighting the functional plasticity of RNase III enzymes [8]. Rnt1p orthologue functions in non-RNAi dsRNA processing related to the maturation of noncoding RNA and surveillance of mRNAs are discussed in further detail in the next sections.

Biochemical and structural analyses have shown that *S. cerevisiae* Rnt1p specifically binds and cleaves dsRNAs capped by tetraloops exhibiting an NGNN sequence [9–12], as long as the sequence of the loop adopts the AGNN-type fold and does not form a GNRA- or UNCG-type fold [9,12,13]. An exception in the NGNN sequence was described for one substrate [14,15], but this loop sequence was found to adopt an AGNN-type fold upon binding of the Rnt1p dsRBD [11]. Bioinformatics analyses have shown that some of the budding yeast Rnt1p enzymes have evolved to conserve that specificity [16]. It is currently unknown how the dual-function Dcr1 enzymes present in the *Candida* clade recognize both specific stem-loop substrates, some of which contain the NGNN

tetraloops and others which do not, such as the generic substrates of the RNAi pathway. Interestingly, the *Candida* clade also contains catalytically inactive Dicer-like proteins termed Cdl (*Candida* Dicer-like). This protein family contains the same domain architecture as the Dcr1 protein family, but all six active site residues have been mutated to inactive amino acids (Fig. 10.1B). Other conserved residues within the catalytic domain have been retained in Cdl proteins, suggesting that the catalytic domain retains the overall domain fold and may serve a noncatalytic function [8]. As this protein family retains both dsRBDs, it is possible that it has a role in binding dsRNA *in vivo*. Notably, RNase III cleavage requires two catalytic domains assembled on the dsRNA substrate, so class I and class IIIb enzymes must form intermolecular dimers for cleavage to occur. It is possible that heterodimers may form in species with more than one RNase III enzyme and that the catalytically inactive Cdl1 proteins may interact with the active Dcr1 enzymes and regulate binding and/or cleavage activity.

3. RIBOSOMAL RNA PROCESSING AND RNA POLYMERASE I TRANSCRIPTIONAL TERMINATION

One of the first discovered physiological functions of fungal RNases III was the cleavage of the precursor of ribosomal RNA to generate the 35S pre-rRNA [17,18]. This cleavage occurs at a stem-loop structure located downstream from the 25S rRNA in the 3′ external transcribed spacer (ETS) region (Fig. 10.2), and this stem-loop is found in the rDNA of many Hemiascomycetes [8,16]. This processing function is conserved throughout fungi [16], but in the case of fungal species such as *Candida*, which lack a classical Rnt1p-like enzyme, this function is fulfilled by Dcr1, which is the only active RNase III enzyme in this organism [8]. A function for RNase III in the processing of the pre-rRNA seems highly conserved across eukaryotes, as RNase III homologues have been shown to be involved in pre-rRNA processing in human cells [19] and in plants [20].

This processing of the pre-rRNA occurs cotranscriptionally, as *S. cerevisiae* Rnt1p was shown to localize at the site of transcription of the ribosomal DNA and can be crosslinked to the ribosomal DNA [21]. In addition, Rnt1p can interact with several RNA polymerase I (Pol I) subunits [22], suggesting that a physical interaction with individual subunits of the polymerase in the context of elongation might help bring Rnt1p to its processing site once the polymerase has passed the region of the 25S rRNA.

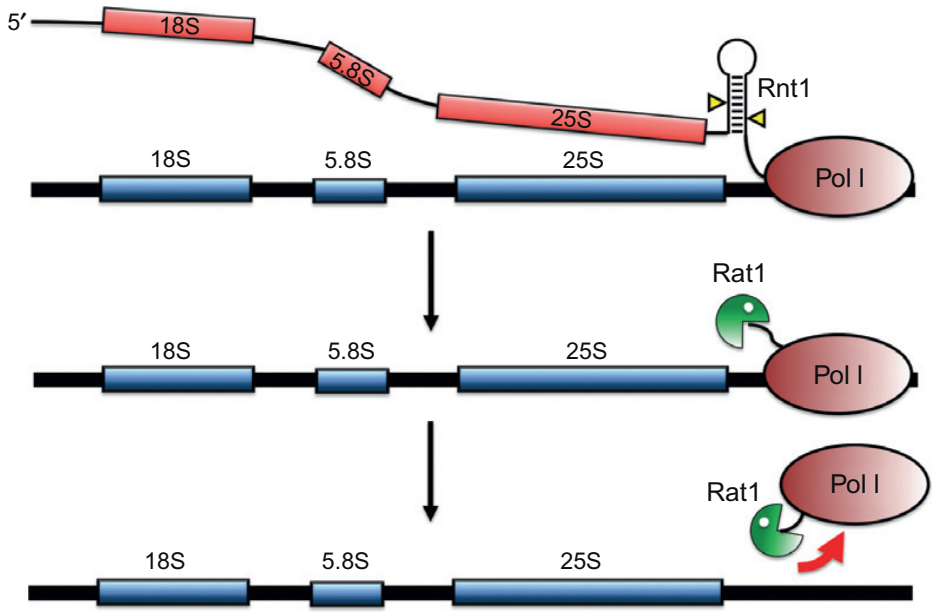


Figure 10.2 Cotranscriptional rRNA processing is coupled to Pol I termination. Cleavage of nascent pre-rRNA precursor by Rnt1p generates the 35S precursor and provides an entry site for the Rat1p exonuclease to degrade the downstream nascent RNA and simultaneously release Pol I from the template.

In the case of *Schizosaccharomyces pombe*, processing at the 3′-end of the pre-rRNA by the RNase III Pac1 occurs within a large complex called the RAC complex [23], which influences the cleavage efficiency by Pac1.

RNase III not only initiates processing of the pre-rRNA [24], but it also triggers termination of transcription by RNA Pol I, as evidenced by the fact that Rnt1p-deficient *S. cerevisiae* strains are defective in transcription termination [25–27]. The current model, as depicted in Fig. 10.2, proposes that Rnt1p cleavage in the 3′-ETS introduces an entry site for the 5′–3′ exonuclease Rat1p, which would then catch up with the elongating RNA Pol I and act as a "torpedo" to trigger dissociation of the polymerase from the rDNA, resulting in termination [27,28]. Additional evidence for the fact that the processing activity of RNase III influences rDNA transcription was provided by the observation that the absence of RNase III leads to an increase in the number of transcriptionally active ribosomal DNA repeats and to an opening of repressed rDNA chromatin [22]. While this appears to be an attempt by the cell to compensate for the severe defects in pre-rRNA processing in the absence of Rnt1p, the precise mechanism that results in the change in rDNA chromatin architecture remains unknown.

While important to promote efficient production of ribosomal RNAs, the function of RNases III in pre-rRNA processing is not absolutely essential, as many fungal species are viable in the absence of RNase III activities. While this might be due to the duplication of genes encoding RNase III enzymes for some of these species, species such as *S. cerevisiae*, express only one member of this class of enzymes. In the absence of Rnt1p cleavage, a second fail-safe termination event is known to take place, which results in the production of an elongated form of the 25S ribosomal RNA transcript [26]. Thus, the production of functional rRNAs in the absence of RNase III activity reflects the ability of other ribonucleases to fulfill its 3′-processing functions, albeit less efficiently than if 3′-end trimming is preceded by Rnt1p endonucleolytic cleavage.

4. SMALL NUCLEAR RNAs PROCESSING

Small nuclear RNAs comprise five transcripts (U1, U2, U4, U5, and U6), one of which is transcribed by RNA Pol III (U6) and the others by RNA Pol II. Most of these RNAs, with the exception of U6, are processed at their 3′-ends by multiple redundant 3′-processing pathways, some of which involve cleavage of the 3′-extended small nuclear RNA (snRNA) precursor by RNase III (Fig. 10.3). The first hint of a function for fungal RNases III in snRNA processing was suggested by genetic suppression data showing that overexpression of the *S. pombe* RNase III Pac1 was sufficient to suppress defects in snRNA biogenesis in this species [29]. Direct evidence for a function for RNase III in the processing of snRNAs was first shown by the reconstitution of the U5 snRNA processing pathway *in vitro* [30]. U5 snRNA exists in *S. cerevisiae* in two forms, which differ at their 3′-end. The longer form, U5L, is processed in an Rnt1p-dependent pathway, while the short form, U5S, is processed in an Rnt1p-independent manner [30]. Cleavage of a stem-loop structure located close to the mature 3′-end of U5L generates an entry site for 3′–5′ exonucleases to generate the mature 3′-end. This final trimming is carried out by the nuclear exosome [31] and/or the Rex exonucleases [32]. The 3′-to-5′ exonuclease activity of the exosome is likely stimulated by polyadenylation of the Rnt1p cleavage product [33]. Exonuclease progression is blocked at the mature 3′-end by the binding of the Sm proteins [34], which ensure the production of the correct 3′-end. This pathway has proven to be similar for the U1, U2, and U4 snRNAs in *S. cerevisiae* [31,34,35], except that these snRNAs exhibit only one major 3′-end. For each of these snRNAs, cleavage by Rnt1p initiates one

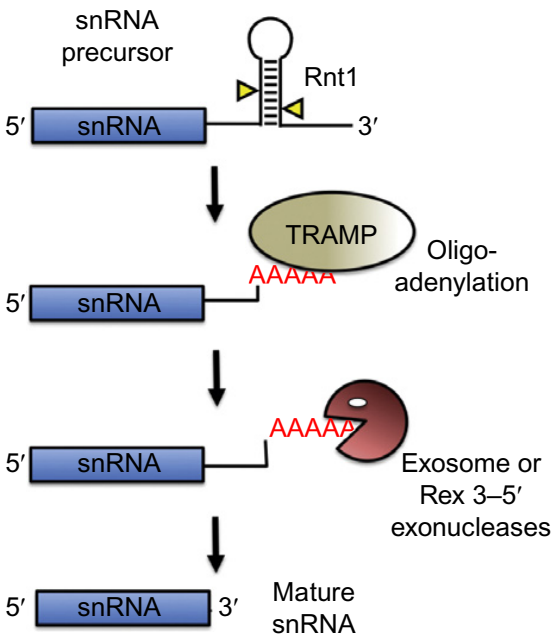


Figure 10.3 Small nuclear RNA processing by RNase III Rnt1p. Pol II-transcribed U1, U2, U4, and U5 snRNA precursors contain Rnt1p target stem-loops downstream of the mature 3′-end. Cleavage by Rnt1 provides an entry site for Rrp6 and the nuclear exosome, and/or the Rex 3′-to-5′ exonucleases, to produce the mature 3′-end. The exonuclease activity of the exosome is stimulated by TRAMP-mediated oligoadenylation of the Rnt1 cleavage product. This pathway is similar for the U3 snoRNA.

possible processing pathway, but alternative 3′-processing pathways ensure the production of functional snRNAs in the absence of Rnt1p cleavage. For instance, the cleavage and polyadenylation machinery can generate polyadenylated but functional U2 snRNA when Rnt1p cleavage activity is inhibited [35], and Rnt1p-independent pathways have also been described for the processing of the U1 snRNA [34].

The function of RNase III in snRNA 3′-end processing is probably conserved in many fungi, at least in species evolutionarily close to *S. cerevisiae*. The presence of potential stem-loop structures that obey the cleavage specificity of Rnt1p was predicted based on the analysis of sequences downstream from snRNA genes in some Hemiascomycetes genomes [16]. Furthermore, the involvement of RNase III in snRNA processing was demonstrated experimentally for *S. pombe* Pac1 for U2 [36] and for U4 in *Candida* [8]. In *S. pombe*, cleavage of the precursor by Pac1p also promotes termination of transcription by RNA Pol II [37]. However, the quantitative contribution of RNase III to snRNA 3′-end processing might vary depending upon the species. For instance, in *Candida*, the major form of

the U5 snRNA is the U5S form [38], which, assuming functional conservation with *S. cerevisiae*, does not involve RNase III cleavage. Only a small amount of U5L can be detected in *Candida* [38], suggesting that the contribution of RNase III to U5 processing might be minor in this species in comparison to *S. cerevisiae*. In addition, in several species close to *S. cerevisiae* such as *S. dairensis* and *S. kluyveri*, only one form of U5 can be detected by Northern blot [39], which corresponds in size to the short form of U5. Thus, it appears that in these yeast species, the stem-loop structure that forms the Rnt1p recognition site might have been lost or is no longer efficiently recognized by the RNase III homologue, indicative of rapid evolution of RNA processing pathways in these species. Alternatively, it is possible that RNase III cleavage intermediates can be more readily trimmed to the shorter form in species other than *S. cerevisiae*.

5. FUNCTIONS IN SMALL NUCLEOLAR RNAs PROCESSING

Small nucleolar RNAs (snoRNAs) are noncoding RNAs involved in the processing and modification of ribosomal RNAs. They are grouped in two distinct families, the box C/D family, which catalyzes methylation of 2′-hydroxyls of the pre-rRNA precursor, and the box H/ACA family, which catalyzes the modification of uridines into pseudouridines in various RNAs (reviewed in Refs. [24] and [40]). Both families of snoRNAs are bound *in vivo* by four core proteins, different in each family, to form complexes termed small nucleolar ribonucleoproteins (snoRNPs), where one protein provides the enzymatic activity and the snoRNA provides target specificity via the guide sequence (reviewed in Refs. [24] and [40]). Expression of these snoRNAs is relatively diverse in fungi, as some of them are expressed from independent transcription units, others from polycistronic transcripts containing multiple snoRNAs in the same precursor, while a few are produced from the introns of host mRNA genes (which is how most snoRNAs are expressed in metazoans; reviewed in Ref. [41]). Because of this diversity in their genomic organization, it is not surprising that RNase III cleavage can have very diverse functions in the processing of snoRNAs. Depending on the location of the target stem-loop structures, cleavage of these precursors can promote 5′- and/or 3′-end processing, and in some cases, releases individual snoRNAs from polycistronic precursors or from introns of pre-mRNA transcripts.

5.1. Processing of independently transcribed snoRNA precursors

For most snoRNAs produced by independent transcription units, the main function of RNase III cleavage is to initiate the 5′-end processing. Cleavage of a stem-loop structure present in the 5′-extension of the precursors typically provides an entry site for 5′–3′ exonucleolytic processing, which is fulfilled by the exonucleases Xrn1p and Rat1p (Fig. 10.4A; [42]). These exonucleases stop at the mature 5′-end, probably because the binding of the snoRNP core proteins to the snoRNA block progression into the mature sequence. This 5′-end processing pathway is frequently found for *S. cerevisiae* box C/D snoRNAs [14,42,43] and less frequent for box H/ACA snoRNAs [43,44]. The conservation of stem-loop structures in other Hemiascomycetes suggests that at least some snoRNAs undergo a similar 5′-end processing pathway in these other fungal species [16]. As shown for other processing pathways, the contribution of Rnt1p cleavage to the production of mature snoRNAs depends on snoRNA species; some of these snoRNAs can be processed relatively efficiently in the absence of Rnt1p [42], showing that the 5′–3′ exonucleases Xrn1p and Rat1p can process the snoRNAs directly from decapped precursors.

For most of these independently transcribed snoRNAs, RNase III cleavage is not a major determinant of 3′-end processing. There are a few exceptions to this rule. One is the *snR40* transcript, for which the target stem-loop that promotes processing loops out the snoRNA sequence [43]. In this case, Rnt1p cleavage initiates both the 5′ and 3′ processing of the snoRNA. Another exception is the U3 snoRNA, which is involved in ribosomal RNA processing [24]. In *S. cerevisiae*, U3 is produced from two independent transcription units in the genome. The U3 snoRNAs undergo capping, splicing [45], and processing at the 3′-end. The U3 3′-processing pathway strongly resembles that of the U1,U2, U4, and U5 snRNAs (Fig. 10.3), as a stem-loop is present downstream from the mature 3′-end, and Rnt1p cleavage introduces an entry site for 3′–5′ processing by the exosome [46], which is stimulated by TRAMP-mediated polyadenylation of the Rnt1-cleaved U3 intermediate [33]. The involvement of RNase III in the 3′-end processing of U3 is conserved in *S. pombe* [47], and in that organism, Pac1 cleavage also dictates the efficiency of termination of U3 transcription. It is unknown why U3 undergoes a processing pathway that differs dramatically from that of other snoRNAs. However, this might be linked to the unique protein composition of the U3 snoRNP compared to the canonical box C/D snoRNPs and to the essential function of U3 in rRNA processing [48].

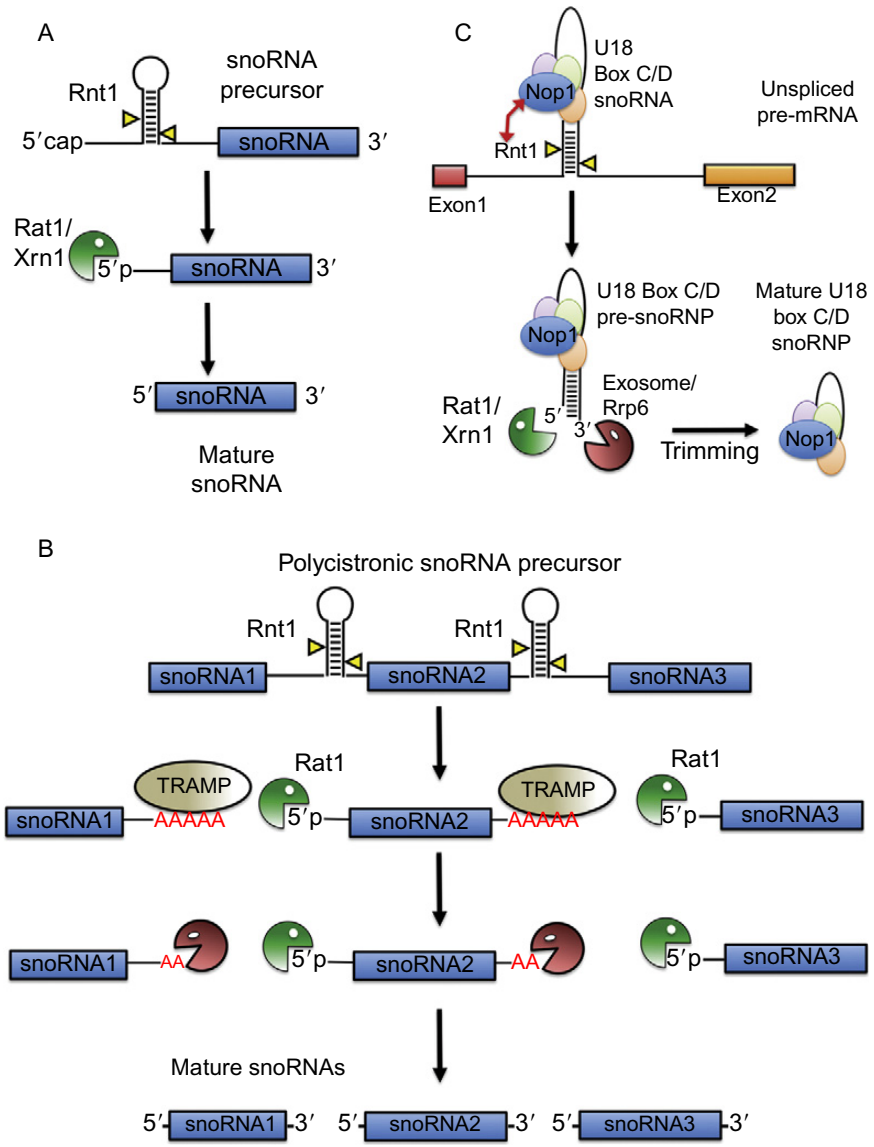


Figure 10.4 Small nucleolar RNA processing by RNase III Rnt1p. (A) Monocistronic snoRNA nascent transcripts typically contain Rnt1p target stem-loops upstream of the mature 5′-end. Rnt1 cleavage provides an entry site for the Rat1 and/or Xrn1p 5′-to-3′ exonucleases to bypass the 5′ cap structure and produce the proper 5′-end. (B) Polycistronic small nucleolar RNAs processing is initiated by the RNase III Rnt1p. Individual snoRNAs within a polycistronic array are released from one another by Rnt1 cleavage of the precursor RNA at target stem-loops in between the mature snoRNA sequences. The released snoRNAs are then trimmed from the 5′- to the 3′-end by Rat1 and/or Xrn1p, and from the 3′- to the 5′-end by Rrp6 and the nuclear exosome, which is likely stimulated by the TRAMP oligoadenylation activity. (C) RNase III Rnt1 processes some intron-embedded snoRNAs. In a pathway alternative to debranching of the lariat splicing product, Rnt1p acts to release some of these box C/D snoRNAs from the intron with concerted cleavage of a stem composed of sequences upstream and downstream of the mature snoRNA. The U18 precursor lacks an Rnt1p tetraloop, and assembly of the Box C/D core protein Nop1p on the intronic snoRNA precursor instead recruits Rnt1p for dsRNA cleavage. After Rnt1p cleavage, exonucleolytic trimming produces both termini of the mature transcript.

5.2. Cleavage of polycistronic transcription units

In yeast, some snoRNAs are produced from polycistronic transcription units, which express several snoRNAs on a single precursor [49]. This type of genomic organization is also conserved in fungi and frequent in plants [50]. For many snoRNAs encoded by polycistronic transcription units, RNase III cleavage serves the purpose of separating the snoRNAs from each other, thereby generating intermediates that are further processed by the 5′–3′ exonucleases Rat1p and Xrn1p and by the exosome to generate the mature 5′- and 3′-ends, respectively (Fig. 10.4B; [14,43,51–53]). Many snoRNAs are encoded as precursors containing only two or three snoRNAs; for these precursors, the presence of one or two stem-loop structures in the precursor is sufficient to separate snoRNA from each other (Fig. 10.4B). However, some precursors are known to contain a larger number of snoRNAs, such as the polycistronic snR72-78 precursor, which generates seven individual snoRNAs from a single precursor [43,52]. In this case, a stem-loop structure which serves as an Rnt1p recognition and cleavage site is usually found in between individual snoRNA sequences [43,52]. In some particular cases, the architecture of the stem-loops seems to have been adapted to maximize the processing function of Rnt1p using the smallest number of recognition sites. For instance, the processing of the three snoRNAs snR55, snR57, and snR61 from their tricistronic precursor requires only one Rnt1p tetraloop recognition site [14]. The short stem in which this tetraloop is located can stack onto two distinct longer stems, one of which loops out snR55, while the other separates snR57 from snR61 [14]. Thus, one tetraloop recognition site in the precursor can mediate two distinct cleavage events, separating three snoRNA transcripts from each other. As observed for other snRNA and snoRNA processing events, the presence of tetraloop signals in polycistronic snoRNA precursors is conserved in many Hemiascomycetes [16], suggesting that RNase III processing of polycistronic snoRNA precursors is a common feature in fungi.

5.3. Processing of intron-encoded snoRNAs

In mammalian cells, almost all snoRNAs are contained within introns, and the processing of most of these snoRNAs is splicing dependent and occurs by debranching of the lariat structures of the introns by the debranching enzyme and subsequent exonucleolytic trimming from both directions [41]. In *S. cerevisiae*, a few snoRNAs are also found within introns, and their processing pathway was thought to occur mostly through debranching as

well [53,54], with speculations that random cleavage in the lariat structure could also contribute to processing [54]. However, *S. cerevisiae* RNase III can provide an alternative or complementary mechanism of initiating the processing of snoRNAs from introns (Fig. 10.4C). The first example described was the U18 snoRNA, for which Rnt1p cleavage occurs in an intronic stem-loop structure that loops out the snoRNA sequence [55]. The target stem-loop is a noncanonical structure that lacks a *bona fide* AGNN tetraloop, and the binding of Rnt1p to this noncanonical site is thought to be assisted by an interaction with the snoRNP protein Nop1p, which binds the snoRNA and recruits Rnt1p to this cleavage site [55]. The observation that Rnt1p can cleave and initiate the processing of U18 provides an explanation for the observation that a substantial amount of mature U18 snoRNA can be produced when the debranching enzyme Dbr1p is inactivated [53]. Thus, although the level of U18 snoRNA is not affected in a strain lacking Rnt1p, this enzyme can clearly provide an alternative processing pathway to splicing and debranching.

Two other examples of processing of intron-encoded snoRNAs by Rnt1p cleavage were described for the snR39 and snR59 snoRNAs [14]. Interestingly, the quantitative contribution and cleavage mechanism of Rnt1p for the processing of these intron-encoded snoRNAs seem to vary depending on the individual snoRNA. For snR39, which is encoded in the second intron of the ribosomal protein gene *RPL7A*, Rnt1p cleavage plays only a minor role in the maturation of this snoRNA, as the production of most of the mature *snR39* is strongly decreased in a strain lacking the debranching enzyme Dbr1p. However, the production of mature snR39 is completely blocked in a *dbr1Δ rnt1Δ* double mutant, showing that Rnt1p cleavage contributes to some extent to the production of the mature snoRNA [14]. Interestingly, cleavage by Rnt1p occurs only on the lariat form of the excised intron, but not on the unspliced precursor form [14]. Thus, as described above for U18, Rnt1p function in the processing of snR39 seems to act as an alternative to Dbr1p in initiating the processing of the snoRNA from the excised lariat introns. By contrast, the cleavage by Rnt1p to initiate processing of snR59 occurs *only* on the unspliced precursor form of the host transcript *RPL7B* and this cleavage initiates degradation of the entire pre-mRNA [14]. Thus, as shown for cleavage in other unspliced precursors such as *RPS22B* (see below), cleavage by Rnt1p is mutually exclusive to splicing of the pre-mRNA, suggesting a competition between production of mature mRNA and production of mature snoRNA

from the same precursor by Rnt1p. It is interesting to note that, while snR39 and snR59 are encoded in introns of duplicated genes coding for the same ribosomal protein (the Rpl7 paralogues), differences in processing pathways have emerged.

6. (PRE)-mRNA SURVEILLANCE, DEGRADATION, AND REGULATION

Because most of the *S. cerevisiae* RNase III localizes in the nucleus [21,56], and in particular, at the site of transcription of the rRNA, it was initially thought that the function of fungal RNase III was limited to the processing of nuclear noncoding RNAs. However, multiple studies have shown that fungal RNases III can also cleave a number of mRNAs. This cleavage serves two main purposes. First, it can provide a way to discard unprocessed or improperly processed mRNAs. Additionally, cleavage can provide an alternative mechanism to degrade mRNAs or limit their accumulation in conditions in which they should be repressed.

A role for fungal RNase III in (pre)-mRNA degradation was first discovered by the identification of target stem-loop structures in the intronic sequences of the ribosomal protein genes *RPL18A* and *RPS22B* [57]. For these transcripts, Rnt1p cleavage can serve two purposes (Fig. 10.5). The first function is to cleave the lariat introns produced by the splicing reaction, providing a mechanism of degradation of these lariats complementary to that provided by the debranching enzyme [57]. The second function is to directly compete with splicing or degrade unspliced pre-mRNAs that have escaped the splicing pathway [57]. As described above, cleavage by Rnt1p provides an entry site for exonucleolytic degradation by the exosome and the 5′–3′ exonucleases Xrn1p and Rat1p. Interestingly, the relative dependence upon the nuclear Rat1p or the cytoplasmic Xrn1p exonucleases for degradation of the 3′ cleavage products seem to vary between *RPS22B* and *RPL18A* [57], suggesting that Rnt1p cleavage intermediates can be exported to the cytoplasm or retained in the nucleus with varying efficiencies. This function in the surveillance of unspliced pre-mRNA species was also demonstrated for the *MAT*a1 transcript [58]. Inhibition of Rnt1p cleavage was found to result in an increase in mature mRNA levels [57,58]. This observation is consistent with the idea that cleavage of the pre-mRNAs by Rnt1p occurs very early in the biogenesis pathway and competes with splicing, thus titrating a fraction of the precursors away from the splicing pathway. This could potentially constitute a means of regulating the

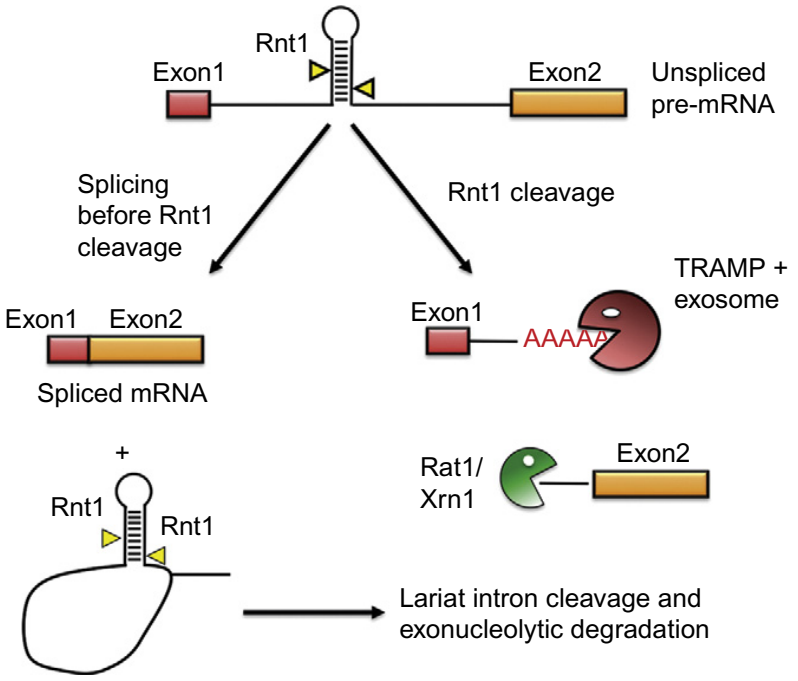


Figure 10.5 RNase III regulation of unspliced pre-mRNAs and lariat degradation. Various intron-containing protein coding genes in *S. cerevisiae* contain intronic Rnt1p target stem-loops. Rnt1p cleavage serves to limit the accumulation of the unspliced precursor and also acts redundantly with the debranching enzyme Dbr1p to degrade the lariat released after splicing of these precursors. Competition between splicing and Rnt1p cleavage regulates levels of the mature mRNA in at least one case, the *RPS22B* ribosomal protein mRNA.

expression of these transcripts; however, it is unknown if the cleavage of these unspliced pre-mRNAs by Rnt1p can be regulated during specific growth conditions. The surveillance function for Rnt1p in the degradation of unspliced pre-mRNAs can be partially redundant with those of other degradation systems, such as degradation by Rat1p or by the nonsense-mediated decay (NMD) pathway. For instance, degradation of *MAT*a1 pre-mRNAs seems to be the most effective only when both Rat1p and Rnt1p are active [58], and the full accumulation of unspliced *RPS22B* pre-mRNAs is observed only when both the Rnt1p stem-loop structure and the NMD system are inactivated [59]. The existence of multiple degradation pathways that localize to different compartments highlights the importance of limiting the accumulation and aberrant translation of unspliced pre-mRNAs.

In addition to these functions in the surveillance of unprocessed mRNAs, a large number of mature mRNAs appear to be cleaved by RNase III in *S. cerevisiae*. These include the mRNA encoding the transcriptional

repressor Mig2p [60], mRNAs coding for proteins involved in iron uptake and assimilation [61] and the acireductone dioxygenase [62], mRNAs encoding proteins associated with the telomerase complex [63], and mRNAs encoding the Swi4p and Hsl1p proteins involved in the cell wall stress response [64]. As described previously, cleavage by RNase III generates entry sites for exonucleases to complete their degradation [61,65]. The lack of cleavage of these pre-mRNAs has been associated with a number of cellular phenotypes [64], including sensitivity to high iron concentrations [61]. However, it is not entirely clear to what extent some of these cellular phenotypes are directly due to the lack of cleavage of these mRNAs in the *rnt1* deletion strain or whether the general growth defect of this strain contributes indirectly to some of these phenotypes.

It does not appear that the genomic context or accessory sequences contribute in a major way to RNase III cleavage on mRNAs, as various stem-loop structures can be transposed onto plasmid-borne reporter mRNAs, which results in the cleavage and degradation of these mRNAs [65,66]. However, cleavage efficiency can vary depending on the stem-loop/reporter combination [65], suggesting that for some of these combinations, the presence of competing secondary structures may impede proper cleavage activity. The ability of Rnt1p to cleave these structures independently from their genomic context has been used to develop synthetic gene expression reporters. Stem-loop structures with various cleavage efficiencies can be inserted into reporter mRNAs to generate a range of reporters to fine-tune gene expression [67] as well as to design gene reporters that are responsive to the binding of metabolites to the reporter mRNA [66].

It is unclear whether other fungal homologues of Rnt1p also cleave specific mRNAs. Overexpression of *S. pombe* Pac1p inhibits mating and sporulation [68], quite possibly by cleaving mRNA(s) that are required for the proper completion of the mating and sporulation processes. However, despite this initial genetic observation, the identity of these mRNAs has yet to be determined. Besides this observation, it is unknown whether RNase III contributes to mRNA degradation independently from of the RNA interference pathway in other fungal species outside *S. cerevisiae*.

7. RNA POLYMERASE II TERMINATION

Transcription termination of RNA Pol II has emerged as an important step in gene expression that can directly affect the fate of the transcribed RNA. Furthermore, proper partitioning of transcription units via efficient

termination is crucial to avoid transcriptional interference and aberrant production of overlapping sense–antisense transcripts [69]. Initial studies had shown that cleavage by fungal RNase III downstream from small RNA genes could trigger termination of the RNA Pol II for these ncRNA transcription units, suggesting that RNase III cleavage might serve a general role in triggering termination of RNA Pol II [37,47]. Recent work has highlighted the role of Rnt1p in providing a more widespread fail-safe mode of RNA Pol II termination, in situations where a failure to cleave at the poly(A) site results in transcription of a downstream Rnt1-target stem-loop [70,71]. Cotranscriptional cleavage by Rnt1p is followed by the binding of the 5′-to-3′ exonuclease Rat1p on the now unprotected 5′-end of the nascent transcript (Fig. 10.6). Upon reaching the elongating RNA Polymerase, Rat1p promotes the release of Pol II from the template [71], in a mechanism similar to the one proposed for the "torpedo" model of transcription termination of RNA Pol I by Rat1p [27,28]. Lack of cleavage at the poly(A) site can be the result of weak

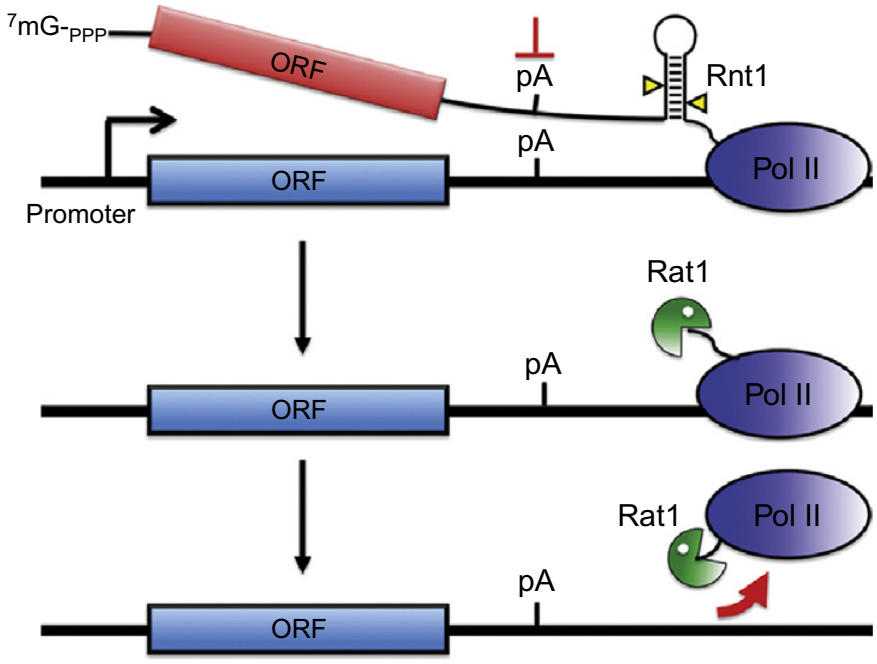


Figure 10.6 Fail-safe transcription termination of RNA Polymerase II by RNase III. Cleavage and polyadenylation of Pol II transcripts can fail to occur due to *cis*-acting elements (e.g., a weak polyadenylation (pA) signal) or *trans*-acting factors (e.g., RNA-binding proteins). In many cases, Rnt1p target stem-loops exist downstream of these pA signals and elicit cotranscriptional cleavage by Rnt1. This is followed by recruitment of Rat1 5′-to-3′ exonuclease, resulting in degradation of the nascent transcript and the release of Pol II from the template.

cis termination or polyadenylation signals, or the presence of *trans*-acting factors inhibiting recognition of the poly(A) site [70]. This effect on termination could also explain the presence of extended species of mRNAs showing target stem-loop structures when Rnt1p is inactivated [61,62]. It has been suggested that this function for RNase III in transcription termination may be a general feature of eukaryotic gene expression [72]. Indeed, the RNase III Dicer-like 4 of *Arabidopsis thaliana* has been shown to be required for transcription termination of the FCA gene involved in chromatin silencing [73]. This highlights the ability of Dicer enzymes to have RNAi-independent functions and suggests that different classes of RNase III may have independently evolved to provide fail-safe termination throughout eukaryotes.

8. CONCLUSIONS AND PERSPECTIVES

In this chapter, we have reviewed the diverse functions of fungal RNase III enzymes in the processing of noncoding RNAs and in the degradation of mRNAs. We have excluded from this discussion the role of this class of enzymes in the RNAi and siRNA pathways, as these will be described extensively in Volume 32. While this class of enzymes plays important functions in the processing of many noncoding RNAs (rRNA, snRNA, and snoRNA), it is clear that functional redundancies are found in almost all processing pathways described, such that inactivation of these enzymes usually results in only a partial inhibition of the production of the mature species. Similar redundancies are also observed for many of the pre-mRNA surveillance functions associated with Rnt1p, as other degradation pathways mediated by Dbr1p, Rat1p, or NMD can effectively compensate for the absence of Rnt1p cleavage. While it is clear that the redundancy is rampant in RNase III-mediated processing and surveillance functions, it is increasingly evident that RNase III cleavage has evolved additional specialized roles in gene expression that are important under specific conditions. For example, in conditions of high iron levels or cell wall stress, RNAse III cleavage is critical for preventing the excessive accumulation of target mRNAs, where high levels of protein expression from these mRNAs would be detrimental.

Most of the examples of RNase III-mediated RNA processing and degradation derive from the work done with the RNase III Rnt1p from the model Hemiascomycete *S. cerevisiae*. While it is clear that some of the processing pathways are conserved in other Hemiascomycetes, the available

evidence suggests that there is also significant variation in the quantitative contributions of RNase III to these various processing pathways, even in closely related species. This variation probably reflects the existence of multiple processing and degradative pathways for individual transcripts, thus allowing for a rapid evolutionary loss of a particular processing pathway. Future work using other model yeast and fungi should further reveal the extent to which RNA processing pathways have been conserved or altered throughout fungal evolution.

Some cellular phenotypes linked to the absence of these enzymes still lack a clear molecular explanation. For example, inactivation of *S. cerevisiae* Rnt1p leads to cell-cycle defects, but these defects can be rescued by the expression of a catalytically inactive mutant [56]. This result suggests the existence of cleavage-independent functions of RNase III, at least for the *S. cerevisiae* enzyme, but it is unknown what molecular functions of the enzyme are directly responsible for these defects. The cell-cycle rescue was shown to be dependent on trafficking of Rnt1p from the nucleolus to the nucleoplasm at the G2/M transition of the cell cycle, suggesting that Rnt1p function may be regulated by the subcellular localization pattern of this enzyme. It has been speculated that this phenomenon may also regulate the Rnt1p cleavage activity on mRNAs, which would likely require the recruitment of Rnt1p from the nucleolus to the nucleoplasm for their cleavage [64]. Future work will elucidate the ways in which fungal RNAse III activity is regulated and whether this is conserved across species. It is clear that RNase III enzymes have evolved to provide cells with an additional tool to fine-tune gene expression and perform quality control. The presence of simple dsRNA structures embedded throughout the transcriptome provides a rapid and potent means of regulating gene expression by multiple mechanisms.

ACKNOWLEDGEMENTS

This work was supported by NIH Grant GM61518 to G. F. C.

K.R. was supported by a Whitcome Fellowship and Ruth L. Kirschtein National Research Service Award GM007185.

REFERENCES

[1] Robertson HD, Webster RE, Zinder ND. A nuclease specific for double-stranded RNA. Virology 1967;32(4):718–9.

[2] Robertson HD, Webster RE, Zinder ND. Purification and properties of ribonuclease III from Escherichia coli. J Biol Chem 1968;243(1):82–91.

[3] Dennis PP. Ancient ciphers: translation in Archaea. Cell 1997;89(7):1007–10.

[4] Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 2008;36 (Web Server issue):W465–9.
[5] Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, et al. Structural basis for double-stranded RNA processing by Dicer. Science 2006;311(5758):195–8.
[6] MacRae IJ, Doudna JA. Ribonuclease revisited: structural insights into ribonuclease III family enzymes. Curr Opin Struct Biol 2007;17(1):138–45.
[7] Weinberg DE, Nakanishi K, Patel DJ, Bartel DP. The inside-out mechanism of Dicers from budding yeasts. Cell 2011;146(2):262–76.
[8] Bernstein DA, Vyas VK, Weinberg DE, Drinnenberg IA, Bartel DP, Fink GR. Candida albicans Dicer (CaDcr1) is required for efficient ribosomal and spliceosomal RNA maturation. Proc Natl Acad Sci USA 2012;109(2):523–8.
[9] Chanfreau G, Buckle M, Jacquier A. Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc Natl Acad Sci USA 2000;97 (7):3142–7.
[10] Wu H, Henras A, Chanfreau G, Feigon J. Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Proc Natl Acad Sci USA 2004;101(22):8307–12.
[11] Wang Z, Hartman E, Roy K, Chanfreau G, Feigon J. Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. Structure 2011;19(7):999–1010.
[12] Nagel R, Ares Jr. M. Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5′-AGNN-3′ tetraloop. RNA 2000;6(8):1142–56.
[13] Wu H, Yang PK, Butcher SE, Kang S, Chanfreau G, Feigon J. A novel family of RNA tetraloop structure forms the recognition site for Saccharomyces cerevisiae RNase III. EMBO J 2001;20(24):7240–9.
[14] Ghazal G, Ge D, Gervais-Bird J, Gagnon J, Abou Elela S. Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol Cell Biol 2005;25(8):2981–94.
[15] Ghazal G, Elela SA. Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. J Mol Biol 2006;363(2):332–44.
[16] Chanfreau G. Conservation of RNase III processing pathways and specificity in hemiascomycetes. Eukaryot Cell 2003;2(5):901–9.
[17] Elela SA, Igel H, Ares Jr. M. RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell 1996;85(1):115–24.
[18] Kufel J, Dichtl B, Tollervey D. Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3′ ETS but not the 5′ ETS. RNA 1999;5(7):909–17.
[19] Wu H, Xu H, Miraglia LJ, Crooke ST. Human RNase III is a 160-kDa protein involved in preribosomal RNA processing. J Biol Chem 2000;275(47):36957–65.
[20] Comella P, Pontvianne F, Lahmy S, Vignols F, Barbezier N, Debures A, et al. Characterization of a ribonuclease III-like protein required for cleavage of the pre-rRNA in the 3'ETS in Arabidopsis. Nucleic Acids Res 2008;36(4):1163–75.
[21] Henras AK, Bertrand E, Chanfreau G. A cotranscriptional model for 3′-end processing of the Saccharomyces cerevisiae pre-ribosomal RNA precursor. RNA 2004;10 (10):1572–85.
[22] Catala M, Tremblay M, Samson E, Conconi A, Abou Elela S. Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. Mol Cell Biol 2008;28(2):619–29.
[23] Spasov K, Perdomo LI, Evakine E, Nazar RN. RAC protein directs the complete removal of the 3′ external transcribed spacer by the Pac1 nuclease. Mol Cell 2002;9 (2):433–7.

[24] Venema J, Tollervey D. Ribosome synthesis in Saccharomyces cerevisiae. Annu Rev Genet 1999;33:261–311.
[25] Prescott EM, Osheim YN, Jones HS, Alen CM, Roan JG, Reeder RH, et al. Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. Proc Natl Acad Sci USA 2004;101:6068–73.
[26] Braglia P, Kawauchi J, Proudfoot NJ. Co-transcriptional RNA cleavage provides a failsafe termination mechanism for yeast RNA polymerase I. Nucleic Acids Res 2011;39(4):1439–48.
[27] Kawauchi J, Mischo H, Braglia P, Rondon A, Proudfoot NJ. Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. Genes Dev 2008;22(8):1082–92.
[28] El Hage A, Koper M, Kufel J, Tollervey D. Efficient termination of transcription by RNA polymerase I requires the 5′ exonuclease Rat1 in yeast. Genes Dev 2008;22 (8):1069–81.
[29] Rotondo G, Gillespie M, Frendewey D. Rescue of the fission yeast snRNA synthesis mutant snm1 by overexpression of the double-strand-specific Pac1 ribonuclease. Mol Gen Genet 1995;247(6):698–708.
[30] Chanfreau G, Elela SA, Ares Jr. M, Guthrie C. Alternative 3′-end processing of U5 snRNA by RNase III. Genes Dev 1997;11(20):2741–51.
[31] Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J 1999;18(19):5399–410.
[32] van Hoof A, Lennertz P, Parker R. Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. EMBO J 2000;19(6):1357–65.
[33] Egecioglu DE, Henras AK, Chanfreau GF. Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA 2006;12(1):26–32.
[34] Seipelt RL, Zheng B, Asuru A, Rymond BC. U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway. Nucleic Acids Res 1999;27 (2):587–95.
[35] Abou Elela S, Ares Jr. M. Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA. EMBO J 1998;17(13):3738–46.
[36] Zhou D, Frendewey D, Lobo Ruppert SM. Pac1p, an RNase III homolog, is required for formation of the 3′ end of U2 snRNA in Schizosaccharomyces pombe. RNA 1999;5(8):1083–98.
[37] Nabavi S, Nazar RN. Cleavage-induced termination in U2 snRNA gene expression. Biochem Biophys Res Commun 2010;393(3):461–5.
[38] Mitrovich QM, Guthrie C. Evolution of small nuclear RNAs in S. cerevisiae, C. albicans, and other hemiascomycetous yeasts. RNA 2007;13(12):2066–80.
[39] Roiha H, Shuster EO, Brow DA, Guthrie C. Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive size variation. Gene 1989;82(1):137–44.
[40] Henras AK, Soudet J, Gerus M, Lebaron S, Caizergues-Ferrer M, Mougin A, et al. The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci 2008;65 (15):2334–59.
[41] Filipowicz W, Pogacic V. Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 2002;14(3):319–27.
[42] Lee CY, Lee A, Chanfreau G. The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of S. cerevisiae box C/D snoRNAs. RNA 2003;9 (11):1362–70.
[43] Chanfreau G, Legrain P, Jacquier A. Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. J Mol Biol 1998;284(4):975–88.

[44] Hiley SL, Babak T, Hughes TR. 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 2005;33(9):3048–56.
[45] Myslinski E, Segault V, Branlant C. An intron in the genes for U3 small nucleolar RNAs of the yeast Saccharomyces cerevisiae. Science 1990;247(4947):1213–6.
[46] Kufel J, Allmang C, Chanfreau G, Petfalski E, Lafontaine DL, Tollervey D. Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. Mol Cell Biol 2000;20(15):5415–24.
[47] Nabavi S, Nazar RN. Pac1 endonuclease and Dhp1p 5′ → 3′ exonuclease are required for U3 snoRNA termination in Schizosaccharomyces pombe. FEBS Lett 2010;584 (15):3436–41.
[48] Venema J, Tollervey D. Processing of pre-ribosomal RNA in Saccharomyces cerevisiae. Yeast 1995;11(16):1629–50.
[49] Samarsky DA, Fournier MJ. A comprehensive database for the small nucleolar RNAs from Saccharomyces cerevisiae. Nucleic Acids Res 1999;27(1):161–4.
[50] Brown JW, Clark GP, Leader DJ, Simpson CG, Lowe T. Multiple snoRNA gene clusters from Arabidopsis. RNA 2001;7(12):1817–32.
[51] Chanfreau G, Rotondo G, Legrain P, Jacquier A. Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. EMBO J 1998;17 (13):3726–37.
[52] Qu LH, Henras A, Lu YJ, Zhou H, Zhou WX, Zhu YQ, et al. Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. Mol Cell Biol 1999;19(2):1144–58.
[53] Petfalski E, Dandekar T, Henry Y, Tollervey D. Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. Mol Cell Biol 1998;18 (3):1181–9.
[54] Ooi SL, Samarsky DA, Fournier MJ, Boeke JD. Intronic snoRNA biosynthesis in Saccharomyces cerevisiae depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. RNA 1998;4(9):1096–110.
[55] Giorgi C, Fatica A, Nagel R, Bozzoni I. Release of U18 snoRNA from its host intron requires interaction of Nop1p with the Rnt1p endonuclease. EMBO J 2001;20 (23):6856–65.
[56] Catala M, Lamontagne B, Larose S, Ghazal G, Elela SA. Cell cycle-dependent nuclear localization of yeast RNase III is required for efficient cell division. Mol Biol Cell 2004;15(7):3015–30.
[57] Danin-Kreiselman M, Lee CY, Chanfreau G. RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol Cell 2003;11(5):1279–89.
[58] Egecioglu DE, Kawashima TR, Chanfreau GF. Quality control of MATa1 splicing and exon skipping by nuclear RNA degradation. Nucleic Acids Res 2012;40(4):1787–96.
[59] Sayani S, Janis M, Lee CY, Toesca I, Chanfreau GF. Widespread impact of nonsense-mediated mRNA decay on the yeast intronome. Mol Cell 2008;31(3):360–70.
[60] Ge D, Lamontagne B, Elela SA. RNase III-mediated silencing of a glucose-dependent repressor in yeast. Curr Biol 2005;15(2):140–5.
[61] Lee A, Henras AK, Chanfreau G. Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. Mol Cell 2005;19(1):39–51.
[62] Zer C, Chanfreau G. Regulation and surveillance of normal and 3′-extended forms of the yeast aci-reductone dioxygenase mRNA by RNase III cleavage and exonucleolytic degradation. J Biol Chem 2005;280(32):28997–9003.
[63] Larose S, Laterreur N, Ghazal G, Gagnon J, Wellinger RJ, Elela SA. RNase III-dependent regulation of yeast telomerase. J Biol Chem 2007;282(7):4373–81.
[64] Catala M, Aksouh L, Abou Elela S. RNA-dependent regulation of the cell wall stress response. Nucleic Acids Res 2012 May 10. [Epub ahead of print].

[65] Meaux S, Lavoie M, Gagnon J, Abou Elela S, van Hoof A. Reporter mRNAs cleaved by Rnt1p are exported and degraded in the cytoplasm. Nucleic Acids Res 2011;39 (21):9357–67.
[66] Babiskin AH, Smolke CD. Engineering ligand-responsive RNA controllers in yeast through the assembly of RNase III tuning modules. Nucleic Acids Res 2011;39 (12):5299–311.
[67] Babiskin AH, Smolke CD. Synthetic RNA modules for fine-tuning gene expression levels in yeast by modulating RNase III activity. Nucleic Acids Res 2011;39 (19):8651–64.
[68] Iino Y, Sugimoto A, Yamamoto M. S. pombe pac1+, whose overexpression inhibits sexual development, encodes a ribonuclease III-like RNase. EMBO J 1991;10 (1):221–6.
[69] Kuehner JN, Pearson EL, Moore C. Unravelling the means to an end: RNA polymerase II transcription termination. Nat Rev Mol Cell Biol 2011;12(5):283–94.
[70] Ghazal G, Gagnon J, Jacques PE, Landry JR, Robert F, Elela SA. Yeast RNase III triggers polyadenylation-independent transcription termination. Mol Cell 2009;36 (1):99–109.
[71] Rondon AG, Mischo HE, Kawauchi J, Proudfoot NJ. Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol Cell 2009;36(1):88–98.
[72] Nabavi S, Nazar RN. Fail-safe termination elements: a common feature of the eukaryotic genome? FASEB J 2010;24(3):684–8.
[73] Liu F, Bakht S, Dean C. Cotranscriptional role for Arabidopsis DICER-LIKE 4 in transcription termination. Science 2012;335(6076):1621–3.

CHAPTER 3

Structure of a Yeast RNase III dsRBD Complex with a Noncanonical RNA Substrate Provides New Insights into Binding Specificity of dsRBDs

Structure of a Yeast RNase III dsRBD Complex with a Noncanonical RNA Substrate Provides New Insights into Binding Specificity of dsRBDs

Zhonghua Wang,[1,2,3] Elon Hartman,[1,2,3] Kevin Roy,[1,2] Guillaume Chanfreau,[1,2] and Juli Feigon[1,2,*]
[1]Department of Chemistry and Biochemistry
[2]Molecular Biology Institute
P.O. Box 951569, University of California, Los Angeles, CA 90095-1569, USA
[3]These authors contributed equally to this work
*Correspondence: feigon@mbi.ucla.edu
DOI 10.1016/j.str.2011.03.022

SUMMARY

dsRBDs often bind dsRNAs with some specificity, yet the basis for this is poorly understood. Rnt1p, the major RNase III in *Saccharomyces cerevisiae*, cleaves RNA substrates containing hairpins capped by A/uGNN tetraloops, using its dsRBD to recognize a conserved tetraloop fold. However, the identification of a Rnt1p substrate with an AAGU tetraloop raised the question of whether Rnt1p binds to this noncanonical substrate differently than to A/uGNN tetraloops. The solution structure of Rnt1p dsRBD bound to an AAGU-capped hairpin reveals that the tetraloop undergoes a structural rearrangement upon binding to Rnt1p dsRBD to adopt a backbone conformation that is essentially the same as the AGAA tetraloop, and indicates that a conserved recognition mode is used for all Rnt1p substrates. Comparison of free and RNA-bound Rnt1p dsRBD reveals that tetraloop-specific binding requires a conformational change in helix α1. Our findings provide a unified model of binding site selection by this dsRBD.

INTRODUCTION

Eukaryotic members of the RNase III family of endoribonucleases cleave double-stranded RNA (dsRNA) targets involved in a variety of gene expression pathways (Conrad and Rauhut, 2002; Lamontagne et al., 2001), including the maturation of precursor ribosomal RNA (rRNA) (Elela et al., 1996; Henras et al., 2004; Kufel et al., 1999), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) processing (Conrad and Rauhut, 2002; Lamontagne et al., 2001), and RNA interference (RNAi) and microRNA (miRNA) processing (Ketting et al., 2001; Knight and Bass, 2001; Lee et al., 2003b). Rnt1p, the major RNase III present in *Saccharomyces cerevisiae*, plays an essential role in the processing of rRNA, snRNAs, and snoRNAs (Chanfreau et al., 1998a, 1998b; Elela et al., 1996; Kufel et al., 1999) in budding yeast. Rnt1p is also important for mRNA quality control, cleaving intronic sequences of unspliced pre-mRNAs (Danin-Kreiselman et al., 2003). A Rnt1p target site in the mRNA coding for the essential telomerase protein Est1p has been proposed to be important for maintenance of telomere length through regulation of Est1p expression (Larose et al., 2007). Several studies have indicated that Rnt1p may play a role in transcription termination by cleavage of nascent transcripts (Catala et al., 2008; El Hage et al., 2008; Ghazal et al., 2009; Prescott et al., 2004). Drosha and Dicer, other members of the RNase III family, are involved in miRNA processing and RNAi (Ketting et al., 2001; Lee et al., 2003b). Although the RNAi pathway has been evolutionarily lost in *S. cerevisiae*, it has been found in closely related yeast species such as *S. castellii* and *Kluyveromyces polysporus*. Introduction of *S. castellii* Dicer and Argonaute into *S. cerevisiae* resulted in a functional RNAi pathway (Drinnenberg et al., 2009). Furthermore, budding yeast Dicers are more closely related to Rnt1p than to canonical Dicer, highlighting the importance of this representative of the RNase III family of enzymes.

All RNase III enzymes contain one or two conserved endonuclease domains (endoNDs). Two endoNDs form an intra- or intermolecular dimer to create a large catalytic valley for Mg^{2+}-dependent catalysis (Gan et al., 2005, 2008; Ji, 2006, 2008; MacRae and Doudna, 2007; Nicholson, 1999; Sun et al., 2005). Most RNase III enzymes also have one or more double-stranded RNA binding domains (dsRBDs). The N-terminal region of RNase III enzymes is variable across the family, and may include one or more additional domains with different functions. Rnt1p, like the bacterial RNase III enzymes, contains one endonuclease domain and one dsRBD. However, Rnt1p also has an N-terminal domain unique to yeast RNase IIIs whose function is uncertain but which may be required for the stabilization of Rnt1p homodimers (Lamontagne et al., 2000). RNase IIIs found in budding yeasts have a domain structure similar to that of Rnt1p, whereas most Drosha and Dicer enzymes have one or two dsRBDs, two endoNDs, and an N-terminal accessory domain (MacRae and Doudna, 2007; Nowotny and Yang, 2009).

The dsRBD is the second most abundant RNA binding domain. This domain specifically recognizes dsRNA (Bycroft et al., 1995; Doyle and Jantsch, 2002; Hall, 2002; Kharrat et al., 1995) and has a conserved $\alpha\beta\beta\beta\alpha$ fold (Bycroft et al., 1995; Kharrat et al., 1995; Nanduri et al., 1998; Ramos et al., 2000; Ryter and Schultz, 1998; Wu et al., 2004). The first crystal structure of a dsRBD/dsRNA complex revealed that helix α1, the

β3-α2 loop, and the β1-β2 loop interact with the sugar-phosphate backbone of successive minor, major, and minor grooves, respectively, along one face of an RNA helix, without any apparent base pair specificity (Ryter and Schultz, 1998). The recognition of dsRNA by dsRBDs plays an important role in the catalytic cleavage or modification of dsRNAs by many RNases (Gan et al., 2008; Ghazal et al., 2009; Ji, 2006, 2008; MacRae and Doudna, 2007; Sun et al., 2005; Wu et al., 2004) and RNA modification enzymes such as ADAR (Stefl et al., 2010; Yamashita et al., 2011). Based on the structures of bacterial RNase III/dsRNA complexes (Gan et al., 2006), dsRBDs are thought to contribute primarily to specificity of binding to dsRNA but not to target site selection (Nicholson, 1999; Shi et al., 2011), although there are two base contacts that have recently been proposed to be sequence specific (Stefl et al., 2010). In contrast, Rnt1p dsRNA substrates are capped by A/uGNN tetraloops located 14–16 bp away from the cleavage site (Chanfreau et al., 2000; Nagel and Ares, 2000), and specificity for these substrates resides in the dsRBD (Chanfreau et al., 2000; Lamontagne and Elela, 2004; Lamontagne et al., 2003; Nagel and Ares, 2000; Wu et al., 2004). Structural studies have revealed that A/uGNN tetraloops adopt a conserved fold (Lebars et al., 2001; Wu et al., 2001, 2004), the shape of which is recognized on the minor groove side by Rnt1p dsRBD (Wu et al., 2004). Surprisingly, the dsRBD has no direct contacts to the conserved A and G bases, which point into the major groove, but rather its helix α1 fits snugly into the minor groove side of the tetraloop and the top of the stem, interacting with the sugar-phosphate backbone and nonconserved 3′ bases. To date, this is the only known example of dsRBD binding specificity through terminal loop recognition. Rnt1p dsRBD is also unusual in containing an additional helix α3, which has been proposed to contribute to recognition of the tetraloop indirectly by stabilizing helix α1 (Leulliot et al., 2004; Wu et al., 2004).

A genome-wide search in *S. cerevisiae* for snoRNA substrates of Rnt1p identified the noncanonical snoRNA substrate snR48 (Ghazal et al., 2005). Unlike most Rnt1p substrates, snR48 contains a Rnt1p recognition site consisting of an AAGU-capped hairpin. It was proposed that the AAGU tetraloop adopts a different fold from the canonical fold of AGNN tetraloops (Gaudin et al., 2006; Ghazal and Elela, 2006), and that Rnt1p distinguishes between these two different "classes" of tetraloop-hairpin substrates using different networks of protein-RNA interactions. To investigate the molecular basis for the recognition of the AAGU hairpin and to gain further understanding of substrate-specific recognition by Rnt1p, we determined the NMR solution structure of Rnt1p dsRBD in complex with a dsRNA hairpin capped by an AAGU tetraloop and investigated in vitro and in vivo cleavage and dsRBD binding. We find that when in complex with the dsRBD, the AAGU tetraloop undergoes a structural rearrangement to adopt a backbone fold and interactions that are essentially the same as those in the complex of the dsRBD with an AGAA hairpin. Comparison of the structures of the complexes to previously determined solution and crystal structures of the free dsRBD showed that the dsRBD helix α1 undergoes a conformational change upon binding to both AAGU and AGAA tetraloops. Taken together, our results provide new insights into substrate-specific recognition by dsRBDs and provide a structural framework for a conserved general mode of RNA substrate recognition by Rnt1p.

RESULTS

The AAGU Hairpin Binds to and Is Efficiently Cleaved by Rnt1p in the Context of the snR47 Stem Sequence

We investigated, using NMR spectroscopy, the interaction of Rnt1p dsRBD with a 32 nt RNA hairpin containing a 14 bp stem and capped by the AAGU tetraloop (AAGU hairpin) found in the snoRNA snR48 precursor. To facilitate comparisons between this complex (dsRBD/AAGU) and one with a canonical A/uGNN tetraloop, we used the same stem sequence as in the previously determined structure of a Rnt1p dsRBD/snR47h complex (dsRBD/AGAA), where the hairpin sequence was derived from the snoRNA snR47 precursor (Wu et al., 2004). The 2 bp below the tetraloop are the same for snR47 and snR48, but otherwise the stem sequences differ, except for the first and fourth base pairs in our hairpins. Complex formation was initially monitored by chemical shift changes observed in the ^{1}H-^{15}N HSQC spectra of the dsRBD upon the addition of AAGU (see Figure S1 available online). The dsRBD is at or near fast exchange with the hairpin on the NMR timescale, but the complex is saturated at a small excess of RNA, as observed for the dsRBD/AGAA complex, and amide chemical shift changes are very similar (Wu et al., 2004).

To compare the relative binding affinity of Rnt1p dsRBD to AGAA and AAGU hairpins, a series of ^{1}H-^{15}N HSQC spectra of the dsRBD were acquired at various ratios of added AGAA, AAGU, and UUCG hairpins under high-salt conditions. Previous studies have shown that the dsRBD does not specifically recognize a UUCG-capped hairpin (Chanfreau et al., 2000), so this hairpin was used as a comparison for nonspecific binding of the dsRBD to double-stranded RNA. Apparent K_D values calculated from global fitting of the HSQC titration data, assuming a one-site binding model, are 34.1 ± 2.9, 30.1 ± 3.8, and 280 ± 24.6 μM for the AAGU, AGAA, and UUCG hairpins, respectively (Figures 1A–1C). These values indicate that under the NMR conditions, Rnt1p dsRBD binds with the same affinity to both the AAGU and AGAA hairpins. In contrast, nonspecific binding to dsRNA capped by a UUCG tetraloop is about an order of magnitude weaker. This result is consistent with binding assays using full-length Rnt1p, which show it binds with a 5- to 10-fold weaker affinity to non-AGNN-containing tetraloops than to canonical tetraloop-containing substrates (Chanfreau et al., 2000; Nagel and Ares, 2000).

Previously, Rnt1p was reported to require a defined sequence in the stem region of snR48-derived AAGU hairpin substrates for optimal cleavage (Gaudin et al., 2006; Ghazal et al., 2005). To confirm that the snR47-AAGU hybrid hairpin is a good substrate for Rnt1p, we performed Rnt1p cleavage assays with substrates derived from snR47, capped by AGAA, AAGU, or UUCG tetraloops (snR47-AGAA, snR47-AAGU, and snR47-UUCG, respectively) (Figure 2A), under single-turnover conditions (Figure 2B). Both snR47-AAGU and snR47-AGAA showed specific cleavage, whereas snR47-UUCG was almost unreactive over the assayed reaction time. The fact that the snR47-UUCG substrate shows only a 10-fold reduction of binding (Figure 1C) but is almost completely refractory to Rnt1p cleavage confirms that, as shown

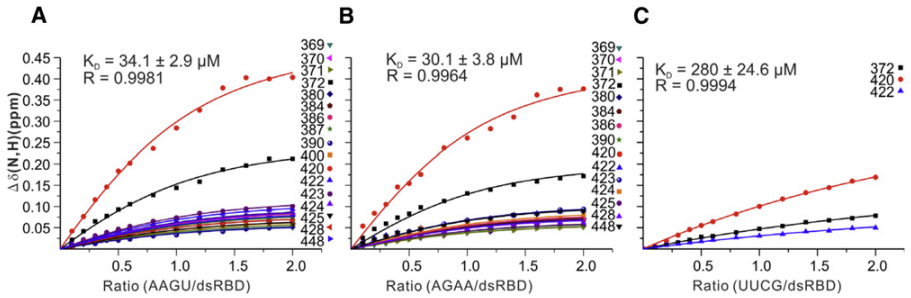


Figure 1. snR47-Derived dsRNA Capped by an AAGU Tetraloop Binds to the Rnt1p dsRBD

HSQC titration curves showing ^{15}N and ^{1}H chemical shift change $\Delta\delta(N,H) = [(\Delta\delta(H_N))^2 + (\Delta\delta(N_H)/4)^2]^{1/2}$ as a function of the concentration ratio for the titration of dsRBD with the (A) AAGU, (B) AGAA, and (C) UUCG hairpins. The continuous lines show the curves for all amides, with chemical shift changes >0.05 ppm fitted globally to a one-site binding model.

previously (Chanfreau et al., 2000; Nagel and Ares, 2000), the differences in binding affinity between different tetraloop-containing stems are not sufficient to explain the strong cleavage discrimination between these substrates. snR47-AAGU also shows more additional cleavage products than snR47-AGAA: one corresponds to an intermediate cleaved in one strand only (band below substrate band in Figure 2B), and the others correspond to some alternate cleavage sites. In vivo, cleavage at any of these alternative sites would still be expected to result in correct subsequent processing, because each of these cleaved intermediates is expected to be used at similar efficiencies by the Rat1p exonuclease (Lee et al., 2003a). Overall, cleavage of the snR47-AAGU substrate occurred at a slightly slower rate than for snR47-AGAA, but the relative cleavage rates for this noncanonical substrate is comparable to the previously reported relative rates for canonical Rnt1p substrates containing AGAA and UGAA tetraloops (Wu et al., 2001). These results indicate that the AAGU tetraloop in the context of the snR47 stem sequence is sufficient to support cleavage at a rate comparable to A/uGNN tetraloop hairpin substrates.

To further investigate the ability of Rnt1p to recognize and process a substrate carrying an AAGU tetraloop, we analyzed the processing of snR47 mutant derivatives with different types of tetraloops in vivo. We introduced AAGU in place of AGAA in

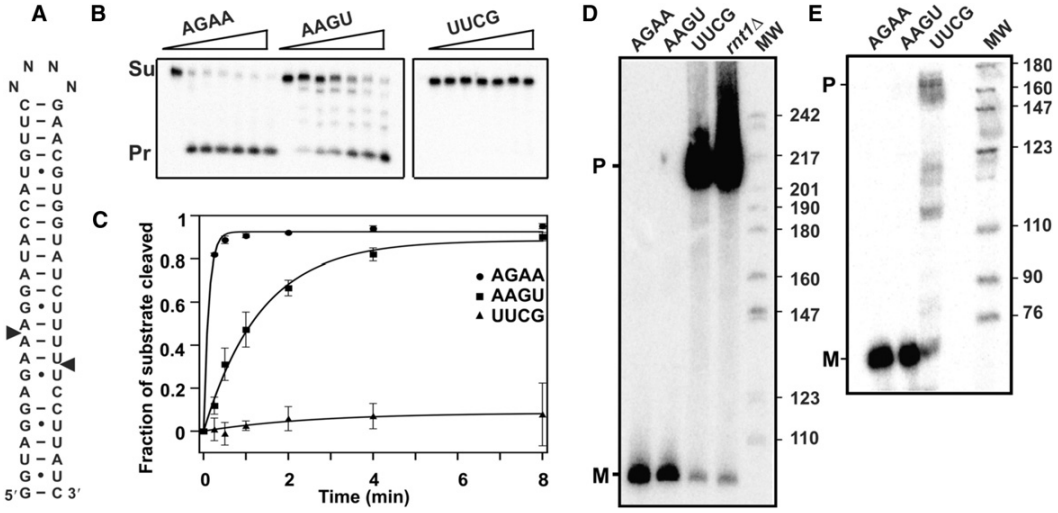


Figure 2. dsRNA Capped by an AAGU Tetraloop Is an Efficient Substrate for Rnt1p

(A) Sequence of snR47-derived RNA substrates, where NNNN is AGAA, AAGU, or UUCG, with the Rnt1p cleavage site indicated by arrowheads.

(B and C) Single-turnover cleavage kinetics for snR47-AGAA, snR47-AAGU, and snR47-UUCG.

(B) Phosphorimages of polyacrylamide gels of RNA from the cleavage reactions. Bands corresponding to the full-length substrate (Su) and the cleavage product (Pr) are indicated.

(C) Plot of fraction of substrate cleaved versus time. Error bars are the standard deviation for three experiments.

(D) Northern blot analysis of snR47 snoRNA expression carrying normal and mutant tetraloop sequences. Strains expressing snR47 with the wild-type (AGAA), AAGU, or UUCG loop sequences and a strain inactivated for Rnt1p (*rnt1Δ*) were analyzed. MW, molecular weight marker (MspI-digested pBR322); P, unprocessed precursor; M, mature snR47 snoRNA.

(E) Primer extension analysis of snR47 snoRNA expression carrying normal and mutant tetraloop sequences. Legends are as in (D).

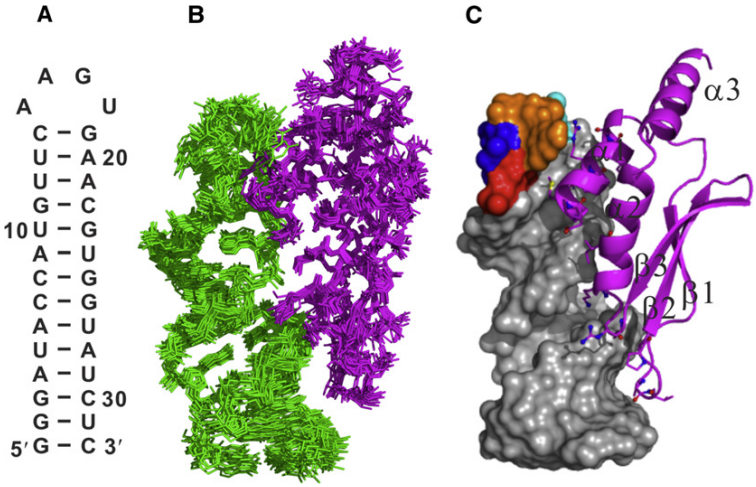


Figure 3. Solution Structure of the Rnt1p dsRBD/AAGU Complex

(A) Sequence and secondary structure of the AAGU hairpin used in the NMR studies.

(B) Superposition of the 16 lowest-energy structures, showing all heavy atoms. The dsRBD is in magenta and the AAGU hairpin is in green.

(C) Lowest-energy structure of the complex. The RNA is shown as a solvent-accessible surface and the protein as ribbons, with the amino acids at the protein-RNA interface shown as ball and sticks, with oxygen in red and nitrogen in blue. The nucleotides in the AAGU tetraloop are colored 5′ to 3′ red (A), blue (A), orange (G), and cyan (U).

the stem-loop sequence upstream from the *snR47* (Chanfreau et al., 1998a) chromosomal locus by homologous recombination (see Experimental Procedures). As controls for inefficient processing and complete loss of Rnt1p function, we generated a strain carrying a UUCG tetraloop and used a strain inactivated for Rnt1p (*rnt1Δ*), respectively. As shown by northern blot in Figure 2D, the AAGU mutant snoRNA precursor was processed as efficiently as the wild-type precursor in vivo, showing high levels of mature snoRNA and no apparent unprocessed precursor accumulation. In contrast, samples extracted from the UUCG tetraloop mutant or from the *rnt1Δ* strain exhibited little or no mature snoRNA and a strong accumulation of unprocessed precursors (Figure 2D). Primer extension analysis of the wild-type (AGAA) and AAGU or UUCG mutant strains confirmed the efficient 5′ end processing of the AAGU mutant and the strong processing defect of the UUCG mutant (Figure 2E). These results indicate that a hairpin capped by an AAGU tetraloop sequence can serve as an efficient recognition site for Rnt1p in vivo, even when placed on a substrate stem that normally contains a canonical AGNN tetraloop. In conclusion, AAGU tetraloop hairpins can be efficiently recognized by Rnt1p in vitro and in vivo, regardless of the stem sequence that they cap.

Overview of the Solution Structure of the Rnt1p dsRBD/AAGU Hairpin Complex

Protein and RNA resonances in the Rnt1p dsRBD/AAGU complex were assigned following previously established protocols (Wu et al., 2005). The protein-RNA interface was well defined by 43 intermolecular NOEs assigned from 2D filtered/edited NOESY (Peterson et al., 2004) (Figure S1). The structure of the dsRBD/AAGU complex has backbone root-mean-square deviations (rmsds) to the mean of 0.55 ± 0.11 Å and 0.64 ± 0.15 Å for the dsRBD and AAGU hairpin, respectively (Figure 3A and Table 1). The dsRBD adopts the standard α1β1β2β3α2 fold (Doyle and Jantsch, 2002), with the additional helix α3 packed against and stabilizing the C-terminal end of helix α1, as previously observed (Leulliot et al., 2004; Wu et al., 2004). The dsRBD binds to one face of the RNA and interacts primarily with the sugar-phosphate backbone in three successive regions: the tetraloop minor groove and top 2 bp with helix α1, the stem major groove with the N-terminal end of helix α2 and the β3-α2 loop, and the stem minor groove with the β1-β2 loop (Figure 3B). All of the bases in the tetraloop are in the *anti* conformation, and the RNA stem forms an A-form helix.

Analysis of RDCs Indicates that the dsRBD Adopts the Same Conformation in Both the AAGU and AGAA Complexes

We compared the structures of dsRBDs in the dsRBD/AAGU and dsRBD/AGAA complexes using residual dipolar couplings (RDCs) (Cornilescu et al., 1998; Lipsitz and Tjandra, 2004). Structure calculations for the dsRBD/AAGU complex included 83 ^{1}H-^{15}N RDCs in the final refinement step. For comparison purposes, we measured and analyzed a comparable set of 81 RDCs for the dsRBD/AGAA complex, which had been previously determined and refined with 43 RDCs (Wu et al., 2004), and recalculated the structure. We first evaluated the quality of the structures by back-calculating the RDCs from the RDC-refined dsRBD/AAGU and dsRBD/AGAA structures using the program PALES (Zweckstetter and Bax, 2000) (Figures 4A and 4B). The quality factors (Q) are 11.5% and 9.4% for dsRBD/AAGU and dsRBD/AGAA, respectively, indicating an excellent agreement between the structures and their RDC data (Cornilescu et al., 1998). We next evaluated how similar the structures of the dsRBDs in the two complexes were to each other by back-calculating the set of RDCs for each complex using the structures of the other complex. Results for the lowest-energy structure in each structure ensemble are shown in Figures 4C and 4D. R factors for the back-calculated dsRBD/AGAA and dsRBD/AAGU complexes are 0.98, and Q factors are <20% for both cases. Similar results are observed for each set of individual structures in the structure ensembles (Figure S2). These values indicate that the dsRBD conformations in the two complexes are highly similar. The dsRBDs in the two complexes have an rmsd between the two ensembles of 1.18 ± 0.32 Å for all heavy atoms (Figure 4E), which is within experimental error of the pairwise rmsds of each ensemble.

The AAGU Tetraloop in the Complex Adopts a Backbone Fold Similar to that of the AGAA Tetraloop

The solution structure of the free AAGU tetraloop is substantially different from the A/uGNN fold (Gaudin et al., 2006) (Figure 5C;

Table 1. Structural Statistics of the Rnt1p dsRBD/AAGU Hairpin Complex

Distance and Dihedral Restraints	Protein	RNA
Total NOE restraints	2095	695
Intraresidue	842	243
Sequential	498	325
Medium (i+2 to i+4)	405	10
Long-range (>i+4)	350	117
Intermolecular NOE restraints	45	
Hydrogen-bond restraints	80	72
RDC restraints	83	
Dihedral-angle restraints	120	225
Structure Statistics (16 Lowest-Energy Structures)		
Number of NOE violations >0.2 Å	0	
Number of NOE violations >0.5 Å	0	
Number of dihedral violations >5°	0	
Number of RDC violations >2 (Hz)	0	
Rmsd of RDC (Hz)	0.878 ± 0.045	
Rmsd from ideal covalent geometry		
Bond lengths (Å)	0.0034 ± 0.0001	
Bond angles (°)	0.7725 ± 0.0057	
Impropers (°)	0.4700 ± 0.0178	
Rmsd from the mean structure	Backbone	Heavy Atoms
Protein (366–448) (Å)	0.54 ± 0.10	0.90 ± 0.14
RNA (3–14, 19–30) (Å)	0.59 ± 0.11	0.55 ± 0.09
Complex (3–14, 19–30, 366–448) (Å)	0.70 ± 0.11	0.85 ± 0.11
Ramachandran statistics		
Most favored regions (%)	78.0	
Additional allowed regions (%)	18.6	
Generously allowed regions (%)	2.1	
Disallowed regions (%)	1.3	

see below). Surprisingly, however, in the dsRBD/AAGU complex, the backbone fold and shape of the minor groove of the tetraloop are highly similar to that of the AGAA tetraloop (Figures 5A and 5B), with an rmsd for all backbone heavy atoms of 1.06 Å for the lowest-energy structures (Figure 5B). As is the case for the AGNN tetraloops (Wu et al., 2004), the first two bases in the AAGU tetraloop, A15 and A16, stack on each other and point into the major groove away from the Rnt1p dsRBD helix α1 in the minor groove. The last two bases, G17 and U18 on the 3′ side, point into the minor groove, and their backbones interact with helix α1 of the dsRBD (Figure 6). Although all point into the minor groove, the positions of these bases, which are not conserved in AGNN tetraloops, differ somewhat among tetraloops.

Although the backbone fold is the same for the AGAA and AAGU tetraloops in complex with the dsRBD, the second base (A16) is in the *anti* conformation and its position is quite different from that of the *syn* G in the AGAA tetraloop. The *syn* conformation positions the G amino group within hydrogen-bonding distance of one of the nonbridging oxygen atoms on its 5′ phosphate group, and therefore presumably stabilizes its stacking on the A (or U). An A in the second position does not have a proton donor at the same position as the G, and therefore cannot hydrogen bond to the backbone even in the *syn* conformation. Thus, in the complex, the stabilization conferred by a hydrogen bond to the backbone from a *syn* G at position 2 of the tetraloop does not appear to be required for the AAGU tetraloop to adopt the same backbone conformation as an AGAA tetraloop.

In the free AAGU tetraloop, the first two As (A15 and A16) stack on each other and point into the major groove and the backbone turns after the second A, as is the case for the bound AAGU tetraloop; however, the position of the backbone before the turn is substantially different (Figure 5C). On the 3′ side of the loop, the Watson-Crick face of G17 points into the minor groove and the base is nearly coplanar with A16, while U18 points up above G17 and out into the major groove (Gaudin et al., 2006). In contrast, in the dsRBD/AAGU complex, the base of U18 is in the minor groove, and the positions of both G17 and U18 are significantly different, with the base of G17 pointing into the minor groove from near the top of the tetraloop and U18 below G17. Thus, the AAGU tetraloop in the hairpin undergoes a large conformational change upon binding to the dsRBD, to a conformation that presents a minor groove surface and backbone contacts to helix α1 that are highly similar to those in the AGAA tetraloop.

Because the dsRBD has the same binding affinity for the two hairpins but the AAGU tetraloop undergoes a larger conformational change, we measured the binding by isothermal titration calorimetry (Figure S3). Although the two complexes have the same ΔG, the dsRBD/AAGU complex has larger negative values for both ΔH and ΔS. For the dsRBD/AAGU and dsRBD/AGAA complexes, $\Delta H = -1.42 \times 10^4$ versus -0.91×10^4 kcal/mol and $\Delta S = -21.8$ and -3.94 kcal/molK, respectively. These results indicate that although the AAGU/dsRBD complex undergoes a larger change in enthalpy upon complex formation, this is offset by a compensatory decrease in entropy.

Comparison of the Protein-RNA Interfaces in the Rnt1p dsRBD/AAGU and dsRBD/AGAA Complexes

In both the dsRBD/AAGU and dsRBD/AGAA complexes, the β1-β2 loop contacts the stem minor groove of base pairs 2–5, the N-terminal end of helix α2 inserts into the major groove between base pairs 5 and 10, and helix α1 specifically recognizes the minor groove of the tetraloop and top 2 bp (Figures 3 and 6; Table S1). Almost all of the contacts are to the phosphodiester backbone. Detailed comparison of the dsRBD/AAGU and dsRBD/AGAA complexes shows that the interactions between the protein and RNA stem are nearly identical (Figure 6; Figure S4). In the minor groove of the 3 bp adjacent to the tetraloop, the D367 side-chain carboxyl group interacts with A20 2′ OH and A21 2′ OH through potential direct or water-mediated hydrogen bonds, and the side chain of K371 forms potential hydrogen bonds to A20 2′ OH and A21 O4′ for both complexes. On the 3′ side of the tetraloops, the guanidinium group of the R372 stacks onto the ring of the base G17 (in the AAGU tetraloop) or A17 (in the AGAA tetraloop) in the complexes (Figure 6; Figure S4). On the 5′ side of the tetraloop, the nonpolar side chain of M368 stacks onto the A15 ribose, and its sulfur group forms a potential water-mediated hydrogen bond with C14 O2. In the dsRBD/AGAA complex, the R372 guanidinium group and S376 OH form hydrogen bonds with the 2′ OH of A17 and A18, respectively (Figure S4), whereas in the AAGU complex the R372

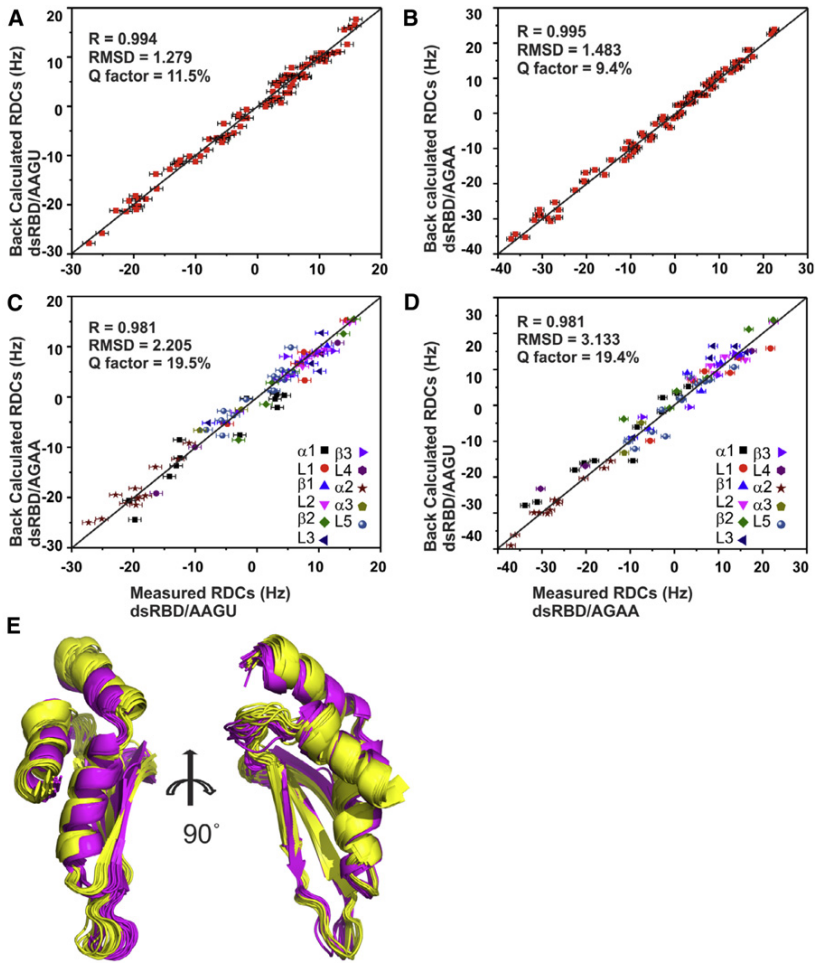


Figure 4. Correlation Plots between Experimental and Back-Calculated RDCs for the dsRBD/AAGU and dsRBD/AGAA Hairpin Complexes

(A) dsRBD/AAGU hairpin versus itself.

(B) dsRBD/AGAA hairpin versus itself.

(C) dsRBD/AAGU hairpin versus dsRBD/AGAA hairpin.

(D) dsRBD/AGAA hairpin versus dsRBD AAGU hairpin.

(A–D) R factors, rmsds, and Q values are shown with the plots. Residues in (C) and (D) are shown as follows: α1 (367–378), L1 (379–385), β1 (386–390), L2 (391–399), β2 (400–405), L3 (406–411), β3 (412–417), L4 (418–420), α2 (421–432), L5 (433–434), and α3 (435–448). Error bars are ±1 Hz, which is the standard deviation of the RDC measurement.

(E) Superpositions of the ensembles of the 16 lowest-energy structures of the dsRBD in the dsRBD/AAGU hairpin (magenta) and dsRBD/AGAA hairpin (gold) complexes. The structures were aligned using the secondary-structure elements α1, β1, β2, β3, α2, and α3.

guanidinium group forms a hydrogen bond with the 2′ OH of A15 and G17, and S376 OH forms a hydrogen bond with the 2′ OH of U18 (Figure 6B). These small differences can be attributed to the difference in sequence at these positions. Neither the tetraloop sequence nor the top 2 bp are conserved and, consistent with this, no specific contacts to the bases were observed. Mutation of M368, which has a potential interaction with C14 O2, to alanine, does not affect binding (Henras et al., 2005). Thus, the dsRBD recognizes the A/uGNN and AAGU tetraloops in the same way, by shape-specific binding.

Conformational Changes in the dsRBD upon Binding to Target RNA

Because the structure of the free Rnt1p dsRBD has been reported (Leulliot et al., 2004), we were able to examine any conformational changes that take place in the dsRBD upon binding to the AAGU hairpin as well as the AGAA hairpin in detail. We acquired a set of RDC data for the free dsRBD in solution. Of the NMR solution and two crystal structures (from one asymmetric unit) reported, the experimental RDCs fit best to the crystal structure of chain A (Protein Data Bank [PDB] ID code 1T4O) (Figure S5), so this structure was used for comparison to our bound complexes. To quantitatively compare the structures of the free and AAGU tetraloop-bound dsRBDs, the experimental RDCs from the secondary-structure elements β1, β2, β3, α1, and α2 for the dsRBD/AAGU complex were plotted versus the RDCs calculated from the crystal structure (Figure 7A). Helix α3 was excluded from this analysis, because it adopts three different orientations in the solution and two crystal structures (Leulliot et al., 2004). The correlation gives a Q factor of 32%, but when the RDCs from helix α1 are deleted from the analysis, the Q factor decreases to 16%. Similar results were obtained for dsRBD bound to the AGAA hairpin (Figure 7B). When experimental RDCs from helix α1 only are compared (free dsRBD versus dsRBD/AAGU complex and dsRBD/AAGU complex versus dsRBD crystal chain A), poor correlations are obtained (Q = 32% and 49%, respectively) (Figure S5). Taken together, these data indicate that there is a significant change in helix α1 when the dsRBD binds to target RNA, consistent with structural differences observed by direct comparison of the structures as described below.

Comparison of the structures of the free and AAGU hairpin-bound dsRBD revealed that all of the regions of the dsRBD that interact with the RNA show significant changes in position between the free and bound dsRBD (Figure 7C). The β1-β2 loop, which inserts into the stem minor groove in the complex, points away ~6 Å in the free dsRBD. The N-terminal end of helix α2 and the β3-α2 loop also shift to insert into the major groove. Helix α1 rotates about 18° (Figure 7E) and bends slightly from L374 to S376 (Figure 7D) to fit into the convex surface of the tetraloop. In the free dsRBD, helix α1 begins at N369, whereas in the complex it begins at L366. Side chains of M368, R372,

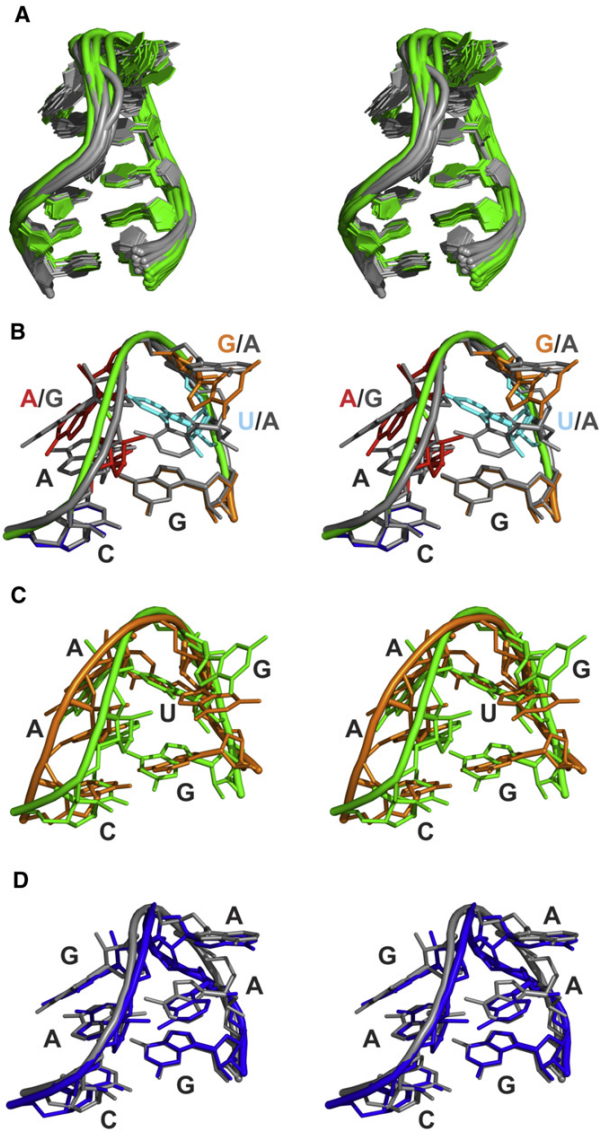


Figure 5. Comparison of the Structures of the AAGU and AGAA Tetraloops in the Free and dsRBD-Bound States

Stereo views are shown.

(A) Superposition of the ensemble structures (16 lowest-energy) of the dsRBD-bound AAGU (green) and dsRBD-bound AGAA (gray) tetraloops and the four stem base pairs below the tetraloops. The structures were aligned using residues 11–22.

(B) Superpositions of the lowest-energy structures of the dsRBD-bound AAGU tetraloop and dsRBD-bound AGAA tetraloop (rmsd 0.975 Å). Nucleotides for dsRBD-bound AAGU are shown in red (A), orange (G), dark blue (C), and cyan (U).

(C) Free AAGU tetraloop (gold) and dsRBD-bound AAGU (green) tetraloop (rmsd 1.64 Å).

(D) Free AGAA tetraloop (blue) and dsRBD-bound AGAA (gray) tetraloop (rmsd 0.99 Å).

Superpositions for (B)–(D) include the tetraloops and closing base pair. Rmsds are for backbone heavy atoms.

and S376 all shift position to align along one face of the helix to form van der Waals interactions and hydrogen bonds to the 2′ OH in the tetraloop minor groove (Figures 6A and 7D). Thus, helix α1 undergoes a change in helix length and bend and rotates 18° when it binds to the dsRNA hairpin substrate.

DISCUSSION

Although most dsRBDs bind to dsRNA, the finding that the binding of Rnt1p dsRBD to A/uGNN hairpins is a major determinant of target selection provides the first clear example of a binding specificity for a dsRBD (Chanfreau et al., 1998b, 2000; Nagel and Ares, 2000; Wu et al., 2004). Structural studies revealed that helix α1 recognizes the specific shape of this broad class of tetraloops (Wu et al., 2004). Thus, the discovery of a second class of tetraloops that did not conform to this minimal consensus and had a different free tetraloop structure led to the proposal that Rnt1p bound these substrates in a different way (Gaudin et al., 2006). Comparison of the dsRBD/AAGU structure, reported here, with the dsRBD/AGAA structure revealed that the AAGU hairpin has the same backbone fold in the complex as the AGNN tetraloops, and the dsRBD interactions and RDCs are the same for both complexes. We conclude that a conserved recognition mode is used for all Rnt1p substrates, regardless of their terminal loop sequences.

Conformational analysis of the free Rnt1p dsRBD (Leulliot et al., 2004) versus the dsRBD in the dsRBD/AAGU and dsRBD/AGAA (Wu et al., 2004) complexes revealed that helix α1 has a significant change in conformation upon binding to the tetraloop. We previously compared the structure of Rnt1p dsRBD in complex with the AGAA hairpin to that of a nonspecific complex of Xlrpba dsRBD with dsRNA (Wu et al., 2004). We noted that the two dsRBDs had a difference of ~15° in the orientation of helix α1 which positions the Rnt1p dsRBD helix α1 to fit perfectly into the minor groove of the AGNN tetraloop and the top of the stem without changing the spacing of contacts to the minor groove and major groove, 1 and 0.5 turns away, respectively. Interestingly, the structure of the free Rnt1p dsRBD is similar to the structure of Xlrpba dsRBD in complex with dsRNA, with an rmsd of 0.41 Å (Figure S6A). Thus, the conformational change in helix α1 may be a key factor in the specific recognition of Rnt1p substrates.

Conformational Change in the AAGU Tetraloop upon dsRBD Binding

For the AGAA tetraloop, the positions of the bases in the free versus bound are very similar, although there is some change in the backbone on the 3′ side of the loop (Figure 5D). Because the structures of the hairpins capped by AGAA, AGUU, and UGCA tetraloops, which are all substrates for Rnt1p, all had a similar fold with a *syn* G (Lebars et al., 2001; Wu et al., 2001) and this fold was retained in the dsRBD/AGAA complex (Wu et al., 2004), it was proposed that the dsRBD recognized the conserved shape of the tetraloop. It was therefore surprising to find that for the AAGU hairpin, the positions of the bases and the backbone trajectory both change significantly in the complex (Figure 5C). Thus, it appears that the AAGU tetraloop and helix α1 of the dsRBD cooperatively fold to form a specific complex with a conserved tetraloop fold. In the complexes, these two

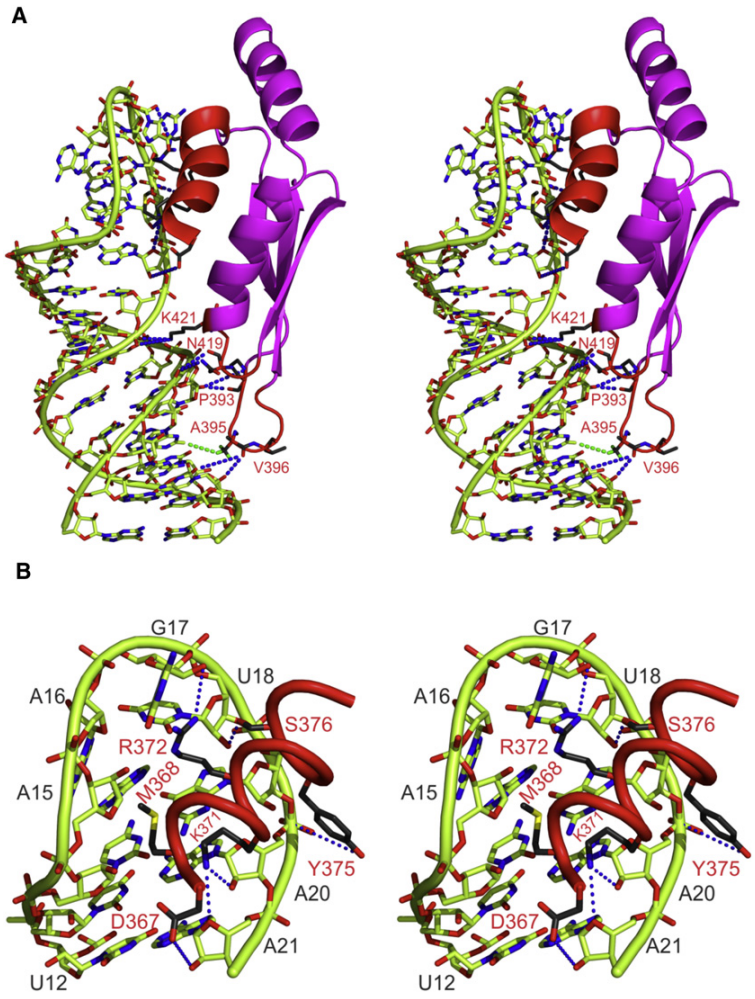


Figure 6. Interactions between the Rnt1p dsRBD and RNA

Stereo views of (A) the lowest-energy structure of the dsRBD/AAGU complex and (B) helix α1 and side-chain interactions in the tetraloop and top 3 bp. dsRBD helix α, the β1-β2 loop, and the β3-α2 loop are red and the rest of the dsRBD is in magenta, the amino acid side chains that interact with RNA are shown as gray sticks, oxygen is in red, and nitrogen is in blue. The RNA is shown as green sticks with oxygens in red. Potential direct and water-mediated hydrogen bonds are indicated by blue dashed lines between heavy atoms. A hydrophobic interaction between Ala395 methyl and A4 H2 is shown with a green dashed line. The interaction surfaces of dsRBD are shown in red.

different tetraloops provide a rare example of two distinct RNA sequences that adopt the same functional fold (Zhang et al., 2010).

Whereas the conformational changes of the free versus bound AAGU tetraloop are larger than for the AGAA tetraloop, both the free and bound tetraloops have features in common that are likely essential for recognition and binding. In all cases, the backbone turns after the second nucleotide, and the position of the backbone in the turn is the same. On the 5′ side of the tetraloop, the first two bases point into the major groove and are stacked on each other. In the complexes, these two bases have no contacts to the dsRBD and the third base is positioned above the binding site. Finally, we note that the ACAA tetraloop has been proposed to have a similar conformation to the AGAA (Staple and Butcher, 2003). However, hairpins capped by ACAA are not cleaved by Rnt1p (Wu et al., 2001). In the ensemble of ACAA tetraloop structures (Staple and Butcher, 2003), about half have a backbone conformation at the turn that is very different from the AGAA and AAGU tetraloops, such that helix α1 would not be able to insert into the minor groove.

Comparison with Other dsRBD/RNA Complexes

Although the dsRBD is the second most abundant family of RNA recognition motifs, structures of only a few dsRBDs in complex with RNA have been solved. There are now six proteins for which the structures of both the free dsRBD and the dsRBD in complex with RNA have been reported. In addition to Rnt1p dsRBD (Leulliot et al., 2004; Wu et al., 2004; and this work), these include Staufen dsRBD (Bycroft et al., 1995; Ramos et al., 2000), TAR RNA binding protein 2 (TRBP2) (Yamashita et al., 2011), *Arabidopsis* HYL1 dsRBD (Yang et al., 2010), ADAR2 dsRBD1 and dsRBD2 (Stefl et al., 2006, 2010), and *Aquifex aeolicus* RNase III (Gan et al., 2006, 2008; Ramos et al., 2000; Ryter and Schultz, 1998). All of the free dsRBDs, with the exception of ADAR2 dsRBD1 and dsRBD2, superimpose well on each other and have virtually the same angle of helix α1 relative to the other secondary-structure elements (Figure S6). Furthermore, the conformations of the free and RNA-bound dsRBDs of HYL1, TRBP, Staufen, and *A. aeolicus* RNase III are the same, respectively, indicating that helix α1 does not change its conformation upon binding RNA. Of the complexes solved to date, only the dsRBD of Rnt1p and dsRBDs of ADAR2 have different helix α1 positions in complex with RNA relative to the free dsRBD (Figures S6I and S6J). The dsRBDs of ADAR2 have recently been shown to bind dsRNA in a sequence-specific manner, with base recognition via the minor groove from one amino acid each on helix α1 and the β1-β2 loop (Stefl et al., 2010). These two dsRBDs undergo relatively large conformational changes upon RNA binding, similar to Rnt1p. However, in contrast to Rnt1p dsRBD, the position of helix α1 in the free ADAR2 dsRBD1 and dsRBD2 is different compared to Xlrpba and the other dsRBDs (Figure S6).

Rnt1p requires specific tetraloop structures for substrate cleavage both in vivo and in vitro, whereas *A. aeolicus* RNase III, the homolog of *Escherichia coli* RNase III, cleaves dsRNA in vitro with little apparent sequence specificity. Crystal structures of *A. aeolicus* RNase III in complex with RNA have revealed

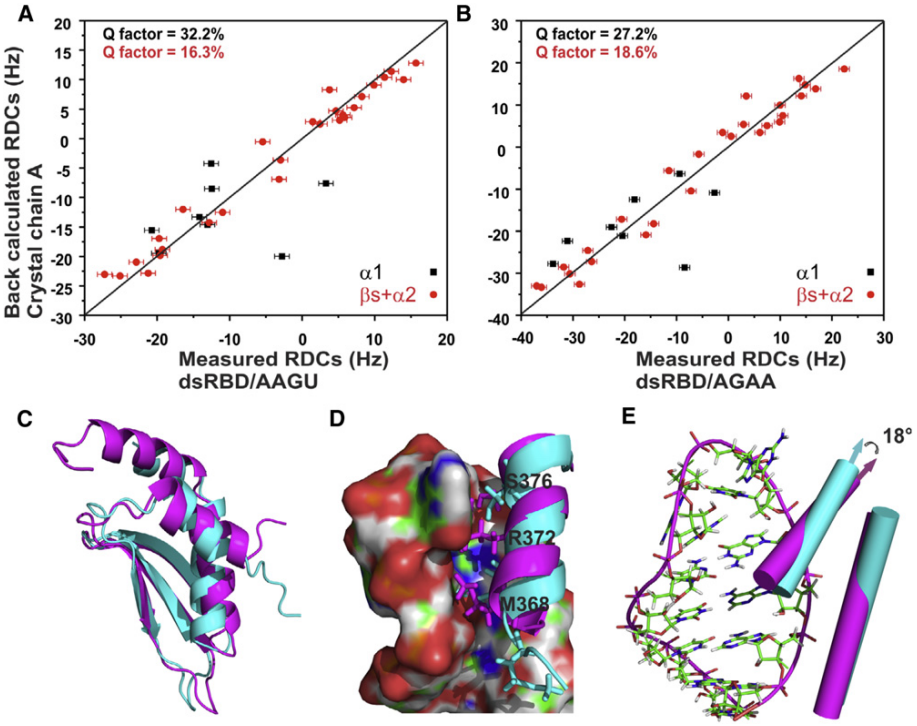


Figure 7. Comparison of the Structures of Free dsRBD and RNA-Bound dsRBD

(A and B) Correlation plots between experimental RDCs for the dsRBD secondary-structure elements in the (A) dsRBD/AAGU complex versus the RDCs calculated for the structure of the free dsRBD (PDB ID code 1T4O chain A) and (B) dsRBD/AGAA complex versus the RDCs calculated for the structure of the free dsRBD. For the correlation plots, the order tensor was determined from secondary-structure elements α2, β1, β2, and β3. RDC values are shown as black squares (α1) and red circles (α2, β1, β2, β3). Q factors were calculated for secondary-structure elements α1, α2, β1, β2, and β3 (black numbers) and for α2, β1, β2, and β3 (red numbers). Error bars are ±1 Hz, which is the standard deviation of the RDC measurement.

(C) Superposition of the free dsRBD and the dsRBD bound to the AAGU hairpin. Superposition is on the secondary-structure elements α2, β1, β2, and β3.

(D) Close-up view showing the interaction of α1 with the minor groove of the AAGU tetraloop and comparison to the free dsRBD. The helices are shown as ribbons, and the conserved side chains M368, R372, and S376 in the free and bound dsRBD are shown as sticks. The RNA is shown as a solvent-accessible surface.

(E) Comparison of the angle of helix α1 in the free and dsRBD-bound AAGU complex. The free dsRBD is in cyan and the bound dsRBD in the AAGU hairpin complex is in magenta.

that, in addition to non-sequence-specific contacts to the backbone, three bases have direct contacts to the dsRBD, two in helix α1 and one in the β1-β2 loop (Gan et al., 2006). One of these, Q157, is conserved in all bacterial RNase IIIs, and deletion of it abolished cleavage and binding. The equivalent residue in Rnt1p does not contact the RNA. The other two residues, including *A. aeolicus* RNase III Q161 in helix α1, have been proposed to give rise to sequence-specific binding (Stefl et al., 2010). Of the sequence- or tetraloop-specific dsRBD/RNA complexes solved to date, *A. aeolicus* RNase III is the only example where there is no significant change in the orientation of the dsRBD helix α1 upon binding to RNA.

In conclusion, our results show that the noncanonical AAGU tetraloop adopts a canonical fold upon binding to the dsRBD and that reorientation of helix α1 plays a major role in substrate-specific recognition. We propose that the Rnt1p dsRBD initially binds nonspecifically to dsRNA and scans along the RNA until it reaches an A/uGNN or AAGU tetraloop. Helix α1 is locked into position by the tetraloop fold like a ball in a glove, allowing subsequent positioning of the active site of Rnt1p at the cleavage site 14–16 bp away.

EXPERIMENTAL PROCEDURES

NMR Sample Preparation

The Rnt1p dsRBD, consisting of residues 366–453, was expressed as a glutathione transferase (GST) fusion protein and purified essentially as described (Wu et al., 2004), except for the addition of 1 mM DTT to the gel-filtration purification step. Details of the purification are given in Supplemental Experimental Procedures. NMR samples were ~1 mM dsRBD in 20 mM sodium phosphate (pH 6.5), 150 mM NaCl, 1 mM DTT. For NMR binding studies, the 32 nt AGAA (Figure 3A), AAGU, or UUCG hairpins were prepared by in vitro transcription using His_6-tagged mutant T7 polymerase (P266L) (Guillerez et al., 2005) with a synthetic DNA template and purified on denaturing gels as described (Wu et al., 2001). Unlabeled, uniformly ^{13}C,^{15}N-labeled, and A-, U-, G-, or C-^{13}C,^{15}N-labeled AAGU hairpins were used for structure determination of the dsRBD/AAGU complex. The dsRBD/AAGU complex was prepared at a 1.1:1 ratio (RNA:protein) by adding the dsRBD to the RNA under dilute conditions followed by concentration in NMR buffer to 1 mM complex.

NMR Spectroscopy and Structure Calculations

All NMR spectra were recorded at 25°C on Bruker DRX 500 and 600 MHz spectrometers, except for 2D NOESY spectra of exchangeable proton resonances of RNA, which were recorded at 10°C. The assignments of the Rnt1p dsRBD in the complex were derived from the analysis of 3D CBCANH, 3D CBCA(CO)NH, 3D HCCH-TOCSY, 3D HCCH-COSY, 3D ^{13}C-NOESY-HSQC,

and 3D ^{15}N-NOESY-HSQC experiments (Grzesiek and Bax, 1993; Kay et al., 1994; Schleucher et al., 1994) acquired on ^{13}C,^{15}N-labeled dsRBD in complex with unlabeled AAGU hairpin. The assignments of the AAGU hairpin were derived from 2D HCCH-COSY, 3D HCCH-TOCSY, 2D NOESY, 2D TOCSY (Cromsigt et al., 2001), and a suite of 2D filtered/edited NOESY (Peterson et al., 2004) using unlabeled dsRBD with A-, -G-, -U-, or -C-^{13}C,^{15}N-labeled AAGU. Finally, intermolecular NOEs were derived from 2D filtered/edited NOESY experiments as described (Peterson et al., 2004). One-bond ^{1}H-^{15}N RDCs were measured from HSQC-IPAP experiments (Ottiger et al., 1998) in the presence and absence of C12E5/hexanols (Ruckert and Otting, 2000) on a 600 MHz spectrometer. A total of 84, 81, and 83 RDCs were obtained for the free dsRBD, dsRBD/AGAA complex (Wu et al., 2004), and dsRBD/AAGU complex, respectively. Structure calculations were performed essentially as described (Peterson et al., 2004), and details are given in Supplemental Experimental Procedures. For comparison purposes, the dsRBD/AGAA complex was re-refined with the larger set of RDCs (81 versus 43).

Determination of Apparent K_D from ^{1}H-^{15}N HSQC Chemical Shift Titrations

The AAGU, AGAA, and UUCG hairpins were individually titrated into 0.1 mM ^{15}N-labeled dsRBD samples prepared in 500 μl high-salt NMR buffer (20 mM sodium phosphate [pH 6.5], 300 mM NaCl, 1 mM DTT) up to RNA:protein ratios of 2:1 (Figure 1). The higher-salt conditions (300 mM NaCl), instead of the 150 mM NaCl used for the structure studies, were used in order to minimize nonspecific binding. ^{1}H-^{15}N HSQC spectra were recorded for each titration point. The apparent dissociation constant K_D was obtained from changes in weighted chemical shift differences $\Delta\delta(N,H) = [(\Delta\delta(H_N))^2 + (\Delta\delta(N_H)/4)^2]^{1/2}$, assuming a one-site binding model. The titration curves were fitted globally using the following equation (Fielding, 2007):

$$\Delta\delta(N,H) = \left[\Delta\delta(N,H)_{max}/2P_0\right]\left[(P_0 + L_0 + K_D) - \left[(P_0 + L_0 + K_D)^2 - 4P_0L_0\right]^{1/2}\right],$$

where $\Delta\delta(N,H)_{max}$ is the average chemical shift difference between the free and bound forms, and P_0 and L_0 are the total concentration of dsRBD and AAGU hairpin, respectively.

Cleavage Kinetics Assays

Full-length Rnt1p was expressed with an N-terminal His_6 tag in BL-21 (DE3) Gold cells and purified on a GE Healthcare HisTrap Ni^{2+}-affinity column, followed by anion-exchange (HiTrap Q) and gel-filtration (HiLoad S75) chromatography. Purified Rnt1p was concentrated to ~5 mg/ml. For kinetics assays, 52 nt RNA hairpins snR47-AGAA, snR47-AAGU, or snR47-UUCG (Figure 2A) were ^{32}P end labeled with T4 polynucleotide kinase. Cleavage reactions were prepared under single-turnover conditions with 25 nM RNA and 1 μM Rnt1p in 20 mM Tris (pH 8.0), 150 mM NaCl at 25°C, and reactions were initiated by adding $MgCl_2$ to a final concentration of 5 mM. Ten microliter aliquots were removed at time points of 0.25, 0.5, 1, 2, 4, and 8 min, and quenched with 10 μl of formamide gel-loading buffer with 20 mM EDTA. Samples were run on a 10% denaturing polyacrylamide gel (19:1 acrylamide:bisacrylamide), dried, and imaged on a Molecular Imager FX Pro Plus (Bio-Rad). Bands in the gel image corresponding to the uncleaved fraction of the substrate were quantified using ImageJ (NIH). Plotted values are the average of three experiments.

In Vivo Analysis of Tetraloop Mutants

Tetraloop mutants (AAGU or UUCG) were inserted into the tetraloop upstream from the snR47 snoRNA gene using the delitto perfetto method (Storici et al., 2001). A core *Kan^R-URA3* cassette was inserted between the second and third positions of the snR47 snoRNA tetraloop, and double-stranded DNA oligonucleotides were used to excise the core sequence and introduce the AAGU or UUCG sequence. Genomic DNA sequences were confirmed by sequencing. Strains were grown in YPD and harvested, and northern blot and primer extension analysis were performed as described (Chanfreau et al., 1998a; Henras et al., 2005).

ACCESSION NUMBERS

Coordinates and restraints for the 16 lowest-energy structures of the Rnt1p dsRBD/AAGU complex have been deposited in the Protein Data Bank under ID code 2LBS, and chemical shifts have been deposited in the BioMagResBank under accession number 17574.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and one table and can be found with this article online at doi:10.1016/j.str.2011.03.022.

ACKNOWLEDGMENTS

This work was supported by NIH grants GM37254 (to J.F.) and GM61518 to (G.C.). K.R. was supported by a Whitcome Fellowship and Ruth L.Kirschtein National Research Service Award GM007185. The authors thank Qi Zhang for helpful advice.

Received: December 29, 2010
Revised: March 1, 2011
Accepted: March 31, 2011
Published: July 12, 2011

REFERENCES

Bycroft, M., Grunert, S., Murzin, A.G., Proctor, M., and St Johnston, D. (1995). NMR solution structure of a dsRNA binding domain from *Drosophila* staufen protein reveals homology to the N-terminal domain of ribosomal protein S5. EMBO J. *14*, 3563–3571.

Catala, M., Tremblay, M., Samson, E., Conconi, A., and Abou Elela, S. (2008). Deletion of Rnt1p alters the proportion of open versus closed rRNA gene repeats in yeast. Mol. Cell. Biol. *28*, 619–629.

Chanfreau, G., Legrain, P., and Jacquier, A. (1998a). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. J. Mol. Biol. *284*, 975–988.

Chanfreau, G., Rotondo, G., Legrain, P., and Jacquier, A. (1998b). Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. EMBO J. *17*, 3726–3737.

Chanfreau, G., Buckle, M., and Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. Proc. Natl. Acad. Sci. USA *97*, 3142–3147.

Conrad, C., and Rauhut, R. (2002). Ribonuclease III: new sense from nuisance. Int. J. Biochem. Cell Biol. *34*, 116–129.

Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998). Validation of protein structure from anisotropic carbonyl chemical shifts in a dilute liquid crystalline phase. J. Am. Chem. Soc. *120*, 6836–6837.

Cromsigt, J., van Buuren, B., Schleucher, J., and Wijmenga, S. (2001). Resonance assignment and structure determination for RNA. Methods Enzymol. *338*, 371–399.

Danin-Kreiselman, M., Lee, C.Y., and Chanfreau, G. (2003). RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol. Cell *11*, 1279–1289.

Doyle, M., and Jantsch, M.F. (2002). New and old roles of the double-stranded RNA-binding domain. J. Struct. Biol. *140*, 147–153.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., and Bartel, D.P. (2009). RNAi in budding yeast. Science *326*, 544–550.

Elela, S.A., Igel, H., and Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. Cell *85*, 115–124.

El Hage, A., Koper, M., Kufel, J., and Tollervey, D. (2008). Efficient termination of transcription by RNA polymerase I requires the 5′ exonuclease Rat1 in yeast. Genes Dev. *22*, 1069–1081.

Fielding, L. (2007). NMR methods for the determination of protein-ligand dissociation constants. Prog. Nucl. Magn. Reson. Spectrosc. *51*, 219–242.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2005). Intermediate states of ribonuclease III in complex with double-stranded RNA. Structure *13*, 1435–1442.

Gan, J., Tropea, J.E., Austin, B.P., Court, D.L., Waugh, D.S., and Ji, X. (2006). Structural insight into the mechanism of double-stranded RNA processing by ribonuclease III. Cell *124*, 355–366.

Gan, J., Shaw, G., Tropea, J.E., Waugh, D.S., Court, D.L., and Ji, X. (2008). A stepwise model for double-stranded RNA processing by ribonuclease III. Mol. Microbiol. *67*, 143–154.

Gaudin, C., Ghazal, G., Yoshizawa, S., Elela, S.A., and Fourmy, D. (2006). Structure of an AAGU tetraloop and its contribution to substrate selection by yeast RNase III. J. Mol. Biol. *363*, 322–331.

Ghazal, G., and Elela, S.A. (2006). Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. J. Mol. Biol. *363*, 332–344.

Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol. Cell. Biol. *25*, 2981–2994.

Ghazal, G., Gagnon, J., Jacques, P.E., Landry, J.R., Robert, F., and Elela, S.A. (2009). Yeast RNase III triggers polyadenylation-independent transcription termination. Mol. Cell *36*, 99–109.

Grzesiek, S., and Bax, A. (1993). The importance of not saturating H_2O in protein NMR. Application to sensitivity enhancement and NOE measurements. J. Am. Chem. Soc. *115*, 12593–12594.

Guillerez, J., Lopez, P.J., Proux, F., Launay, H., and Dreyfus, M. (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc. Natl. Acad. Sci. USA *102*, 5958–5963.

Hall, K.B. (2002). RNA-protein interactions. Curr. Opin. Struct. Biol. *12*, 283–288.

Henras, A.K., Bertrand, E., and Chanfreau, G. (2004). A cotranscriptional model for 3′-end processing of the *Saccharomyces cerevisiae* pre-ribosomal RNA precursor. RNA *10*, 1572–1585.

Henras, A.K., Sam, M., Hiley, S.L., Wu, H., Hughes, T.R., Feigon, J., and Chanfreau, G.F. (2005). Biochemical and genomic analysis of substrate recognition by the double-stranded RNA binding domain of yeast RNase III. RNA *11*, 1225–1237.

Ji, X. (2006). Structural basis for non-catalytic and catalytic activities of ribonuclease III. Acta Crystallogr. D Biol. Crystallogr. *62*, 933–940.

Ji, X. (2008). The mechanism of RNase III action: how Dicer dices. Curr. Top. Microbiol. Immunol. *320*, 99–116.

Kay, L.E., Xu, G.Y., and Yamazaki, T. (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. J. Magn. Reson. *109*, 129–133.

Ketting, R.F., Fischer, S.E., Bernstein, E., Sijen, T., Hannon, G.J., and Plasterk, R.H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. Genes Dev. *15*, 2654–2659.

Kharrat, A., Macias, M.J., Gibson, T.J., Nilges, M., and Pastore, A. (1995). Structure of the dsRNA binding domain of *E. coli* RNase III. EMBO J. *14*, 3572–3584.

Knight, S.W., and Bass, B.L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. Science *293*, 2269–2271.

Kufel, J., Dichtl, B., and Tollervey, D. (1999). Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3′ ETS but not the 5′ ETS. RNA *5*, 909–917.

Lamontagne, B., and Elela, S.A. (2004). Evaluation of the RNA determinants for bacterial and yeast RNase III binding and cleavage. J. Biol. Chem. *279*, 2231–2241.

Lamontagne, B., Tremblay, A., and Abou Elela, S. (2000). The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage. Mol. Cell. Biol. *20*, 1104–1115.

Lamontagne, B., Larose, S., Boulanger, J., and Elela, S.A. (2001). The RNase III family: a conserved structure and expanding functions in eukaryotic dsRNA metabolism. Curr. Issues Mol. Biol. *3*, 71–78.

Lamontagne, B., Ghazal, G., Lebars, I., Yoshizawa, S., Fourmy, D., and Elela, S.A. (2003). Sequence dependence of substrate recognition and cleavage by yeast RNase III. J. Mol. Biol. *327*, 985–1000.

Larose, S., Laterreur, N., Ghazal, G., Gagnon, J., Wellinger, R.J., and Elela, S.A. (2007). RNase III-dependent regulation of yeast telomerase. J. Biol. Chem. *282*, 4373–4381.

Lebars, I., Lamontagne, B., Yoshizawa, S., Aboul-Elela, S., and Fourmy, D. (2001). Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing. EMBO J. *20*, 7250–7258.

Lee, C.Y., Lee, A., and Chanfreau, G. (2003a). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of *S. cerevisiae* box C/D snoRNAs. RNA *9*, 1362–1370.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003b). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Leulliot, N., Quevillon-Cheruel, S., Graille, M., van Tilbeurgh, H., Leeper, T.C., Godin, K.S., Edwards, T.E., Sigurdsson, S.T., Rozenkrants, N., Nagel, R.J., et al. (2004). A new α-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. EMBO J. *23*, 2468–2477.

Lipsitz, R.S., and Tjandra, N. (2004). Residual dipolar couplings in NMR structure analysis. Annu. Rev. Biophys. Biomol. Struct. *33*, 387–413.

MacRae, I.J., and Doudna, J.A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. Curr. Opin. Struct. Biol. *17*, 138–145.

Nagel, R., and Ares, M., Jr. (2000). Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5′-AGNN-3′ tetraloop. RNA *6*, 1142–1156.

Nanduri, S., Carpick, B.W., Yang, Y., Williams, B.R., and Qin, J. (1998). Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. EMBO J. *17*, 5458–5465.

Nicholson, A.W. (1999). Function, mechanism and regulation of bacterial ribonucleases. FEMS Microbiol. Rev. *23*, 371–390.

Nowotny, M., and Yang, W. (2009). Structural and functional modules in RNA interference. Curr. Opin. Struct. Biol. *19*, 286–293.

Ottiger, M., Delaglio, F., and Bax, A. (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. J. Magn. Reson. *131*, 373–378.

Peterson, R.D., Theimer, C.A., Wu, H., and Feigon, J. (2004). New applications of 2D filtered/edited NOESY for assignment and structure elucidation of RNA and RNA-protein complexes. J. Biomol. NMR *28*, 59–67.

Prescott, E.M., Osheim, Y.N., Jones, H.S., Alen, C.M., Roan, J.G., Reeder, R.H., Beyer, A.L., and Proudfoot, N.J. (2004). Transcriptional termination by RNA polymerase I requires the small subunit Rpa12p. Proc. Natl. Acad. Sci. USA *101*, 6068–6073.

Ramos, A., Grunert, S., Adams, J., Micklem, D.R., Proctor, M.R., Freund, S., Bycroft, M., St Johnston, D., and Varani, G. (2000). RNA recognition by a Staufen double-stranded RNA-binding domain. EMBO J. *19*, 997–1009.

Ruckert, M., and Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. J. Am. Chem. Soc. *122*, 7793–7797.

Ryter, J.M., and Schultz, S.C. (1998). Molecular basis of double-stranded RNA-protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. EMBO J. *17*, 7505–7513.

Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S.J., Sorensen, O.W., and Griesinger, C. (1994). A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. J. Biomol. NMR *4*, 301–306.

Shi, Z., Nicholson, R.H., Jaggi, R., and Nicholson, A.W. (2011). Characterization of *Aquifex aeolicus* ribonuclease III and the reactivity epitopes of its pre-ribosomal RNA substrates. Nucleic Acids Res. *39*, 2756–2768.

Staple, D.W., and Butcher, S.E. (2003). Solution structure of the HIV-1 frameshift inducing stem-loop RNA. Nucleic Acids Res. *31*, 4326–4331.

Stefl, R., Xu, M., Skrisovska, L., Emeson, R.B., and Allain, F.H. (2006). Structure and specific RNA binding of ADAR2 double-stranded RNA binding motifs. Structure *14*, 345–355.

Stefl, R., Oberstrass, F.C., Hood, J.L., Jourdan, M., Zimmermann, M., Skrisovska, L., Maris, C., Peng, L., Hofr, C., Emeson, R.B., and Allain, F.H. (2010). The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. Cell *143*, 225–237.

Storici, F., Lewis, L.K., and Resnick, M.A. (2001). In vivo site-directed mutagenesis using oligonucleotides. Nat. Biotechnol. *19*, 773–776.

Sun, W., Pertzev, A., and Nicholson, A.W. (2005). Catalytic mechanism of *Escherichia coli* ribonuclease III: kinetic and inhibitor evidence for the involvement of two magnesium ions in RNA phosphodiester hydrolysis. Nucleic Acids Res. *33*, 807–815.

Wu, H., Yang, P.K., Butcher, S.E., Kang, S., Chanfreau, G., and Feigon, J. (2001). A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III. EMBO J. *20*, 7240–7249.

Wu, H., Henras, A., Chanfreau, G., and Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. Proc. Natl. Acad. Sci. USA *101*, 8307–8312.

Wu, H., Finger, L.D., and Feigon, J. (2005). Structure determination of protein/RNA complexes by NMR. Methods Enzymol. *394*, 525–545.

Yamashita, S., Nagata, T., Kawazoe, M., Takemoto, C., Kigawa, T., Güntert, P., Kobayashi, N., Terada, T., Shirouzu, M., Wakiyama, M., et al. (2011). Structures of the first and second double-stranded RNA-binding domains of human TAR RNA-binding protein. Protein Sci. *20*, 118–130.

Yang, S.W., Chen, H.Y., Yang, J., Machida, S., Chua, N.H., and Yuan, Y.A. (2010). Structure of *Arabidopsis* HYPONASTIC LEAVES1 and its molecular implications for miRNA processing. Structure *18*, 594–605.

Zhang, Q., Kim, N.K., Peterson, R.D., Wang, Z., and Feigon, J. (2010). Structurally conserved five nucleotide bulge determines the overall topology of the core domain of human telomerase RNA. Proc. Natl. Acad. Sci. USA *107*, 18761–18768.

Zweckstetter, M., and Bax, A. (2000). Prediction of sterically induced alignment in a dilute liquid crystalline phase: aid to protein structure determination by NMR. J. Am. Chem. Soc. *122*, 3791–3792.

CHAPTER 4

Intrinsic Dynamics of an Extended Hydrophobic Core in the *S. cerevisiae* RNase III dsRBD Contributes to Recognition of Specific RNA Binding Sites

Intrinsic Dynamics of an Extended Hydrophobic Core in the *S. cerevisiae* RNase III dsRBD Contributes to Recognition of Specific RNA Binding Sites

Elon Hartman†, Zhonghua Wang†, Qi Zhang, Kevin Roy, Guillaume Chanfreau and Juli Feigon

Department of Chemistry and Biochemistry, *and Molecular Biology Institute, P.O. Box 951569, University of California, Los Angeles, Los Angeles, CA 90095-1569, USA*

Correspondence to Juli Feigon: *feigon@mbi.ucla.edu*
http://dx.doi.org/10.1016/j.jmb.2012.11.025
Edited by A. G. Palmer III

Abstract

The *Saccharomyces cerevisiae* RNase III enzyme Rnt1p preferentially binds to double-stranded RNA hairpin substrates with a conserved (A/u)GNN tetraloop fold, via shape-specific interactions by its double-stranded RNA-binding domain (dsRBD) helix α1 to the tetraloop minor groove. To investigate whether conformational flexibility in the dsRBD regulates the binding specificity, we determined the backbone dynamics of the Rnt1p dsRBD in the free and AGAA hairpin-bound states using NMR spin-relaxation experiments. The intrinsic microsecond-to-millisecond timescale dynamics of the dsRBD suggests that helix α1 undergoes conformational sampling in the free state, with large dynamics at some residues in the α1–β1 loop (α1–β1 hinge). To correlate free dsRBD dynamics with structural changes upon binding, we determined the solution structure of the free dsRBD used in the previously determined RNA-bound structures. The Rnt1p dsRBD has an extended hydrophobic core comprising helix α1, the α1–β1 loop, and helix α3. Analysis of the backbone dynamics and structures of the free and bound dsRBD reveals that slow-timescale dynamics in the α1–β1 hinge are associated with concerted structural changes in the extended hydrophobic core that govern binding of helix α1 to AGAA tetraloops. The dynamic behavior of the dsRBD bound to a longer AGAA hairpin reveals that dynamics within the hydrophobic core differentiate between specific and nonspecific sites. Mutations of residues in the α1–β1 hinge result in changes to the dsRBD stability and RNA-binding affinity and cause defects in small nucleolar RNA processing *in vivo*. These results reveal that dynamics in the extended hydrophobic core are important for binding site selection by the Rnt1p dsRBD.
© 2012 Elsevier Ltd. All rights reserved.

Introduction

RNase III enzymes process double-stranded RNA (dsRNA) substrates for many non-coding RNA precursors, including pre-rRNAs, pre-small nucleolar RNAs (snoRNAs), and pre-small nuclear RNAs, as well as microRNA and small interfering RNA.[1–3] RNase III family members typically have one or two dsRNA-binding domains (dsRBDs) and one or two endonuclease domains (endoNDs), which cleave dsRNA substrates as a dimer. Each endoND cleaves the backbone of one RNA strand via a two-Mg^{2+} catalytic mechanism, leaving a two-nucleotide 3′ overhang on processed RNAs, a defining feature of RNase III cleavage.[4,5] In *Saccharomyces cerevisiae*, Rnt1p is the only characterized RNase III enzyme and is involved in the processing of the pre-rRNA precursor[6,7] and of the precursors of many snoRNAs[8–10] and small nuclear RNAs.[11–14] For most of these non-coding RNAs, Rnt1p cleavage provides a site for subsequent processing by the Rat1p or Xrn1p exonucleases or the exosome.[14–17] Rnt1p activity is also important for the quality control of mRNA, processing unspliced mRNAs.[18,19] Rnt1p cleavage can influence transcription termination by cleaving stem–loop structures that are found downstream from normal polyadenylation signals.[20,21] Finally, Rnt1p cleavage limits the expression of a

number of mRNAs.[15,22–24] Thus, Rnt1p activity controls the production of a large number of cellular transcripts. Rnt1p has a characteristic substrate specificity, cleaving the dsRNA stem of (A/u)GNN tetraloop hairpins 14 and 16 bp from the conserved tetraloop on its RNA targets.[25,26] Selective binding by the dsRBD to (A/u)GNN tetraloop hairpins, a unique feature of Rnt1p, determines target site selection.[26]

Although *S. cerevisiae* does not have RNAi machinery, other budding yeasts carry out RNAi with a Dicer, called Dcr1, which is evolutionarily related to Rnt1.[27] Dcr1 resembles Rnt1 in having a single endoND that dimerizes intermolecularly, unlike other eukaryotic Dicers, which have two tandem endoNDs that dimerize intramolecularly. The Dcr1 endoND is followed by a dsRBD, but has an additional dsRBD separated by a long linker sequence. How these dsRBDs contribute to substrate recognition and processing is unknown, although the endoND-adjacent dsRBD in Dcr1 is required for small interfering RNA processing. Intriguingly, *Candida albicans* Dcr1 has been found to carry out both RNAi and Rnt1 functions.[28]

Canonical dsRBDs have an αββββα secondary-structure motif and interact with a broad range of dsRNA substrates. Residues in helix α1, the β1–β2 loop, and helix α2 mediate interactions with successive RNA minor, major, and minor grooves on one face of the duplex, respectively.[29] The dsRBDs generally recognize dsRNA without any additional substrate specificity, a binding mode typified by the crystal structure of the Xlrbpa dsRBD in complex with A-form dsRNA.[30] In contrast, the structure of human ADAR dsRBD in complex with dsRNA revealed that this and other dsRBDs, notably *Aquifex aeolicus* RNase III dsRBD, can have some sequence specificity for their dsRNA substrates through hydrophobic contacts between dsRBD side chains and nucleotide bases.[31] Additionally, some dsRBDs have a canonical dsRBD fold but do not independently bind to dsRNA with high affinity, such as the human Drosha dsRBD.[32]

The Rnt1p dsRBD is unique among dsRBDs studied to date in recognizing RNA hairpins capped by a tetraloop with the consensus sequence (A/u)GNN,[25] through structure-specific recognition of the tetraloop fold by helix α1, with no base-specific contacts.[33] Binding of the Rnt1p dsRBD to the conserved tetraloop fold is required for correct substrate cleavage,[25] although cleavage independently from the presence of the tetraloop can be observed *in vitro* in specific conditions.[24,26] The structure of the Rnt1p dsRBD differs from canonical dsRBDs in having an additional C-terminal helix α3 that has been proposed to contribute to specific recognition of Rnt1p substrates by indirectly reshaping the RNA binding surface.[33,34] Our recent structure of the dsRBD bound to an AAGU tetraloop hairpin,[35] a specific but non-canonical substrate,[8,36] showed that the dsRBD employs a single binding mode for AGAA and AAGU tetraloop hairpins, with the AAGU tetraloop adopting the same shape as the AGAA tetraloop upon binding by the dsRBD. The identification of a single binding mode for two substrates with dissimilar sequences and conformations in the free state provided further evidence for the structure-specific, rather than sequence-specific, nature of the interaction between the Rnt1p dsRBD and target RNAs. This study further showed that conformational changes in the tetraloop-binding helix α1 are important for allowing the dsRBD to adopt the bound conformation.[35]

The dynamic properties of biomolecules often contribute to their biological functions by enabling conformational changes necessary for binding and catalysis. Moreover, conformational flexibility can allow proteins to sample functionally important alternative conformations.[37,38] Here, we have investigated the intrinsic backbone dynamics of the Rnt1p dsRBD using NMR ^{15}N spin-relaxation measurements. Further, we have examined the relationship between dsRBD dynamics and structural changes that occur upon binding to AGAA tetraloop hairpins. Slow-timescale dynamics of the dsRBD indicate that helix α1, which interacts with the tetraloop in the complex, undergoes conformational sampling in the free state, with particularly large dynamics at a hinge within the α1–β1 loop. Upon binding to RNA, dynamics at the α1–β1 hinge are partially quenched. We have determined the solution structure of the free dsRBD for the same construct previously used for the structures of Rnt1p dsRBD/RNA complexes, enabling precise comparison between free and bound states. Changes in the structure and dynamics of the dsRBD upon binding to an AGAA hairpin substrate for regions distal to the binding face reveal a network of hydrophobic residues within α1, the α1–β1 loop, and α3 with specific dynamic properties that facilitate binding to specific tetraloops. Mutation of individual residues in the α1–β1 hinge causes changes in dsRBD conformation and stability and results in defects in snoRNA processing *in vivo*. These results show that the intrinsic dynamics of the dsRBD contributes to the selection of specific tetraloop-hairpin substrates by Rnt1p and that helix α1, the α1–β1 loop, and helix α3 cooperatively contribute to regulation of the dynamics of the RNA-binding region of the dsRBD through interactions within an extended hydrophobic core.

Results

Solution structure of the Rnt1 dsRBD$_{366–453}$

Three structures of the free Rnt1p dsRBD have been reported: a solution structure [Protein Data

Bank (PDB) ID: 1T4N; residues 364–450] and two crystal structures from one asymmetric unit (PDB ID: 1T4O; construct includes residues 364–471; crystal structure chain A, residues 362–443; and crystal structure chain B, residues 361–448).[34] Helix α3 has a different length and orientation in each of these structures, and this heterogeneity was inferred to reflect dynamics for this helix in solution. Helix α3, unique to the Rnt1p dsRBD, was proposed to contribute to specific RNA binding by reshaping the RNA-binding surface of the dsRBD through steric effects on helix α1 and the α1–β1 loop. However, in the crystal structures, helix α3 of chain A terminates at residue 443 due to disorder in the crystal, and the position of helix α3 of chain B is affected by crystal packing. In the solution structure, there are three nonnative residues beyond 447 in helix α3. We previously acquired residual dipolar couplings (RDCs) for the free dsRBD to determine which of the reported structures most closely reflects the conformation of the dsRBD in solution. We showed that the measured RDCs for the free dsRBD correlate best to the back-calculated RDCs for chain A of the crystal structure, although its helix α3 is shorter than in the other reported structures of the free and bound dsRBD.[35] However, the large difference between the measured and back-calculated RDCs for helix α3 in the free dsRBD suggested that none of the structures of the free protein accurately describe the orientation of helix α3 in solution (Supplementary Fig. 1).

In order to be able to completely describe the structural changes in the Rnt1p dsRBD upon RNA substrate binding, we determined the solution structure of the free dsRBD (residues 366–453), including an extensive set of RDCs (Fig. 1a and b). This is the same construct used for the solution structures of the dsRBD/RNA hairpin complexes.[33,35] The structures of the dsRBD are well converged, with a backbone RMSD to the mean of 0.56 ± 0.11 (Fig. 1a and Table 1). Comparison of our solution structure with the previously determined structures of Rnt1p dsRBD shows that the positions of the three β-strands and helix α2, which comprise a hydrophobic core common to all dsRBDs, are nearly identical (RMSD < 1.3; RMSD for the α2 and the β-sheets between crystal structures is ~0.8 Å, and that between the NMR structure is ~1.2 Å) (Fig. 1c). The β1–β2 loop shows evidence of flexibility in all of the reported structures, based on high *B*-factors in the crystal structures and a larger range of conformations in the solution ensembles, especially for the previously determined solution structure.[34] This is consistent with our characterization of the dynamics discussed below. However, there are significant differences in the orientations of helix α1, helix α3, and the α1–β1 loop (Fig. 1d and e), particularly between the two solution structures. In our solution structure of the $dsRBD_{366–453}$, residue I448 is part of

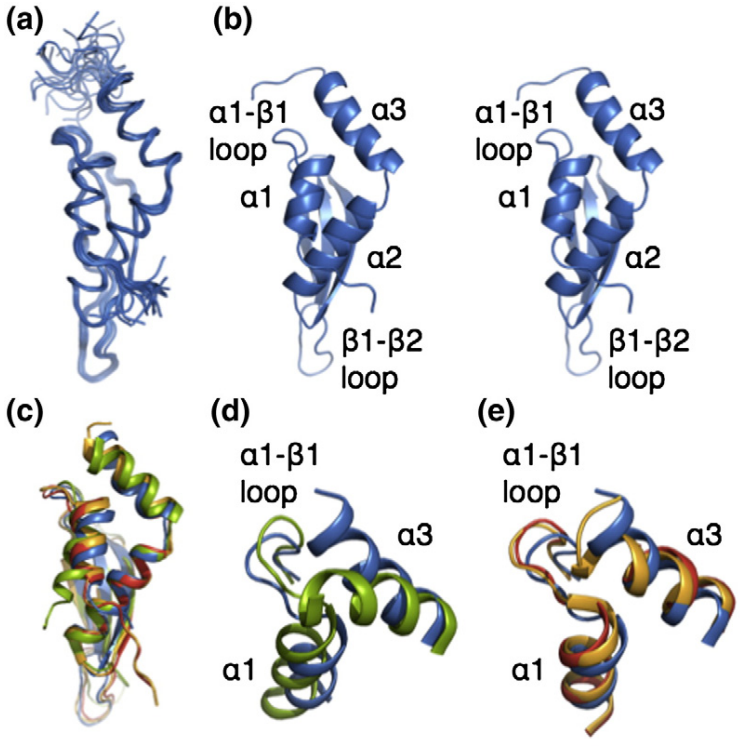


Fig. 1. Solution structure of the Rnt1p $dsRBD_{366–453}$. (a) Superposition of the 20 lowest-energy structures of the free dsRBD. (b) Stereoview of the lowest-energy structure of the free dsRBD. (c) Comparison of the dsRBD determined here (blue) with previously determined solution (1T4N; residues 364–447) (green) and crystal (IT4O; residues 362–471) chain A (red) and chain B (orange) structures. (d and e) Comparison of helix α1, the α1–β1 loop, and helix α3 in (d) the solution structures of the $dsRBD_{366–453}$ and $dsRBD_{364-450}$ (1T4N) and (e) the solution structure of the $dsRBD_{366–453}$ and the crystal structures chain A and chain B. In all structures, helix α1 begins at residue 369. Helix α3 has three nonnative residues beyond residue 447 in the solution structure 1T4N. Helix α3 ends at 443 in the crystal structure (IT4O) chain A, at 448 in the crystal structure (IT4O) chain B, and at 448 in our solution structure. Structures in (c) to (e) are aligned on α2, β1, β2, and β3.

Table 1. Structural statistics of Rnt1p dsRBD

Distance and dihedral restraints	
Total NOE restraints	2068
Intraresidue	581
Sequential	472
Medium (i+2 to i+4)	489
Long range (>i+4)	526
Hydrogen bond restraints	62
RDC restraints	55
Dihedral angle restraints	138
Structure statistics (20 lowest-energy structures)	
No. of NOE violations >0.2 Å	0
No. of dihedral violations >5°	0
No. of RDC violations >2 Hz	0
RMSD of RDC (Hz)	0.20±0.02
RMSD from ideal covalent geometry	
Bond lengths (Å)	0.001±0.0001
Bond angles (°)	0.313±0.005
Impropers (°)	0.255±0.008
RMSD from the mean structure (Å)	
Backbone (residues 366–448)	0.56±0.11
Heavy atoms (residues 366–448)	1.05±0.09
Ramachandran statistics	
Most favored regions (%)	77.3
Additional allowed regions (%)	19.9
Generously allowed regions (%)	2.4
Disallowed regions (%)	0.4

helix α3, which is the nonnative Ala in the previously determined solution structure. Interestingly, as shown below, I448 has one of the largest chemical shift changes upon binding to RNA substrate (see Fig. 6). The overall fold of the α1–β1 loop is the same for our solution structure and the two crystal structures, although the position of the loop is different (Fig. 1e). Detailed analysis of our solution structure of Rnt1p dsRBD reveals interactions among residues from helices α1 and α3 and the α1–β1 loop that constitute an extended hydrophobic core not present in other dsRBDs. In canonical dsRBDs, residues in helix α1 and the α1–β1 loop are typically solvent exposed. While contiguous with the hydrophobic core common to all dsRBDs, formed by contacts among helix α2 and the β-sheets, this extended hydrophobic core constitutes a distinct internal network of hydrophobic interactions, indicating a potential functional role in tetraloop-specific recognition by the Rnt1p dsRBD.

Structural comparison of free and RNA-bound Rnt1p dsRBD reveals concerted changes in the extended hydrophobic core

Comparison of the solution structure of the free $dsRBD_{366–453}$ with the dsRBD in complex with AGAA (Fig. 2) and AAGU tetraloop hairpins confirms the previously described conformational changes in the dsRBD at the RNA-binding interface that were based on comparison to crystal structure chain A[33,35] and provides additional details. Upon binding to the tetraloop minor groove, helix α1 is extended three residues at its N-terminus, rotates 18°, bends between residues L374 and S376, and translates toward the RNA. This reorientation of helix α1 is required for shape-specific binding to the tetraloop minor groove, which is different from the minor groove of A-form RNA. The β1–β2 loop, which interacts with the stem minor groove one helical turn away, moves toward the RNA by about 6 Å compared to its position in the free dsRBD. In the intervening major groove, helix α2 and the β3–α2 loop shift positions for side-chain interactions with the phosphodiester backbone. The side chains of the interacting residues all change positions.

In addition to these conformational changes for residues at the RNA-binding interface, the solution structure of Rnt1p $dsRBD_{366–453}$ reveals specific changes in the positions of some residues at the

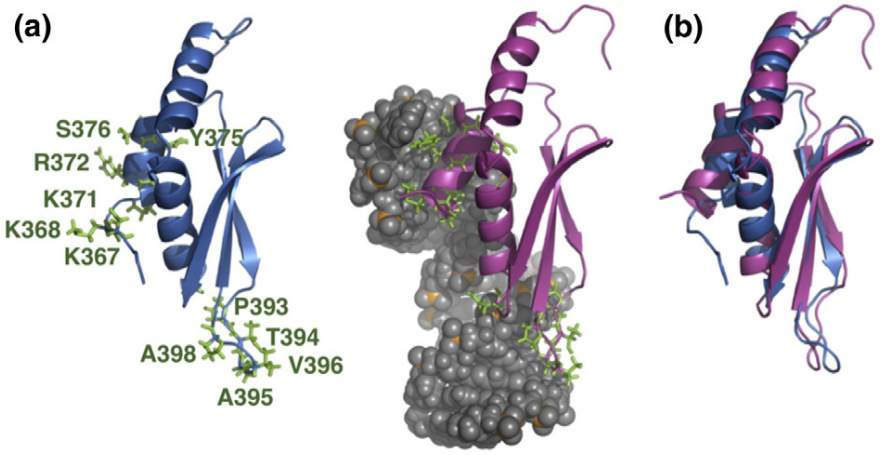


Fig. 2. Comparison of the free RDC-refined dsRBD (this work) with the RDC-refined dsRBD/AGAA complex (PDB ID: 2LUP). (a) Comparison of free and RNA-bound dsRBD, with side chains of residues that interact with the RNA shown. (b) Overlay of free and RNA-bound dsRBD structures, aligned on α2, β1, β2, and β3.

interface between helix α1 and helix α3, which are distal from the protein–RNA interface. Superposition of the free and RNA-bound dsRBD on the core α2, β1, β2, and β3 elements reveals that helix α1, helix α3, and the α1–β1 loop all change positions significantly between the free and bound states (Fig. 3a). However, when the free and bound dsRBDs are aligned on α1 and α3, it becomes clear that these changes are concerted; that is, the backbones of all three of these elements are nearly superimposed, indicating that they all translate in space together (Fig. 3b). The concerted movement of helix α1 and helix α3 upon RNA binding includes some reorientation of side chains in the hydrophobic core (Fig. 3c–e). The I378 side chain is in the *trans* rotamer conformation in the free dsRBD but in the *gauche*– conformation in the RNA-bound state (Fig. 3c and d). This side-chain rotation may be necessary to maintain close hydrophobic contacts between α1 and α3 in the complex. The Y380 ring rotates to a position perpendicular to its position in the free state (Fig. 3e). The backbone of the α1–β1 loop moves from its position in the free dsRBD by about 4 Å to accommodate the changes in position of α1 and the α1–β1 side chains in the bound dsRBD. These changes in side-chain conformation allow the extended hydrophobic core to maintain most of the hyrophobic contacts in the bound state.

Backbone dynamics of the free dsRBD

To investigate whether the concerted conformational changes in the extended hydrophobic core contribute to tetraloop-specific recognition and, more generally, how conformational flexibility within the free dsRBD affects substrate specificity and binding, we investigated the backbone dynamics of free and RNA-bound dsRBDs using NMR spin-relaxation experiments (Fig. 4). The measured ^{15}N relaxation data, R_1, R_2, and {^{1}H}–^{15}N heteronuclear nuclear Overhauser enhancement (NOE) were analyzed using the Lipari–Szabo model-free formalism to obtain a quantitative description of the backbone dynamics, where the order parameter (S^2) and the internal correlation time (τ_e) describe the amplitude and timescale of backbone dynamics, respectively. In addition, to fully describe internal motions, model-free analysis also includes a term to account for chemical exchange at the microsecond-to-millisecond timescale (R_{ex}). Overall, the S^2 values obtained from model-free analysis indicate that the dsRBD is relatively rigid for all structured residues at the picosecond-to-nanosecond timescale, with an average S^2 value of 0.86 (Fig. 5a). The single exception is residue N399, which has an S^2 value of 0.6 and is adjacent to a proline (P398) in the β1–β2 loop. We were unable to determine relaxation parameters for residue D397 on the other side of the proline due to spectral overlap. The conformational flexibility in the β1–β2 loop, as evidenced by the low S^2 value of N399, is consistent with the multiple conformations for the β1–β2 loop in the NMR structure ensembles and high *B*-factors in the crystal structures of the free dsRBD.[34]

Slow-timescale motions, as reflected by the inclusion of an R_{ex} term during model-free analysis,

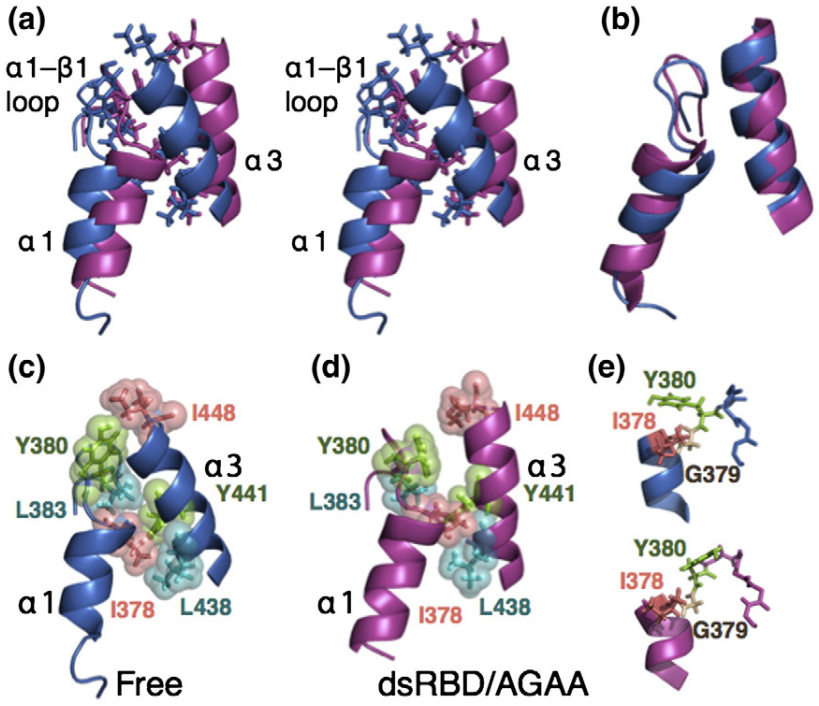


Fig. 3. The α1–α3 extended hydrophobic interface. (a) The hydrophobic interface between helices α1 and α3 in the free dsRBD (blue) and dsRBD/AGAA complex (magenta). Side chains are shown for hydrophobic residues contributing to the interface. The structures are aligned on α2, β1, β2, and β3 (not shown). Space-filling model of the hydrophobic core in the (b) free dsRBD and (c) dsRBD/AGAA complex. (d) The dsRBD and dsRBD/AGAA complex aligned on α1 and α3, illustrating the similar relative orientation of the helices in free and bound states. (e) The dynamic hinge in the α1–β1 loop, comprising residues 378–380, for free (top) and bound (bottom) dsRBDs.

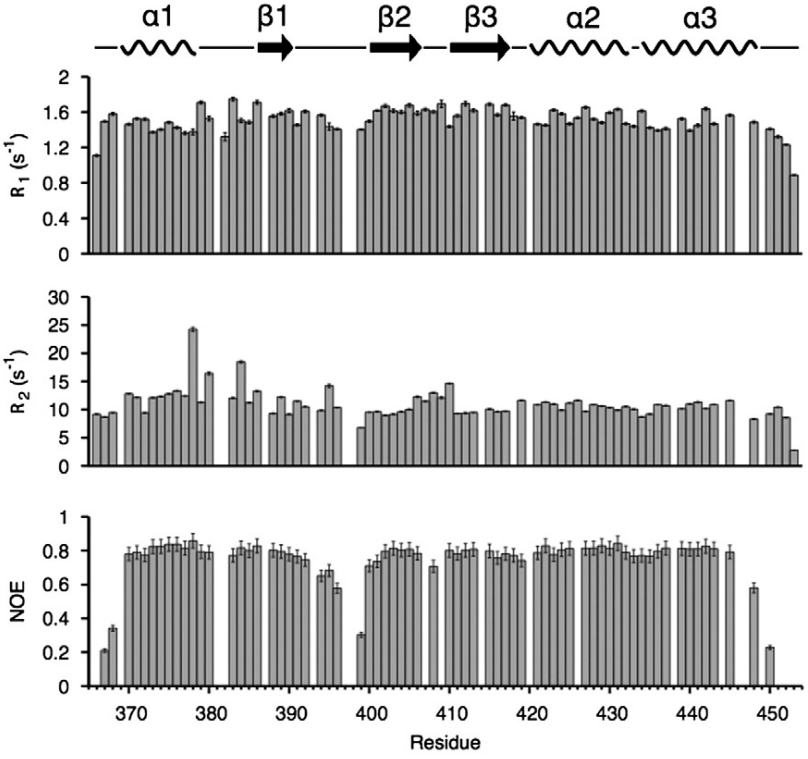


Fig. 4. NMR spin-relaxation parameters. R_1, R_2, and heteronuclear NOE values for the free dsRBD.

are present in some residues in helix α1, the α1–β1 loop, strand β1, the β1–β2 loop, the β2–β3 loop, and residues Y441 and R445 in helix α3 (Fig. 5a and b). The R_{ex} values in the β1–β2 loop and the end of β1 are consistent with the observed flexibility in this region and proposed P393 *cis–trans* isomerization.[34] Within the extended hydrophobic core, a cluster of residues in helix α1 and the α1–β1 loop exhibit notable R_{ex}. One of these residues, I378, has an unusually large value for microsecond-to-millisecond timescale exchange, with an R_{ex} value of 15 s^{-1}. I378 is the C-terminal residue in helix α1 and its hydrophobic side chain is part of the extended hydrophobic core. This large R_{ex} could be due to I378 undergoing jumps between the *trans* and *gauche–* rotamers.[39] Residue R384, which is in the α1–β1 loop, also has a high R_{ex} value (~8 s^{-1}) (Fig. 5a and b). Based on the conformational changes in residues within the α1–β1 loop (Fig. 3e) and the large slow-timescale dynamics for residues 378–380 (Fig. 5a), we identify these residues as a dynamic hinge that we propose allows conformational sampling by helix α1. Since I378 is in the *trans* and *gauche–* conformations in the free dsRBD and RNA-bound dsRBD, respectively, this would suggest that the hinge samples the bound conformation. The β2–β3 loop has low *B*-factors in the crystal structure and is well defined in the solution structures, but there is R_{ex} for some residues. Examination of the structure of the free dsRBD reveals that the residues in the β2–β3 loop that exhibit R_{ex} interact with residues in helix α3 and the α1–β1 loop and would be sensitive to conformational changes in the extended hydrophobic core. The presence of a dynamic hinge and chemical exchange in the extended hydrophobic core imply that the tetraloop-binding helix α1 samples the bound state in the free dsRBD.

Dynamics of the dsRBD in the dsRBD/AGAA complex

To determine whether specific RNA substrate binding changes the microsecond-to-millisecond dynamics observed in the free dsRBD, we collected R_1, R_2, and heteronuclear NOE values for the dsRBD/AGAA hairpin complex whose structure was previously reported[33] (Supplementary Fig. 2). The AGAA hairpin, consisting of a 14-bp dsRNA stem capped by an AGAA tetraloop, is a model substrate derived from the Rnt1p recognition motif in the Rnt1p pre-snoRNA substrate snR47. This RNA provides a minimal binding site for the dsRBD, with only 2–3 bp extending below the interaction of the β1–β2 loop in the minor groove. The average R_1 and R_2 values on the dsRBD/AGAA hairpin complex are lower and higher, respectively, than those of the free protein, as would be expected for the increased molecular weight of the complex. Heteronuclear NOE values indicate that the dsRBD in the complex

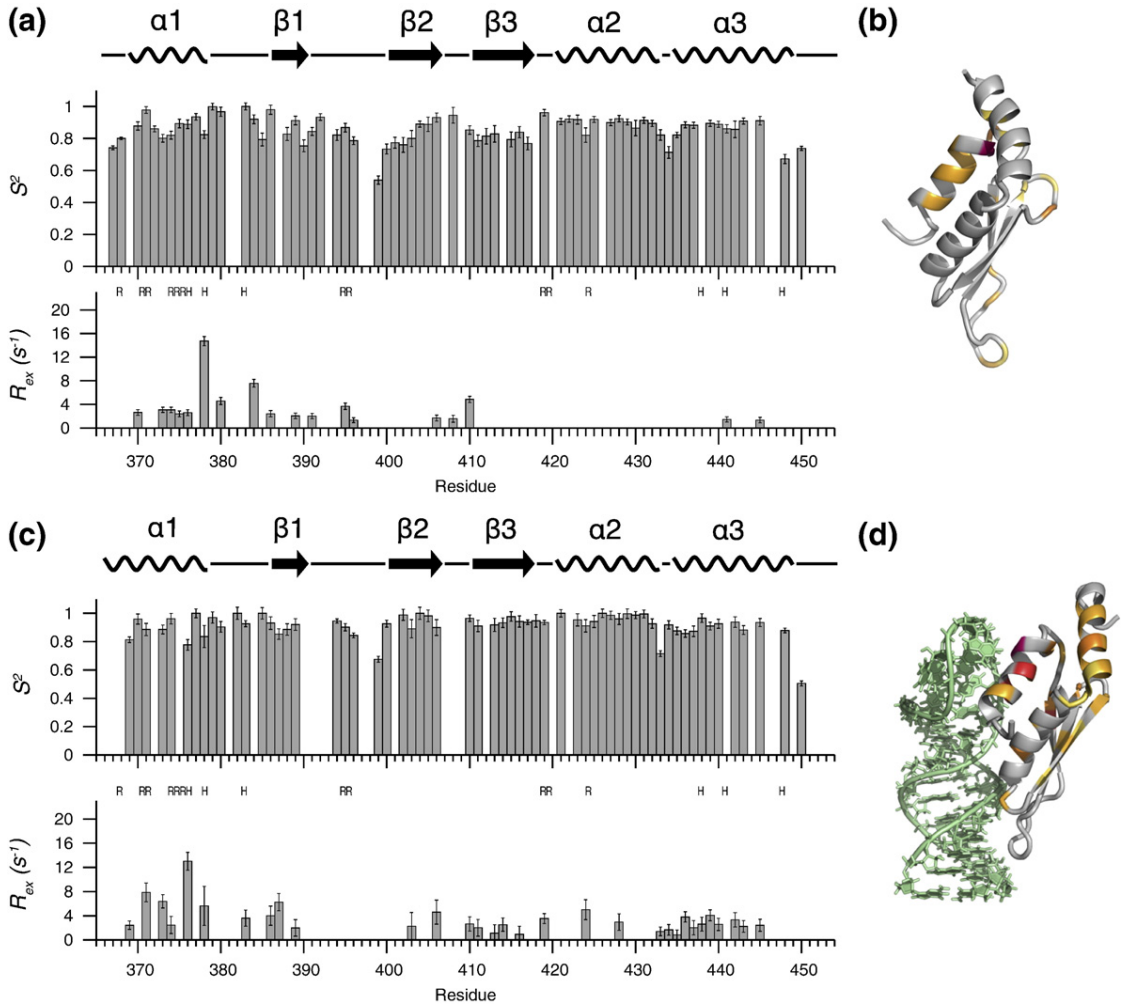


Fig. 5. Fast- and slow-timescale dynamics of the free dsRBD and the dsRBD/AGAA complex at 150 mM NaCl. (a) S^2 and R_{ex} model-free parameters for the free dsRBD. (b) Structure of free dsRBD with residues that show R_{ex} highlighted. (c) S^2 and R_{ex} model-free parameters for the dsRBD/AGAA complex. (d) Structure of dsRBD in the dsRBD/AGAA complex with residues that show R_{ex} highlighted.

is rigid overall, except for the N- and C-termini and the β1–β2 loop. Several residues in the β1–β2 loop have heteronuclear NOE values between 0.4 and 0.6, indicating that this loop remains flexible in the dsRBD/AGAA hairpin complex.

Binding of the dsRBD to the AGAA hairpin results in an overall increase in the S^2 values of most of the protein residues (average increase of 0.15) (Fig. 5a and c). Exceptions are small decreases (<0.1) for helix α1 residues K371 and S376, which interact with the minor groove of the AGAA tetraloop, Y380 and L383 in the extended hydrophobic core in the α1–β1 loop, and the single residue R433 between α2 and α3 (Fig. 5c). R_{ex} values increase for most of the residues in helix α1, with particularly large increases for K371 and S376, which contact the RNA backbone. Helix α3, which has only two residues with R_{ex} in the free dsRBD, also shows R_{ex} for most residues. In contrast, the dynamic hinge residues I378 and Y380, both of which exhibit slow-timescale motions in the free protein, have lower R_{ex} values in the complex. Dynamics in helix α1 likely reflect flexibility at the protein–RNA interface. For helix α3, the uniform increase in R_{ex} could originate from propagation of the dynamics in helix α1 via the extended hydrophobic core and/or from an increase in entropy of the dsRBD in the bound state, an effect that has been observed in other RNA-binding proteins.[40] A395 and V396, near P398 in the β1–β2 loop, which contact the minor groove of the dsRNA stem, also have lower R_{ex} values in the complex (Fig. 5c and d). The decrease in slow-

timescale motions for residues in the dynamic hinge and the β1–β2 loop indicates that some slow-timescale dynamics present in the free dsRBD are quenched upon binding to RNA.

Helix α1 residue S376 has no R_{ex} term in the free dsRBD but has the largest R_{ex} value in the dsRBD/AGAA complex. Helix α1 bends at S376 to insert into the minor groove, and the S376 side chain contacts the RNA backbone on the 3′ side of the AGAA tetraloop. This correlation between changes in structure and dynamics suggests that the dynamic properties of S376 might have a functional role in allowing the dsRBD to adopt the bound conformation. Alternatively, chemical exchange at S376 might be caused by exchange between the specifically and nonspecifically bound states, reflecting the role of this residue in recognizing the backbone of the tetraloop.

To verify that the observed R_{ex} is attributable only to the intrinsic dynamics of the dsRBD and not to nonspecific protein–protein interactions or to exchange between the free and bound state,[41] we measured R_2 values at concentrations of 1 mM and 0.5 mM for the dsRBD/AGAA complex and the free dsRBD (Supplementary Fig. 3). In the absence of these possible additional contributions to chemical exchange, R_2 values and NMR linewidths would be expected to be the same at both concentrations. In both cases, R_2 values and NMR linewidths for two protein concentrations are nearly identical, indicating that the dynamics determined by model-free analysis arise only from the intrinsic dynamics of the dsRBD and not from other possible contributions to chemical exchange.

In summary, two distinct changes in dsRBD dynamics in the extended hydrophobic core are observed upon substrate binding. First, there is a general increase in slow-timescale dynamics for residues in helices α1 and α3 that is associated with concerted changes in the extended hydrophobic core. Second, there is a decrease in slow-timescale dynamics for residues in the α1–β1 hinge and the β2–β3 loop, due to "locking in" of helix α1 by shape-specific binding to the tetraloop minor groove.

Ionic strength dependence of dynamics for the dsRBD in the dsRBD/AGAA complex

Previous NMR titration and isothermal titration calorimetry experiments revealed that the dsRBD can bind to the AGAA hairpin both specifically and nonspecifically at 150 mM NaCl, with saturation of the RNA at a protein:RNA ratio of 2:1.[35] The relaxation data discussed above for the dsRBD/AGAA complex were measured at a protein:RNA ratio of 1:1.1 and were expected to primarily reflect values for the dsRBD bound to the specific site. To further confirm this, we investigated the binding and dynamics of the dsRBD/AGAA hairpin complex at 300 mM NaCl (Supplementary Fig. 4). At this salt concentration, nonspecific binding should be minimal. Chemical shift mapping for the dsRBD upon RNA binding at 300 mM NaCl revealed chemical shift changes similar in pattern to those for the complex at 150 mM NaCl, but with a much smaller magnitude (Fig. 6). This is consistent with the lower binding affinity of the dsRBD for RNA at a higher salt concentration, as measured by NMR titration and

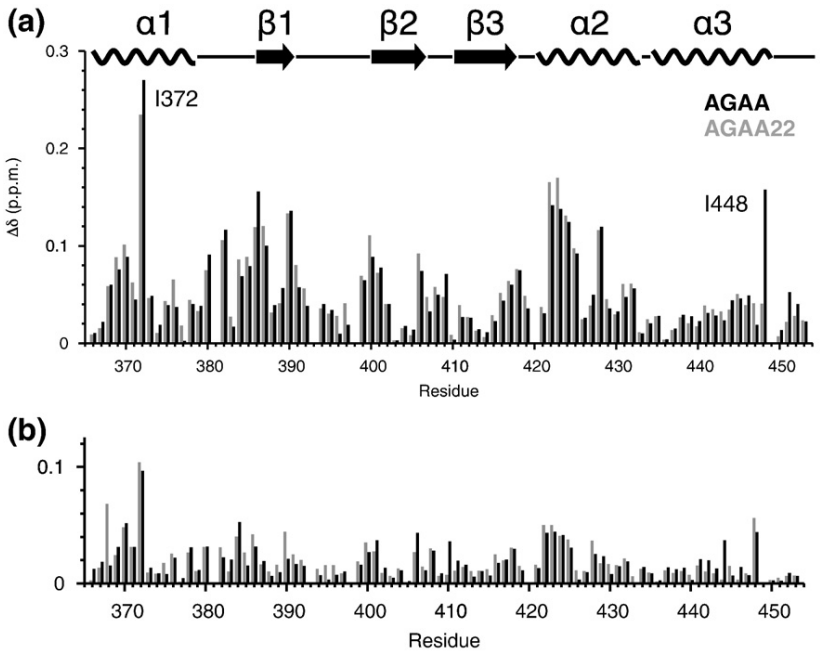


Fig. 6. The effect of substrate length and salt concentration on chemical shift. Chemical shift mapping of the dsRBD at 150 mM (a) and 300 mM (b) NaCl bound to a tetraloop hairpin RNA with a 14-bp stem (AGAA) (black) and a tetraloop hairpin RNA with a 22-bp stem (AGAA22) (gray).

isothermal titration calorimetry.[35] At 300 mM NaCl, the dsRBD exhibits R_{ex} values similar overall to those observed at 150 mM NaCl (compare Fig. 7a and b with Fig. 5c and d). This observation is consistent with a single, specific binding site on the AGAA hairpin. Hence, the dsRBD is fully bound to the AGAA hairpin at the specific binding site under the conditions used for spin-relaxation experiments at a high concentration of the complex and 150 mM or 300 mM NaCl.

Dynamics of the dsRBD in the presence of both specific and nonspecific binding sites

As discussed above, the dsRNA construct was designed such that it has a minimal binding site for the dsRBD. To investigate whether there is a difference in dsRBD dynamics when both nonspecific and specific binding sites are present, we collected NMR spin-relaxation data for Rnt1p dsRBD in complex with an AGAA tetraloop hairpin with a 22-bp stem (AGAA22) at 300 mM NaCl (Supplementary Fig. 5). AGAA22 has a stem that is 8 bp longer than that of the AGAA hairpin (14 bp), allowing for nonspecific binding to the longer dsRNA stem in addition to specific binding site at the AGAA tetraloop. Because of the longer stem, Rnt1p dsRBD can potentially exchange between the specific site and nonspecific sites on AGAA22. AGAA22 more closely reflects native conditions for substrate binding by Rnt1p dsRBD, as the stem length is the same as the stem in the pre-snR47 snoRNA (excluding a single bulge). In general, the values for chemical exchange, as described by R_{ex}, are

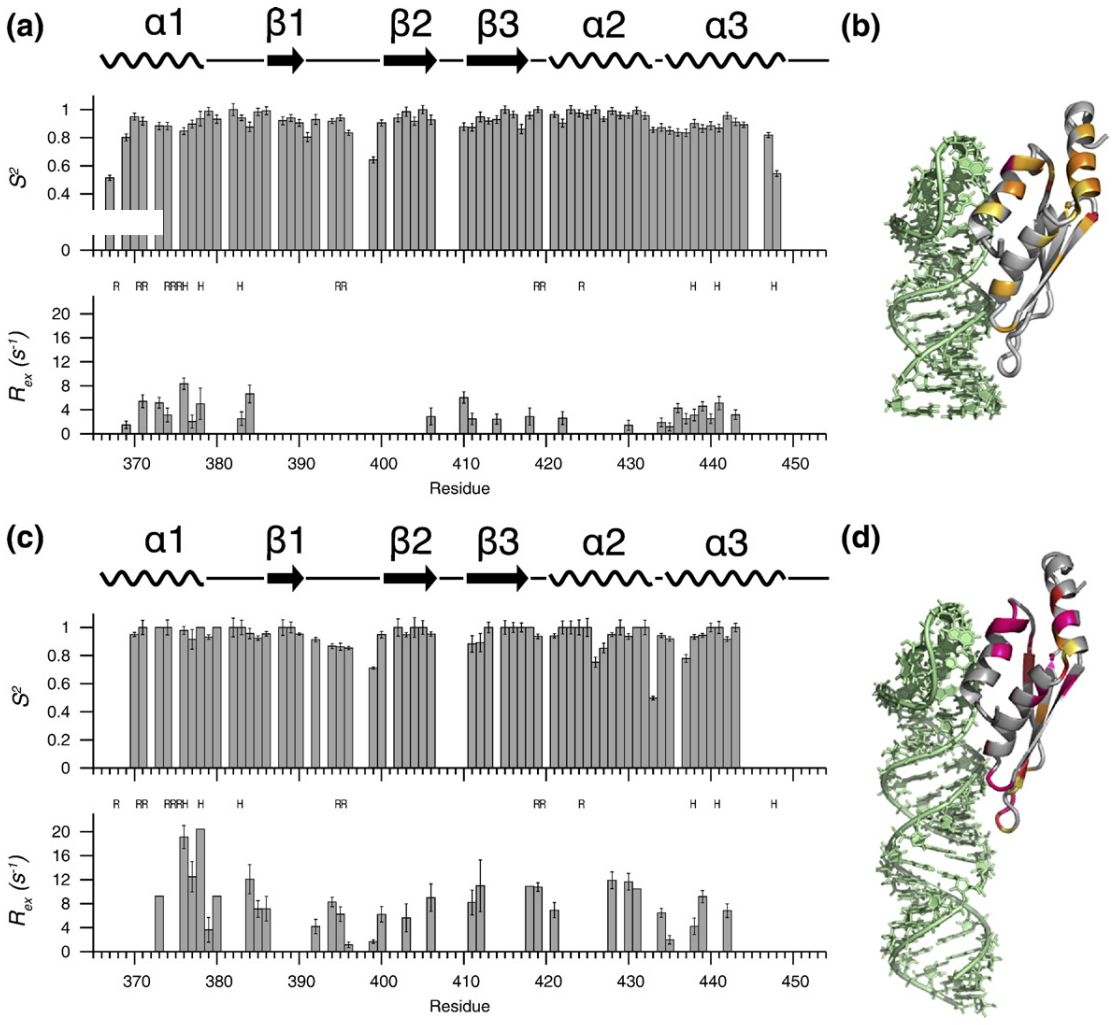


Fig. 7. Fast- and slow-timescale dynamics of dsRBD/AGAA and dsRBD/AGAA22 complexes at 300 mM NaCl. (a and b) S^2 and R_{ex} model-free parameters for the (a) dsRBD/AGAA complex and (b) dsRBD/AGAA26 complex. (c and d) R_{ex} values mapped onto the structure of the (c) dsRBD/AGAA and (d) dsRBD/AGAA26 complex.

significantly larger for dsRBD/AGAA22 than for dsRBD/AGAA. Slow-timescale dynamics for dsRBD/AGAA22 are present for residues in helices α1 and α3 in or near the extended hydrophobic core, residues throughout the RNA-binding interface, including helix α2 and the β1–β2 loop, and in β2 and β3. In dsRBD/AGAA, there was no R_{ex} for any residues in the β1–β2 loop, while in dsRBD/AGAA22, most of the β1–β2 loop residues show R_{ex} (Fig. 7c and d). The R_{ex} in the β1–β2 loop is both larger and present in more residues than in the free dsRBD (Fig. 5a). Under the experimental conditions of 300 mM NaCl and excess RNA, the dsRBD is relatively selective for specific binding; thus, the additional contribution to R_{ex} arising from the presence of additional nonspecific binding sites can be attributed to exchange between tetraloop (specific) and stem (nonspecific) binding sites. It is notable that all of the elements of the dsRBD that interact with the minor and major groove of the dsRNA stem show more conformational exchange than when only a specific binding site is available. The additional protein dynamics for the dsRBD/AGAA22 complex, compared to the dsRBD/AGAA complex, also reveal the significance of conformational changes in the extended hydrophobic core, in addition to residues at the RNA binding interface, in binding site selection. We conclude that the difference in R_{ex} for the complex with AGAA22 *versus* AGAA reflects some nonspecific binding to the dsRNA on the longer substrate. Furthermore, once helix α1 locks in to the tetraloop, the rest of the dsRBD locks into place, resulting in a decrease of R_{ex} in the β1–β2 loop for AGAA *versus* AGAA22.

Hydrophobic interactions with the α1–β1 loop maintain dsRBD stability

To extend insights from our characterization of dsRBD structure and dynamics, we further investigated the importance of residues in the α1–β1 loop for RNA binding by generating four dsRBD mutants with single mutations in the α1–β1 loop: I378A, G379P, G379A, and Y380A. These three residues are part of the α1–β1 hinge in the extended hydrophobic core. The side chains of I378 and Y380 change position between the free and RNA-bound states and maintain hydrophobic contacts with residues in helix α3. The ϕ and ψ angles for G379 also change between free and bound states, due to conformational changes in the hydrophobic core (Fig. 3c–e). The $^{1}H–^{15}N$ heteronuclear single quantum coherence (HSQC) spectrum of I378A was poorly dispersed (Supplementary Fig. 6), and the CD spectrum showed no evidence for secondary structure (Supplementary Fig. 7), indicating that the majority of the protein is unfolded at 25 °C. However, in freshly prepared protein samples, there appears to be about 10% folded protein based on the $^{1}H–^{15}N$ HSQC. Addition of RNA to I378A results in some chemical shift changes indicative of binding for the peaks from the folded protein, but fewer than for the wild-type dsRBD, and the protein unfolds over time. Thus, we conclude that the mutation I378A destabilizes the protein and may also lower RNA binding affinity. Since I378 interacts with residue Y441, which is in helix α3 and part of the extended hydrophobic core, we tested the importance of this interaction by making a Y441A mutation. Y441A is also unstable in solution and precipitates after about 20 min at 25 °C, and $^{1}H–^{15}N$ HSQC spectra indicate that it is unfolded prior to precipitation. We note that the mutation R445A, in helix α3, was previously shown to destabilize the extended hydrophobic core of the dsRBD.[33,34] This residue is close to the α1–β1 loop residue S382 and shows R_{ex} in both the free and bound dsRBD.

For the G379P mutant, CD spectra indicate that the T_m decreases by ~6 °C and that melting is less cooperative. The $^{1}H–^{15}N$ HSQC (Supplementary Fig. 7) has chemical shift changes throughout helices α1 and α3, and none of the resonances for the α1–β1 loop are observed (Supplementary Figs. 6 and 8). Analysis of the backbone chemical shifts for dsRBD G379P indicates that the C-terminal end of helix α1 and all of helix α3 are altered relative to the wild-type dsRBD (Supplementary Fig. 8). Addition of the AGAA hairpin to G379P up to a 2:1 excess of RNA resulted in almost no changes in the $^{1}H–^{15}N$ HSQC spectrum, indicating that the G379P substitution essentially abrogates binding of the dsRBD to RNA. G379A had a substantially altered $^{1}H–^{15}N$ HSQC spectrum (Supplementary Fig. 6) and is unstable, as the protein precipitated after 20 min at 25 °C. However, in the presence of the AGAA hairpin, the G379A mutant gives $^{1}H–^{15}N$ HSQC spectra, indicating that G379A forms a stable complex (Supplementary Fig. 6). Lastly, we found that Y380A degrades during expression, implying that the mutation of this residue also significantly destabilizes the protein.

In summary, all of the mutations in the α1–β1 loop and α3 destabilize the extended hydrophobic core to some extent and have variable effects on RNA binding. For the G379P mutation, changes to the extended hydrophobic core completely disrupt RNA binding, although this mutation has the smallest effect on dsRBD stability. For G379A, binding to RNA helps stabilize the folded state of the dsRBD. Because all of the residue substitutions in the extended hydrophobic core affect protein stability, we were not able to assess their effects on dynamics independently. Nevertheless, these results support a central structural role for the α1–β1 loop and extended hydrophobic core in maintaining dsRBD stability.

Effect of α1–β1 loop mutations on snoRNA processing *in vivo*

To determine whether the mutations in the α1–β1 hinge have an effect on cleavage of Rnt1p substrates *in vivo*, we introduced the single mutations I378A, G379A, and G379P into the RNT1 gene and examined by Northern blot the effect of these mutations on the processing of snR36 and snR47 snoRNAs *in vivo*. For comparison, we included in the analysis strains harboring a previously studied mutation in helix α1 that affect processing (K371A),[33] a catalytically inactive mutant (E320K), and an RNT1 deletion (rnt1Δ). Unlike the previously studied K371A mutant, all three α1–β1 loop mutants exhibit temperature-sensitive growth defects (Fig. 8a and b). The growth defects for strains bearing the I378A and G379A mutations are comparable and relatively modest, while the G379P strain had a growth defect comparable to the rnt1Δ strain. This is consistent with the *in vitro* results that showed that while stable, dsRBD G379P does not bind RNA.

The strain bearing the I378A mutation shows an inhibition of snoRNA processing comparable to the K371A mutation, with a slight processing defect for snR47 and a more pronounced defect for snR36 (Fig. 8c). snR36 was previously observed to be more sensitive than snR47 to mutations in the Rnt1p dsRBD, because of the presence of a large bulge after the fourth base pair below the tetraloop.[42] The G379A strain exhibited only minor effects on snoRNA processing *in vivo*. Although it has a growth defect comparable to I378A, the dsRBD is stabilized by binding to RNA, which may explain the difference in effect on snoRNA processing. The G379P strain, in contrast, showed severe processing defects for both substrates. The processing defects in strains bearing the I378A, G379P, and G379A mutations are consistent with NMR and CD results that indicate that these mutations introduce changes in stability and RNA-binding affinities of the dsRBDs. We conclude that mutations in the α1–β1 hinge, all of which destabilize the extended hydrophobic core and affect RNA binding to different extents, compromise the function of Rnt1p *in vivo*.

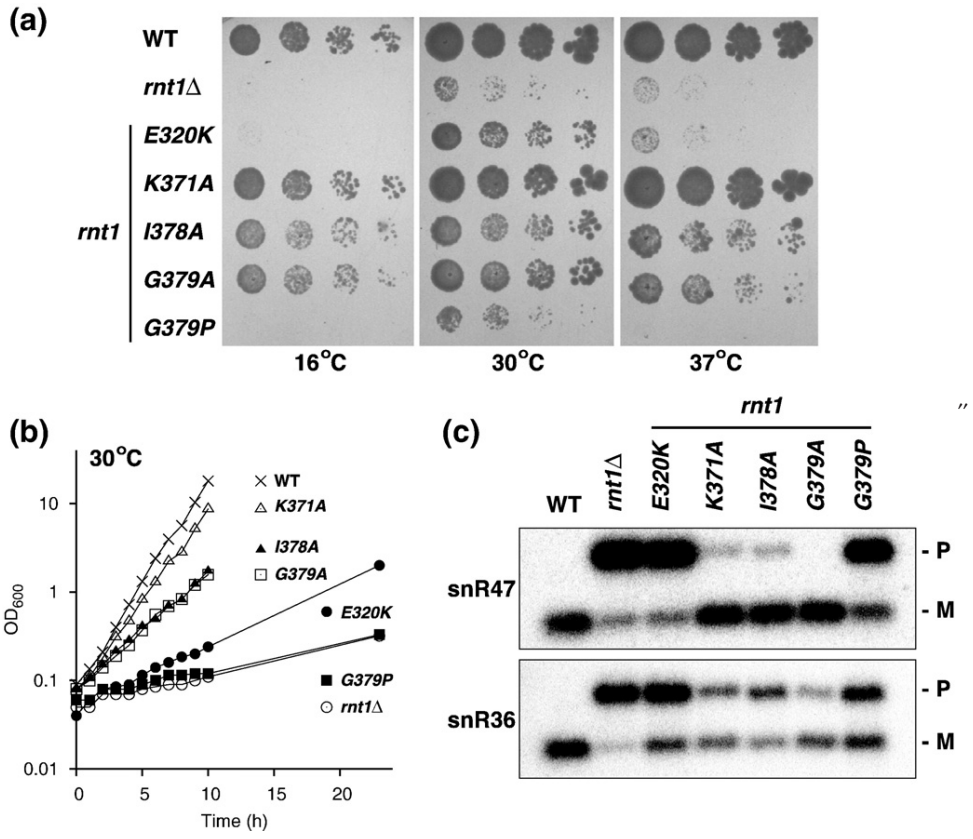


Fig. 8. *In vivo* analysis of dsRBD α1–β1 loop mutants. (a) Growth of wild-type dsRBD and dsRBD hinge mutants, with serial dilutions at 16, 30, and 37 °C. (b) Growth curve for the wild-type dsRBD and dsRBD hinge mutants at 30 °C. (c) Northern blot analysis of snR47 and snR46 snoRNA processing for wild-type dsRBD and dsRBD hinge mutants, showing unprocessed precursor (P) and mature snoRNAs (M).

Discussion

dsRBDs recognize dsRNA primarily by interactions with the phosphodiester backbone of successive minor, major, and minor grooves via the β1–β2 loop, the N-terminal end of helix α2, and helix α1, respectively. The dsRBD of Rnt1p is unusual in that helix α1 recognizes a tetraloop through shape-specific recognition of its minor groove.[33] The Rnt1p dsRBD also has an additional helix α3, which packs against the α1–β1 loop to form a distinctive extended hydrophobic core. All dsRBDs have a conserved hydrophobic core, with residues contributed by the C-terminal residues of helix α1, helix α2, and strand β3. Hydrophobic interactions among these residues in the core of the protein stabilize the folded conformation of the dsRBD.[29,43] The unique hydrophobic interface between helices α1 and α3 in the Rnt1p dsRBD is contiguous with the conserved hydrophobic core. Through a detailed analysis of the structures and dynamics of the free and bound dsRBDs, we have shown that this extended hydrophobic core plays an essential role in enabling defined conformational changes associated with RNA substrate recognition. Residues in the α1–β1 loop that interact with helix α3 to form the extended hydrophobic core constitute a dynamic hinge that allows a concerted change in the positions of helix α1 and helix α3 between the free and bound states, a key feature of substrate recognition by the Rnt1p dsRBD (Fig. 3). The importance for RNA binding of residues in the extended hydrophobic core, which includes the α1–β1 hinge, is further supported by the results of mutagenesis of individual residues on snoRNA processing *in vivo* (Fig. 8).

Helix α3 has been previously proposed to contribute to specific RNA binding indirectly by affecting the length and orientation of helix α1 in the free protein.[34] However, the orientation of helix α1 in the free dsRBD is the same as other dsRBDs that bind to dsRNA nonspecifically. Our results show that the helix α3 contributes to substrate-specific binding by participating in the reorientation of helix α1 in the bound state through concerted structural changes in the extended hydrophobic core.

The free dsRBD samples multiple conformations

Backbone dynamics of the free dsRBD obtained via model-free analysis of NMR spin-relaxation data reveal extensive slow-timescale dynamics primarily localized in the RNA-binding interface and extended hydrophobic core, including helix α1, the α1–β1 loop, and the β1–β2 loop, as well as limited dynamics in helix α3. Our comparison of free and bound structures shows that upon binding of the RNA substrate, the α1–β1 loop changes conformation to allow helices α1 and α3 to undergo concerted changes in orientation and side-chain position in order for helix α1 to be able to bind to the minor groove of the substrate tetraloop (Figs. 2 and 3). The β1–β2 loop also translocates about 6 Å to bind to the minor groove of dsRNA one helical turn away from the tetraloop (Fig. 2). Thus, the dynamic behavior of the free dsRBD on both fast and slow timescales is associated with conformational changes within the extended hydrophobic core that accompany substrate recognition. This localized flexibility supports the notion that conformational adaptation upon substrate binding is enabled by the dynamics of the free dsRBD. Mutations of residues in the α1–β1 loop affect the stability and RNA-binding properties of the dsRBD, revealing that the interactions of the α1–β1 loop with helix α3 are essential for dsRBD stability and function.

The observed dynamics of the α1–β1 loop in the free dsRBD and the concerted movement of helices α1 and α3 led us to propose that residues 378–380 in the α1–β1 loop serve as a dynamic hinge enabling conformational exchange between the free and bound states. One possible model for the contribution of the α1–β1 hinge is that hinge dynamics on the microsecond-to-millisecond timescale backbone facilitate conformational sampling by helix α1. Moreover, hinge dynamics would facilitate concerted movement of helices α1 and α3 upon binding because hinge residues I378 and Y380 are also part of the extended hydrophobic core. These dynamics, along with those of the β1–β2 loop, are partially quenched upon binding to the specific site on target substrates (Fig. 9). We cannot exclude, however, that the dsRBD experiences a combination of conformational selection and induced fit to achieve its final bound conformation.

Previous studies of the contribution of protein dynamics to RNA recognition have shown that high-affinity binding to RNA is generally associated with the presence of extensive R_{ex} throughout an RNA-binding domain. NMR relaxation studies of the two dsRBDs of protein kinase R indicated that residues that directly interact with the RNA and throughout helix α1, sheet β1, and helix α2 have slow-timescale motions for PKR dsRBD1, which binds to dsRNA with high affinity. In contrast, there are but few such motions for PKR dsRBD2, which has weaker binding affinity for dsRNA.[44] Dynamics within the PKR dsRBD1 were proposed to allow for adaptation to nonuniform RNA substrates. The observed R_{ex} within PKR dsRBD1 and Rnt1p dsRBD are different in both the distribution and extent of slow-timescale dynamics. Thus, dsRBDs with different dsRNA substrates can have different dynamic modes despite having similar structures in the free state. Here, we have shown the first example where dynamic properties of a dsRBD are associated with defined structural changes in the protein that take place upon binding to RNA.

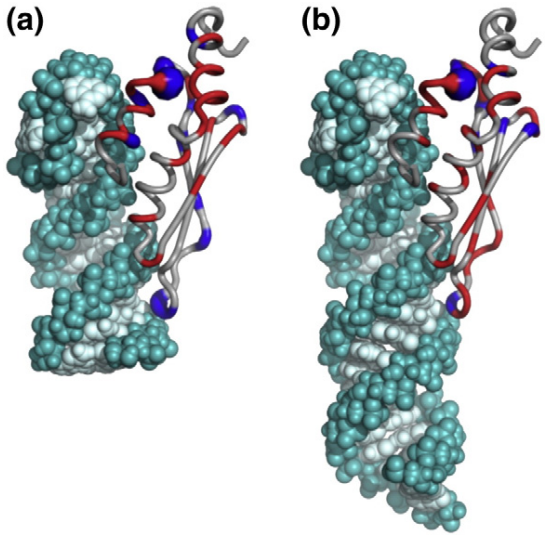


Fig. 9. dsRBD dynamics are associated with concerted structural changes necessary for binding. Change in R_{ex} for free *versus* RNA-bound dsRBD for (a) the dsRBD/AGAA complex and (b) the dsRBD/AGAA22 complex. The model of the dsRBD/AGAA22 complex with dsRBD bound specifically at the tetraloop is based on the dsRBD/AGAA complex. ΔR_{ex} (R_{ex} free dsRBD minus R_{ex} dsRBD in complex with RNA) is shown in red (increase in R_{ex}) and blue (decrease in R_{ex}). Relative magnitude of the change in R_{ex} is represented by the tube width.

Substrate binding induces changes in backbone dynamics

In complex with RNA, dynamics in Rnt1p dsRBD are present throughout the extended hydrophobic core and RNA-binding interface, corresponding to binding-induced conformational changes. Changes in R_{ex} between free and AGAA tetraloop hairpin-bound dsRBDs are shown in Fig. 9. Residues within the α1–β1 hinge that are dynamic in the free dsRBD are partly or completely quenched, while residues within helices α1 and α3 have even larger slow-timescale dynamics in the complex. Residues at the N-terminus of helix α1 in the complex fold only upon binding to RNA,[35] and these residues also have a decrease in slow timescale dynamic upon binding. Two residues within the β1–β2 loop also have lower R_{ex} in the complex than in the free dsRBD, which is consistent with the stable interaction of this loop with the stem minor groove.

An increase of the prevalence of chemical exchange upon substrate binding appears to be a common feature of nucleic acid binding proteins, as well as proteins involved in protein–protein interactions.[31,45] Typically, this increase in R_{ex} is observed to occur at sites distant from the interface with the partner molecular and implies an indirect role for conformational flexibility in binding. The increase in R_{ex} is often distributed throughout the protein and is not well correlated with specific structural changes. In contrast, for Rnt1p dsRBD, the observed increase in R_{ex} for regions of the protein distal from the RNA binding surface, particularly in the C-terminus of helix α3, is correlated with conformational changes associated with RNA binding. The counterintuitive increase in dynamics for regions of the dsRBD distal to the RNA-binding face reveals that a broad network of residues within the dsRBD contributes to conformational adaptation to the specific tetraloop binding site. The results of these experiments highlight the importance of protein conformational flexibility, particularly within the extended hydrophobic core, in binding of the Rnt1p dsRBD to RNA tetraloop hairpin substrates.

Dynamics reflect binding site exchange on a long substrate

Because the dsRBD binds to the minimal substrate (AGAA) at 300 mM NaCl only at the specific tetraloop site, slow-timescale dynamics are likely to be limited to intrinsic dynamics of the bound state, reflecting intrinsic conformational entropy of the dsRBD in complex with RNA. However, exchange between specific and nonspecific sites becomes significant in the dsRBD/AGA22 complex, as the longer stem allows for nonspecific binding away from the tetraloop. In the dsRBD/AGAA22 complex, chemical exchange values are quantitatively higher and are present throughout the dsRBD. Hence, dynamics within the dsRBD in the dsRBD/AGAA22 complex suggest that the dsRBD distinguishes between specific and nonspecific complexes after binding to the dsRNA substrate through conformational exchange. Increased dynamics throughout the extended hydrophobic core and RNA-binding interface, compared to the dsRBD/AGAA complex (Fig. 9b), suggests that flexibility within the RNA-bound dsRBD remains important for selection of the specific binding site even after the dsRBD is bound to RNA. Moreover, elevated R_{ex} values of helix α1 and the β1–β2 loop would be expected to be present at the RNA-binding interface for the dsRBD/AGAA22 complex if the dsRBD searches between the specific tetraloop site and nonspecific stem region, since contacts to the RNA stem in the nonspecific complex would not be identical with the specific complex.

Comparison to budding yeast Dicer

Budding yeast Dicers have two dsRBDs: the first (dsRBD1) is located immediately adjacent to the endoND, as in Rnt1p, and the second (dsRBD2) is at the C-terminus of the protein and separated from dsRBD1 by a long intervening sequence with no known structural motifs. The *Saccharomyces castellii* Dcr1 dsRBD1, but not dsRBD2, was shown to

be necessary for efficient processing of long dsRNA substrates to 23-nt fragments and for specificity of the enzyme for dsRNA over single-stranded RNA.[46] Dcr1, however, does not have tetraloop specificity. A superimposition of the crystal structure of the free *Kluyveromyces polysporus* Dcr1 dsRBD1 and the solution structure of the free Rnt1p dsRBD (Supplementary Fig. 9) indicates that the two dsRBDs have the same overall conformation, including a short helix α3 in Dcr1, but differ in regions that are important for specific RNA binding by Rnt1p dsRBD. Notably, in Dcr1, there are no interactions between the α1–β1 loop and helix α3 because helix α3 is shorter in the Dcr1 dsRBD1, and the α1–β1 loop adopts a different conformation compared to the Rnt1p dsRBD. Although Dcr1 dsRBD1 has some conserved hydrophobic residues in the α1–β1 loop (e.g., L278 and I280 in *K. polysporus* Dcr1 correspond to Rnt1p dsRBD hinge residues I378 and Y380), it does not appear to have an extended hydrophobic core. L278 cannot undergo the rotameric change that we see for I378, and the I280 side chain is oriented toward the outside of the protein. These residues do not interact with helix α3 and do not appear to constitute an analogous hinge. Interestingly, *K. polysporus* Dcr1 L275 has hydrophobic interactions with the conserved residue Y341, which is in the loop extending past helix α3 in the dsRBD1. We speculate that the absence in Dcr1 dsRBD1 of an extended hydrophobic core involving the α1–β1 loop results in the loss of tetraloop specificity for *K. polysporus* and *S. castellii* Dcr1. Nevertheless, structural features within this region that are unique to the Dcr1 dsRBD may affect its binding affinity to Dcr1 substrates. Interestingly, *C. albicans* Dcr1 is able to carry out both Rnt1 and Dcr1 functions,[28] suggesting that its dsRBD1 may retain the structural features necessary for tetraloop recognition, including the extended hydrophobic core.

Materials and Methods

NMR sample preparation

The Rnt1p dsRBD (residues 366–453 of Rnt1p) and single-residue mutants were expressed as glutathione *S*-transferase (GST) fusion proteins containing a thrombin cleavage site using the pGEX-2T vector (GE Healthcare) in BL21 (DE3) Gold cells (Stratagene).[33] The ^{15}N-labeled and ^{13}C,^{15}N-labeled GST-dsRBD fusion proteins were expressed at 30 °C and 37 °C, respectively, for 16 h in M9 minimal media containing 1 g/L ^{15}N ammonium chloride and 1 g/L ^{13}C glucose. GST-dsRBD was purified using a GSTrap 4B glutathione Sepharose column (GE Healthcare), followed by a HiLoad 26/60 Superdex 75 pg (S75) gel-filtration column (GE Healthcare). GST-dsRBD was cleaved with 10 units of thrombin (GE Healthcare) per milligram of fusion protein for 24 h at a concentration of about 1 mg/mL in a buffer containing 20 mM Tris–HCl, pH 7.5, 150 mM NaCl, and 2.5 mM $CaCl_2$. Cleaved dsRBD was purified on an S75 gel-filtration column in NMR buffer (20 mM sodium phosphate, pH 6.5, 150 mM NaCl, and 1 mM DTT) and concentrated to about 1 mM.[35]

RNA samples were prepared by *in vitro* transcription from a synthetic double-stranded DNA template using mutant T7 RNA polymerase (P266L)[47] and purified on a 15% denaturing polyacrylamide gel containing 8 M urea as previously described.[48] RNA samples were electroeluted, further purified on a HiTrap Q anion-exchange column (GE Healthcare), exchanged into NMR buffer using an Amicon Ultra centrifugal filter, and concentrated to about 1 mM. RNAs were then refolded by heating to 95 °C and slow cooling to 4 °C. dsRBD/RNA complexes for NMR spectroscopy were prepared by adding RNA to protein at a 1.1:1 ratio of RNA:protein and concentrated to 0.8–1 mM.

NMR spectroscopy for structure calculations

NMR spectra for structure determination were recorded on Bruker DRX 500- and 600-MHz spectrometers at 25 °C. The assignments for the Rnt1p dsRBD were derived from analysis of 3D CBCANH, 3D CBCA(CO)NH, 3D HCCH-total correlated spectroscopy, 3D HCCH-correlated spectroscopy, 3D ^{13}C-NOE spectroscopy-HSQC, and 3D ^{15}N-NOE spectroscopy-HSQC experiments[49–51] acquired on ^{13}C,^{15}N-labeled dsRBD. One-bond ^{1}H–^{15}N RDCs were measured from HSQC–in-phase/antiphase experiments[52] in the presence and absence of the RDC alignment medium C12E5/Hexanol[53] on the 600-MHz spectrometer. A total of 84 RDCs were obtained for the free dsRBD. For structure determination of the free dsRBD, a total of 2068 experimental distance restraints were obtained from NOE intensities and classified as strong (1.8–3.0 Å), medium (1.8–4.5 Å), and weak (1.8–6.0 Å). A total of 138 dihedral angles were derived from TALOS.[54] Structures were calculated using the NIH-Xplor package[55] following standard protocols. Briefly, the calculation started from the extended protein in random orientations. The protein was then folded during 40,000 steps of high-temperature dynamics with a time step of 0.002 fs. The structures were cooled down using 75 ps of slow cooling from 2000 K to 100 K. The final structures were obtained after refinement with 52 RDCs (only RDCs from secondary-structure elements were used for structure calculations) during a second slow cooling from 1200 K to 100 K. The axial (−30 Hz) and rhombic (0.52) components of the alignment tensor were derived from a grid-search procedure.[56] The force constant for the RDCs was gradually increased from 0.001 to 0.2 kcal mol^{-1} Hz^{-2}. The 20 lowest-energy structures were selected, and the structures were analyzed using MOLMOL[57] and PyMOL.[58]

NMR spectroscopy for spin-relaxation experiments

R_1, R_2, and ^{1}H–^{15}N NOE values were measured for the free dsRBD and the dsRBD/AGAA and dsRBD/AGAA22 complexes at 20 °C on a Bruker DRX 600-MHz spectrometer. R_1 experiments used the following time delays: for the free dsRBD, 41, 161, 299, 299, 437, and 644 ms; for the dsRBD/RNA complexes, 46, 207, 207, 575, 575, and 989 ms. R_2 rates were determined with Carr–Purcell–

Meiboom–Gill experiments, with the following time delays: for the free dsRBD, 11.2, 22.4, 22.4, 44.8, 67.2, 67.2, and 89.6 ms; for the dsRBD/RNA complexes, 11.2, 22.4, 22.4, 33.5, 33.5, and 44.8 ms. Spectra were processed using NMRPipe/NMRDraw, and peak intensities were obtained using NMRView. Relaxation rates were determined by fitting the expression for relaxation decay, $I(R) = I_0 e^{-Rt}$, to the peak intensities using in-house software.

Model-free analysis of relaxation data

Initial estimates of the rotational correlation time and the diffusion tensor for the free dsRBD$_{366–453}$ (2LUQ; reported here) and RDC-refined dsRBD/AGAA complex (PDB ID: 2LUP) were obtained using the program HYDRONMR[59,60] and were subsequently optimized using the program ModelFree 4.20[61] prior to model selection. Relaxation parameters were interpreted using the Lipari–Szabo model-free formalism to obtain values for motional parameters describing the dynamic behavior of backbone amide bond vectors.[62,63] ModelFree[61] was used to fit relaxation data for each residue to one of five increasingly complex models using optimized initial estimates of the diffusion tensor and correlation time, where model 1 includes the parameter S^2_s; model 2, S^2_s and τ_e; model 3, S^2_s and R_{ex}; model 4, S^2_s, τ_e, and R_{ex}; and model 5, S^2_s, S^2_f, and τ_e. Following model selection for all residues, global and internal parameters were optimized with a grid-search algorithm using an axially symmetric diffusion tensor (Supplementary Tables 1 and 2). For model-free analysis, bond lengths of 1.02 Å and CSA values of −160 p.p.m. were used. Using a bond length of 1.04 Å results in small changes in the values of model-free parameters (<5%) but does not change the outcome of our analysis. To confirm that the R_{ex} terms that we observe reflect backbone chemical exchange rather than diffusion anisotropy, we checked that the most significant N–H bond vectors for the free dsRBD are not aligned with the long axis of the diffusion tensor by calculating the angle between the N–H bond and the diffusion tensor *z*-axis, as defined by the fitted diffusion tensor obtained after model-free analysis.

In vivo analysis of RNT1 hinge mutants

All strains were derived from the BMA64 background. The rnt1::TRP deletion mutant and RNT1 K371A dsRBD mutant were described previously.[9,42] The dsRBD hinge mutants (I378A, G379A, G379P) and the catalytic mutant (E320K) were constructed using the delitto perfetto method.[64] A strain carrying the CORE KanR-URA3 cassette at position S376 was transformed with double-stranded DNA oligonucleotides to excise the CORE sequence and introduce the appropriate mutation in the hinge (I378A, G379A, or G379P), while the E320K mutant was produced from a strain with the CORE KanR-URA3 insertion at position E320. Genomic DNA sequences were confirmed by sequencing. Strains were grown in YPD and RNA was harvested and analyzed by Northern blotting as previously described[9] with the following modifications: 10 μg of RNA was denatured with glyoxal, run on 1× BPTE 2% agarose gels as previously described,[65] and transferred to Hybond-N+ membranes (GE Healthcare).

Accession numbers

Coordinates for the 20 lowest-energy structures of the Rnt1p dsRBD$_{366–453}$ have been deposited in the PDB under accession code 2LUQ, and chemical shifts have been deposited in the Biological Magnetic Resonance Data Bank under accession code 18535.

Acknowledgements

This work was supported by National Institutes of Health grants GM37254 and GM48123 to J.F. Qi Zhang was a Baltimore Family Fund Postdoctoral Fellow of the Life Sciences Research Fund. We thank Anni Zhao for help with preparation of the hydrophobic core mutants.

Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2012.11.025

Received 18 September 2012;
Received in revised form 17 November 2012;
Accepted 20 November 2012
Available online 28 November 2012

Keywords:
NMR;
Rnt1;
RNA binding domain;
protein–RNA;
spin relaxation

† E.H. and Z.W. contributed equally to this work.

Present address: Q. Zhang, Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, 120 Mason Farm Road, Campus Box 7260, Chapel Hill, NC 27599-7260, USA.

Abbreviations used:
dsRNA, double-stranded RNA; dsRBD, double-stranded RNA-binding domain; snoRNA, small nucleolar RNA; endoND, endonuclease domain; PDB, Protein Data Bank; RDC, residual dipolar coupling; NOE, nuclear Overhauser enhancement; GST, glutathione *S*-transferase.

References

1. Ketting, R. F., Fischer, S. E., Bernstein, E., Sijen, T., Hannon, G. J. & Plasterk, R. H. (2001). Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev.* **15**, 2654–2659.

2. Knight, S. W. & Bass, B. L. (2001). A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science*, **293**, 2269–2271.
3. He, L. & Hannon, G. J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**, 522–531.
4. Ji, X. (2006). Structural basis for non-catalytic and catalytic activities of ribonuclease III. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 933–940.
5. MacRae, I. J. & Doudna, J. A. (2007). Ribonuclease revisited: structural insights into ribonuclease III family enzymes. *Curr. Opin. Struct. Biol.* **17**, 138–145.
6. Kufel, J., Dichtl, B. & Tollervey, D. (1999). Yeast Rnt1p is required for cleavage of the pre-ribosomal RNA in the 3′ ETS but not the 5′ ETS. *RNA*, **5**, 909–917.
7. Elela, S. A., Igel, H. & Ares, M., Jr. (1996). RNase III cleaves eukaryotic preribosomal RNA at a U3 snoRNP-dependent site. *Cell*, **85**, 115–124.
8. Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J. & Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol. Cell. Biol.* **25**, 2981–2994.
9. Chanfreau, G., Legrain, P. & Jacquier, A. (1998). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.* **284**, 975–988.
10. Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J. *et al.* (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature*, **425**, 415–419.
11. Chanfreau, G., Elela, S. A., Ares, M., Jr. & Guthrie, C. (1997). Alternative 3′-end processing of U5 snRNA by RNase III. *Genes Dev.* **11**, 2741–2751.
12. Abou Elela, S. & Ares, M., Jr. (1998). Depletion of yeast RNase III blocks correct U2 3′ end formation and results in polyadenylated but functional U2 snRNA. *EMBO J.* **17**, 3738–3746.
13. Seipelt, R. L. & Peterson, M. L. (1995). Alternative processing of IgA pre-mRNA responds like IgM to alterations in the efficiency of the competing splice and cleavage-polyadenylation reactions. *Mol. Immunol.* **32**, 277–285.
14. Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E. & Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399–5410.
15. Lee, C. Y., Lee, A. & Chanfreau, G. (2003). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of *S. cerevisiae* box C/D snoRNAs. *RNA*, **9**, 1362–1370.
16. Petfalski, E., Dandekar, T., Henry, Y. & Tollervey, D. (1998). Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* **18**, 1181–1189.
17. Qu, L. H., Henras, A., Lu, Y. J., Zhou, H., Zhou, W. X., Zhu, Y. Q. *et al.* (1999). Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.* **19**, 1144–1158.
18. Danin-Kreiselman, M., Lee, C. Y. & Chanfreau, G. (2003). RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol. Cell*, **11**, 1279–1289.
19. Egecioglu, D. E., Kawashima, T. R. & Chanfreau, G. F. (2012). Quality control of MATa1 splicing and exon skipping by nuclear RNA degradation. *Nucleic Acids Res.* **40**, 1787–1796.
20. Ghazal, G., Gagnon, J., Jacques, P. E., Landry, J. R., Robert, F. & Elela, S. A. (2009). Yeast RNase III triggers polyadenylation-independent transcription termination. *Mol. Cell*, **36**, 99–109.
21. Rondon, A. G., Mischo, H. E., Kawauchi, J. & Proudfoot, N. J. (2009). Fail-safe transcriptional termination for protein-coding genes in *S. cerevisiae*. *Mol. Cell*, **36**, 88–98.
22. Lavoie, M., Ge, D. & Abou Elela, S. (2012). Regulation of conditional gene expression by coupled transcription repression and RNA degradation. *Nucleic Acids Res.* **40**, 871–883.
23. Ge, D., Lamontagne, B. & Elela, S. A. (2005). RNase III-mediated silencing of a glucose-dependent repressor in yeast. *Curr. Biol.* **15**, 140–145.
24. Zer, C. & Chanfreau, G. (2005). Regulation and surveillance of normal and 3′-extended forms of the yeast aci-reductone dioxygenase mRNA by RNase III cleavage and exonucleolytic degradation. *J. Biol. Chem.* **280**, 28997–29003.
25. Chanfreau, G., Buckle, M. & Jacquier, A. (2000). Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III. *Proc. Natl Acad. Sci. USA*, **97**, 3142–3147.
26. Nagel, R. & Ares, M., Jr. (2000). Substrate recognition by a eukaryotic RNase III: the double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5′-AGNN-3′ tetraloop. *RNA*, **6**, 1142–1156.
27. Drinnenberg, I. A., Weinberg, D. E., Xie, K. T., Mower, J. P., Wolfe, K. H., Fink, G. R. & Bartel, D. P. (2009). RNAi in budding yeast. *Science*, **326**, 544–550.
28. Bernstein, D. A., Vyas, V. K., Weinberg, D. E., Drinnenberg, I. A., Bartel, D. P. & Fink, G. R. (2012). *Candida albicans* Dicer (CaDcr1) is required for efficient ribosomal and spliceosomal RNA maturation. *Proc. Natl Acad. Sci. USA*, **109**, 523–528.
29. Nanduri, S., Carpick, B. W., Yang, Y., Williams, B. R. & Qin, J. (1998). Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. *EMBO J.* **17**, 5458–5465.
30. Ryter, J. M. & Schultz, S. C. (1998). Molecular basis of double-stranded RNA–protein interactions: structure of a dsRNA-binding domain complexed with dsRNA. *EMBO J.* **17**, 7505–7513.
31. Stefl, R., Oberstrass, F. C., Hood, J. L., Jourdan, M., Zimmermann, M., Skrisovska, L. *et al.* (2010). The solution structure of the ADAR2 dsRBM-RNA complex reveals a sequence-specific readout of the minor groove. *Cell*, **143**, 225–237.
32. Mueller, G. A., Miller, M. T., Derose, E. F., Ghosh, M., London, R. E. & Hall, T. M. (2010). Solution structure of the Drosha double-stranded RNA-binding domain. *Silence*, **1**, 2.
33. Wu, H., Henras, A., Chanfreau, G. & Feigon, J. (2004). Structural basis for recognition of the AGNN tetraloop RNA fold by the double-stranded RNA-binding domain of Rnt1p RNase III. *Proc. Natl Acad. Sci. USA*, **101**, 8307–8312.

34. Leulliot, N., Quevillon-Cheruel, S., Graille, M., van Tilbeurgh, H., Leeper, T. C., Godin, K. S. *et al.* (2004). A new alpha-helical extension promotes RNA binding by the dsRBD of Rnt1p RNAse III. *EMBO J.* **23**, 2468–2477.
35. Wang, Z., Hartman, E., Roy, K., Chanfreau, G. & Feigon, J. (2011). Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. *Structure*, **19**, 999–1010.
36. Ghazal, G. & Elela, S. A. (2006). Characterization of the reactivity determinants of a novel hairpin substrate of yeast RNase III. *J. Mol. Biol.* **363**, 332–344.
37. Henzler-Wildman, K. & Kern, D. (2007). Dynamic personalities of proteins. *Nature*, **450**, 964–972.
38. Boehr, D. D., Nussinov, R. & Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796.
39. Massi, F., Grey, M. J. & Palmer, A. G., 3rd. (2005). Microsecond timescale backbone conformational dynamics in ubiquitin studied with NMR R1rho relaxation experiments. *Protein Sci.* **14**, 735–742.
40. Ravindranathan, S., Oberstrass, F. C. & Allain, F. H. (2010). Increase in backbone mobility of the VTS1p-SAM domain on binding to SRE-RNA. *J. Mol. Biol.* **396**, 732–746.
41. Maynard, C. M. & Hall, K. B. (2010). Interactions between PTB RRMs induce slow motions and increase RNA binding affinity. *J. Mol. Biol.* **397**, 260–277.
42. Henras, A. K., Sam, M., Hiley, S. L., Wu, H., Hughes, T. R., Feigon, J. & Chanfreau, G. F. (2005). Biochemical and genomic analysis of substrate recognition by the double-stranded RNA binding domain of yeast RNase III. *RNA*, **11**, 1225–1237.
43. Kharrat, A., Macias, M. J., Gibson, T. J., Nilges, M. & Pastore, A. (1995). Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J.* **14**, 3572–3584.
44. Nanduri, S., Rahman, F., Williams, B. R. & Qin, J. (2000). A dynamically tuned double-stranded RNA binding mechanism for the activation of antiviral kinase PKR. *EMBO J.* **19**, 5567–5574.
45. Arumugam, S., Gao, G., Patton, B. L., Semenchenko, V., Brew, K. & Van Doren, S. R. (2003). Increased backbone mobility in beta-barrel enhances entropy gain driving binding of N-TIMP-1 to MMP-3. *J. Mol. Biol.* **327**, 719–734.
46. Weinberg, D. E., Nakanishi, K., Patel, D. J. & Bartel, D. P. (2011). The inside-out mechanism of Dicers from budding yeasts. *Cell*, **146**, 262–276.
47. Guillerez, J., Lopez, P. J., Proux, F., Launay, H. & Dreyfus, M. (2005). A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl Acad. Sci. USA*, **102**, 5958–5963.
48. Wu, H., Finger, L. D. & Feigon, J. (2005). Structure determination of protein/RNA complexes by NMR. *Methods Enzymol.* **394**, 525–545.
49. Kay, L. E., Xu, G. Y. & Yamazaki, T. (1994). Enhanced-sensitivity triple-resonance spectroscopy with minimal H_2O saturation. *J. Magn. Reson.* **109**, 129–133.
50. Grzesiek, S. & Bax, A. (1993). The importance of not saturating water in protein NMR. Application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594.
51. Schleucher, J., Schwendinger, M., Sattler, M., Schmidt, P., Schedletzky, O., Glaser, S. J. *et al.* (1994). A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. *J. Biomol. NMR*, **4**, 301–306.
52. Ottiger, M., Delaglio, F. & Bax, A. (1998). Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* **131**, 373–378.
53. Ruckert, M. & Otting, G. (2000). Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* **122**, 7793–7797.
54. Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, **13**, 289–302.
55. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. (2003). The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73.
56. Clore, G. M., Gronenborn, A. M. & Tjandra, N. (1998). Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**, 159–162.
57. Koradi, R., Billeter, M. & Wuthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics*, **14**, 29–32.
58. DeLano, W. L. (2002). The PyMOL Molecular Graphics System.
59. Garcia de la Torre, J., Huertas, M. L. & Carrasco, B. (2000). HYDRONMR: prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *J. Magn. Reson.* **147**, 138–146.
60. Bernado, P., Garcia de la Torre, J. & Pons, M. (2002). Interpretation of ^{15}N NMR relaxation data of globular proteins using hydrodynamic calculations with HYDRONMR. *J. Biomol. NMR*, **23**, 139–150.
61. Mandel, A. M., Akke, M. & Palmer, A. G., 3rd. (1995). Backbone dynamics of *Escherichia coli* ribonuclease HI: correlations with structure and function in an active enzyme. *J. Mol. Biol.* **246**, 144–163.
62. Lipari, G. & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559–4570.
63. Lipari, G. & Szabo, A. (1982). Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559.
64. Storici, F., Lewis, L. K. & Resnick, M. A. (2001). In vivo site-directed mutagenesis using oligonucleotides. *Nat. Biotechnol.* **19**, 773–776.
65. Sambrook, J. & Russell, D. W. (2001). Molecular Cloning: A Laboratory Manual, 3rd edit, Cold Spring Harbor Laboratory Press, New York.

CHAPTER 5

Stress-Induced Nuclear RNA Degradation Pathways Regulate Yeast Bromodomain Factor 2 to Promote Cell Survival

OPEN ACCESS Freely available online

PLOS | GENETICS

Stress-Induced Nuclear RNA Degradation Pathways Regulate Yeast Bromodomain Factor 2 to Promote Cell Survival

Kevin Roy, Guillaume Chanfreau*

Department of Chemistry and Biochemistry and the Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California, United States of America

Abstract

Bromodomain proteins are key regulators of gene expression. How the levels of these factors are regulated in specific environmental conditions is unknown. Previous work has established that expression of yeast Bromodomain factor 2 (*BDF2*) is limited by spliceosome-mediated decay (SMD). Here we show that *BDF2* is subject to an additional layer of post-transcriptional control through RNase III-mediated decay (RMD). We found that the yeast RNase III Rnt1p cleaves a stem-loop structure within the *BDF2* mRNA to down-regulate its expression. However, these two nuclear RNA degradation pathways play distinct roles in the regulation of *BDF2* expression, as we show that the RMD and SMD pathways of the *BDF2* mRNA are differentially activated or repressed in specific environmental conditions. RMD is hyper-activated by salt stress and repressed by hydroxyurea-induced DNA damage while SMD is inactivated by salt stress and predominates during DNA damage. Mutations of *cis*-acting signals that control SMD and RMD rescue numerous growth defects of cells lacking Bdf1p, and show that SMD plays an important role in the DNA damage response. These results demonstrate that specific environmental conditions modulate nuclear RNA degradation pathways to control *BDF2* expression and Bdf2p-mediated gene regulation. Moreover, these results show that precise dosage of Bromodomain factors is essential for cell survival in specific environmental conditions, emphasizing their importance for controlling chromatin structure and gene expression in response to environmental stress.

Citation: Roy K, Chanfreau G (2014) Stress-Induced Nuclear RNA Degradation Pathways Regulate Yeast Bromodomain Factor 2 to Promote Cell Survival. PLoS Genet 10(9): e1004661. doi:10.1371/journal.pgen.1004661

Editor: Jeff Coller, Case Western Reserve University, United States of America

Received June 5, 2014; **Accepted** August 11, 2014; **Published** September 18, 2014

Copyright: © 2014 Roy, Chanfreau. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This study was funded by grant GM061518 from the National Institute of General Medical Sciences to GC; KR was supported by a Philip Whitcome Graduate Fellowship and the National Institutes of Health training program T32GM007185. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: guillom@chem.ucla.edu

Introduction

DNA in eukaryotes is wrapped around histone octamers to form nucleosomes [1]. The tails of the histone proteins are subject to a diverse set of chemical modifications, including acetylation, phosphorylation, methylation, and ubiquitination, impacting the majority of DNA-based processes, including transcription, heterochromatin formation, DNA replication, and DNA recombination and repair [2,3]. Non-histone proteins recognize specific tail modifications to mediate the downstream effects [4]. Histone lysine acetylation, one of the best-studied modifications, has important roles in transcription activation, DNA repair and heterochromatin formation [5]. Histone acetylation can increase accessibility of DNA by weakening the interaction between the positively charged histone tail and the nucleosomal DNA [6]. Histone acetylation can also recruit proteins containing bromodomains, which are evolutionarily conserved motifs that recognize acetyl-lysines and play an important role in anchoring chromatin-associated complexes to the nucleosome [7].

S. cerevisiae bromodomain factors 1 and 2 (Bdf1p and Bdf2p) localize throughout the genome at loci enriched for acetylated histones 3 and 4 [5]; [8], where they function in various aspects of transcription initiation and chromatin remodeling [9], as well as protection of euchromatin against heterochromatin spreading [10]. While the *bdf1Δbdf2Δ* double deletion is lethal, both single deletion mutants are viable, indicating that there is at least partial functional redundancy between the two paralogs [11]. In wild-type cells, Bdf1p is present at nearly 3-fold the levels of Bdf2p [12], and each occupies distinct genomic locations [8]. Cells lacking *BDF1* (*bdf1Δ*) show an upregulation of Bdf2 protein (Bdf2p) levels [13] and a redistribution of Bdf2p to the acetylated histones at genomic loci normally bound by Bdf1p [8]. The deletion of *BDF2* affects expression of less than 0.05% of the transcriptome in normal conditions, while the deletion of *BDF1* results in a greater than 2-fold change in the levels of ~15% of all expressed transcripts [10].

In addition to activating TFIID-dependent transcription [11], Bdf1p is part of the SWR1-C chromatin-remodeling complex responsible for replacing histone H2A with the variant H2AZ [14]. The NuA4 histone acetyltransferase acts upstream, depositing an acetyl group on the histone H4, resulting in the recruitment of Bdf1p and SWR1-C [15]. Recent work has revealed that mutations in NuA4 subunits, Bdf1p, SWR1-C, or H2AZ render cells hypersensitive to DNA damage agents, suggesting that histone acetylation-mediated H2AZ deposition plays crucial roles in the

Author Summary

Cells adapt to changes in the environment through modulating gene expression at both the RNA and protein levels. RNA degradation plays a central role in this adaption response, by controlling the stability of specific mRNAs to optimize protein production in different conditions. In this study, we show that the gene encoding Bromodomain Factor 2 (*BDF2*) is tightly regulated according to environmental conditions by two distinct RNA degradation mechanisms. We show that these RNA degradation pathways are critical for cell growth in specific conditions. Our study suggests that environmental modulation of nuclear RNA degradation pathways is a previously unappreciated aspect of gene expression control.

DNA damage response [16]; [17]. Furthermore, cells lacking *BDF1* display a growth defect in normal conditions and hypersensitivity to a wide range of stress conditions [10,16,18]. While cells lacking *BDF2* exhibit no growth defect in normal conditions, they display decreased resistance to the DNA damage agents camptothecin and bleomycin [19]. Additionally, recent studies in fission yeast have demonstrated an important role for Bdf2p in the S-phase stress response and the establishment of heterochromatin boundaries [17,20].

Despite detailed biochemical and functional analysis of the Bdf1p and Bdf2p proteins, little is known about how these factors are regulated in specific environmental conditions. A recent study demonstrated a novel role for the spliceosome in degrading the *BDF2* mRNA, through coupling the splicing of an intron encoded within the *BDF2* ORF to decay by nuclear exonucleases [13]. Spliceosome-mediated decay (SMD) of *BDF2* mRNA was found to be dependent on Bdf1p-mediated recruitment of the spliceosome to the *BDF2* locus, and its proposed biological function was to maintain the homeostasis of Bdf2p levels in normal growth conditions. Here we demonstrate that *BDF2* expression is subject to extensive control by two distinct nuclear RNA degradation pathways, and that each pathway is regulated in different stress conditions to maintain cellular fitness. These results also establish that tight control of the levels of Bromodomain factors is important for cell survival in specific environmental conditions, underscoring their importance for regulating gene expression during stress.

Results

The *S.cerevisiae* RNase III Rnt1p cleaves the *BDF2* mRNA *in vivo* and *in vitro*

Recent studies have revealed the presence of cryptic splicing signals within *S.cerevisiae* transcripts [13,21,22]. In contrast to the canonical role of the spliceosome in promoting gene expression, the usage of these splice sites by the spliceosome promotes degradation of these transcripts [13]. We previously showed that the yeast orthologue of RNase III (Rnt1p) cleaves stem-loop structures within introns of various transcripts to regulate the levels of the spliced transcripts [23]. To test whether these recently identified cryptic introns contain cleavage signals for Rnt1p, we performed a computational RNA secondary structure screen for these intron-containing transcripts using m-Fold [24]. We analyzed the predicted secondary structures for canonical Rnt1p cleavage signals, which consist of double-stranded RNA (dsRNA) hairpin structures capped by NGNN or AAGU tetraloops [25,26]. This analysis revealed the presence of a canonical Rnt1p stem-loop within the open reading frame (ORF) of the yeast gene encoding bromodomain factor 2 (*BDF2*) (Fig. 1A). This 40 base-long stem-loop is situated within the ORF-embedded intron of the *BDF2* mRNA, 1165 nucleotides downstream of the 5′-SS and 298 nucleotides upstream of the annotated 3′-SS ([13], Fig. 1B). The existence of this stem-loop structure was confirmed *in vivo* in a transcriptome-wide study of RNA secondary structure that employed dimethyl sulfate (DMS) modification of single-stranded RNA coupled to deep sequencing [27]. Consistent with its potential degradation by Rnt1p, microarray analysis showed that the *BDF2* transcript is upregulated 2-fold in cells lacking Rnt1p (*rnt1Δ*; [28]). Furthermore, recent studies have found that *BDF2* mRNA exhibits a high transcription rate in conjunction with a short half-life that is similar to the histone mRNAs, suggesting that RNA degradation may play a major role in its regulation ([29], see Discussion). These initial observations suggested that *BDF2* expression is highly regulated at the level of RNA degradation and raised the possibility that Rnt1p may directly control *BDF2* expression.

To investigate if the upregulation of *BDF2* mRNA in *rnt1Δ* is specifically due to the loss of Rnt1p catalytic activity, we compared *BDF2* transcripts levels in wild-type (WT) and *rnt1Δ* strains, and in a strain expressing the catalytically inactive *rnt1 E320K* mutant (Fig. 1C, lanes 1–2, 7–8). This analysis revealed a substantial increase of *BDF2* mRNA upon the loss of Rnt1p catalytic function. To formally demonstrate cleavage of *BDF2* by Rnt1p *in vivo*, we analyzed *BDF2* species in strains carrying mutations in various exonucleases in order to stabilize the 3′ product of Rnt1p cleavage, which is normally subject to decay by 5′-3′ exonucleases due to the presence of an unprotected 5′-end with an exposed monophosphate. The *xrn1Δ* background was used to inactivate the major pathway of cytoplasmic 5′ to 3′ degradation. To inactivate the nuclear 5′ to 3′ exonuclease Rat1p, we used the temperature-sensitive *rat1-1* mutant. In a parallel approach, we deleted the *MET22* gene in the *xrn1Δ* background. Upon a shift to medium lacking methionine, this strain accumulates 3′-phosphoadenosine-5′-phosphate (pAp), a metabolite that inhibits cellular 5′ to 3′ exonuclease activities [30]. We utilized a probe hybridizing to the 3′UTR of *BDF2* mRNA and found that both methods of inactivating the 5′ to 3′ exonucleases revealed a substantial accumulation of a species migrating faster than the full-length *BDF2* mRNA (Fig. 1C, lanes 3,4,9,10). This species was no longer detected upon disruption of *RNT1* in these strains (Fig. 1C, lanes 5,6,11,12) suggesting that it is the downstream (3′) product of Rnt1p cleavage.

We detected a slower migrating form of *BDF2* mRNA in samples from *rnt1* mutant backgrounds (Fig1C, asterisk). Previous studies had identified two different transcription start sites for *BDF2* corresponding to two distinct nucleosome free regions (NFRs) [31–33]. To confirm that two different forms of *BDF2* mRNA differing by their 5′-UTR were present, we performed prolonged electrophoresis in high percentage gels (2% agarose). This enabled the detection of two closely migrating forms of the *BDF2* mRNA (denoted *BDF2-L* and *BDF2-S*), the longer of which was up-regulated in the *met22Δ* background after a shift to medium lacking methionine (Fig. 1D, upper panel, lanes 4,6). This analysis confirmed that *BDF2-S* is the predominant form in wild-type cells under normal conditions. We subsequently used probes targeted to the 5′UTR of the transcript arising from the upstream transcription start site (discussed below), to demonstrate that *BDF2-L* is indeed a 5′-extended form of *BDF2-S*.

We next tested whether Rnt1p could directly cleave the *BDF2* mRNA *in vitro* by incubating recombinant Rnt1p enzyme with total RNAs extracted from a strain expressing the catalytic mutant *rnt1 E320K* (Fig. 1D, lanes 7–8). The 3′-UTR probe detected a

single band from the *in vitro* cleavage reaction, and this band co-migrated with the band stabilized *in vivo* in the *xrn1Δ* background (Fig. 1D, upper panel, lanes 3,4, and 8). Next, we utilized a 5′UTR probe to detect the upstream product of Rnt1p cleavage. This probe was designed to bind predominantly to *BDF2-L*, and with only a short stretch of complementarity to the 5′-end of *BDF2-S*. This resulted in an enhanced signal for *BDF2-L* relative to *BDF2-S*, consistent with *BDF2-L* having a 5′-extension. Furthermore, *in vitro* cleavage of total RNAs by Rnt1p gave rise to two different 5′-products, the longer of which was also observed as an *in vivo* cleavage product in the *xrn1Δmet22Δ* background (Fig. 1D, lower panel, lane 4). This was surprising as the 5′ cleavage product generated by Rnt1p exhibits an exposed 3′-OH, lacks a stabilizing poly(A) tail, and is subject to degradation by 3′ to 5′ exonucleases [23,34]. The stabilization of this species upon depletion of 5′ to 3′ exonuclease activity may be an indirect consequence of 3′ to 5′ exonucleases being saturated by RNA substrates in these conditions. We tested the ability of Rnt1p to cleave the mRNA encoding the *BDF2* paralog (*BDF1*), and other targets of spliceosome-mediated decay [13] identified in our computational screen, but found no evidence for *in vivo* or *in vitro* cleavage of these transcripts (Fig. S1).

To precisely identify the Rnt1p cleavage site on *BDF2* mRNA, we performed primer extension analysis using a primer hybridizing 70 nucleotides downstream of the stem-loop structure. We detected two primer extension stops from *in vitro* cleavage with recombinant Rnt1p (Fig. 1E), corresponding to 14 and 15 bases below the tetraloop for the 5′ and 3′ cleavage sites, respectively (Fig. 1A and 1E, arrows 1 and 2). Primer extension analysis using RNAs extracted from the 5′ to 3′ exonuclease mutants detected the same site *in vivo* for the 5′ cleavage, and a 3′ cleavage site 17 nucleotides below the tetraloop (Fig. 1E lane 3 and 1A, arrow 3). The distances from the cleavage site to the tetraloop are consistent with the known enzymatic properties of the enzyme *in vitro* and *in vivo* ([25] [35]). Importantly, these cleavage sites were no longer detected when Rnt1p was inactivated in the 5′ to 3′ exonuclease mutants (Fig. 1E, lane 4), showing that they are dependent on Rnt1p activity.

To confirm that this stem loop structure is required for Rnt1p cleavage, we constructed a plasmid-borne version of *BDF2* fused at the C-terminus to GFP, and disrupted the top six base pairs with synonymous mutations (Fig. 1A). We transformed plasmids expressing either the wild-type *BDF2* mRNA or the stem mutant into the *xrn1Δ rat1-1* background and tested for the presence of an Rnt1p cleavage product from these constructs. We utilized a GFP probe to avoid hybridization with endogenous *BDF2* mRNA, and detected the Rnt1p degradation intermediate from the WT *BDF2* construct, which was not observed after disruption of the Rnt1p stem loop (stem mutant, Fig. 1F). These results confirmed that Rnt1p degrades *BDF2* mRNA *in vivo* by recognition of a canonical Rnt1p-target stem loop.

Mutation of the Rnt1p-target stem loop upregulates *BDF2-S* and stabilizes the intron-exon 2 degradation intermediate of spliceosome-mediated decay (SMD)

A previous study demonstrated that cytoplasmic nonsense-mediated decay (NMD) and nuclear degradation systems can have partially redundant roles in the degradation of various unspliced pre-mRNAs, such that only upon inactivation of both systems does substantial accumulation of these species occur [36]. We found a slight increase in the *BDF2-L* form upon deletion of the gene coding for the NMD helicase Upf1p (*upf1Δ*; Fig. 1G), consistent with the presence of multiple upstream ORFs in the extended 5′ UTR that would elicit recognition of a premature-termination codon and NMD in accordance with the *faux*-3′ UTR model [37]. To test whether NMD cooperates with SMD or Rnt1p in the regulation of *BDF2* mRNA, we mutated the *BDF2* 5′-splice site (5′-SS) or the Rnt1p-target stem loop at the *BDF2* chromosomal copy in otherwise wild-type or *upf1Δ* backgrounds (Fig. 1G). During the course of this study, RT-PCR analysis of full-length *BDF2* mRNA and its spliced products generated by SMD revealed that mutating the 5′-side of the top six base pairs of the Rnt1p-target stem loop from UUCAUC to CAGCAG inadvertently created a new 3′ splice site, due to the introduction of two YAG sequences proximal to a polypyrimidine tract, $(UUC)_3$ (Fig. S2). As an alternative approach to inactivating Rnt1p cleavage *in cis*, we deleted twelve nucleotides at the top of the stem-loop structure, resulting in the in-frame removal of four codons (UCA-UCU-GAU-GAU) encoding an SSDD amino acid sequence (red letters, Fig. 1A). Both mutations of the Rnt1p stem loop resulted in slightly increased *BDF2-S* mRNA levels (Fig. 1F). These stem-loop mutations did not noticeably affect *BDF2-L* levels in the *upf1Δ* background, suggesting that NMD is the primary degradation system regulating *BDF2-L*. Strikingly, we observed a band migrating faster than *BDF2-S* in the stem-loop mutants. This band was also detectable at low levels in WT cells, but not in the 5′-SS mutant. This band matches the size expected for the linearized intron–exon2 product of SMD, which would be generated by debranching of the lariat intermediate generated by the first splicing step of the *BDF2* mRNA [13,21]. To definitively demonstrate the identity of this band as the intron-exon2 species, we performed a probe-walking experiment with probes hybridizing to exon1, the intron, and exon2 (Fig. S3). Only the intronic and exon2 probes detected this band, while all three probes detected the full-length *BDF2* mRNA (Fig. S3). It is remarkable that the intron-exon2 species is readily detectable in the absence of exonuclease mutations, as it contains an exposed 5′-phosphate and should be subject to rapid 5′ to 3′ decay. The substantial accumulation of this species in stem-loop mutants suggests that a large fraction of *BDF2* mRNA is subject to SMD in the absence of Rnt1p cleavage.

Rrp6 and the nuclear retention factors Mlp1/2 regulate *BDF2-S*

We next tested whether the nuclear exosome co-factor Rrp6p and NMD have functional overlap in regulating the *BDF2* mRNA. We found that while both repress *BDF2-L*, there is no additive increase upon inactivation of both degradation systems (Fig. S4A, right panel). The loss of Rrp6p also resulted in increased *BDF2*-S levels, as well as in stabilization of the SMD and Rnt1p cleavage products (Fig. S4A, left panel). The Mlp proteins have previously been implicated in the nuclear retention of unspliced mRNAs [38]. To test the hypothesis that these proteins mediate the nuclear retention of the unspliced-like full-length *BDF2* mRNA, we inactivated Mlp1p and Mlp2p, both of which function to prevent the export of intron-containing mRNA from the nucleus in a manner that depends on 5′-splice site recognition [38] [36]. The *mlp1Δmlp2Δ* strain exhibited an upregulation of *BDF2-S*, suggesting that the leakage of *BDF2-S* out of the nucleus results in diminished nuclear decay (Fig. S4B). Mlp1/2 had no effect on the NMD-sensitive *BDF2-L* transcript, consistent with more efficient export and cytoplasmic degradation of this species. In the *rrp6Δ* background, the inactivation of Mlp1/2 factors did not affect the levels of full-length *BDF2* mRNA or its degradation intermediates.

Inactivation of Rnt1p-mediated decay (RMD) of *BDF2* rescues the salt sensitivity of the *bdf1Δ* mutant and its inability to grow on non-fermentable carbon

In non-fermentable carbon sources or hyperosmotic stress conditions, *BDF2* levels are limiting for growth in *bdf1Δ* cells,

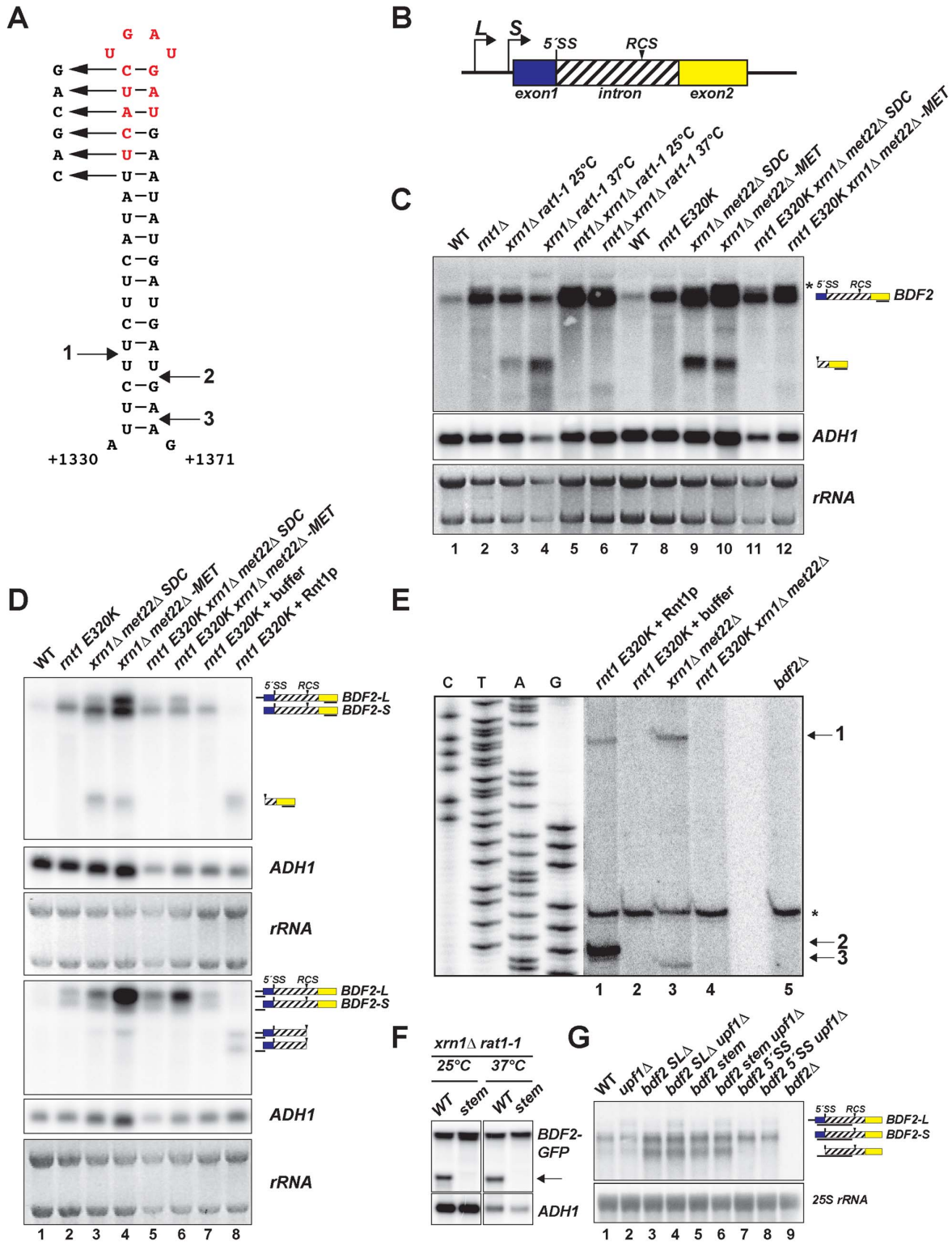


Figure 1. ***BDF2*** **mRNA contains a canonical Rnt1p target stem-loop within its intron and is cleaved by Rnt1p** ***in vivo*** **and** ***in vitro*****. A**. The predicted RNA secondary structure for the *BDF2* transcript is shown for positions +1330 to +1371 of the open reading frame (chr IV: 332354–332395). **B**. Two different transcription start sites give rise to short and long forms of *BDF2* (*-S* and *-L*). The 5′-splice site (5′-SS) and the intronic RNase III cleavage site (RCS) positions are shown relative to exon 1 (blue box) and exon 2 (yellow box). The *BDF2* intron (chr IV: 331189-332695) encodes amino

acids 56–557 of the *BDF2* open reading frame, and is depicted as a hashed box to distinguish it from the typical representation of non-coding introns. **C**. Northern blot detecting the full-length *BDF2* transcript and the Rnt1p degradation product with a riboprobe targeting *BDF2* exon2. The nuclear 5′ to 3′ exonuclease Rat1p was inactivated by shifting the thermo-sensitive *rat1-1* strain from 25°C to the non-permissive temperature (37°C) for 2.5 hours (lanes 3–6), or by shifting the *met22Δ* strain from synthetic-dextrose complete (SDC) media to medium lacking methionine (-MET) for 12 hours (lanes 9–12). *ADH1* and the ethidium bromide-stained rRNAs of the gel prior to transfer are shown as loading controls. **D**. Lanes 1–6 are from (C). RNA from the *rnt1 E320K* catalytic mutant was incubated with the presence of buffer only (lane 7) or with recombinant Rnt1p (lane 8). The same samples were loaded in duplicate series and probed for either exon2 (top panel) or the 5′- UTR region (bottom panel). The black line beneath each *BDF2* RNA species indicates the relative position of the probe. **E**. Primer extension mapping of Rnt1p cleavage sites on *BDF2* mRNA *in vitro* and *in vivo*. A primer was designed to hybridize to the *BDF2* transcript downstream of the Rnt1p cleavage site, at position +1436 to +1457 in the open reading frame (chr IV: 332460-332481). The sequencing ladder and lanes 1–5 represent the alignment from two different exposures of the same gel. RNA from the *rnt1 E320K xrn1Δ met22Δ* strain was incubated in the presence of recombinant Rnt1p (lane 1), or buffer only (lane 2) to demonstrate *in vitro* cleavage sites. Lanes 3 and 4 consist of the *xrn1Δ met22Δ* and *rnt1 E320K xrn1Δ met22Δ* strains shifted to –MET to demonstrate *in vivo* Rnt1p cleavage sites. Lane 5 consists of primer extension on RNA from the *bdf2Δ* strain, demonstrating that the asterisked band present in all lanes is a non-specific primer extension product. The Rnt1p cleavage sites are numbered 1 (*in vivo and in vitro*), 2 (*in vitro*), and 3 (*in vivo*) and are mapped onto the stem-loop in (A). **F**. The *BDF2* ORF was cloned into the pUG23 vector with a C-terminal GFP fusion (depicted as a black box) and transformed into the *xrn1Δ rat1-1* background. The top six base pairs of the Rnt1p target stem-loop were disrupted using synonymous base substitutions as shown in (A). The strains were shifted to the non-permissive temperature as in (C). The *BDF2-GFP* fusion transcript was detected with a probe to the *GFP* ORF. The arrow points to the 3′ Rnt1p cleavage product. **G**. Northern blot analysis of WT *BDF2*, 5′-SS mutant (*bdf2 5′-SS*) and Rnt1 stem-loop mutants (*bdf2 SLΔ* and *bdf2 stem mut*) in the wild-type and *upf1Δ* backgrounds. An exon1-intron riboprobe detects both *-L and -S* forms of *BDF2*, as well the intron-exon2 cleavage product of spliceosome-mediated decay.
doi:10.1371/journal.pgen.1004661.g001

as overexpression of *BDF2* mRNA suppresses *bdf1Δ* growth defects in a dose-dependent manner [18]. The presence of two distinct nuclear degradation systems acting on the *BDF2* mRNA prompted us to investigate the physiological importance of these degradation pathways in these environmental conditions. A previous study found that blocking SMD of *BDF2* mRNA by mutating the 5′-SS improved the growth of *bdf1Δ* in the presence of 0.6 M NaCl [13]. To compare the consequences of SMD inactivation and RMD inactivation on *bdf1Δ* growth phenotypes, we deleted *BDF1* in the *BDF2* 5′-SS mutant (*bdf2 5′SS mut*), the stem-loop mutants (*bdf2 SLΔ* and *bdf2 stem*), and the double mutant (*bdf2 5′SS mut+stem*). In synthetic minimal medium (SDC), but not rich medium (YPD), we found that inactivation of RMD by the stem-loop deletion (*bdf2 SLΔ*) resulted in a modest growth enhancement over wild-type *BDF2* (Fig. 2A, top panel). By contrast, the 5′-SS mutant grew at the same rate as the WT in both of these conditions, consistent with previous observations using plasmid-borne *BDF2* [13].

At 0.2 M and 0.4 M NaCl, the inactivation of RMD resulted in a slight rescue of the *bdf1Δ* growth defect, with SMD inactivation showing no effect (Fig. 2A, top panel). The simultaneous inactivation of SMD and RMD showed no additional effect over the stem loop deletion strain in these conditions. Surprisingly, at 0.6 M NaCl, the inactivation of RMD resulted in a substantial growth rescue (Fig. 2A, middle panel). The stem loop mutation (which incidentally introduced an additional 3′-SS) resulted in a reduced rescue compared to the stem loop deletion. Inactivating the 5′-SS in the stem mutant resulted in a similar rescue to that of the stem loop deletion, suggesting that the presence of the new 3′-SS site enhanced the SMD of *BDF2* mRNA in these conditions. Importantly, the growth rescue observed upon the deletion of the stem-loop confirmed that the removal of the SSDD amino acid sequence from the Bdf2p protein did not perturb protein function, validating the use of this mutant to assay the phenotypic consequences of increased *BDF2* expression due to RMD inactivation.

We next tested different carbon sources, and found that inactivation of RMD, but not SMD, restored the growth of *bdf1Δ* cells on the non-fermentable carbon sources ethanol and glycerol (Fig. 2B, top panel). Furthermore, inactivation of RMD conferred a substantial growth rescue in galactose, for which optimal growth requires simultaneous respiration and fermentation [39]. By contrast, disabling the 5′-SS failed to rescue growth in galactose, and slightly exacerbated the growth defect of *bdf1Δ*. It is possible that regulation by SMD occurs at a different phase of the cell cycle than RMD, and that inactivating SMD results in toxic levels of Bdf2p at a particular stage in the presence of galactose. On the other hand, RMD inactivation might increase levels of Bdf2p at a stage where more Bdf2p is beneficial. Overall, these results suggest that the underlying basis for *bdf1Δ* growth defects in hyper-osmotic stress and respiratory conditions is the degradation of *BDF2* mRNA by RNase III.

Simultaneous inactivation of the RMD and SMD pathways of *BDF2* suppresses the hypersensitivity of the *bdf1Δ* mutant to elevated temperatures, high salt and lithium stress

While the inactivation of *BDF2* RMD rescued *bdf1Δ* growth defects in moderate levels of salt stress, we found that only the simultaneous inactivation of RMD and SMD allowed the *bdf1Δ* strain to grow at elevated temperatures in non-fermentable carbon (34°C and 37°C), or in high salt stress (1.2M NaCl) (Fig. 2B). We also tested lithium chloride stress, as it confers not only osmotic stress but also lithium ion toxicity. Strains lacking Bdf1p are sensitive to 0.1 M lithium chloride and hyper-accumulate lithium ions [40,41]. The simultaneous inactivation of RMD and SMD of *BDF2* mRNA enabled growth of *bdf1Δ* at 0.3 M lithium chloride (Fig. 2B, bottom right). These results suggest that in these conditions, both SMD and RMD play significant roles in limiting *BDF2* expression, as inactivation of either degradation pathway alone is insufficient to enhance the growth of *bdf1Δ* cells. Furthermore, these results suggested that either *bdf1Δ* cells require more *BDF2* mRNA to survive in stress conditions than is required for growth in normal conditions, or that these stress conditions enhance the activity of the nuclear degradation pathways of *BDF2* mRNA, resulting in reduced levels of *BDF2* mRNA relative to normal growth conditions. As cells lacking *BDF1* are strictly dependent on *BDF2* for survival, we hypothesized that stress-induced degradation of *BDF2* mRNA might be responsible for the phenotypes detected in the *bdf1Δ* strain.

RMD limits *BDF2* expression in osmotic stress

The previous growth assays suggest that RMD plays an important role in regulating *BDF2* expression particularly in salt stress conditions. To understand how *BDF2* expression behaves in these conditions, we monitored *BDF2* mRNA levels in wild-type cells after a shift from normal to high salt medium. Within 8 minutes of a shift to high osmolarity, we observed a marked drop

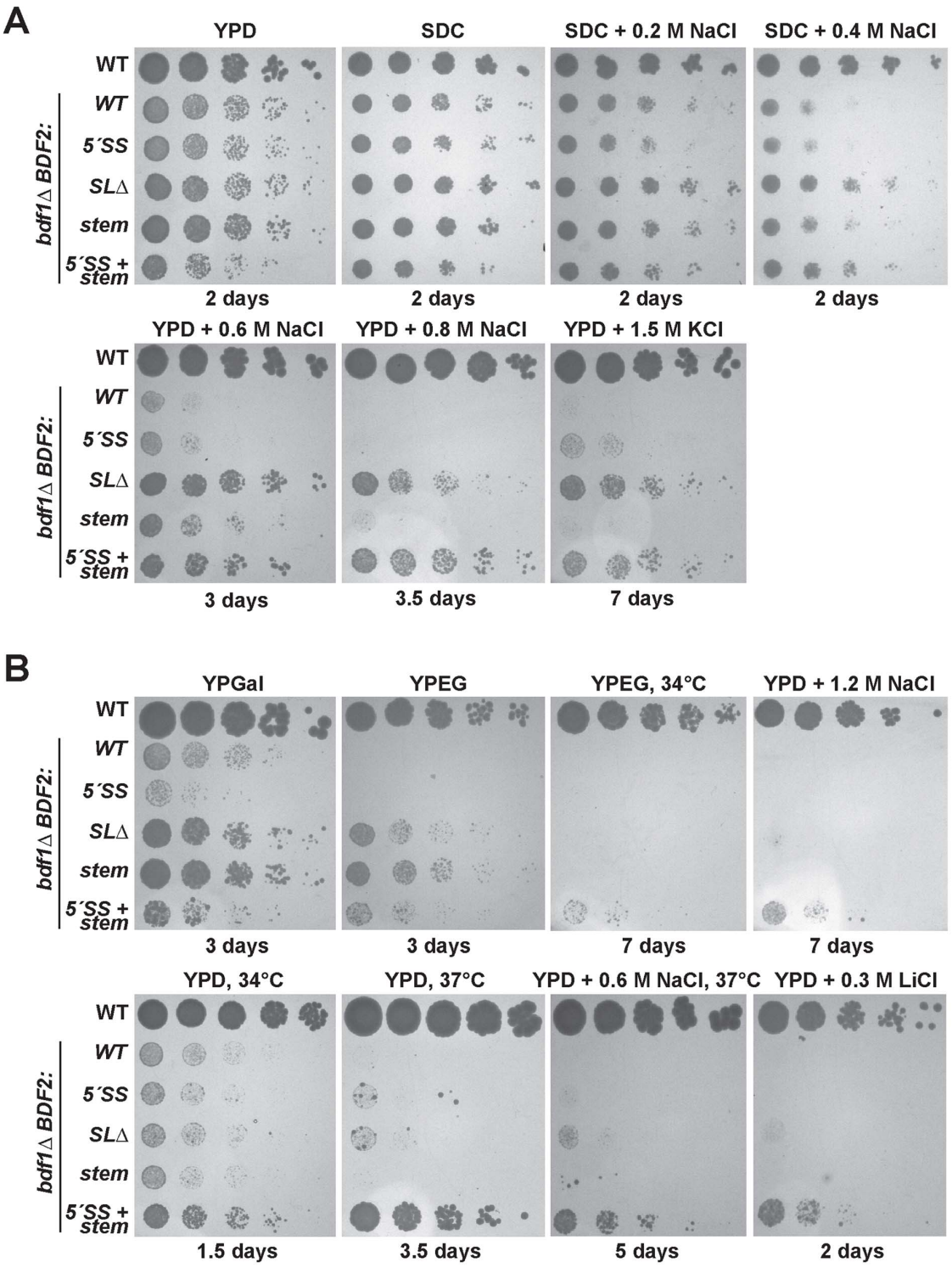


Figure 2. Mutation of the Rnt1p target stem-loop and 5′-splice site suppresses *bdf1Δ* growth defects in various environmental conditions. The top row of spot dilutions in each panel is the wild-type (WT) parent strain. The rest are from the *bdf1Δ* background with either *WT BDF2*, mutations in the 5′-splice site (*5′-SS*), deletion of the Rnt1p stem-loop (*SLΔ*), mutation of the Rnt1p stem-loop (*stem*), or both (*5′-SS+stem*). Unless indicated otherwise, growth was conducted at 30°C for the indicated number of days. **A**. Inactivating Rnt1p cleavage rescues *bdf1Δ* salt sensitivity. The differences between the *BDF2* stem mutation (*stem*) and the deletion of the stem (*SLΔ*) are due to an additional *3′-SS* introduced by the stem mutation. **B**. Inactivating Rnt1p cleavage rescues the inability of *bdf1Δ* to grow on non-fermentable carbon (YPEG) and improves growth on galactose (YPGal). Simultaneous inactivation of both Rnt1p- and spliceosome-mediated decay rescues growth at elevated temperatures and in lithium stress.
doi:10.1371/journal.pgen.1004661.g002

in the levels of the *BDF2* mRNA (Fig. 3A). By contrast, *BDF1* and *RNT1* mRNA levels increased in the first 8 minutes of osmotic shock. After 1 hour, *BDF2* levels were still repressed, while *BDF1* and *RNT1* levels had returned to pre-treatment levels. To assess *BDF2* expression and the levels of its RMD degradation intermediates after prolonged growth in high salt conditions, we monitored the levels of *BDF2* mRNA in wild-type and *rrp6Δ* strains during steady state growth in either normal medium or high salt (Fig. 3B). To test whether the Mlp1/2 factors might mediate increased nuclear retention and RMD of *BDF2* mRNA during osmotic stress, we also examined the *mlp1Δmlp2Δ* and *mlp1Δmlp2Δrrp6Δ* strains. Under normal growth conditions, we observed a stabilization of the 5′ product of RMD in the *rrp6Δ* strain, as expected (Fig. 3B). Remarkably, full-length *BDF2* was undetectable under steady-state growth in high-salt conditions, while the RMD cleavage products persisted. This suggested that transcription of the *BDF2* gene continues after prolonged exposure to osmotic stress, but that the transcripts are continuously degraded by SMD and RMD such that full-length *BDF2* mRNA is no longer detected. Inactivation of the Mlp1/2 factors had no effect on the levels of full-length *BDF2* mRNA or its degradation intermediates, suggesting that the increased targeting of *BDF2* mRNA by Rnt1p is independent of these nuclear retention factors.

To test whether the salt sensitivity of *bdf1Δ* is a consequence of increased RMD of *BDF2* mRNA, we examined *BDF2* mRNA levels in the *bdf1Δ* background after a prolonged shift to high salt conditions. This revealed a significant reduction of *BDF2* mRNA levels for the *WT BDF2* and the 5′-SS mutant after six hours (Fig. 3C). Notably, we detected the intron-exon2 product in *BDF2 WT* and the stem loop deletion strains in normal medium, as observed previously for mutations inactivating RMD in an otherwise wild-type *BDF1* background (Fig. 1F). This demonstrates that in normal growth conditions, spliceosome-mediated decay degrades a large fraction of *BDF2* mRNA in a manner that does not require Bdf1p, contrary to previous observations [13]. By two hours of shift to high salt, this SMD degradation intermediate disappeared and the RMD degradation intermediate was now detected in both the *WT* and the 5′-SS mutants. Strikingly, *BDF2* mRNA levels in the stem-loop deletion strain actually increased during this time course (Fig. 3C). This demonstrates that *BDF2* mRNA repression during osmotic stress is due to increased RMD rather than transcriptional repression. Furthermore, this confirms that *bdf1Δ*-salt sensitivity is a consequence of increased RMD and loss of *BDF2* full-length mRNA in osmotic stress, rather than a specific requirement for Bdf1p in activating osmotic stress response genes.

To examine how RMD affects the regulation of *BDF2* mRNA in the initial phase of the salt stress response, we monitored the levels of full-length *BDF2* mRNA and the SMD and RMD degradation products in the wild-type strain and strains carrying the 5′-SS and stem loop mutations after 20 minutes of osmotic shock (Fig. 3D). In normal medium, we detected the intron-exon2 product resulting from SMD in the stem-loop mutants. Significantly, we detected the accumulation of the RMD cleavage intermediate in WT and the 5′-SS mutant after 20 minutes in high salt (Fig. 3D). Consistent with the results from the *bdf1Δ* background, full-length *BDF2* mRNA increased in the stem-loop mutants, but not the 5′-SS mutant, suggesting that osmotic stress-induced transcripts are normally rapidly degraded by RMD. A previous study had found that the transcription rate of *BDF2* mRNA increases 10 minutes after osmotic shock in 0.4 M NaCl [42]. Our data show that this increase in transcription rate is counter-acted by an even greater increase in *BDF2* mRNA degradation through the RMD pathway.

To confirm that the increased abundance of the RMD cleavage intermediate is not due to a decrease in its degradation rate, we monitored levels of the cleavage product in the *rrp6Δ* strain, as the nuclear exosome is the primary factor involved in the decay of this cleavage intermediate. The RMD cleavage intermediate increased substantially after 20 minutes in osmotic shock in this context, showing that the increased abundance of this cleavage intermediate is due to hyper-activation of RMD (Fig. 3E). We confirmed that the *in vivo* cleavage bands induced in salt stress match the migration of the bands produced by *in vitro* Rnt1p cleavage (Fig. 3F). These results indicate that yeast RNase III plays a hitherto unappreciated role in the regulation of gene expression during osmotic stress.

We next tested the impact of an extended shift in 0.6 M NaCl on *BDF2* mRNA levels, and found a substantial reduction of full-length *BDF2* transcripts by one hour in both the WT and 5′-SS mutant (Fig. 3G). Strikingly, the *bdf2 SLΔ* mutant exhibited an increase in full-length *BDF2* mRNA levels throughout the first hour of high salt exposure, consistent with the increased transcription rate reported previously [42]. Interestingly, the SMD intron-exon2 product, present in *bdf2 SLΔ* in normal conditions, was absent after 15 minutes of salt stress (Fig. 3G). By one hour, this product re-appeared, suggesting that osmotic stress results in a transient deactivation of SMD (Fig. 3G). The dramatic decrease in *BDF2* mRNA levels detected in both the WT and the 5′-SS mutant, but not in the *bdf2 SLΔ* mutant, demonstrates that RMD is the primary mechanism controlling *BDF2* mRNA expression in salt shock conditions. The increase in RMD activity for the *BDF2* mRNA is not simply a consequence of decreased competition by SMD, resulting in increased flux of *BDF2* mRNA through RMD, because the 5′-SS mutation in normal salt conditions does not phenocopy the RMD hyper-activation that occurs during salt stress (Fig. 3G). Indeed, the *BDF2* mRNA harboring the 5′-SS mutation displays a similar profile in salt stress as WT *BDF2*, where by 60 minutes the RMD degradation product is present at higher levels than full-length *BDF2* mRNA (Fig. 3G). Overall, these data confirm that RMD of *BDF2* mRNA is hyper-activated in salt stress.

Next we investigated the impact of *BDF2* RMD on Bdf2p-regulated genes during osmotic stress. A previous study found that the loss of Bdf2p resulted in the upregulation of 20 transcripts at least 2-fold [10]. These transcripts were enriched in stress-responsive genes, including genes involved in carbohydrate metabolism and the heat shock response. We hypothesized that Bdf2p might normally repress these genes, and that an excess of Bdf2p, due to the inactivation of RMD, might inhibit the activation of these transcripts in response to osmotic stress. To test this hypothesis, we monitored the expression of various genes that were reported to be upregulated in *bdf2Δ* (Fig. 3G and S5). We found a significant decrease in the induction of the *GPH1* transcript in *bdf2 SLΔ* mutant after two hours, but not in the WT, *5′-SS* mutant, or *bdf2Δ* strains. *GPH1* encodes glycogen phosphorylase and is critical for preventing the over-accumulation of glycogen during various stress responses and in stationary phase [43]. Our data suggests that during osmotic stress, Rnt1p must repress *BDF2* mRNA levels in order for cells to fully induce *GPH1* expression.

Inactivating the SMD of *BDF2* sensitizes the *bdf1Δ* mutant to DNA replication fork stress-inducing agents hydroxyurea and camptothecin

Cells lacking *BDF1* are hypersensitive to the DNA damage agents methyl methanesulfonate (MMS) and hydroxyurea (HU) [16]; [44]. The broad range of phenotypes rescued by inactivating

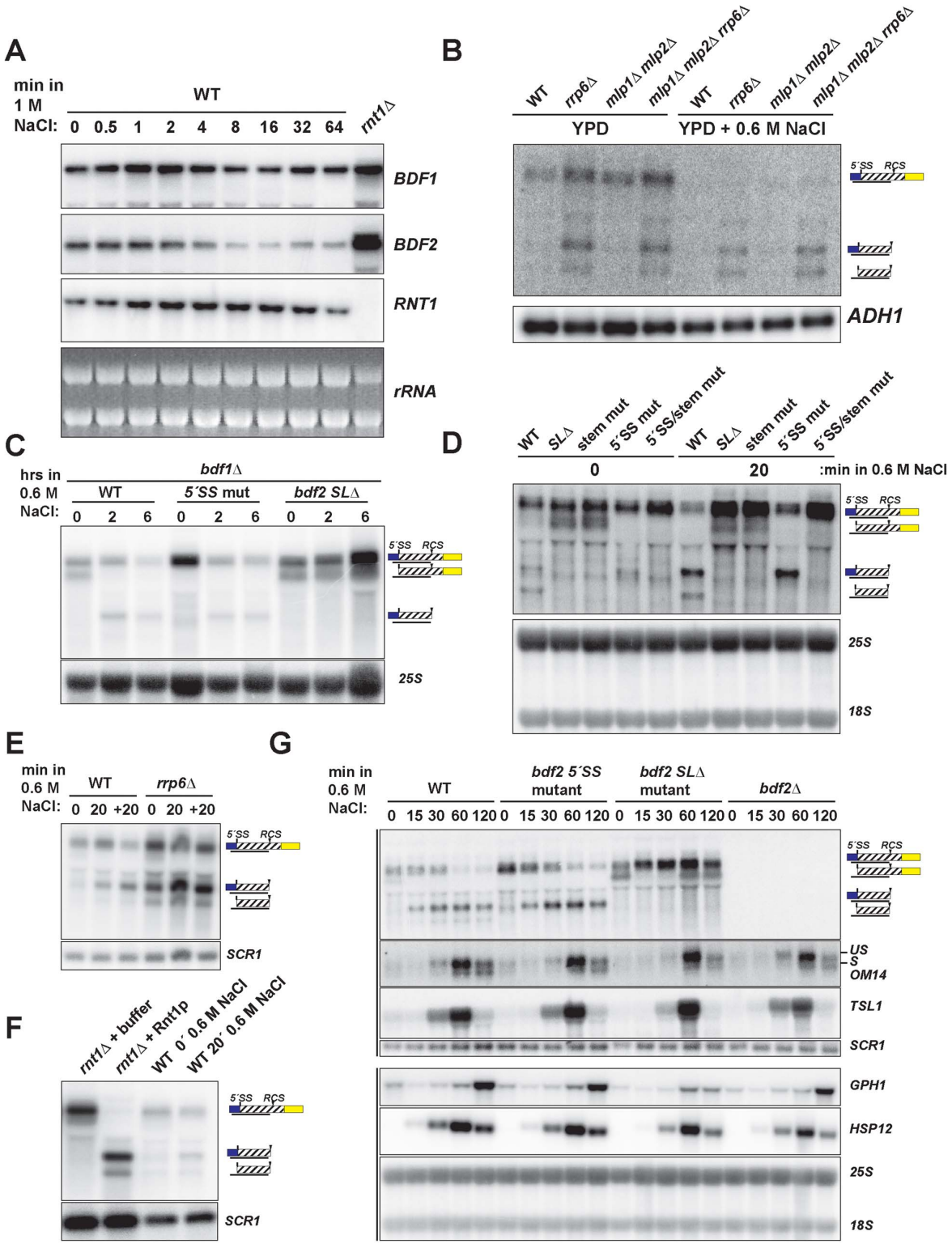


Figure 3. Hyper-activation of Rnt1p cleavage of *BDF2* mRNA in osmotic stress. A. The wild-type strain was exposed to high salt stress (1 M NaCl) and harvested at the indicated time points by dispensing cultures directly into −80°C pre-chilled ethanol. *BDF1* and *BDF2* were detected with riboprobes targeted to their respective 3′-UTRs. *RNT1* was detected with a riboprobe targeted to its open reading frame. A strain deleted for *RNT1* (*rnt1Δ*) is included as a control (far right lane). The ethidium bromide-stained rRNAs are shown as loading controls. **B**. The indicated strains were

grown in either normal conditions (YPD) or high salt conditions (YPD+0.6 M NaCl) under steady state growth. *BDF2* mRNA and its degradation intermediates were detected with a riboprobe targeting the exon1-intron region. *ADH1* is a loading control. **C**. *bdf1Δ* strains carrying either *WT BDF2*, the 5′-SS mutation (*5′-SS mut*), or the deletion of the Rnt1p target stem loop (*bdf2 SLΔ*), were shifted from normal to high salt conditions for two and six hours. The 25S rRNA was detected with an oligonucleotide probe as a loading control. **D**. Wild-type *BDF1* strains with *WT BDF2, bdf2 SLΔ*, the mutation of the stem-loop (*stem mut*), *5′-SS mut*, or both mutations (*5′-SS/stem mut*) were exposed to high salt for 20 minutes. The 25S and 18S rRNAs probed with radiolabeled oligonucleotides are shown as loading controls. **E**. WT and the nuclear exosome co-factor *RRP6* mutant (*rrp6Δ*) were shifted for 20 minutes to 0.6 M NaCl (20), and then shifted back to normal medium for another 20 mintues (+20). *SCR1* is a loading control. **F**. RNA from the *rnt1Δ* strain was incubated with buffer or recombinant Rnt1p to demonstrate that the cleavage products observed in salt stress in the wild-type background co-migrate with the *in vitro* Rnt1p cleavage products. *SCR1* is a loading control. **G**. Wild-type and the specified *BDF2* mutants were shifted to 0.6 M NaCl for the time points indicated. An exon2 probe for the stress-induced, intron-containing *OM14* transcript detects both spliced and unspliced species. *TSL1*, *GPH1*, and *HSP12* were detected with riboprobes targeted to their respective open reading frames. Vertical black lines on the left side indicate panels that were part of the same blot. *SCR1* and the 25S and 18S rRNAs are loading controls.
doi:10.1371/journal.pgen.1004661.g003

RMD and SMD of *BDF2* mRNA led us to expect a rescue of *bdf1Δ* growth defects in DNA damaging conditions as well. Startlingly, inactivation of *BDF2* SMD resulted in enhanced toxicity of HU and camptothecin (CPT) for the *bdf1Δ* strain (Fig. 4A). Both HU and CPT result in the destabilization of replication forks and subsequent induction of double-strand breaks (DSBs) through different mechanisms [45], [46]. Interestingly, the stem loop mutation resulted in a slight rescue of the growth defect conferred by the 5′-SS mutation in 20 μM CPT, and to a lesser extent in 20 mM HU (Fig. 4A). It is possible that SMD and RMD predominate at different stages of the cell cycle, and in the presence of DNA replication stress, additional molecules of Bdf2p gained at one stage by inactivating RMD can compensate for the increased toxicity of excess Bdf2p due to SMD inactivation at another stage. Significantly, these results reveal that SMD plays an important role in the context of DNA replication stress. The lack of an effect from RMD inactivation in 50 mM HU led us to hypothesize that RMD activity on *BDF2* mRNA might be decreased in these conditions, and that the flux of *BDF2* mRNA degradation through SMD or RMD could be dependent upon specific environmental conditions.

SMD predominates over RMD of *BDF2* mRNA in HU-induced DNA damage

The toxic effect of HU observed for the *bdf1Δ* strain harboring the 5′-SS mutation suggested that SMD protects against HU toxicity by preventing an excess accumulation of Bdf2p protein. Furthermore, the lack of toxicity of the stem-loop mutations indicated that SMD of the *BDF2* mRNA might predominate over RMD in DNA damage conditions. Bdf2p protein levels were reported to increase in response to the DNA damage agents MMS and HU [47]. To determine if the increase in protein is due to an increase at the mRNA level, we monitored *BDF2* mRNA expression after exposure to these DNA damage agents. Both treatments resulted in a transient increase in *BDF2* mRNA in the first hour, with slight differences in the kinetics, followed by a decrease to normal levels after two hours (Fig. 4B).

To further investigate how SMD and RMD affect the levels of *BDF2* mRNA in the *bdf1Δ* background, we exposed the 5′-SS and stem-loop mutants to 50 mM HU for two hours. There was no significant change in the levels of full-length *BDF2* mRNA or the intron-exon2 SMD degradation intermediate throughout this time course (Fig. 4C). Nonetheless, the time course demonstrated that SMD of *BDF2* mRNA remains active in the presence of HU. Because of the lack of an effect of the stem mutations on *bdf1Δ* growth in HU, we predicted that Rnt1p cleavage activity might decrease in HU. To test how RMD on *BDF2* mRNA behaves during HU exposure in the wild-type *BDF1* background, we treated wild-type, the 5′-SS mutant, the stem-loop deletion mutant, and *rrp6Δ* strains with 200 mM HU and monitored the expression of *BDF2* mRNA and its degradation products. Consistent with reduced RMD, there was a substantial drop in the levels of the RMD intermediate within 30 minutes of HU treatment in the *rrp6Δ* background (Fig. 4D). Furthermore, the intron-exon2 SMD product, which is normally degraded by Rnt1p due to the presence of the Rnt1p target stem loop in the intron, actually increased in the WT strain throughout the HU treatment. Together these results indicate that RMD of *BDF2* mRNA is inhibited during DNA replication stress, and that the increased flux of *BDF2* mRNA through SMD is important to protect *bdf1Δ* cells from the toxic effects of excess Bdf2p accumulation.

Discussion

BDF2 expression is predominantly regulated at the level of RNA decay

In this study we demonstrate that two major nuclear RNA degradation pathways, Rnt1p- and spliceosome-mediated decay, limit the expression of *BDF2* mRNA, and that *BDF2* expression is tightly regulated at the post-transcriptional level in specific environmental conditions. Strikingly, in normal growth conditions, the *BDF2* gene exhibits a high transcription rate (0.51 RNA molecules per minute per cell) in the top 4.6% of all RNA polymerase II (Pol II)-transcribed genes [29], with an mRNA stability that is among the lowest of all *S.cerevisiae* Pol II transcripts (~5 minute half life, lowest 1.1% [48]). During osmotic stress, a condition where *BDF2* mRNA levels rapidly drop, the Pol II transcription rate across the *BDF2* gene paradoxically increases. However, RNase III-mediated degradation (RMD) over-compensates for the increased transcription, and *BDF2* mRNA levels drop rapidly. Simultaneously, there is a transient decrease in spliceosome-mediated decay (SMD) of *BDF2* mRNA during salt stress. A decrease in spliceosome activity during osmotic stress is consistent with a recent report that intron-containing ribosomal protein pre-mRNAs, the predominant substrates for the spliceosome, accumulate as unspliced transcripts in osmotic stress conditions when their degradation by nonsense-mediated decay is inactivated [49]. Therefore, both nuclear and cytoplasmic RNA degradation systems have evolved to degrade intron-containing transcripts that accumulate during osmotic stress. By contrast, SMD of *BDF2* mRNA plays a critical role during DNA replication stress, as evidenced by the toxicity of the 5′-SS mutation to cells lacking *BDF1* when grown in the presence of DNA double-strand break agents (Fig. 4A). Conversely, we found that RMD of *BDF2* mRNA is diminished during hydroxyurea exposure and its inactivation confers no phenotype towards *bdf1Δ* cells in HU or CPT (Fig. 4A,D). Interestingly, there was not a significant effect of the 5′-SS mutation on *BDF2* mRNA levels in HU during the time window we examined (Fig. 4C,D). It is possible that an effect of HU on *BDF2* mRNA levels in the 5′SS mutant requires longer exposure times to HU. It is also possible that the 5′SS mutation results in enhanced leakage of the "unspliced-like" *BDF2* mRNA

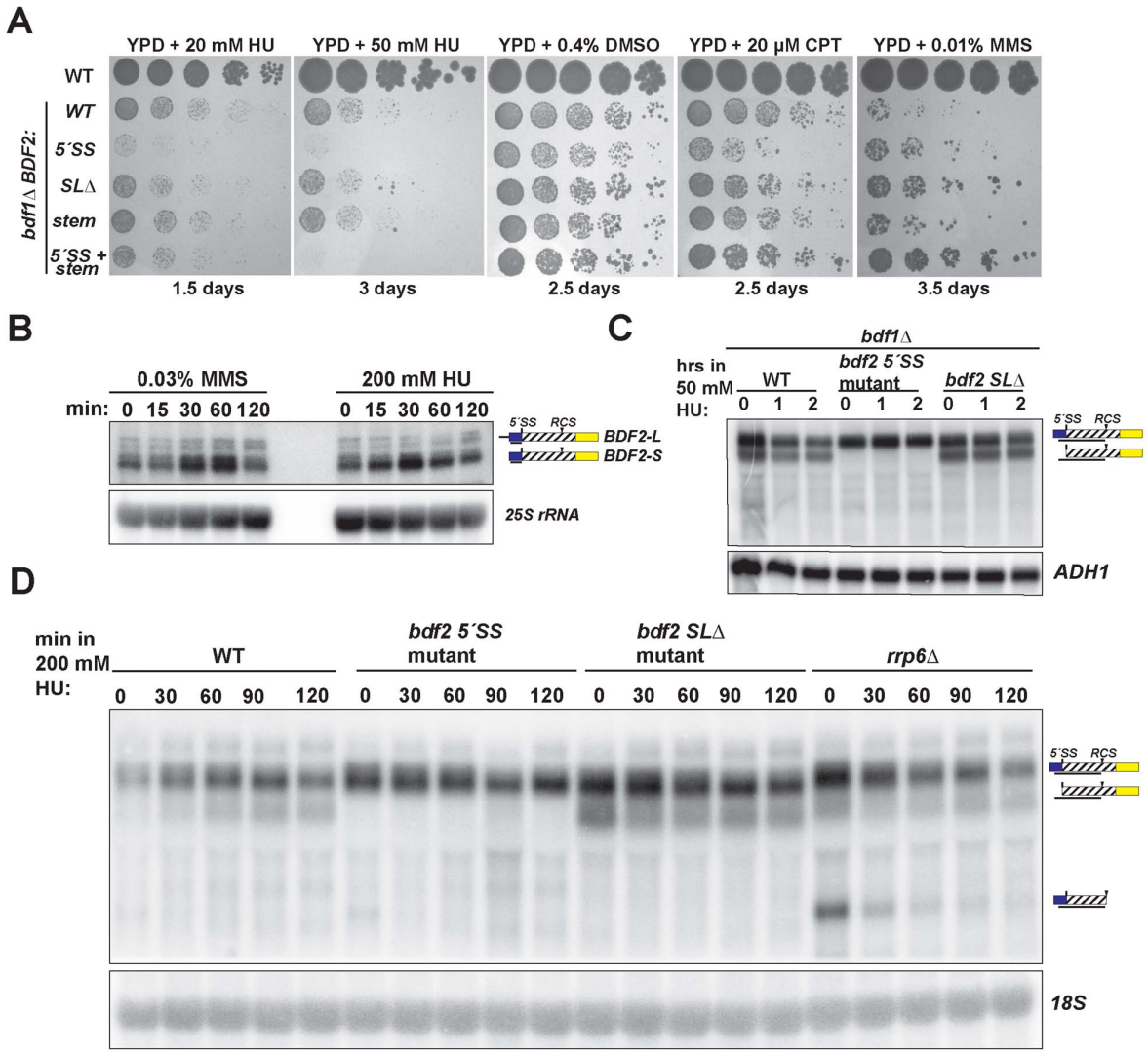


Figure 4. Spliceosome-mediated decay of *BDF2* mRNA predominates over Rnt1p cleavage during DNA replication stress. A. The set of strains from Figure 2 were grown on media containing the indicated concentrations of hydroxyurea (HU), dimethyl sulfoxide (DMSO), which was the vehicle control for camptothecin (CPT), and methyl methanesulfonate (MMS). **B**. The wild-type strain was incubated with the indicated concentrations of MMS or HU for the indicated time points. *BDF2* mRNA was detected with a riboprobe targeting exon1. The 25S rRNA is a loading control. **C**. *bdf1Δ* strains carrying either *WT BDF2*, the 5′-SS mutation (*5′-SS mut*), or the deletion of the Rnt1p target stem loop (*bdf2 SLΔ*), were exposed to 50 mM HU for 1 and 2 hours. The *BDF2* full-length mRNA and SMD intron-exon2 degradation intermediates were detected with an exon1-intron riboprobe. *ADH1* is a loading control. **D**. The indicated strains were exposed to 200 mM HU for the specified time points. *BDF2* mRNA and the SMD and RMD degradation intermediates were detected with same probe as (C). The 18S rRNA was probed with a radiolabeled oligonucleotide as a loading control.
doi:10.1371/journal.pgen.1004661.g004

to the cytoplasm and thus increased levels of Bdf2p translation, as the nuclear retention of intron-containing transcripts by Mlp1p is known to require an intact 5′-splice site [38]. This would be consistent with our finding that Mlp1/2p proteins regulate *BDF2* mRNA levels (Fig. S4B).

Our results reveal that the relative flux of *BDF2* mRNA degradation through RMD or SMD is dependent on the particular environmental condition, such that when one degradation system is inactivated, the other system predominates (Fig. 5). In particular, we found that levels of the SMD degradation intermediate decrease in salt stress, whereas the RMD intermediate increases in abundance. During DNA replication stress, the RMD intermediate decreases while the SMD degradation intermediate increases. These molecular phenotypes are consistent with the growth phenotypes of the *BDF2* 5′-SS and *BDF2* stem loop mutations in the *bdf1Δ* background in these stress conditions. It appears that the *BDF2* gene has evolved to retain the degradation signals for both systems, such that its tight post-transcriptional control is maintained in conditions that block either RMD or SMD. Therefore, the presence of two independent nuclear degradation pathways acting on the *BDF2* mRNA guards against excess accumulation of *BDF2* mRNA across a range of stress conditions.

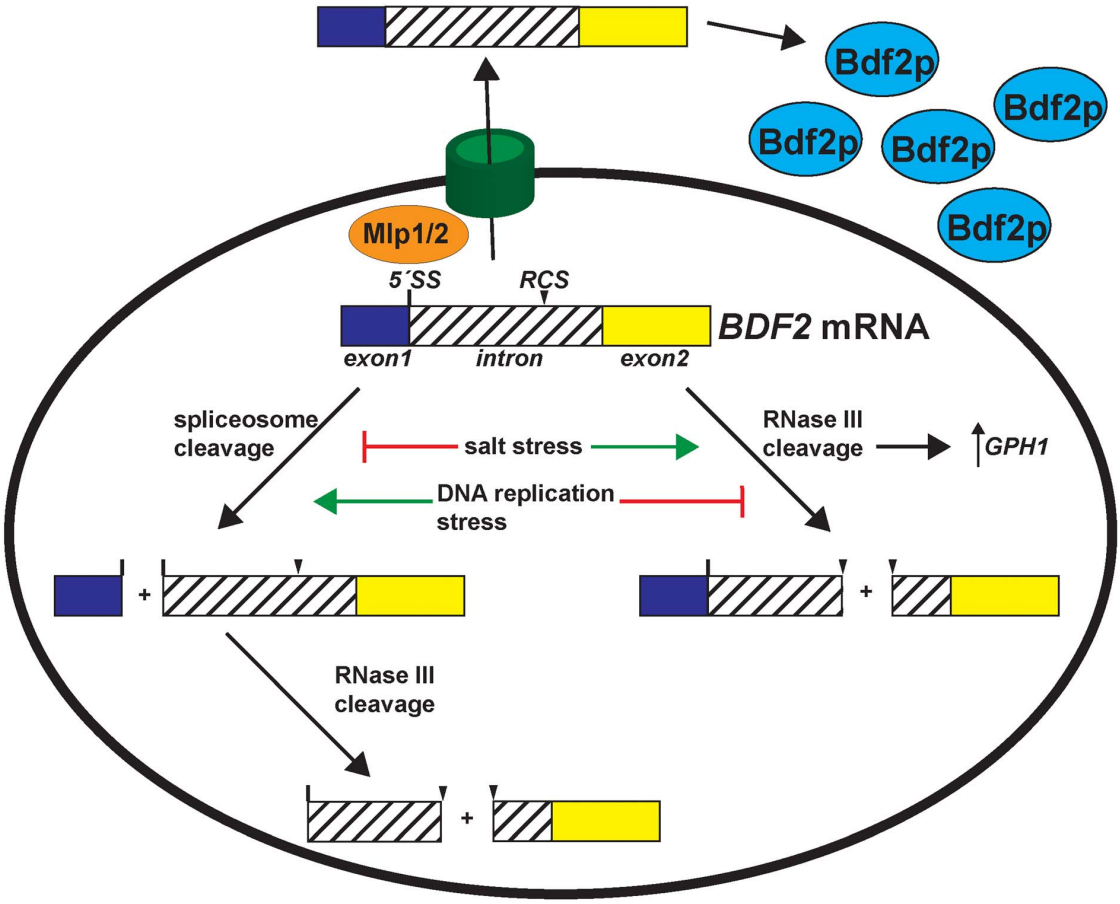


Figure 5. Environmental stress conditions control the expression of Bromodomain Factor 2 mRNA through modulating RNase III and spliceosome-mediated decay. Salt stress transiently deactivates spliceosome-mediated decay and hyper-activates RNase III cleavage of *BDF2*, which is required for the full induction of the stress-induced *GPH1* transcript. DNA replication stress de-activates RNase III cleavage of *BDF2* mRNA, while spliceosome-mediated decay persists in these conditions to repress *BDF2* expression. In addition, the Mlp1/2 nuclear retention factors repress *BDF2* expression. *BDF2* mRNAs that escape nuclear surveillance are exported and translated.
doi:10.1371/journal.pgen.1004661.g005

The observation that *BDF2* mRNA is detectable in WT cells in normal growth conditions suggests that a substantial fraction of *BDF2* mRNA normally escapes nuclear degradation by SMD and RMD. In salt stress, an increase in the absolute activity of RMD leads to a substantial increase in the overall fraction of full-length *BDF2* mRNA subject to nuclear degradation. Furthermore, inhibiting both RMD and SMD has an additive effect compared to inhibiting either pathway alone in specific conditions (Fig. 1D and Fig. 2B), demonstrating that the pathways are non-redundant and independently regulate *BDF2* mRNA in a condition-specific manner (Fig. 5).

RMD and SMD modulate Bromodomain factor 2 expression to maintain cellular fitness in response to adverse environmental conditions

Despite detailed biochemical work on the binding affinities of Bdf1p and Bdf2p for different acetylated histone tail peptides [50], and extensive chromatin immunoprecipitation studies mapping genomic binding sites [8], it is largely unknown how Bdf1p and Bdf2p achieve their specificity *in vivo* and how these factors are regulated in different conditions. We found that during osmotic stress, degradation of *BDF2* mRNA is required for the full induction of glycogen phosphorylase (*GPH1*, Fig. 3F), suggesting that Bdf2p may normally repress a subset of salt-stress induced genes. We also found that inactivating SMD or RMD in a wild-type background resulted in sensitivity to lithium chloride (Fig. S6), suggesting that decreasing *BDF2* mRNA levels via nuclear RNA decay constitutes an important mechanism for adapting gene expression to lithium ion stress.

The regulation of *BDF2* expression by these two RNA degradation mechanisms controls highly specific phenotypes. Strikingly, the inactivation of SMD, but not RMD, sensitized *bdf1Δ* cells to hydroxyurea (HU) and camptothecin (CPT), but not methyl methanesulfonate (MMS) (Fig. 4A). While all three agents stall replication fork progression, only HU and CPT induce *in vivo* double-strand DNA breaks (DSBs) [51]. This suggests that an excess of Bdf2p in the absence of Bdf1p is toxic specifically in the presence of DSBs. This finding is consistent with a report that the fission yeast homolog of *BDF2* is toxic to cells with unstable replication forks, as the deletion of *BDF2* suppressed the HU sensitivity of numerous S-phase checkpoint mutants [17]. It remains to be determined whether the toxicity of Bdf2p is due to a

direct effect on chromatin structure, or rather due to its effects on gene expression in DNA damage conditions. Notably, the human homolog of Bdf2p, *BRD4*, has been shown to activate cell death in response to DNA damage by promoting chromatin condensation [52]. It is therefore possible that at least some of the functions of Bdf2p in the DNA damage response have been conserved from yeast to humans. Significantly, we found that *BDF2* is required for resistance of wild-type cells to DMSO (Fig. S6), which has been implicated in causing DNA damage by an unknown mechanism [53]. Therefore, Bdf2p is beneficial in specific conditions, and can play antagonizing roles in the DNA damage response depending on the specific type of DNA damage.

Overall, our study highlights how environmental stress conditions can differentially control RNA degradation systems to regulate gene expression (Fig. 5), and emphasizes the importance of modulating bromodomain factors expression in the cellular response to environmental changes. The ability of the cell to degrade or stabilize specific transcripts by environmentally-controlled RNA degradation pathways mirrors the control that transcriptional activators and repressors can provide in response to stress conditions. We speculate that the presence of SMD and/or RMD degradation signals in other mRNAs may allow for broad post-transcriptional fine-tuning of gene expression in conditions where the activity of these degradation systems is modulated. Future work will unravel precisely what mechanisms regulate RNase III and spliceosome-mediated decay in different environmental conditions, as well as uncover the repertoire of transcripts targeted by these pathways.

Materials and Methods

RNA secondary structure screen

Transcripts targeted by spliceosome-mediated decay were obtained from Supplementary Table 2 from [13]. The 5′ and 3′ UTRs for each transcript were obtained from [31] from *Saccharomyces Genome Database*. If the transcript did not have an annotated UTR, the open reading frame start or stop coordinates were used. Each sequence was first split into 200 base fragments in 100 base overlapping steps. mFold [24] was used to predict the fold of each 200 base sequence, and sequences capable of adopting stem-loop structures with an NGNN or AAGU tetraloop were identified for further study.

Strains

The BMA64 (mating type a) background was used to generate the mutant strains used in this study (Table S1). The indicated deletion mutants were constructed by the lithium acetate/PEG method [54] by replacing the open reading frames with PCR products encoding either *S. cerevisiae TRP1, S. kluyveri HIS5*, or the *KANMX6* gene [55,56]. Strains containing the *rat1-1* mutation were previously described [28]. Mutations were introduced into the endogenous *BDF2* gene by the *delitto perfetto* method [57]. First, the CORE cassette encoding the *URA3* and *KANMX6* genes was inserted at Rnt1p-target stem loop (chrIV:332367-80). G418 resistant colonies were screened for successful integration of the CORE cassette by PCR. For the second step, we used the pUG23-*BDF2* plasmid (described below), containing either the WT or various *BDF2* mutations, as a template to generate high-fidelity PCR products (NEB Phusion). Primers were designed so that the resulting PCR products contained >100 bases of homology flanking the mutation sites (Table S1). 200 µl of each PCR product were precipitated in 1 ml ethanol and 40 µl 3 M sodium acetate pH 5.2. Pellets were washed in 70% ethanol and resuspended in 34 µl water and used for excision of the CORE cassette by standard yeast transformation techniques [54]. Transformations were plated overnight on YPD, and then replica-plated to 5-FOA to select for loss of *URA3*. Colonies were screened for sensitivity to G418 to confirm loss of the CORE cassette, and mutations were confirmed by PCR amplification of the *BDF2* locus followed by sequencing (Laragen, Inc).

Plasmids

The *BDF2* open reading frame without the stop codon was amplified by high-fidelity PCR (NEB Phusion) from WT genomic DNA with the forward primer containing the SpeI restriction site and the reverse primer containing the ClaI restriction site (Table S1). The insert was cloned into the pUG23 vector and 5′-splice site and stem-loop mutations were constructed with the Quick-Change site-directed mutagenesis kit (Agilent Technologies, Inc). All plasmids and mutations were confirmed by sequencing (Laragen, Inc.).

Media and culturing

Unless otherwise indicated, cultures were grown in YPD (1% yeast extract, 2% bacto-peptone, 2% dextrose) at 30°C at 200 rpm. SDC (.67% yeast nitrogen base with ammonium sulfate, 2% dextrose, and 0.2% complete amino acid mixture) was used to culture strains containing the *met22Δ* deletion, and where indicated these cultures were pelleted, washed once with SD medium lacking methionine (SD-MET), and resuspended in SD-MET for 12 hours. Strains containing the pUG23 vector were grown in SD-URA-MET. Strains containing the *rat1-1* mutation were grown at 25°C and shifted to pre-warmed flasks containing the same medium at 37°C. All cultures were harvested by centrifugation at 4000 rpm for 2 minutes, washed in deionized water, and spun down in microcentrifuge tubes. The supernatant was removed and pellets were flash-frozen in liquid nitrogen and stored at −80°C.

RNA extraction and northern blot analysis

RNA extractions were performed by the addition of 400 µl acid-washed glass beads, 500 µl phenol:chloroform:isoamyl alcohol (25:24:1, pH 6.7, Fisher Scientific), and 500 µl of RNA-SDS buffer (50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM EDTA, 2% SDS) to the frozen pellets. Samples were vortexed for 1 minute, heated in a 65°C water bath for 5 minutes, and vortexed for an additional minute. Samples were spun at 13,200 rpm for 5 minutes, and 450 µl of the supernatant was transferred to a new tube containing 450 µl of phenol:chloroform:isoamyl alcohol. Samples were vortexed for 1 minute, and spun at 15,000 rpm for 5 minutes. 400 µl of supernatant were precipitated in 1 ml EtOH and 40 µl of 3M NaOAc pH 5.2, and incubated at −80°C for 30 minutes. Samples were spun at 15,000 rpm for 5 minutes, and pellets were washed with 70% EtOH and resuspended in nuclease-free water (Ambion). *In vitro* cleavage of 50 µg total RNA with 4 pmol recombinant Rnt1p was performed at 30°C for 20 minutes as described [58,59]. 1 volume of RNA was denatured in 5 volumes of glyoxal buffer [60% DMSO (Sigma-Aldrich), 8% glyoxal w/v (Sigma-Aldrich), 5% glycerol, 40 µg/ml ethidium bromide, 1× BPTE pH 6.5 (10 mM PIPES (Sigma-Aldrich), 30 mM Bis-Tris (Sigma-Aldrich), 10 mM EDTA pH 8.0)]. Samples were vortexed and denatured at 55°C for one hour. Samples were chilled on ice for 10 minutes, and 10× BPTE-RNA loading dye (0.25% xylene cyanol, 0.25% bromophenol blue, 5% glycerol, 1× BPTE) was added to each sample before loading on 2% agarose gels with 1× BPTE buffer. Electrophoresis was performed at 4 V/cm while mixing the buffer

with magnetic stir bars. Gels were visualized and washed for 10 minutes in deionized water, 20 minutes in 75 mM NaOH, an additional 10 minutes in deionized water, and 10 minutes in 10× SSPE (1.5 M NaCl, 100 mM sodium phosphate, 100 mM EDTA, pH 7.4). Gels were passively transferred in 10× SSPE to positively charged nylon membranes (Hybond-N^+, GE Healthcare Life Sciences). Blots were cross-linked with the auto cross-link function (120,000 joules/cm^2) on the Stratalinker UV Crosslinker 2400 (Stratagene), and stored in 2× SSPE at 4°C. RNA probes were constructed by *in vitro* transcription with the T3 RNA polymerase (Promega) and α-^{32}P- UTP (Perkin-Elmer) following the manufacturer's protocol, with the exception that radioactive UTP was used instead of radioactive CTP. Templates were constructed by PCR using genomic DNA and primers hybridizing to the target gene (Table S1). The reverse primer contained at its 5′-end the T3 promoter sequence (5′-AATTAACCCTCACTAAAGGGA-3). Blots were pre-hybridized in Church's buffer (1% BSA, 1 mM EDTA, 0.5 M $NaPO_4$ pH 7.2, 7% SDS) at 65°C for 1 hour, and transcription reactions were diluted in 100 µl of deionized water and added directly to the hybridization bottles. Hybridizations were conducted overnight at 65°C, and blots were washed for 20 minutes with 2× SSPE+0.1% SDS at 65°C, followed by two washes for 20 minutes with 0.1× SSPE+0.1% SDS at 65°C. Blots were exposed to K-screens (Kodak) and scanned with the BioRad FX Phosphorimager.

Primer extension mapping of the Rnt1p cleavage site

20 pmol of *BDF2* downstream RCS reverse primer (5′-CAATTGTTGCAATTCAACCTCC-3′) was labeled with γ-^{32}P-ATP with T4 polynucleotide kinase (NEB) according to the manufacturer's protocol. Primer extension was performed using the SuperScript III reverse transcriptase (Invitrogen/Life Technologies) according to manufacturer's protocol, except that denaturation was performed at 75°C for 4 minutes in a thermocycler, followed by primer annealing at 45°C for 5 minutes. Primer extension was performed at 55°C for 60 minutes, followed by 70°C for 15 minutes to inactivate the enzyme. An equal volume of 2× formamide loading dye (Ambion) was added to the primer extension reactions, heated at 75°C for 3 minutes, and 10 µl was loaded onto 8% poly-acrylamide sequencing gels. To obtain the sequencing ladder, a PCR product was first generated using the reverse primer and the *BDF2* exon1 forward primer (5′-GCACATTCTGCTTTACTGGCAGC-3′) and pUG23-*BDF2* as template. The PCR product was used as a template with the Thermo Sequenase Cycle Sequencing Kit using 0.5 picomoles of labeled primer for each of the 4 dideoxynucleotide sequencing reactions (according to manufacturer's protocol). Gels were dried at 80°C for one hour on a vacuum gel drier (Amersham Pharmacia), and exposed to K-screens (Kodak) and scanned with the BioRad FX Phosphorimager.

RT-PCR analysis

40 µg total RNA of total RNA was treated with DNase I (Ambion) according to manufacturer's protocol. 5 µg of DNase I-treated RNA was reverse-transcribed with random hexamers and M-MLV reverse transcriptase (Invitrogen) according to manufacturer's protocol. 2 µl of cDNA was used as a template in standard 50 µl PCR with the forward primer 5′-AGCAGCGAGAGTAGTAGTAACAAAAAC-3′ and either reverse primer R1 5′-ATCTCTCATAAATTCTTG-CAATAGTCG-3′ or R2 5′-TGCTTCATTGCCTTCCTACG-3′. The following thermo-cycler parameters were used: 94°C 2 min, 35 cycles of 94°C 30 s, 60°C 30 s, 72°C 60 s. RT-PCR products were analyzed on 1× TAE 2% agarose gels and stained with ethidium bromide. The identity of each band was verified by gel extraction of the excised band followed by Sanger sequencing (Laragen, Inc.).

Plate growth assays

Cultures were grown to mid-log phase ($OD_{600\ nm}$ 0.4–0.6) and 1 ml of each culture was spun at 13,200 rpm for 30 seconds. Pellets were resuspended in 1 ml of sterile water, and diluted to $OD_{600\ nm}$ 0.1 followed by four 5-fold serial dilutions in sterile water. 4 µl of each dilution was spotted onto the indicated media. Hydroxyurea was purchased from Alfa-Aesar, and camptothecin and methyl methanesulfonate were purchased from Sigma-Aldrich.

Supporting Information

Figure S1 Northern blot analysis of genes identified as targets of spliceosome-mediated decay (Supp. Table 2 from ([13]) containing RNA sequences capable of folding into canonical Rnt1p target stem-loops. Riboprobes for the indicated genes were designed to target the open reading frames. *SCR1* is shown as a loading control. The *NPL3* readthrough transcript is a positive control for a known Rnt1p mRNA target [60,61].
(TIF)

Figure S2 RT-PCR analysis of full-length *BDF2* mRNA and spliced products generated by spliceosome-mediated decay in wild-type, the strain carrying the disruption of the Rnt1p stem loop (*bdf2 stem mutant*), and the deletion of nuclear exosome co-factor *RRP6* (*rrp6Δ*). (A) The locations of the forward (F) and reverse primers (R1 and R2) are shown relative to the Rnt1p cleavage site (*RCS*), the previously annotated AAG 3′-splice site at +1672 into the ORF [13], and an alternative AAG 3′-splice site at +1595 (identified here by sequencing). The stem mutant introduces two 3′-splice site CAG motifs into the Rnt1p target stem-loop downstream of the $(UUC)_3$ polypyrimidine tract. (B) RT-PCR was performed with the same forward primer and either R1 (top panel) or R2 (bottom panel). The spliced species corresponding to each band is indicated to the right of each band. The band denoted with an asterisk in both panels arises from mis-priming of the F primer 3′end at +1426 in the *BDF2* ORF.
(TIF)

Figure S3 Probe walking to verify the identity of the *BDF2* mRNA degradation intermediates generated by spliceosome-mediated decay and Rnt1p. RNA from the indicated strains was analyzed on the same gel in a triplicate series and transferred to three strips of membranes. Riboprobes were designed to hybridize to exon1 (probe I), the intron (probe II), or exon2 (probe III) of *BDF2* mRNA and blots were aligned during the exposure to compare the migration of each *BDF2* species. The band labeled with the asterisk on the exon1 blot is due to cross-hybridization with the 18S rRNA.
(TIF)

Figure S4 The nuclear exosome co-factor *RRP6*, the nonsense mediated decay (NMD) factor *UPF1*, and the pre-mRNA nuclear retention factors *MLP1/2* participate in the surveillance of *BDF2* transcripts. (**A**) A riboprobe for the *BDF2* exon1 detects the 5′ products of Rnt1p and spliceosome endonucleolytic cleavage (left panel). A riboprobe was designed to hybridize to the 5′ UTR of *BDF2-L*, upstream of the annotated transcription start site for *BDF2-S* (chrIV: 331002), in order to exclusively detect *BDF2-L* and its RMD degradation intermediate (right panel). (**B**) A riboprobe spanning exon 1 and the intronic region up to the Rnt1p cleavage site (*RCS*) (+23 to +1303 into the ORF) detects

full-length *BDF2* as well as Rnt1p cleavage products (top half). The *BDF2-L* full-length transcript and its RMD product were detected with the 5′-UTR probe that does not bind *BDF2-S* (bottom half).
(TIF)

Figure S5 Wild-type and the specified *BDF2* mutants were shifted to 0.6 M NaCl for the time points indicated. *GLK1, HAL5, GRE3*, and *LEU1* were detected with riboprobes targeting the open reading frames of their respective transcripts. SCR1 is a loading control.
(TIF)

Figure S6 The wild-type strain and the specified *BDF2* mutants were grown on plates with the indicated media at 30°C for the indicated number of days. (**A**) Inactivation of RMD or SMD of *BDF2* mRNA has no effect on the growth of the wild-type background in high salt or hyperosmotic stress conditions. (**B**) Inactivation of RMD or SMD of *BDF2* mRNA confers sensitivity to 0.6 M lithium chloride, but not methyl methanesulfonate (MMS) or hydroxyurea (HU). (**C**) Cells lacking *BDF2* are hypersensitive to 2% dimethyl sulfoxide (DMSO), and do not grow on the combination of 2% DMSO and 100 μM camptothecin (CPT).
(TIF)

Table S1 *Saccharomyces cerevisiae* strains, plasmids, and oligonucleotides used in this study. The strain ID, genotype and reference is provided for each strain studied. The plasmids are generated from the pUG23 vector, and contain the described mutations in the *BDF2* ORF. The oligonucleotide sequences used to generate T3 riboprobe templates or radiolabeled oligonucleotide probes are indicated, with a reference to the relevant figures. The oligonucleotides used to generate the plasmids and yeast strains are indicated along with the purpose of each primer pair.
(XLSX)

Acknowledgments

We thank Rebekah Liu and Douglas Zhang for extensive technical assistance, and Jason Gabunilas for helpful comments and critical reading of the manuscript.

Author Contributions

Conceived and designed the experiments: KR GC. Performed the experiments: KR. Analyzed the data: KR GC. Contributed to the writing of the manuscript: KR GC.

References

1. Kornberg RD (1974) Chromatin structure: a repeating unit of histones and DNA. Science 184: 868–871.
2. Kornberg RD, Lorch Y (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell 98: 285–294.
3. Wu J, Grunstein M (2000) 25 years after the nucleosome model: chromatin modifications. Trends Biochem Sci 25: 619–623.
4. Strahl BD, Allis CD (2000) The language of covalent histone modifications. Nature 403: 41–45. doi:10.1038/47412.
5. Kurdistani SK, Grunstein M (2003) Histone acetylation and deacetylation in yeast. Nat Rev Mol Cell Biol 4: 276–284. doi:10.1038/nrm1075.
6. Zentner GE, Henikoff S (2013) Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol 20: 259–266. doi:10.1038/nsmb.2470.
7. Zeng L, Zhou MM (2002) Bromodomain: an acetyl-lysine binding domain. FEBS Lett 513: 124–128.
8. Durant M, Pugh BF (2007) NuA4-directed chromatin transactions throughout the Saccharomyces cerevisiae genome. Mol Cell Biol 27: 5327–5335. doi:10.1128/MCB.00468-07.
9. Krogan NJ, Keogh M-C, Datta N, Sawa C, Ryan OW, et al. (2003) A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. Mol Cell 12: 1565–1576.
10. Ladurner AG, Inouye C, Jain R, Tjian R (2003) Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. Mol Cell 11: 365–376.
11. Matangkasombut O, Buratowski RM, Swilling NW, Buratowski S (2000) Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. Genes Dev 14: 951–962.
12. Ghaemmaghami S, Huh W-K, Bower K, Howson RW, Belle A, et al. (2003) Global analysis of protein expression in yeast. Nature 425: 737–741. doi:10.1038/nature02046.
13. Volanakis A, Passoni M, Hector RD, Shah S, Kilchert C, et al. (2013) Spliceosome-mediated decay (SMD) regulates expression of nonintronic genes in budding yeast. Genes Dev 27: 2025–2038. doi:10.1101/gad.221960.113.
14. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, et al. (2004) A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol 2: E131. doi:10.1371/journal.pbio.0020131.
15. Lu PYT, Lévesque N, Kobor MS (2009) NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. Biochem Cell Biol 87: 799–815. doi:10.1139/O09-062.
16. Chua P, Roeder GS (1995) Bdf1, a yeast chromosomal protein required for sporulation. Mol Cell Biol 15: 3685–3696.
17. Garabedian MV, Noguchi C, Ziegler MA, Das MM, Singh T, et al. (2012) The double-bromodomain proteins Bdf1 and Bdf2 modulate chromatin structure to regulate S-phase stress response in Schizosaccharomyces pombe. Genetics 190: 487–500. doi:10.1534/genetics.111.135459.
18. Fu J, Hou J, Liu L, Chen L, Wang M, et al. (2013) Interplay between BDF1 and BDF2 and their roles in regulating the yeast salt stress response. FEBS J 280: 1991–2001. doi:10.1111/febs.12219.
19. Kapitzky L, Beltrao P, Berens TJ, Gassner N, Zhou C, et al. (2010) Cross-species chemogenomic profiling reveals evolutionarily conserved drug mode of action. Mol Syst Biol 6: 451. doi:10.1038/msb.2010.107.
20. Wang J, Tadeo X, Hou H, Tu PG, Thompson J, et al. (2013) Epe1 recruits BET family bromodomain protein Bdf2 to establish heterochromatin boundaries. Genes Dev 27: 1886–1902. doi:10.1101/gad.221010.113.
21. Harigaya Y, Parker R (2012) Global analysis of mRNA decay intermediates in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 109: 11764–11769. doi:10.1073/pnas.1119741109.
22. Kawashima T, Douglass S, Gabunilas J, Pellegrini M, Chanfreau GF (2014) Widespread Use of Non-productive Alternative Splice Sites in Saccharomyces cerevisiae. PLoS Genet 10: e1004249. doi:10.1371/journal.pgen.1004249.
23. Danin-Kreiselman M, Lee CY, Chanfreau G (2003) RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. Mol Cell 11: 1279–1289.
24. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415.
25. Chanfreau G, Buckle M, Jacquier A (2000) Recognition of a conserved class of RNA tetraloops by Saccharomyces cerevisiae RNase III. Proc Natl Acad Sci U S A 97: 3142–3147. doi:10.1073/pnas.070043997.
26. Wang Z, Hartman E, Roy K, Chanfreau G, Feigon J (2011) Structure of a yeast RNase III dsRBD complex with a noncanonical RNA substrate provides new insights into binding specificity of dsRBDs. Structure 19: 999–1010. doi:10.1016/j.str.2011.03.022.
27. Rouskin S, Zubradt M, Washietl S, Kellis M, Weissman JS (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature 505: 701–705. doi:10.1038/nature12894.
28. Lee A, Henras AK, Chanfreau G (2005) Multiple RNA surveillance pathways limit aberrant expression of iron uptake mRNAs and prevent iron toxicity in S. cerevisiae. Mol Cell 19: 39–51. doi:10.1016/j.molcel.2005.05.021.
29. Pelechano V, Chávez S, Pérez-Ortín JE (2010) A complete set of nascent transcription rates for yeast genes. PLoS ONE 5: e15442. doi:10.1371/journal.pone.0015442.
30. Dichtl B, Stevens A, Tollervey D (1997) Lithium toxicity in yeast is due to the inhibition of RNA processing enzymes. EMBO J 16: 7184–7195. doi:10.1093/emboj/16.23.7184.
31. Yassour M, Kaplan T, Fraser HB, Levin JZ, Pfiffner J, et al. (2009) Ab initio construction of a eukaryotic transcriptome by massively parallel mRNA sequencing. Proc Natl Acad Sci U S A 106: 3264–3269. doi:10.1073/pnas.0812841106.
32. Zhang Z, Dietrich FS (2005) Mapping of transcription start sites in Saccharomyces cerevisiae using 5′ SAGE. Nucleic Acids Res 33: 2838–2851. doi:10.1093/nar/gki583.
33. Lee W, Tillo D, Bray N, Morse RH, Davis RW, et al. (2007) A high-resolution atlas of nucleosome occupancy in yeast. Nat Genet 39: 1235–1244. doi:10.1038/ng2117.
34. Egecioglu DE, Henras AK, Chanfreau GF (2006) Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA 12: 26–32. doi:10.1261/rna.2207206.
35. Liang Y-H, Lavoie M, Comeau M-A, Abou Elela S, Ji X (2014) Structure of a eukaryotic RNase III postcleavage complex reveals a double-ruler mechanism for substrate selection. Mol Cell 54: 431–444. doi:10.1016/j.molcel.2014.03.006.
36. Sayani S, Chanfreau GF (2012) Sequential RNA degradation pathways provide a fail-safe mechanism to limit the accumulation of unspliced transcripts in Saccharomyces cerevisiae. RNA 18: 1563–1572. doi:10.1261/rna.033779.112.

37. Amrani N, Ganesan R, Kervestin S, Mangus DA, Ghosh S, et al. (2004) A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature 432: 112–118. doi:10.1038/nature03060.
38. Galy V, Gadal O, Fromont-Racine M, Romano A, Jacquier A, et al. (2004) Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. Cell 116: 63–73.
39. Fendt S-M, Sauer U (2010) Transcriptional regulation of respiration in yeast metabolizing differently repressive carbon substrates. BMC Syst Biol 4: 12. doi:10.1186/1752-0509-4-12.
40. Chen L, Liu L, Wang M, Fu J, Zhang Z, et al. (2013) Hal2p functions in Bdf1p-involved salt stress response in Saccharomyces cerevisiae. PLoS ONE 8: e62110. doi:10.1371/journal.pone.0062110.
41. Zhao J, Lin W, Ma X, Lu Q, Ma X, et al. (2010) The protein kinase Hal5p is the high-copy suppressor of lithium-sensitive mutations of genes involved in the sporulation and meiosis as well as the ergosterol biosynthesis in Saccharomyces cerevisiae. Genomics 95: 290–298. doi:10.1016/j.ygeno.2010.02.010.
42. Romero-Santacreu L, Moreno J, Pérez-Ortín JE, Alepuz P (2009) Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae. RNA 15: 1110–1120. doi:10.1261/rna.1435709.
43. Favre C, Aguilar PS, Carrillo MC (2008) Oxidative stress and chronological aging in glycogen-phosphorylase-deleted yeast. Free Radic Biol Med 45: 1446–1456. doi:10.1016/j.freeradbiomed.2008.08.021.
44. Chang M, Bellaoui M, Boone C, Brown GW (2002) A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. Proc Natl Acad Sci USA 99: 16934–16939. doi:10.1073/pnas.262669299.
45. Koç A, Wheeler LJ, Mathews CK, Merrill GF (2004) Hydroxyurea arrests DNA replication by a mechanism that preserves basal dNTP pools. J Biol Chem 279: 223–230. doi:10.1074/jbc.M303952200.
46. Tesauro C, Morozzo della Rocca B, Ottaviani A, Coletta A, Zuccaro L, et al. (2013) Molecular mechanism of the camptothecin resistance of Glu710Gly topoisomerase IB mutant analyzed in vitro and in silico. Mol Cancer 12: 100. doi:10.1186/1476-4598-12-100.
47. Tkach JM, Yimit A, Lee AY, Riffle M, Costanzo M, et al. (2012) Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. Nat Cell Biol 14: 966–976. doi:10.1038/ncb2549.
48. Wang Y, Liu CL, Storey JD, Tibshirani RJ, Herschlag D, et al. (2002) Precision and functional specificity in mRNA decay. Proc Natl Acad Sci USA 99: 5860–5865. doi:10.1073/pnas.092538799.
49. Garre E, Romero-Santacreu L, Barneo-Muñoz M, Miguel A, Pérez-Ortín JE, et al. (2013) Nonsense-mediated mRNA decay controls the changes in yeast ribosomal protein pre-mRNAs levels upon osmotic stress. PLoS ONE 8: e61240. doi:10.1371/journal.pone.0061240.
50. Matangkasombut O, Buratowski S (2003) Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. Mol Cell 11: 353–363.
51. Lundin C, North M, Erixon K, Walters K, Jenssen D, et al. (2005) Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. Nucleic Acids Res 33: 3799–3811. doi:10.1093/nar/gki681.
52. Floyd SR, Pacold ME, Huang Q, Clarke SM, Lam FC, et al. (2013) The bromodomain protein Brd4 insulates chromatin from DNA damage signalling. Nature 498: 246–250. doi:10.1038/nature12147.
53. Gaytán BD, Loguinov AV, La Rosa De VY, Lerot J-M, Vulpe CD (2013) Functional genomics indicates yeast requires Golgi/ER transport, chromatin remodeling, and DNA repair for low dose DMSO tolerance. Front Genet 4: 154. doi:10.3389/fgene.2013.00154.
54. Gietz RD, Schiestl RH (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc 2: 31–34. doi:10.1038/nprot.2007.13.
55. Longtine MS, McKenzie A, Demarini DJ, Shah NG, Wach A, et al. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. Yeast 14: 953–961. doi:10.1002/(SICI)1097-0061(199807)14:10<953::AID-YEA293>3.0.CO;2-U.
56. Wach A, Brachat A, Pöhlmann R, Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 10: 1793–1808.
57. Stuckey S, Storici F (2013) Gene knockouts, in vivo site-directed mutagenesis and other modifications using the delitto perfetto system in Saccharomyces cerevisiae. Meth Enzymol 533: 103–131. doi:10.1016/B978-0-12-420067-8.00008-8.
58. Chanfreau G, Rotondo G, Legrain P, Jacquier A (1998) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. EMBO J 17: 3726–3737. doi:10.1093/emboj/17.13.3726.
59. Lamontagne B, Elela SA (2001) Purification and characterization of Saccharomyces cerevisiae Rnt1p nuclease. Meth Enzymol 342: 159–167.
60. Rondón AG, Mischo HE, Kawauchi J, Proudfoot NJ (2009) Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol Cell 36: 88–98. doi:10.1016/j.molcel.2009.07.028.
61. Ghazal G, Gagnon J, Jacques P-E, Landry J-R, Robert F, et al. (2009) Yeast RNase III triggers polyadenylation-independent transcription termination. Mol Cell 36: 99–109. doi:10.1016/j.molcel.2009.07.029.

Figure S1

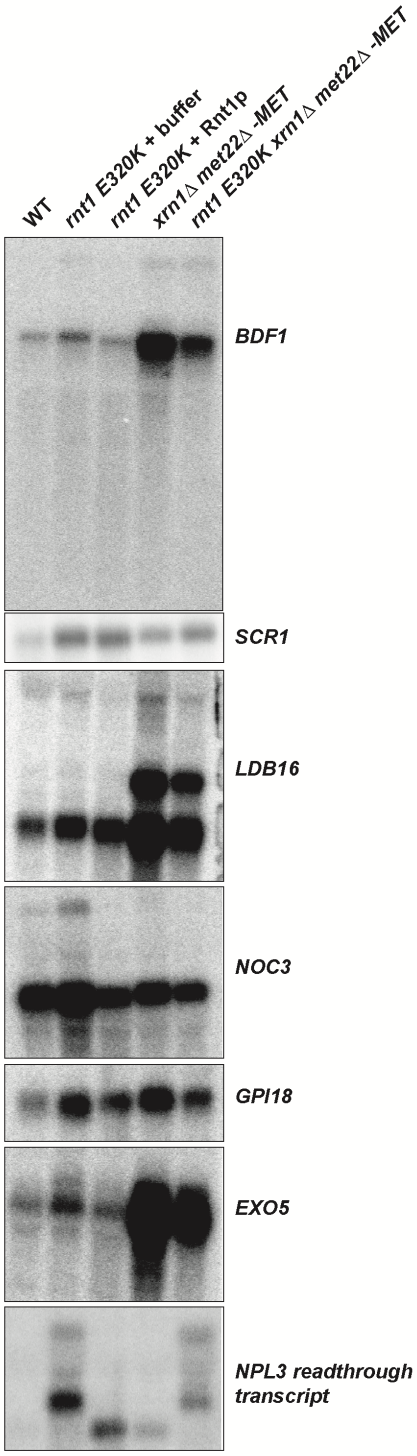

WT
rnt1 E320K + buffer
rnt1 E320K + Rnt1p
xrn1Δ met22Δ -MET
rnt1 E320K xrn1Δ met22Δ -MET
BDF1
SCR1
LDB16
NOC3
GPI18
EXO5
NPL3 readthrough transcript

Figure S2

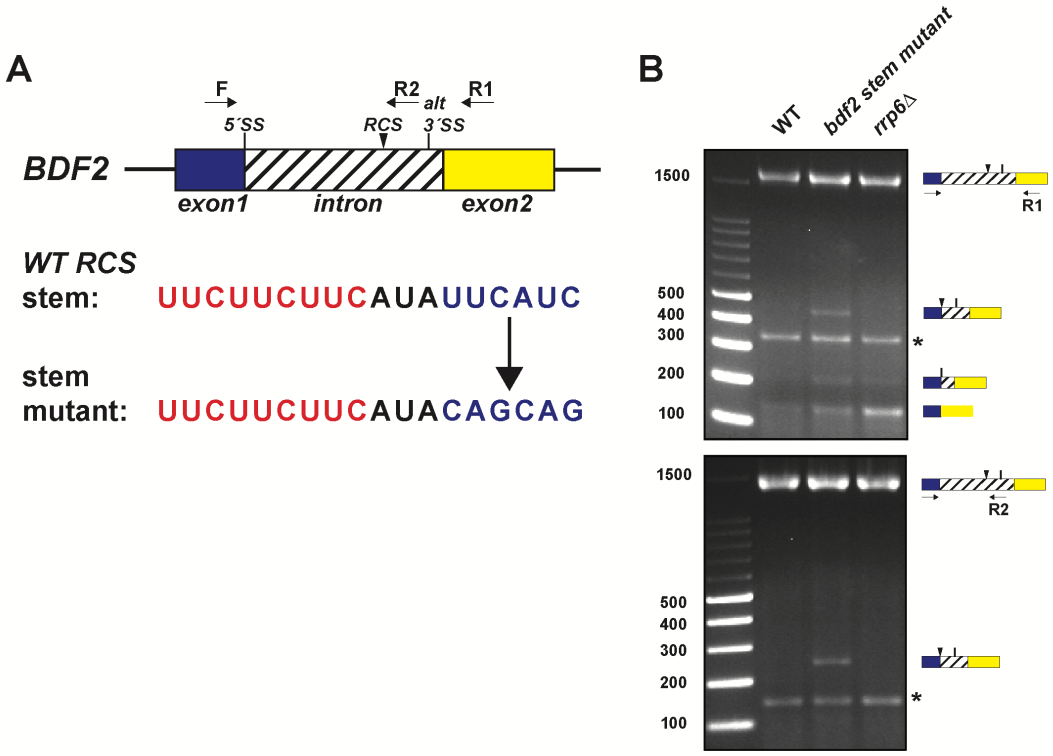

A
F
R2
alt
R1
5´SS
RCS
3´SS
BDF2
exon1
intron
exon2
WT RCS
stem:
UUCUUCUUCAUAUUCAUC
stem
mutant:
UUCUUCUUCAUACAGCAG
B
WT
bdf2 stem mutant
rrp6Δ
1500
500
400
300
200
100
R1
R2

Figure S3

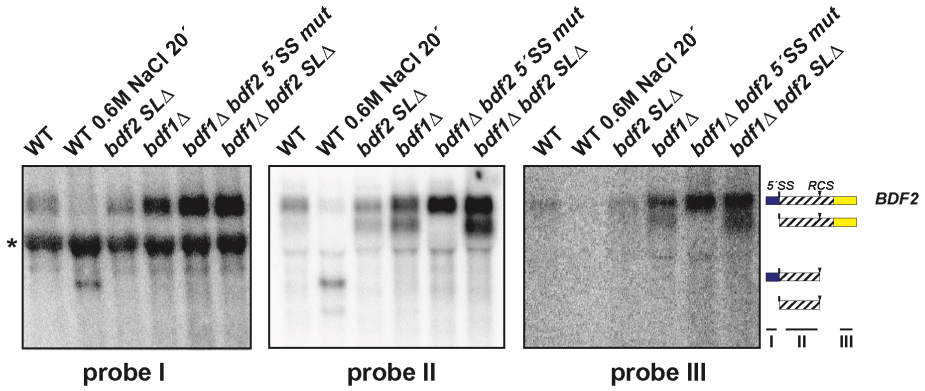
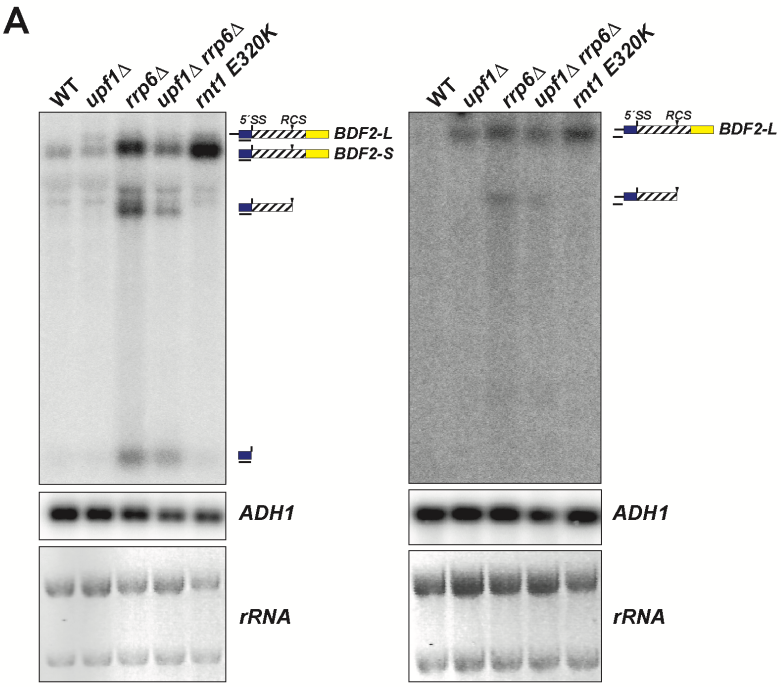

WT
WT 0.6M NaCl 20´
bdf2 SL∆
bdf1∆
bdf1∆ bdf2 5´SS mut
bdf1∆ bdf2 SL∆
*
probe I
probe II
probe III
5´SS
RCS
BDF2
I
II
III

Figure S4


A
WT
upf1Δ
rrp6Δ
upf1Δ rrp6Δ
rnt1 E320K
5´SS
RCS
BDF2-L
BDF2-S
ADH1
rRNA
WT
upf1Δ
rrp6Δ
upf1Δ rrp6Δ
rnt1 E320K
5´SS
RCS
BDF2-L
ADH1
rRNA

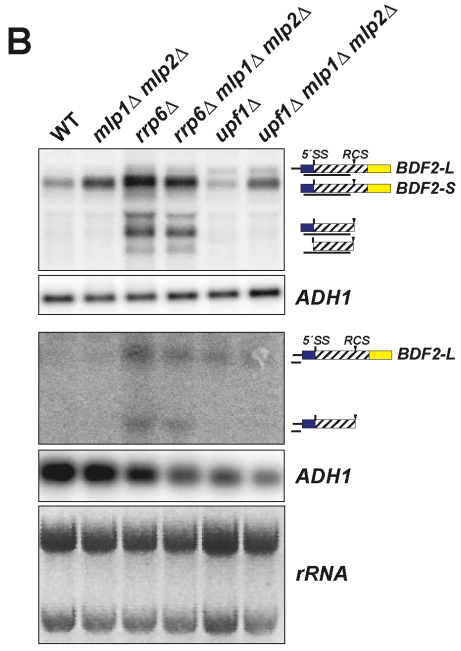

B
WT
mlp1Δ mlp2Δ
rrp6Δ
rrp6Δ mlp1Δ mlp2Δ
upf1Δ
upf1Δ mlp1Δ mlp2Δ
5´SS
RCS
BDF2-L
BDF2-S
ADH1
5´SS
RCS
BDF2-L
ADH1
rRNA

Figure S5

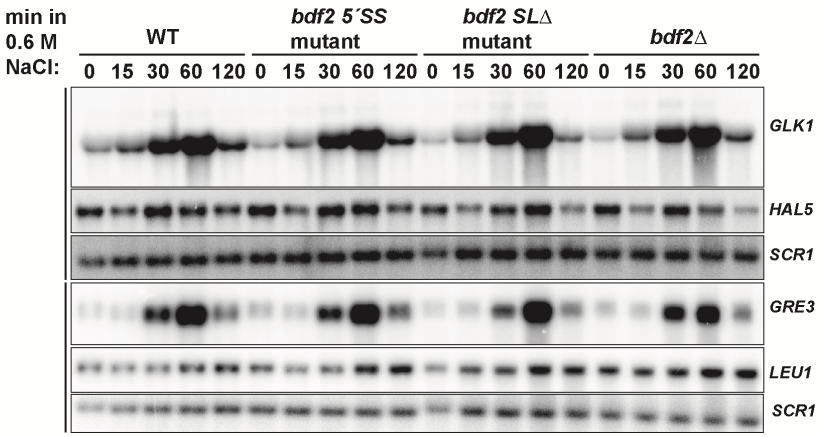
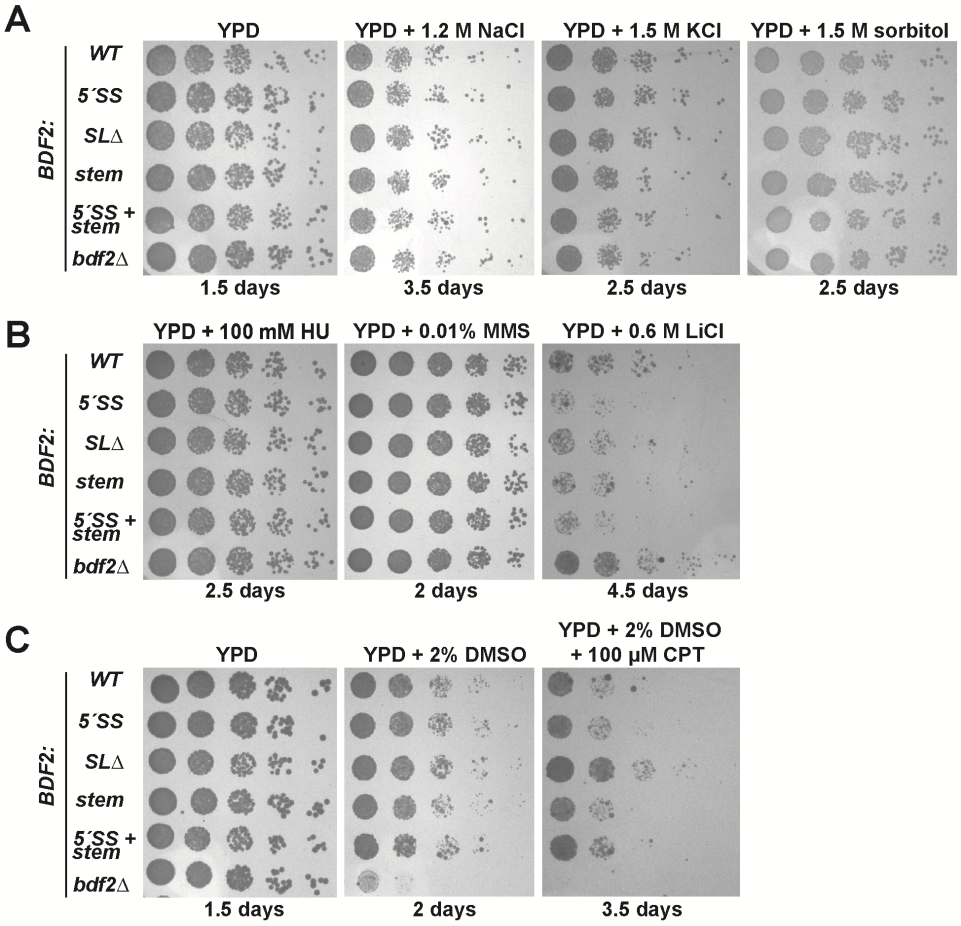

min in 0.6 M NaCl:
WT
bdf2 5´SS mutant
bdf2 SLΔ mutant
bdf2Δ
0 15 30 60 120 0 15 30 60 120 0 15 30 60 120 0 15 30 60 120
GLK1
HAL5
SCR1
GRE3
LEU1
SCR1

Figure S6


A
YPD
YPD + 1.2 M NaCl
YPD + 1.5 M KCl
YPD + 1.5 M sorbitol
BDF2:
WT
5´SS
SLΔ
stem
5´SS + stem
bdf2Δ
1.5 days
3.5 days
2.5 days
2.5 days
B
YPD + 100 mM HU
YPD + 0.01% MMS
YPD + 0.6 M LiCl
BDF2:
WT
5´SS
SLΔ
stem
5´SS + stem
bdf2Δ
2.5 days
2 days
4.5 days
C
YPD
YPD + 2% DMSO
YPD + 2% DMSO + 100 µM CPT
BDF2:
WT
5´SS
SLΔ
stem
5´SS + stem
bdf2Δ
1.5 days
2 days
3.5 days

CHAPTER 6

The Dark Side of Poly(A) Tails Reveals Genome-Wide Mechanisms of Nuclear RNA Processing and Decay

The Dark Side of Poly(A) Tails Reveals Genome-Wide Mechanisms of Nuclear RNA Processing and Decay

Kevin Roy, Jason Gabunilas, Elisabeth Petfalski, Rebekah Liu, Duy Ngo, Marco Morselli, Abigail Gillespie, Doug Zhang, Matteo Pellegrini, David Tollervey, and Guillaume Chanfreau

Abstract

RNA polymerase II (Pol II) transcribes the vast majority of the eukaryotic genome. However, a substantial fraction of Pol II transcription does not produce stable RNA. Various nuclear RNA decay pathways are coupled to the activity of the Trf4/5p-Air2/1p-Mtr4p polyadenylation (TRAMP) complex. TRAMP adds a destabilizing oligo(A) tail to target RNAs, promoting their 3′ end trimming or degradation by the nuclear exosome in a manner dependent on the Rrp6p component of the nuclear exosome. In order to identify the 3′ ends targeted by TRAMP-exosome activity genome-wide, we mapped the global profile of poly(A)+ RNA 3′ ends accumulating in yeast cells lacking Rrp6p. This confirmed snoRNA 3′-extensions as major targets of Rrp6p, and revealed distinct profiles for polyadenylated mature box C/D and box H/ACA snoRNA 3′ ends in the absence of Rrp6p. We demonstrate that the transcriptional activator Reb1p serves not only as a roadblock, but impacts RNA processing by enhancing Nrd1p-Nab3p-Sen1p mediated termination and controlling poly(A) site selection in 3′ untranslated regions of pre-mRNAs. Furthermore, we reveal that a global roadblock of Pol II by TFIIIB at tRNA genes plays a major role in restricting pervasive transcription. In addition, we uncover a previously hidden class of long intergenic non-coding RNAs (lincRNAs) targeted by the yeast nuclear RNase III homolog Rnt1p, and show that co-transcriptional cleavage of lncRNAs by Rnt1p plays an important role to limit pervasive transcription during the salt stress

response. Our results demonstrate how diverse 3′ end processing mechanisms operate to control pervasive transcription and regulate gene expression genome-wide.

Introduction

A variety of chromatin-based modifications repress spurious RNA polymerase II (Pol II) transcription initiation in intragenic, intergenic, and heterochromatic regions (Carrozza et al., 2005; Govind et al., 2010; Kim and Buratowski, 2009; Xue et al., 2015). However, increasing evidence demonstrates that the vast majority of the eukaryotic genome is subject to transcription at some level by Pol II (Berretta and Morillon, 2009; Tisseur et al., 2011). In addition to transcribing stable functional RNAs, Pol II transcribes a vast array of cryptic unstable transcripts (CUTs) with ill-defined functions (Wyers et al., 2005). A substantial fraction of Pol II transcription units produce transcripts that are not observed in wild-type cells due to their rapid decay by a combination of cytoplasmic and nuclear RNA degradation pathways (van Dijk et al., 2011; Wyers et al., 2005). These pervasive transcripts arise from divergent transcription at bidirectional promoters, antisense transcription at termination regions of genes, and at nucleosome-free regions throughout intergenic sequence (Xu et al., 2009; van Dijk et al., 2011; Wyers et al., 2005). A major question emerging from these studies is how these cryptic RNAs are distinguished from *bona fide* mRNAs and selectively destabilized (Porrua and Libri, 2015).

In budding yeast, various mechanisms selectively target cryptic Pol II transcripts for decay, the most well-characterized being the Nrd1p-Nab3p-Sen1p (NNS) termination pathway, the functional homologue of which in metazoans is the RBM7-NEXT pathway (Lubas et al., 2011). In this mode of termination, Nrd1p and Nab3p recognize specific motifs in the nascent RNA and recruit the Sen1p helicase to direct termination (Carroll et al., 2007). Other modes of transcription termination coupled to RNA decay include co-transcriptional cleavage by the yeast

RNase III homolog Rnt1p and roadblock termination by the Reb1p transcriptional activator (Colin et al., 2014; Ghazal et al., 2009; Rondón et al., 2009). The unprotected 3′ ends of RNAs released by these various mechanisms are then targeted by the TRAMP complex, consisting of the Mtr4p helicase, the Trf4p or Trf5p poly(A) polymerase, and the Air1p or Air2p Zn-knuckle RNA binding proteins (Wyers et al., 2005; LaCava et al., 2005; Vanácová et al., 2005). The TRAMP complex tags the 3′ end of the RNA with an oligo-adenosine tail, which recruits the nuclear exosome for rapid 3′-5′ decay of the transcript (Callahan and Butler, 2010). The TRAMP subunit Mtr4p interacts directly with the nuclear exosome co-factor Rrp6p to promote unwinding of the 3′ end of the RNA into the core tunnel of the exosome for 3′ trimming or complete degradation by the Rrp44p/Dis3p catalytic subunit (Schuch et al., 2014).

In the absence of Rrp6p, nuclear exosome-targeted RNA substrates are stabilized, and the distributive poly(A) addition by TRAMP results in a lengthening of the A tails (Wyers et al., 2005). Previous studies uncovered exosome substrates by using oligonucleotide arrays or RNA-seq to analyze RNAs accumulating in a strain lacking Rrp6p (Davis and Ares, 2006; Wyers et al., 2005; Xu et al., 2009). While these studies uncovered broad genomic regions targeted by the nuclear exosome, they lacked the ability to precisely identify 3′ ends targeted by TRAMP. In this study, we sequenced 3′ poly(A)+ ends of RNA accumulating in WT cells and *rrp6Δ* cells, creating a global profile of TRAMP-targeted RNA processing and degradation intermediates. We reveal the genome-wide impact of transcription pauses and roadblocks by Reb1p and TFIIIB. We demonstrate that Reb1p-induced pausing of Pol II enhances termination efficiency by promoting the usage of upstream 3′ processing sites for both the NNS and canonical cleavage and polyadenylation termination pathways. We also demonstrate that the TFIIIB Pol III transcription initiation factor blocks pervasive transcription genome-wide, coupling the pausing of Pol II to TRAMP / nuclear exosome degradation of the arrested 3′ end. These results also uncover genome-wide endonuclease cleavage by the yeast RNase III homolog Rnt1p. To

characterize the genome-wide role of Rnt1p in controlling the transcriptome, we employed a combination of Rnt1p UV crosslinking and analysis of cDNAs (CRAC), computational RNA folding, and high-throughput sequencing of *in vitro* cleaved RNA by recombinant Rnt1 to identify the precise sites of cleavage. Our results show how RNase III degrades both coding and pervasive transcripts to shape the quality and quantity of gene expression.

Results

Direct RNA sequencing (DRS) of poly(A)+ RNA in cells lacking the nuclear exosome component Rrp6 reveals nuclear exosome targets genome-wide

Previous work established that in the absence of the nuclear RNA exosome component Rrp6p, oligo(A) tails on TRAMP substrates are lengthened *in vivo* and can efficiently bind oligo-dT (Wyers et al., 2005). In order to globally assess TRAMP-nuclear exosome targets, we sequenced 3´ ends of poly(A)+ RNAs in wild-type cells and cells deleted for *RRP6* with strand-specific direct RNA sequencing on the Helicos platform. We obtained more than 2.7 million uniquely mapped reads for each sample, with nearly 90% of the poly(A)+ 3´ ends in WT cells mapping to 3´ UTR regions (Fig 6.1A). In the remaining ~10% of reads, we detected 3´ ends mapping to intergenic regions, ORF antisense transcripts, stable unannotated transcripts (SUTs), Xrn1-sensitive unstable transcripts (XUTs), and cryptic unstable transcripts (CUTs). Significantly, the deletion of *RRP6* resulted in a greater than 10-fold upregulation of 3´ ends in numerous categories of pervasive transcripts, including CUTs, SUTs, and XUTs, as well as in 5´ UTR regions and within ORFs, consistent the known role of Rrp6p in transcriptome surveillance and degrading early termination products within ORFs (Fig 6.1A). Overall, 15,298 upregulated poly(A)+ 3´ ends in *rrp6Δ* clustered into 3,395 distinct regions (Fig 6.1B). The class of transcripts with the highest number of clusters was the CUTs, consistent with the known role for

Rrp6 in degrading these transcripts (Wyers et al., 2005). Our approach also uncovered Rrp6p-targeted transcripts arising from unannotated intergenic regions, as well as throughout coding sequence (CDS) and antisense to coding sequence (Fig 6.1B).

Strikingly, the absence of Rrp6p resulted in greater than 20% of all poly(A)+ RNAs arising from the 79 small nucleolar RNAs (snoRNAs) and their downstream flanking regions (Fig 6.1A). snoRNAs are heavily transcribed by Pol II and the majority of snoRNAs undergo extensive processing at both the 5′ and 3′ ends. Nascent snoRNAs undergo transcription termination by the Nrd1p-Nab3p-Sen1p (NNS) pathway followed by exosome trimming (Carroll et al., 2004; Steinmetz et al., 2001). The trimming efficiency is significantly diminished in the absence of Rrp6p, and cells accumulate snoRNAs with extended 3′ ends containg poly(A) tails (Allmang et al., 1999; Grzechnik and Kufel, 2008; van Hoof et al., 2000). Metagene analysis of the downstream genomic regions for all 79 snoRNAs revealed the characteristically broad window of heterogenous 3′ ends generated by NNS termination (Fig 6.1C), as previously demonstrated by tiling array and RNA-seq (Fox et al., 2015; Wyers et al., 2005). In addition, we identified a substantial peak centered at 25 nucleotides downstream of mature 3′ ends (Fig. 6.1D). This result is reminiscent with previous work showing that Rrp6p is required for optimal trimming of the final 30 nucleotides of the 3′ end of the precursor 5.8S rRNA (Briggs et al., 1998). Our result indicates that the majority of snoRNA 3′ processing intermediates follow a similar pathway (Fig 6.1D). As suggested for the 5.8S rRNA, our data is consistent with the model that the nuclear exosome trims RNP 3′ ends by threading the 3′ extensions through the core exosome, where Rrp44p/Dis3p processively hydrolyzes nucleotides until the lid of the exosome encounters the stable RNP 3′ end. This occurs at a distance of ~25 nucleotides from the lid to the Rrp44p/Dis3p active site, whereupon Rrp6p is required for efficient trimming of the final ~25 nucleotides (Makino et al., 2015).

Intriguingly, we found that mature Box C/D snoRNAs, but not Box H/ACA, exhibit poly-adenylated mature 3′ ends in the absence of Rrp6p (Fig 6.1D.) Box H/ACA snoRNAs contain single-stranded 3′ ends protected by the Cbf5p protein, while box C/D 3′ ends are base-paired with their mature 5′ ends (Lapinaite et al., 2013; Zhou et al., 2011). Our results suggest that box C/D 3′ ends, but not the 3′ ends of box H/ACA RNAs, may normally be accessible to the TRAMP complex, and that Rrp6p trims these TRAMP-added poly(A) tails back to their mature 3′ ends. Surveillance by TRAMP and Rrp6p may serve as a quality control mechanism for monitoring the proper RNP structure of Box C/D 3′ ends. In the absence of proper base-pairing between the mature 5′ and 3′ ends of the Box C/D RNA, Rrp6p and the nuclear exosome would then initiate RNP turnover.

The Reb1p DNA binding factor pauses Pol II genome-wide and destabilizes nascent transcripts through degradation by Rrp6p

While the majority of snoRNAs exhibit a broad distribution of 3′ length extensions in the absence of Rrp6p, inspection of our data set revealed that a number of snoRNAs did not show a broad distribution and instead exhibited sharp peaks downstream of the +25 nt poly(A)+ peak. We noticed that many of these peaks were 10-15 nucleotides upstream of the TTACCCG motif, the recognition sequence for the Reb1p DNA-binding protein. Additionally, a computational analysis for enriched motifs downstream of peaks of Rrp6p-targeted clusters revealed TTACCCG as highly enriched (Fig 6.2A). We found the majority of these *rrp6Δ* clusters occur in intergenic regions and 5′ UTRs and in coding sequence, while much fewer were observed in 3′ UTRs (Fig 6.2B). Furthermore, aggregating all poly(A)+ 3′ ends in *rrp6Δ* around all 234 Reb1 consensus binding sites (TTACCCGG) in the genome resulted in a peak ~12 nucleotides upstream and ~15 nucleotides downstream of the motif. This distance is consistent with a collision between the leading edge of elongating Pol II and the footprint of the Reb1p DNA-bound protein (Fig 6.2C) (Colin et al., 2014). Nascent elongating transcript sequencing (NET-

seq) was previously developed to monitor the positions of 3′ ends associated with actively elongating Pol II (Churchman and Weissman, 2011). A previous study showed that the strains lacking the backtracking cleavage stimulatory factor TFIIS (*dst1Δ*) exhibit peaks of pausing near Reb1p sites (Colin et al., 2014). Reb1 can bind to consensus as well as near-consensus motifs (Ju et al., 1990). In order to examine the functional impact of these binding events, we gathered empirically determined Reb1p binding sites by ChIP-exo and high-throughput DNase I footprinting studies (Hesselberth et al., 2009; Rhee and Pugh, 2011). We analyzed all Reb1p near-consensus motifs in these data sets with at most 1 nucleotide deviation from the TTACCCG motif. Meta-gene analysis of the poly(A)+ peaks accumulating in the absence of Rrp6p revealed a striking overlap with the 3′ ends accumulating in the *dst1Δ* strain (Fig. 6.2D). These peaks are conspicuously situated proximally upstream of the boundary of genome-wide DNase I footprinting at these Reb1p sites (Fig. 6.2D) (Hesselberth et al., 2009). These data confirm that the 3′ ends of transcripts in Pol II complexes stalled at Reb1p sites across the genome are indeed targeted by Rrp6p.

Reb1p-induced pausing of Pol II enhances termination efficiency of the Nrd1p-Nab3p-Sen1p (NNS) pathway

Reb1p was previously shown to control pervasive transcription by acting as a roadblock to Pol II elongation, as the nascent transcripts of the arrested Pol II complex are cleared by the action of the TRAMP-exosome (Colin et al., 2014). This study suggested that Rsp5-mediated ubiquitylation of Pol II was necessary to remove Pol II from the template. However, the presence of these motifs proximally downstream of highly transcribed snoRNAs suggested a possible role for Reb1p acting in concert with termination by the NNS pathway. We chose first to examine the Box H/ACA snoRNA snR161, as this snoRNA exhibited the strongest signal proximally upstream of a Reb1p-binding site (Fig. 6.3A). We mutated the TTACCCG motif to TTACAAG, as the mutation of these two C's to A's was previously shown to completely abolish

Reb1p binding (Colin et al., 2014). We found that blocking Reb1p binding led to the accumulation of a broad range of 3´ end extensions of snR161 in the absence of Rrp6p (Fig 6.3B, top panels). Inserting an RNA cleavage signal for the nuclear RNase III homolog Rnt1p in place of the Reb1 binding site led to a reduction in the size distribution, demonstrating that the broad signal represents 3´ extensions (Fig. 6.3B, top panels). The gene downstream of snR161, *PDX3*, is only 100 nucleotides downstream of the Reb1p binding site. Our data suggest that the presence of the Reb1p binding site is important to enhance the termination efficiency by the NNS pathway, which normally occurs over a 200-300 bp window (Fig. 6.1D), to prevent transcriptional readthrough into the *PDX3* promoter. Interestingly, we found that the Reb1p site appears to serve as a bi-directional roadblock, as in its absence antisense transcripts accumulate when Rrp6p is genetically inactivated (Fig. 6.3B). Notably, the antisense transcripts are unaffected by the presence of the sense Rnt1p cleavage signal (RCS) as these signals requires specific RNA secondary structure for Rnt1p recognition (Fig. 6.3B). These results underscore the importance of the Reb1p transcriptional activator in enhancing snoRNA termination efficiency.

We next analyzed the function of an apparent roadblock Reb1p site in the 3´ UTR of *RPL9B* (Fig 2E). *RPL9B* was previously characterized to undergo two distinct modes of termination, one promoting expression from canonical cleavage and polyadenylation at various positions, and the other utilizing the NNS pathway downstream of these positions to auto-regulate *RPL9B* levels in manner dependent on Rpl9A/B protein levels (Gudipati et al., 2012). *RPL9B* was shown to contain multiple different poly(A) sites utilized by the cleavage and polyadenylation machinery, as well as NNS–dependent sites downstream (Gudipati et al., 2012). We analyzed nascent transcript elongating sequencing data for the wild-type and TFIIS mutant *dst1Δ* to profile Pol II occupancy in this region (Churchman and Weissman, 2011). We found evidence of Pol II pausing proximally upstream of the Reb1 binding site, and coinciding

iwth the transcript 3′ ends accumulating in the absence of Rrp6p (Fig. 6.3C). The deletion of *RRP6* results in various growth defects and decreased ribosome biogenesis (Fox et al., 2015). In order to examine the role of the Reb1p binding event in a less pleiotropic context, we employed the anchor-away approach, which involves tagging the protein of interest with a domain that forms a stable tertiary interaction with the tagged abundant ribosomal protein Rpl13Ap upon addition of rapamycin to the growth media (Haruki et al., 2008). The massive flow of ribosomes out of nucleus during yeast log phase growth provides a highly abundant "anchor" to rapidly deplete nuclear proteins. We selectively removed either Rrp6p and/or Reb1p from the nucleus, and found the removal of Reb1p and Rrp6p from the nucleus to result in a lengthening of the NNS-terminated transcripts (Fig. 6.3E). Taken together, these results suggest that termination efficiency by the NNS pathway is strongly enhanced by the Reb1p DNA binding factor.

Reb1p-induced pausing of Pol II controls the selection of alternative cleavage and poly-adenylation sites

Reb1 sites are frequently found in between tandemly or convergently transcribed genes (Hartley and Madhani, 2009). After finding that termination efficiency by the NNS pathway is strongly enhanced by the Reb1p DNA binding factor, we speculated that Reb1p may also influence the cleavage and polyadenylation termination pathway. To test this possibility, we sequenced poly(A)+ 3′ ends of WT cells or Reb1p-FRB tagged cells treated with rapamycin for one hour to deplete Reb1p from the nucleus. Strikingly, the anchor away of Reb1p resulted in a substantial shift downstream for the predominant poly(A) site for various genes exhibiting multiple alternative 3′ ends (Fig. 6.4A). Increased RNA-seq signal in the distal 3′ UTR region confirmed the downstream poly(A) site usage (Fig. 6.4A).

We next examined genome-wide the prevalence of Reb1p binding sites internal to 3′ UTRs and within 100 bp downstream of 3′ UTRs. To determine 3′ UTR boundaries, we

examined transcript isoform sequencing (TIF-seq) of simultaneous 5′ and 3′ transcript ends (Pelechano et al., 2013), and set the 3′ boundary to cover 80% of all ORF-spanning 3′ ends. We found a total of 271 Reb1p sites matching these criteria, and aggregated NET-seq data for these regions (Fig. 6.4B). We found the expected peak in the TFIIS mutant *dst1Δ* ~15 bp upstream of the Reb1 site (Fig. 6.4B). Interestingly, in contrast to the meta-gene analysis on all Reb1p binding sites (Fig. 6.2D), there was no backtracked and paused peak in WT cells at 3′ UTR Reb1p sites. Instead, a minor peak occurred directly at the Reb1p binding site (Fig. 6.4B).

We then normalized the signal for each of the 271 3′ UTRs within a 200 bp window around the Reb1p sites, and calculated the mean position of poly(A) site usage in WT and Reb1p anchor away treated cells. We found a global average lengthening of the 3′ UTRs in these regions by 6 nucleotides after anchor away of Reb1p (Fig. 6.4B). This indicated that Reb1p has an important role in globally influencing 3′ end selection by cleavage and polyadenylation, likely by pausing Pol II in the 3′ regions to promote the use of otherwise weak poly(A) sites upstream. In the absence of Reb1p, stronger poly(A) sites downstream may emerge before the usage of weaker upstream poly(A) sites and be used preferentially.

A role of Reb1p in promoting 3′ end cleavage and polyadenylation contrasts with its previously characterized role as a transcriptional roadblock that couples pausing to nuclear exosome decay of the nascent transcripts (Colin et al., 2014). To examine the relative locations of Reb1p sites across ORF transcripts genomewide, we aggregated all ORF transcripts with Reb1p sites within 100 bp upstream of the transcriptional start site to 100 bp downstream of the 3′ UTR end, with transcript termini accounting for 90% of the ORF transcript boundaries determined from TIF-seq data (Pelechano et al., 2013). To align all ORFs to the same window, we divided ORF regions into 100 bins and placed the Reb1p binding sites accordingly. This revealed a substantial accumulation of Reb1p sites upstream of promoters (Fig 6.5A), as expected (Hartley and Madhani, 2009). However, we also found Reb1p sites throughout ORF

regions, with an enrichment toward the poly(A) site region (Fig 6.5A). To see if the Reb1p sites in these different relative locations across ORF transcript regions behave differently with respect to roadblock termination, we screened these sites for the presence of *rrp6Δ* upregulated poly(A)+ signal in a window from -20 to -10 nucleotides upstream of the Reb1p binding site (Fig 6.5B). This revealed an under-enrichment of Reb1p-proximal Rrp6p-targeted transcripts toward the 3´ end of genes, relative to the first half of genes and upstream and downstream genic regions (Fig 6.5B). This suggests that the context of the Reb1p site relative to the transcription start and poly(A) sites influences the tendency of Reb1p roadblocks to promote nuclear exosome targeting of the nascent transcripts (Fig 6.5B).

The tRNA gene *tG(GCC)E* promotes optimal NNS-dependent termination of snR53 transcription

In addition to finding Reb1p sites downstream of various snoRNA gene units, we found that snR53 transcription appeared to terminate within a short ~100 bp genomic window proximal to the tRNA gene *tG(GCC)E* (Fig 6.6A). Both the deletion and anchor away of *RRP6* revealed poly(A) sites ending near the boundary of the tRNA gene (Fig 6.6A). NET-seq data also confirmed the presence of Pol II ending near the tDNA boundary (Fig 6.6A). To test how the presence of the tRNA transcriptional machinery functionally impacts snR53 termination, we co-depleted from the nucleus both *RRP6* and the Bdp1p subunit of the Pol III initiation factor TFIIIB. Strikingly, this revealed the depletion of poly(A) signal proximal to the tRNA and the accumulation of a readthrough poly(A) peak ~300 bp downstream (Fig 6.6A). To confirm these observations, we probed for mature snR53 and the downstream region by Northern analysis of strains deleted for *RRP6* and/or the *tG(GCC)E* gene (Fig 6.6B). This revealed that indeed optimal snR53 termination by the NNS pathway depends on the presence of the tRNA gene downstream. Taken together with our earlier results of Reb1p roadblock enhancement of NNS termination, we conclude that the NNS pathway is generally enhanced by obstacles in the

chromatin, and that particular genes have evolved to strictly depend on these obstacles for efficient NNS termination.

Genome-wide roadblock by Pol III initiation factor TFIIIB restricts interference by cryptic Pol II transcription

The presence of tRNAs was previously suggested to block cryptic intergenic transcription at various loci (Korde et al., 2014; Wang et al., 2014). To test the genome-wide prevalence of this phenomenon, we aggregated all *rrp6Δ* reads in a 500 bp window around tRNA genes, and found a prominent peak centered 65 bp upstream of the mature 5′ end of the tRNA genes (Fig. 6.7A). As transcription by Pol III typically initiates -13 bp upstream of the mature 5′ end of tRNAs, and TFIIIB binds at -35 bp upstream of the transcription start site (Grove et al., 2002), the 65 bp peak likely represents a collision of the leading edge of elongating Pol II with the TFIIIB complex. We again analyzed NET-seq data for evidence of Pol II pausing, and found substantial peaks representing paused and backtracked Pol II at 65 and 70 bp positions upstream of tRNA genes, respectively (Fig. 6.7B). Interestingly, we found that the major TRAMP-exosome targeted peaks coincide with the position of stalled Pol II, indicating that nascent Pol II transcripts stalled at TFIIIB sites can either undergo backtracking and TFIIS cleavage, or undergo Rrp6p targeted degradation. We conclude a global role for TFIIIB in roadblocking cryptic transcription by Pol II, and coupling Pol II stalling to nuclear exosome degradation of the nascent transcript.

We next examined the physiological importance of the TFIIIB-roadblocks. We found substantial accumulation of a 3′ end poly(A)+ peak upstream of the tRNA gene *tN(GUU)P* upon 1 hour anchor away of Rrp6p (Fig. 6.7A). This peak was significantly reduced with the simultaneous anchor away of Rrp6p and Bdp1p, indicating that this peak indeed arises from a TFIIIB roadblock (Fig. 6.7C). *tN(GUU)P* is situated downstream and antisense to the *MEP3*

gene. RNA-seq of cells anchored away for Bdp1p revealed accumulation of readthrough transcripts antisense to *MEP3*, which we confirmed by Northern blot (Fig. 6.7C). To test the impact of this tDNA gene on the expression of *MEP3* in steady state growth, we deleted the *tN(GUU)P* gene and examined the levels of *MEP3* and its antisense transcript (Fig. 6.7D). We found a substantial accumulation of *MEP3* antisense transcripts upon deletion of the tRNA gene (Fig 3D), corroborating our results with the Bdp1p anchor away. This suggested a possible function for the tRNA gene in protecting *MEP3* from transcriptional interference from the downstream antisense promoter. Accordingly, we detected a significant down-regulation of *MEP3* transcripts upon the deletion of *tN(GUU)P* (Fig. 6.7D). *MEP3* encodes an ammonium transporter that facilitates low-affinity but high mass transport of ammonium ion in ammonium replete conditions (Marini et al., 1997). The *MEP3* paralog *MEP1* mediates higher-affinity ammonium transport but a lower maximum transport rate (Marini et al., 1997). Notably, ammonium has been found to be toxic to aging yeast cells (Santos et al., 2012), and the deletion of *MEP3* was previously found to extend chronological lifespan (Powers et al., 2006). In order to test a physiological role for the tDNA roadblock in promoting expression of *MEP3*, we monitored the growth of WT, *mep1Δ*, *mep3Δ*, and *tN(GUU)PΔ* from log-phase into the diauxic shift (Fig. 6.7E). Strikingly, we found that cells containing either a deletion *MEP3* or the tRNA gene downstream, but not *MEP1*, grew significantly faster during the exit from log phase (Fig. 6.7E). This indicates that the roadblock termination of antisense transcription by the *tN(GUU)P* gene plays a major role in controlling *MEP3* expression.

Degradation of *SMF1* mRNA by Rnt1p protects cells against cadmium stress

The majority of snoRNAs undergo 5´ endonucleolytic cleavage by nuclear RNase III homolog Rnt1p, the upstream products of which are degraded by the nuclear exosome (Chanfreau et al., 1998; Ghazal et al., 2005; Lee et al., 2003). Previous studies demonstrated that Rnt1p cleavage fragments are poly-adenylated by the TRAMP complex prior to degradation by Rrp6p

(Egecioglu et al., 2006). We therefore reasoned that identification of poly(A)+ cleavage intermediates could be used to identify Rnt1p cleavage products genome-wide. To obtain direct evidence for *in vivo* binding of Rnt1p to these putative RNA targets, we performed UV crosslinking on a strain expressing tagged Rnt1p and then purified RNA-Rnt1p complexes under strict denaturing conditions. In order to discriminate signal due to non-specific background binding to abundant RNA species, we utilized a strain harboring the G379P mutation in the dsRNA binding domain. This mutation was previously shown to abolish high-affinity binding of Rnt1p to its specific stem-loop targets (Hartman et al., 2013). Comparison of cross-linked RNA fragments enriched at least two fold in the WT Rnt1p over the Rnt1p G379P revealed the identification of 977 specific binding targets (Fig. 6.8A).

Overlapping these peaks with Rrp6p-targeted poly(A)+ clusters revealed the *SMF1* mRNA as a major target for Rnt1p cleavage in the beginning of the *SMF1* open reading frame (Fig. 6.8B). Strikingly, we also found precise overlap of these peaks with 3´ end peaks in the NET-seq data (Fig. 6.8B). This indicates that these Rnt1p cleavage events are likely co-transcriptional, and that the cleavage fragments remain associated with the ternary elongation complex in a stable fashion. To confirm that Rnt1p activity is indeed responsible for these peaks, we prepared total RNA from cells lacking Rnt1p and Rrp6p, and treated these RNAs with purified recombinant Rnt1p or mock conditions. We then *in vitro* poly-adenylated the cleaved or mock treated RNAs with recombinant poly(A) polymerase and performed poly(A)+ site sequencing. This revealed an exact overlap of the NET-seq 3´ ends with the site of Rnt1p cleavage *in vitro* (Fig. 6.8B). *SMF1* encodes a membrane-localized transporter of divalent cations, and exhibits broad specificity for manganese as well as other divalent cations (Supek et al., 1996). In the presence of the divalent toxic metal cadmium, Smf1p is regulated by rapid endocytosis and subsequent degradation at the protein level (Nikko et al., 2008). To test the physiological importance of Rnt1p cleavage of the *SMF1* mRNA in these conditions, we

constructed mutations in the stem loop sequence that disrupted RNA secondary structure but maintained amino acid identity, and also introduced a frameshift mutation as an additional control. We then tested the growth of these strains by spot dilution onto agar medium. No difference in growth rate for these mutations was apparent under normal conditions (Fig. 6.8C). Strikingly, cells unable to cleave *SMF1* by Rnt1p were highly sensitized to cadmium relative to WT cells (Fig. 6.8C). Importantly, the frameshift mutant showed a slight resistance to cadmium relative to WT, indicating that the Rnt1p stem loop mutation results in toxic gain-of-function for *SMF1* (Fig. 6.8C). We conclude that Rnt1p cleavage plays a major role in regulating *SMF1* expression, and complements post-translational regulatory mechanisms to prevent heavy metal toxicity.

Genome-wide co-transcriptional cleavage by Rnt1p controls expression of coding and non-coding transcripts in salt stress

We previously demonstrated a major role for Rnt1p in controlling the expression of *BDF2* mRNA during salt stress (Chapter 5). To extend these results genome-wide, we first tested the ability of a Rnt1p anchor away strategy to mimic the effects of *cis*-acting mutations. We mock treated or rapamycin treated Rnt1p-FRB cells for one hour, and then shifted these cells to high salt stress for one hour (Fig. 6.9A). The anchor away of Rnt1p revealed a substantial upregulation of *BDF2* transcripts, consistent with the genetic inactivation of Rnt1p or its cleavage signal in *BDF2* (Fig. 5.3). Subjecting mock-treated cells to salt stress revealed the expected down-regulation of *BDF2* mRNA, and the emergence of the Rnt1p cleavage product (Fig. 6.9A). Strikingly, the anchor away of Rnt1p prior to salt stress completely rescued the downregulation of *BDF2* mRNA (Fig. 6.9A).

To test how many genes depend on nuclear Rnt1p cleavage for repression during salt stress, we performed high-throughput RNA sequencing on mock-treated or rapamycin treated Rnt1p-FRB strains in the absence or presence of salt stress. Differential expression analysis

revealed 157 genes with two-fold or greater upregulation in Rnt1p-anchor away salt stressed relative to mock-treated salt stressed cells (Fig. 6.9B). Interestingly, many of these genes corresponded to the class of pervasive transcripts known as CUTs, SUTs, and XUTs. We decided to focus on SUT098 as this non-coding transcript exhibited the highest signal amongst the CUTs/SUTs/XUTs in our crosslinking data (Fig. 6.9C). Similarly to *SMF1* mRNA, we found a Rrp6p-targeted peak coinciding with both NET-seq 3´ end peaks, and a prominent peak generated by *in vitro* Rnt1p cleavage (Fig. 6.9C). We then assayed SUT098 by Northern blot and found that two different isoforms of the transcript (Fig. 6.9D, left panel). Upon a shift to salt stress, the shorter form became the more predominant isoform, while a prior treatment with rapamycin abolished this isoform shift (Fig. 6.9D, left panel). We next mutated the cleavage signal by deleting the stem-loop targeted by Rnt1p, and found that the shorter form was abolished by this mutation (Fig. 6.9D, middle panel). We next tested how the shift in 3´-end formation during salt stress impacts expression of the downstream gene, *CHD1*. We observed that normally, *CHD1* expression increases after one hour of high salt exposure (Fig. 6.9D). This increase is completely abolished upon deletion of the SUT098 Rnt1p-targeted stem-loop to the expense of a SUT098 readthrough transcript (Fig. 6.9D). This indicates that alternative 3´-end processing of non-coding RNAs in stress conditions is important for the proper expression of downstream inducible genes.

Discussion

In this work, we profiled the genome-wide distribution of poly(A) tails accumulating upon inactivation of Rrp6p to understand how RNA processing pathways that couple to nuclear exosome degradation impact gene expression. We found that roadblocks by Reb1p and TFIIIB are a major source of exosome targeted poly(A) tails and play a genome-wide role in blocking Pol II transcription, expanding on previous studies (Colin et al., 2014; Korde et al., 2014; Wang et al., 2014). We demonstrated that the context of the Reb1p binding site plays a major role in

influencing the fate of the nascent transcript in the colliding Pol II complex. Our finding that NNS-dependent termination is enhanced by Reb1p collision suggests that the elongation state of the polymerase influences the outcome of the roadblock. We found that Reb1p roadblocks associated with Rrp6p-targeted transcripts exhibit major stalling peaks in *dst1Δ* and backtracked-cleaved peaks in WT cells (Fig. 6.2D). In contrast, Reb1p roadblocks near the 3´ ends of genes which are not associated with Rrp6p-targeted poly(A) tails have less prominent stalling peaks in *dst1Δ* and no detectable backtracked-cleaved peaks in WT cells (Fig. 6.4B). Instead, we found that Reb1p sites in 3´ UTR regions impact the selection of alternative poly(A) sites by the cleavage and polyadenylation machinery, on average favoring upstream sites at the expense of downstream sites (Fig. 6.4B). We propose that cleavage and polyadenylation site selection is dictated by kinetic competition, and that stalling of Pol II by Reb1p allows more time for potentially weaker upstream poly(A) sites to be used. At various genomic loci with downstream stronger poly(A) sites, the depletion of Reb1p reveals a switch in favored poly(A) site usage (Fig. 6.4A).

We have shown that TFIIIB functions as a roadblock to Pol II elongation and triggers exosome degradation of the stalled Pol II transcripts. We demonstrated that the roadblock of cryptic transcription plays an important role in preventing transcriptional interference into downstream genes, using *MEP3* as a model. Lastly, we defined new targets for the yeast RNase III homolog Rnt1p, demonstrating that cleavage of the *SMF1* divalent cation transporter occurs co-transcriptionally and is critical for protecting cells against heavy metal toxicity. We also uncovered a new class of pervasive transcripts targeted by Rnt1p, and showed that Rnt1p cleavage plays a major role in controlling levels of both coding and non-coding transcripts during the salt stress response.

Materials and Methods

Strains

Strains were prepared by standard yeast genetics techniques (Gietz and Schiestl, 2007). *RRP6* was deleted in the BMA64 WT background (mating type alpha) using the *hphMX4* marker. Site specific mutations at Reb1p binding sites, deletions of tRNA genes, and mutations of Rnt1p target stem loops were made using the *delitto perfetto* method (Stuckey and Storici, 2013). Briefly, CORE cassettes were first introduced into the site to be mutated, and the integration of the CORE cassette was confirmed by PCR on genomic DNA. Mutations were then introduced through transformation with oligonucleotides containing the desired modifications. All mutations were confirmed by Sanger sequencing (Laragen, Inc.). The parent strain for anchor away (HHY168, (Haruki et al., 2008)) was first modified by replacing the *natMX6* marker with the *hphMX4* marker by homologous recombination with a PCR product encompassing the TEF promoter and TEF terminator shared by these markers. The indicated anchor away strains were then constructed from the modified parent strain by standard transformation techniques using high-fidelity PCR products generated from either *pFA6a-FRB-kanMX6* or *pFA6a-FRB-HIS3MX6* to tag the gene of interest with the C-terminal FRB module.

Media and culturing

Unless otherwise indicated, cultures were grown in YPD (1% yeast extract, 2% bacto-peptone, 2% dextrose) at 30°C at 200 rpm. For anchor away experiments, FRB-tagged strains were either mock treated with 0.1% v/v of vehicle (90% ethanol, 10% Tween-20) or 1 mg/mL rapamycin to yield a final concentration of 1 µg/mL. All cultures were harvested by centrifugation at 4000 rpm for 2 minutes, washed in deionized water, and spun down in microcentrifuge tubes.

The supernatant was removed and pellets were flash-frozen in liquid nitrogen and stored at -80°C.

Northern analyses and high-throughput RNA sequencing analysis

RNA extractions and Northern blot were performed as described previously (Roy and Chanfreau, 2014). Direct RNA sequencing (DRS) was performed with total RNA from WT and *rrp6Δ* cells using the Helicos single-molecule sequencing system (SeqLL, LLC). Reads were mapped as previously described allowing only uniquely mapped reads (Ozsolak and Milos, 2011). The 5´ ends of DRS reads correspond to one nucleotide upstream of poly(A) sites, as the first non-A residues for DRS are not sequenced (Ozsolak and Milos, 2011). To account for this, 5´ ends of reads were shifted one nucleotide downstream, and tallied for each chromosomal coordinate to generate bedgraphs for chromosomal coordinates and abundances of 3´ ends with poly(A) tails. Mapped positions were first filtered for A/G richness in the immediate genomic region downstream. Poly(A) sites with three or more A nucleotides downstream, and lacking C or T nucleotides in the initial eight nucleotides downstream were flagged as potential internal oligo-dT mis-priming events. Poly(A) sites upregulated in *rrp6Δ* were clustered together if within 20 bp of another upregulated site. Annotating the genomic distribution of upregulated clusters was performed using the annotations provided by the pyCRAC package (Webb et al., 2014). Meta-gene analysis on snoRNA genes, tRNA genes, and Reb1p binding sites was conducted by first obtaining reads across genomic windows around these sites, normalizing the sum of the reads in the window to an arbitrary constant, and then aggregating those normalized reads for all genomic windows. Plots were generated using the matplotlib module of Python 2.7. All anchor away poly(A) site sequencing was conducted using the QuantSeq 3´-end RNA sequencing library preparation kit. Sequencing was performed with the custom sequencing primer (CSP) that includes a 3´ oligo-dT stretch so that the first

sequenced nucleotide corresponds to the poly(A) site (Lexogen, GmbH). Reads were mapped using RNA-STAR to the R64-2-1 version of the *S.cerevisiae* genome obtained from the *Saccharomyces* genome database (SGD) (Dobin et al., 2013). 5´ ends of mapped reads were processed as for the DRS data. Whole transcript RNA sequencing was performed with the Ilumina TruSeq Stranded mRNA library prep kits with sequencing performed on the HiSeq 2000 system in high output mode. Genome browser snapshots were generated from scalable vector graphics files from the Integrated Genomics Viewer (IGV) version 2.3.52.

Motif analyses

The 60 bp of sequences downstream of peaks of Rrp6p-targeted clusters were subjected to analysis by DREME motif discovery with default parameters (both sense and antisense sequence allowed) (Bailey, 2011).

In vitro Rnt1p cleavage coupled to high throughput sequencing of cleavage products

RNA samples were subjected to cleavage by recombinant Rnt1p as previously described (Roy and Chanfreau, 2014). Mock treatments and in vitro cleavage reactions were stopped by standard phenol-chloroform extraction, followed by precipitation in salted ethanol. Purified RNAs were then in vitro polyadenylated by *E. coli* poly(A) polymerase (NEB), and subject to QuantSeq 3´ end library preparation and sequencing as described above.

Cross-linking and analysis of cDNAs (CRAC)

The His6-TEV-ProtA tag with the *URA3* marker was inserted at the C-terminus of *RNT1* in strains carrying WT *RNT1* or the *rnt1 G379P* mutant in the BY4741 background (Hartman et al., 2013). CRAC was performed as previously described (Granneman et al., 2009). Reads were

mapped with RNA-STAR as described above, and mapped reads were processed with the pyCRAC software package (Webb et al., 2014).

Plate growth assays

Cultures were grown to mid-log phase ($OD_{600\ nm}$ 0.4-0.6) and 1 ml of each culture was spun at 13,200 rpm for 30 seconds. Pellets were resuspended in 1 ml of sterile water, and diluted to $OD_{600\ nm}$ 0.05 followed by five three-fold serial dilutions in sterile water. 4 µl of each dilution was spotted onto the indicated media. Cadmium chloride was purchased from Sigma Aldrich.

Figure Legends

Figure 6.1

Genome-wide distribution of poly(A)+ 3´ ends in WT cells and cells lacking the nuclear exosome component Rrp6p

A. Global distribution of poly(A)+ 3´ ends in WT and *rrp6Δ* cells. Total RNA from WT and *rrp6Δ* cells were profiled by direct RNA sequencing (DRS) on an oligo-dT coated flow cell to determine 3´ ends of poly(A)+ RNA. Each sample yielded greater than 2.7 million uniquely mapped reads, which were tallied and grouped into the specified annotation types.

B. Distribution of poly(A)+ clusters upregulated in *rrp6Δ* over WT cells. 3´ ends of poly(A)+ RNA upregulated greater than 5 fold in *rrp6Δ* over WT and within 20 bp of each other were clustered together. The clusters were then grouped into the specified annotation types.

C. Meta-gene analysis of poly(A)+ reads in the snoRNA 3´ regions. Genomic regions 100 bp upstream of snoRNA mature 3´ ends to 500 bp downstream were assayed across all snoRNAs. The total reads for each snoRNA region were normalized to an arbitrary constant, and then summed for all snoRNAs.

D. Meta-gene analysis of poly(A)+ reads at the 3´ ends of different snoRNA classes. Genomic regions 10 bp upstream of snoRNA mature 3´ ends to 50 bp downstream were assayed across all snoRNAs, as well as the box C/D, and box H/ACA snoRNAs. The total reads for each snoRNA region were normalized to an arbitrary constant, and then summed for all snoRNAs.

Figure 6.2

The Reb1p DNA binding factor pauses Pol II genome-wide and destabilizes nascent transcripts through degradation by Rrp6p

A. DREME-motif analysis on regions downstream of Rrp6p-targeted poly(A)+ clusters. 60 bp of sequence downstream from the peaks of Rrp6p-targeted clusters were analyzed for enriched motifs by DREME.

B. Genomic distribution of Rrp6p-targeted poly(A)+ clusters within a 30 bp window upstream of Reb1 motifs.

C. Meta-gene analysis of poly(A)+ reads in *rrp6Δ* cells around the consensus Reb1 motif TTACCCGG. Genomic regions 200 bp upstream and downstream of all occurrences of TTACCCGG were analyzed by meta-gene analysis for each 400 bp region exhibiting at least 10 reads.

D. Meta-gene analysis for RNA polymerase II nascent 3´ ends (NET-seq), DNase I resistant footprints, and poly(A)+ reads in *rrp6Δ* for all experimentally determined Reb1p binding sites. Genomic regions from Reb1p ChIP-exo (Rhee and Pugh, 2011) and genome-wide DNase I footprinting (Hesselberth et al., 2009) were analyzed for sequences containing no more than one mismatch to the Reb1p motif TTACCCG. These chromosomal coordinates were then used in meta-gene analysis for the poly(A)+ reads in *rrp6Δ* (this study), and the 3´ ends of Pol II nascent transcripts in WT cells or cells lacking the back-tracking cleavage stimulatory factor TFIIS (*dst1Δ*) (Churchman and Weissman, 2011). The total reads from both strands around the Reb1p binding sites were normalized to an arbitrary constant. The left-to-right direction was oriented with respect to the strand exhibiting the higher upregulation in *rrp6Δ*, and the same direction was used for the orienting the NET-seq data.

Figure 6.3

Reb1p-induced pausing of Pol II enhances termination efficiency of the Nrd1p-Nab3p-Sen1p (NNS) pathway

A. Poly(A)+ 3′ ends in *rrp6Δ* cells for the 3′ region of snR161. The poly(A)+ 3′ ends in WT and *rrp6Δ* cells for the Crick (-) strand are shown in genome browser format for the genomic region encompassing snR161 and *PDX3*. The scale on the y-axis designates the reads per million mapped reads.

B. Analysis of 3′ extensions and antisense transcripts for snR161. The Reb1p binding site downstream of snR161 was mutated (mut) or replaced with a Rnt1p cleavage site (*Δ*::RCS). These strains were then deleted for *rrp6Δ*. 3′-end processing of snR161 was analyzed by Northern blot with probes to the indicated regions. 18S is shown as a loading control.

C. Reb1p promotes efficient NNS-mediated termination of *RPL9B*. The poly(A)+ 3′ ends in WT and *rrp6Δ* cells and NET-seq reads for the Watson (+) strand are shown in genome browser format for the genomic region around the 3′ end of *RPL9B*. The scale on the y-axis designates the reads per million mapped reads.

D. Northern analysis of *RPL9B* 3′ ends upon rapid nuclear depletion of Rrp6p or Reb1p by anchor away (Haruki et al., 2008). For the *RRP6* and *REB1* genes, the FKBP12-rapamycin-binding (FRB) domain of human mTOR was inserted at the C-terminus in the anchor away parent strain (HHY168) (Haruki et al., 2008). The indicated strains were either mock treated or treated with 1 μg/mL rapamycin (RAPA) for one hour, and *RPL9B* transcripts were analyzed by Northern blot with a 3′ UTR riboprobe upstream of the Reb1p binding site (top panel), or downstream of the Reb1p site (bottom panel). scR1 is shown as a loading control.

Figure 6.4

Reb1p-induced pausing of Pol II controls the selection of alternative cleavage and poly-adenylation sites

A. Analysis of 3´ poly(A)+ ends upon nuclear depletion of Reb1p. Poly(A)+ 3´ ends and RNA-seq profiles for WT or *REB1-FRB* cells treated for one hour with 1 µg/mL rapamycin. The location of the Reb1 binding sites are depicted by orange boxes. The scale on the y-axis designates the reads per million mapped reads for the + strand.

B. Meta-gene analysis for 3´ UTR regions with Reb1p binding sites. The experimentally determined Reb1p binding sites were analyzed for proximity to 3´ UTR regions. 3´ UTR regions were determined from transcript isoform sequencing data (TIF-seq), by selecting 3´ ends from the stop codon and moving downstream that accounted for 90% of the total ORF-spanning signals (Pelechano et al., 2013). The experimentally determined Reb1p binding sites in these regions were centered in a genomic window of 200 bp, and the mean positions of the poly(A)+ ends were determined for WT and cells depleted of Reb1p from the nucleus. The positions relative to the first position of the Reb1 site are indicated with vertical lines. NET-seq data for WT and *dst1Δ* and DNase I footprints were analyzed by meta-gene analysis as performed in Fig. 6.2D.

Figure 6.5

Meta-gene analysis of Reb1 binding locations relative to transcription start sites and poly(A) sites for ORF transcripts

A. All open reading frames (ORFs) were analyzed for the presence of Reb1p binding sites from within 100 bp upstream of the transcriptional start site to 100 bp downstream of the 3´ UTR end. Transcript termini accounting for 90% of the ORF transcript boundaries were determined from

TIF-seq data as described in Fig. 6.4B. Reb1p binding sites in the ORF regions were separated into 100 bins, whereas upstream and downstream regions are shown as base pair distance.
B. ORFs were analyzed as in (A) but screened for the presence of *rrp6Δ* upregulated poly(A)+ signal in a window from -30 to -20 nucleotides upstream of the Reb1p binding site.

Figure 6.6

The tRNA gene *tG(GCC)E* promotes optimal NNS-dependent termination of snR53 transcription

A. Analysis of 3´ poly(A)+ ends in WT and *rrp6Δ* cells, or upon nuclear depletion of Rrp6p or both Rrp6p and the TFIIIB factor Bdp1p. NET-seq data for WT and *dst1Δ* cells demonstrates the overlap between 3´ poly(A)+ ends in accumulating after inactivation of Rrp6p and the nascent 3´ ends of elongating Pol II. The scale on the y-axis designates the reads per million mapped reads for the + strand.
B. Cells were deleted for *RRP6, tG(GCC)E,* or both, and snR53 3´ processing was analyzed by Northern blot with probes to the indicated regions. scR1 is shown as a loading control.

Figure 6.7

Genome-wide roadblock by Pol III initiation factor TFIIIB restricts interference by cryptic Pol II transcription

A. Meta-gene analysis of 3´ poly(A)+ ends in *rrp6Δ* around all tRNA genes. The + strand is shown in blue, and the - strand in green.

B. Meta-gene analysis of 3´ poly(A)+ ends in *rrp6Δ*, nascent 3´ ends in WT and *dst1Δ*, and DNase I footprints in a 100 bp window upstream of all tRNA genes. The - strand is shown as negative values.

C. Analysis of 3´ poly(A)+ ends upon nuclear depletion of Rrp6p or both Rrp6p and the TFIIIB factor Bdp1p. The RNA-seq profiles for WT and the anchor away of Bdp1p are shown. The scale on the y-axis designates the reads per million mapped reads for the + strand, and the RNA-seq profile shows the height of bedgraph-aggregated reads per million mapped reads. The right panel shows a Northern blot with a probe targeting transcripts antisense to *MEP3*. The relative location of the probe, and the direction of transcription start sites are shown at the bottom of the left panel. scR1 is shown as a loading control.

D. The effect of TFIIIB roadblock of downstream antisense transcripts on *MEP3* expression. Cells were deleted for *tG(GCC)E* in duplicate and probed for the antisense *MEP3* transcripts or *MEP3* with probes targeting the regions shown in C. RNA expression for *MEP1,* the paralogue of *MEP3*, is shown as a control for the specificity of the antisense cryptic transcription effect on *MEP3*. scR1 is shown as a loading control.

E. Growth curves for WT, cells lacking *MEP1, MEP3*, or *tG(GCC)E.* The indicated strains were grown in YPD to monitor cell growth through lag phase, log phase, and the diauxic shift.

Figure 6.8

Co-transcriptional cleavage of *SMF1* by Rnt1p protects cells against cadmium stress

A. Rnt1p-RNA crosslinking analysis on cells expressing either WT or binding defective *G379P* versions of tagged Rnt1p. The number of distinct clusters arising from uniquely mapped reads is shown. WT Rnt1p exhibited 977 clusters upregulated at least 2 fold over *G379P*.

B. Rnt1p cleavage of *SMF1* mRNA *in vivo* and *in vitro*. A genome browser view for 3´ poly(A)+ ends in WT and *rrp6Δ*, nascent 3´ ends in WT and *dst1Δ* (NET-seq), Rnt1p *in vitro* cleavage or mock treatment followed by *in vitro* polyadenylation, and Rnt1p-crosslinked RNA reads. The scale on the y-axis designates the reads per million mapped reads for the - strand.

C. Growth analysis of WT, *smf1 null*, or *SMF1* strains with a mutation in the Rnt1p recognition site (stem mutant) in the presence of cadmium. The stem loop region in the *SMF1* ORF was mutated with synonymous codons to disrupt the secondary structure recognized by Rnt1p. As a control, a frameshift mutation was added in addition to the mutated stem loop to generate the *smf1 null*. Cells were grown to log-phase and spotted in serial dilution on YPD plates with the indicated concentrations of cadmium chloride.

Figure 6.9

Genome-wide co-transcriptional cleavage by Rnt1p controls expression of coding and non-coding transcripts in salt stress

A. The effect of Rnt1p anchor away on *BDF2* mRNA expression during salt stress. Cells expressing FRB tagged Rnt1p were mock treated or rapamycin treated for one hour, and then treated with 0.6M NaCl for one hour. *BDF2* expression was analyzed by Northern blot with a probe to the *BDF2* region upstream of the Rnt1p cleavage site. scR1 is shown as a loading control.

B. RNA-seq expression analysis on salt-stressed cells in the presence or absence of Rnt1p-anchor away. Cells were treated as in (A), and then extracted RNAs were analyzed by high-throughput RNA sequencing.

C. Rnt1p cleavage of the long non-coding RNA SUT098 *in vivo* and *in vitro*. A genome browser view for 3´ poly(A)+ ends in WT and *rrp6Δ*, nascent 3´ ends in WT and *dst1Δ* (NET-seq), Rnt1p

in vitro cleavage or mock treatment followed by *in vitro* polyadenylation, and Rnt1p-crosslinked RNA reads. The scale on the y-axis designates the reads per million mapped reads for the + strand.

D. Rnt1p cleavage mediates repression of SUT098 and activation of downstream *CHD1* in salt stress. (Left panel) Cells expressing FRB tagged Rnt1p were mock treated or rapamycin treated for one hour, and then treated with 0.6M NaCl for the indicated time points. SUT098 expression was analyzed by Northern blot. (Right panels) WT cells or a strain carrying a deletion of the Rnt1p recognition stem loop were treated for one hour with 0.6M NaCl. T1 represents the 3´ end generated by Rnt1p cleavage, while T2 represents the downstream 3´ end generated by a different pathway.

Figure 6.1

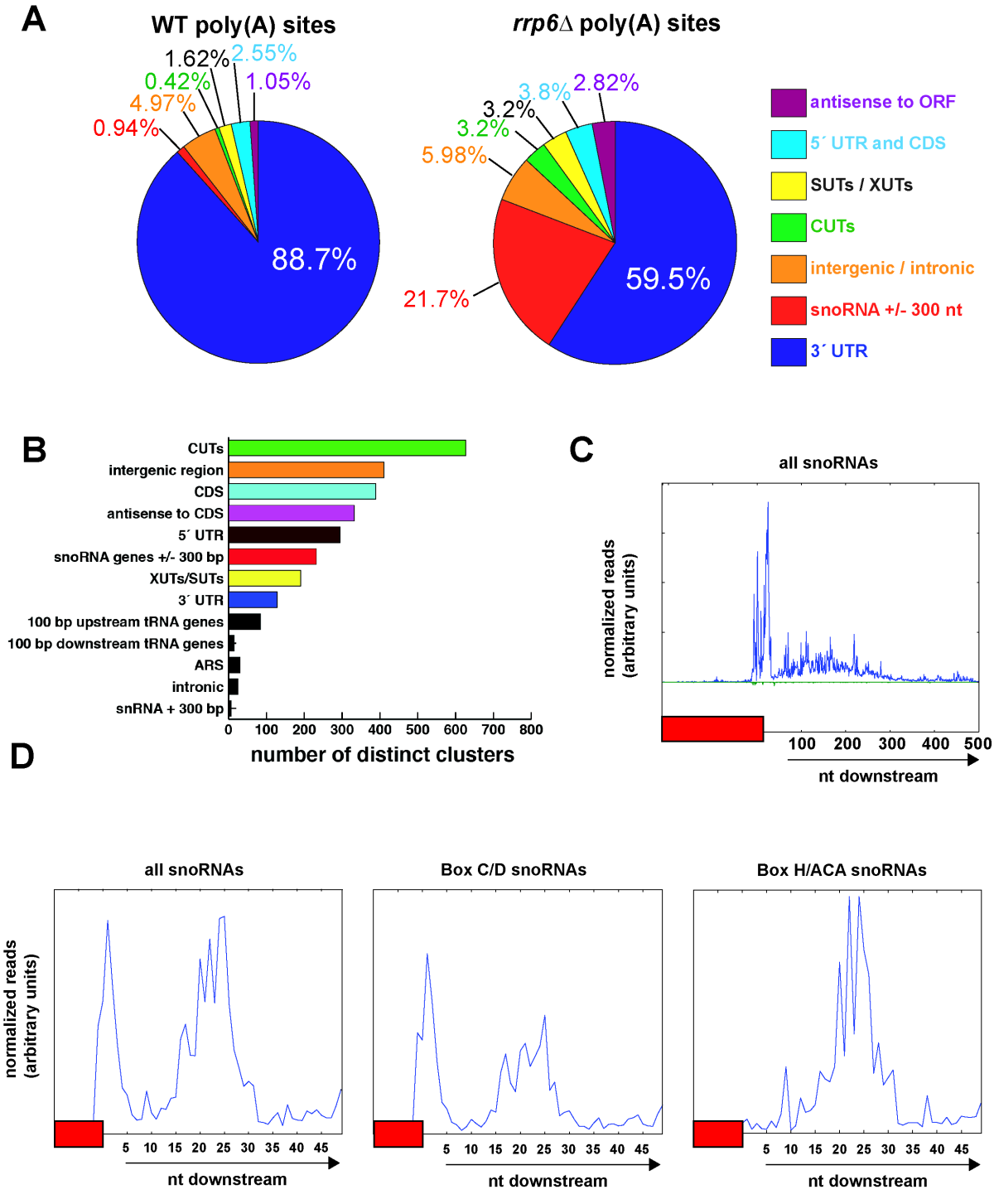
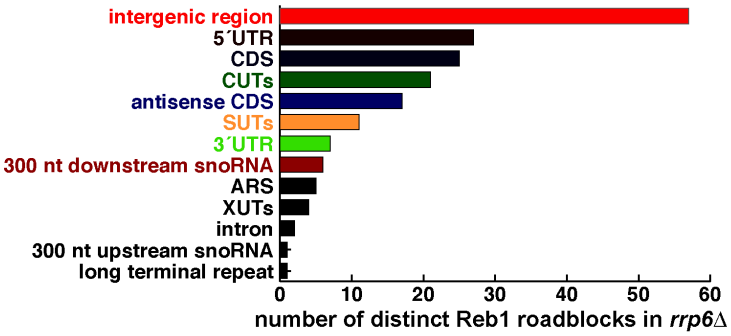

A
WT poly(A) sites
1.62%
2.55%
0.42%
1.05%
4.97%
0.94%
88.7%
rrp6Δ poly(A) sites
3.8%
2.82%
3.2%
3.2%
5.98%
21.7%
59.5%
antisense to ORF
5´ UTR and CDS
SUTs / XUTs
CUTs
intergenic / intronic
snoRNA +/- 300 nt
3´ UTR
B
CUTs
intergenic region
CDS
antisense to CDS
5´ UTR
snoRNA genes +/- 300 bp
XUTs/SUTs
3´ UTR
100 bp upstream tRNA genes
100 bp downstream tRNA genes
ARS
intronic
snRNA + 300 bp
0 100 200 300 400 500 600 700 800
number of distinct clusters
C
all snoRNAs
normalized reads (arbitrary units)
100 200 300 400 500
nt downstream
D
all snoRNAs
Box C/D snoRNAs
Box H/ACA snoRNAs
normalized reads (arbitrary units)
5 10 15 20 25 30 35 40 45
nt downstream

Figure 6.2

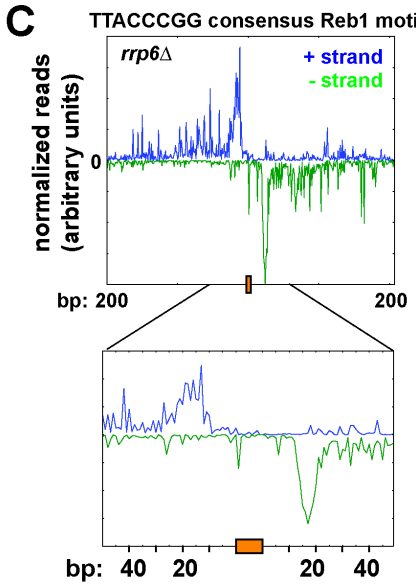
A motifs downstream of *rrp6*Δ poly(A) sites

motif	p-value
TTACCCG	1.3×10^{-15}
(T/C)TCTT(G/C)	4.6×10^{-13}
ATA(T/C)T	2.6×10^{-10}
GNTGGT	1.9×10^{-7}

B



C



D

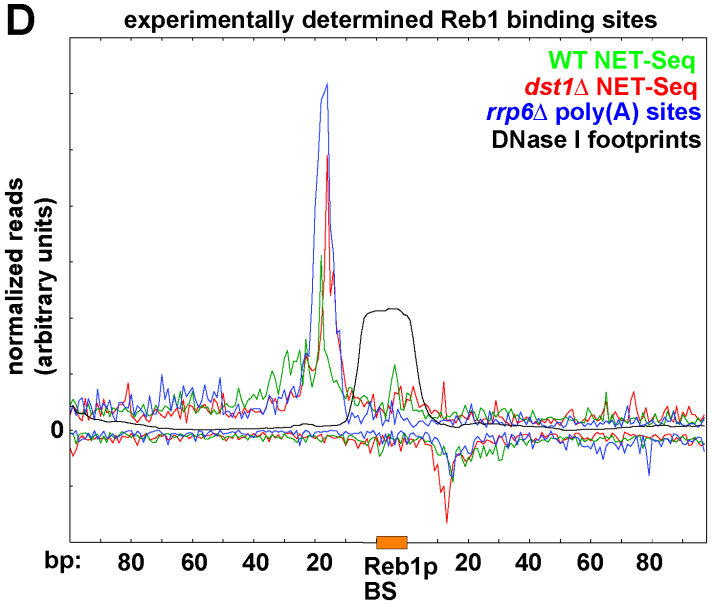


Figure 6.3

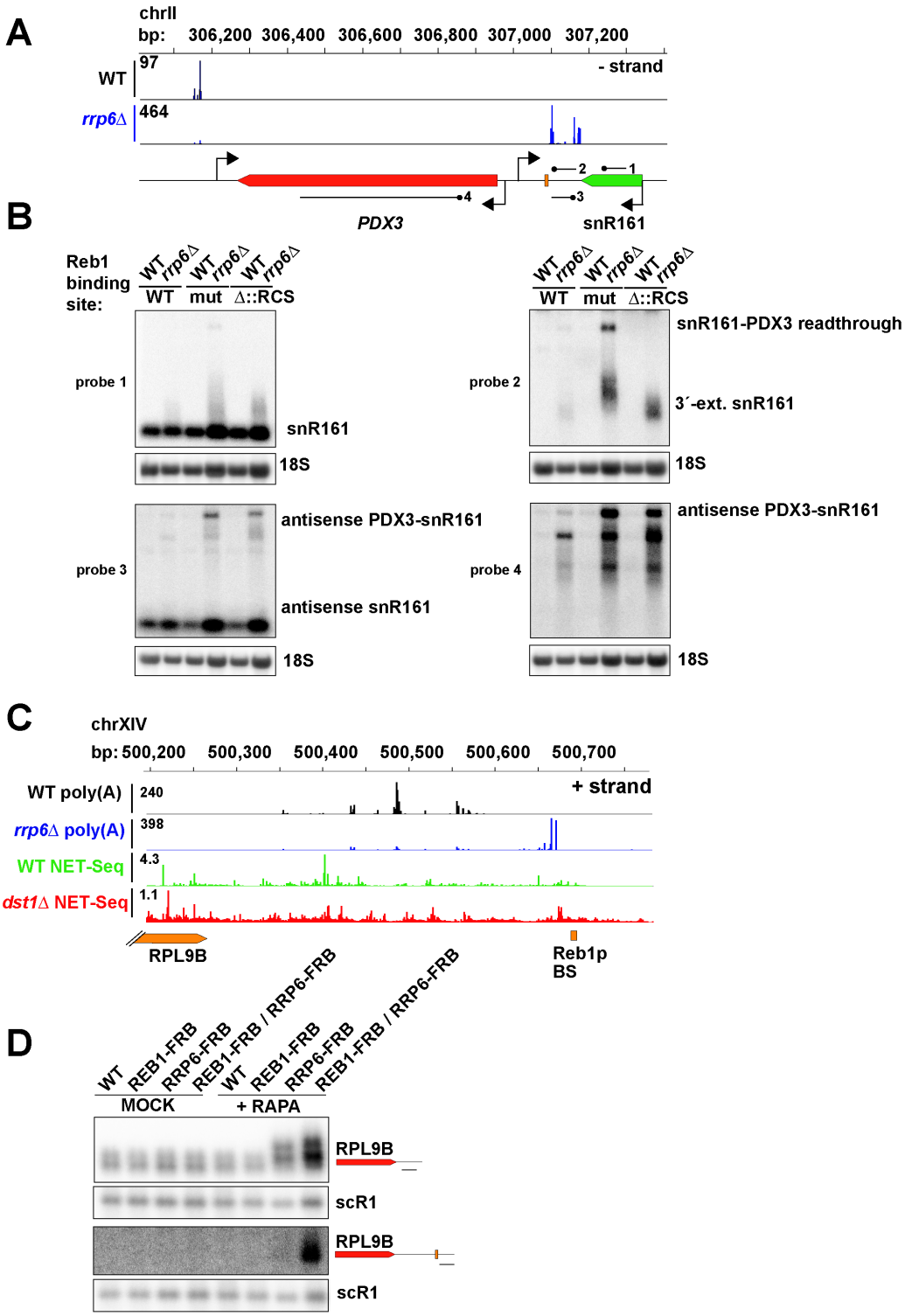

A
chrII
bp: 306,200 306,400 306,600 306,800 307,000 307,200
WT
97
- strand
rrp6Δ
464
2
1
4
3
PDX3
snR161
B
Reb1 binding site:
WT rrp6Δ WT rrp6Δ WT rrp6Δ
WT mut Δ::RCS
probe 1
snR161
18S
probe 2
snR161-PDX3 readthrough
3´-ext. snR161
18S
probe 3
antisense PDX3-snR161
antisense snR161
18S
probe 4
antisense PDX3-snR161
18S
C
chrXIV
bp: 500,200 500,300 500,400 500,500 500,600 500,700
WT poly(A)
240
+ strand
rrp6Δ poly(A)
398
WT NET-Seq
4.3
dst1Δ NET-Seq
1.1
RPL9B
Reb1p BS
D
WT
REB1-FRB
RRP6-FRB
REB1-FRB / RRP6-FRB
MOCK
+ RAPA
RPL9B
scR1
RPL9B
scR1

Figure 6.4

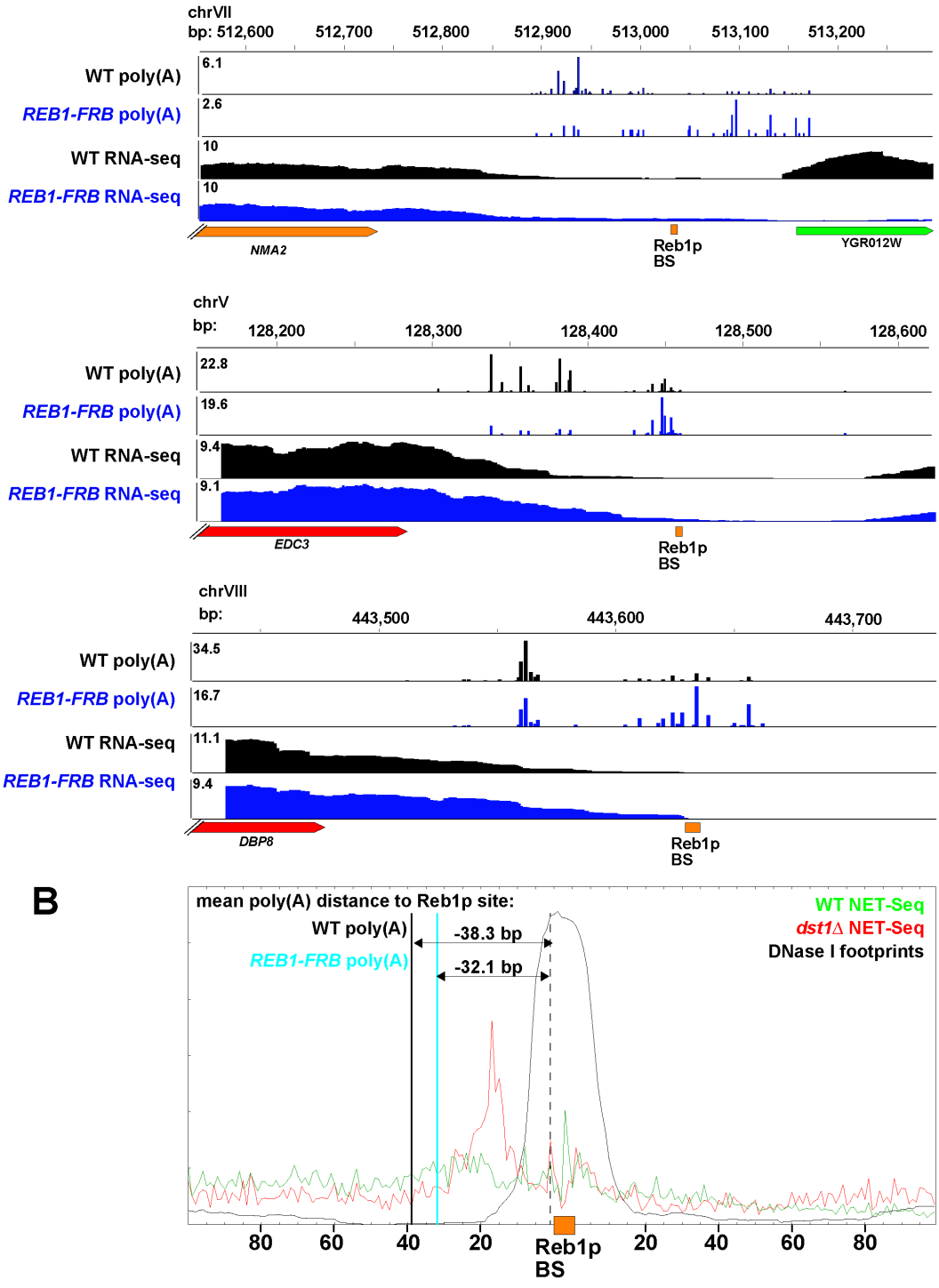

A
chrVII
bp: 512,600
512,700
512,800
512,900
513,000
513,100
513,200
WT poly(A)
REB1-FRB poly(A)
WT RNA-seq
REB1-FRB RNA-seq
6.1
2.6
10
10
NMA2
Reb1p BS
YGR012W
chrV
bp:
128,200
128,300
128,400
128,500
128,600
22.8
19.6
9.4
9.1
EDC3
Reb1p BS
chrVIII
bp:
443,500
443,600
443,700
34.5
16.7
11.1
9.4
DBP8
Reb1p BS
B
mean poly(A) distance to Reb1p site:
WT poly(A)
REB1-FRB poly(A)
-38.3 bp
-32.1 bp
WT NET-Seq
dst1Δ NET-Seq
DNase I footprints
80
60
40
20
Reb1p BS
20
40
60
80

Figure 6.5

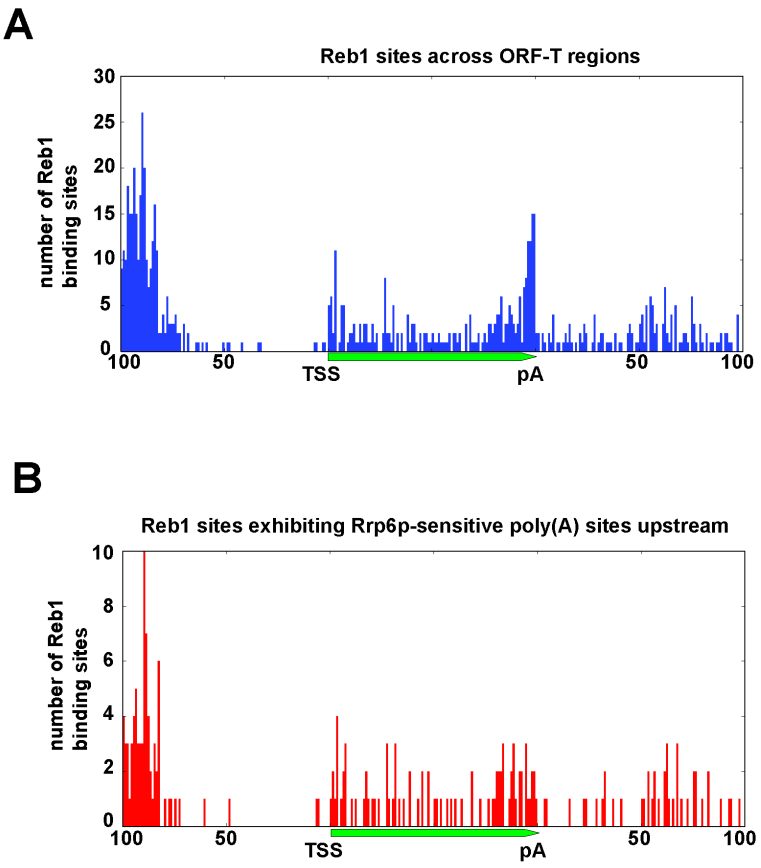

A
Reb1 sites across ORF-T regions
number of Reb1 binding sites
30
25
20
15
10
5
0
100
50
TSS
pA
50
100
B
Reb1 sites exhibiting Rrp6p-sensitive poly(A) sites upstream
number of Reb1 binding sites
10
8
6
4
2
0
100
50
TSS
pA
50
100

Figure 6.6

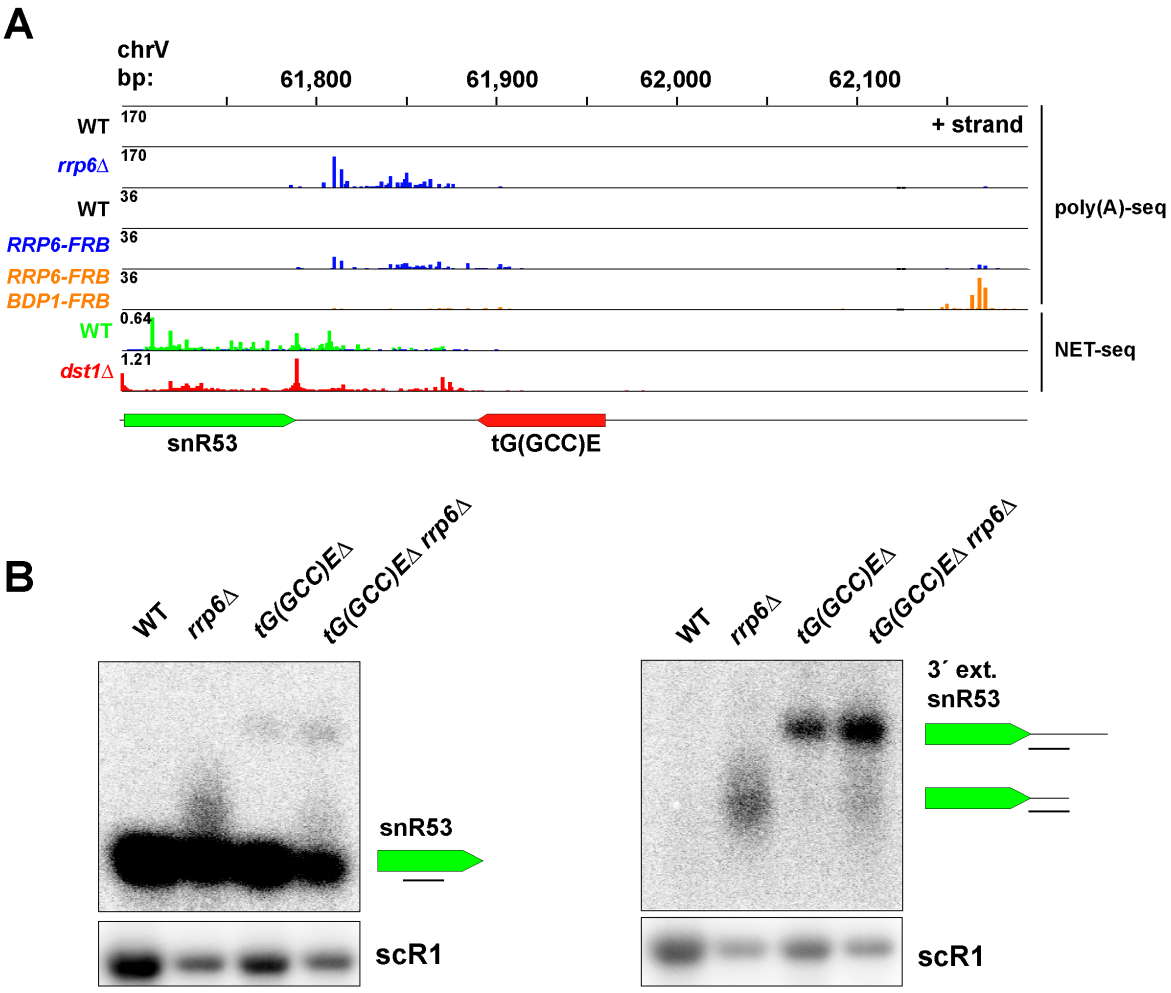

A
chrV
bp:
61,800
61,900
62,000
62,100
WT
170
+ strand
rrp6Δ
170
WT
36
RRP6-FRB
36
RRP6-FRB
36
BDP1-FRB
poly(A)-seq
WT
0.64
dst1Δ
1.21
NET-seq
snR53
tG(GCC)E
B
WT
rrp6Δ
tG(GCC)EΔ
tG(GCC)EΔ rrp6Δ
snR53
scR1
WT
rrp6Δ
tG(GCC)EΔ
tG(GCC)EΔ rrp6Δ
3´ ext.
snR53
scR1

Figure 6.7

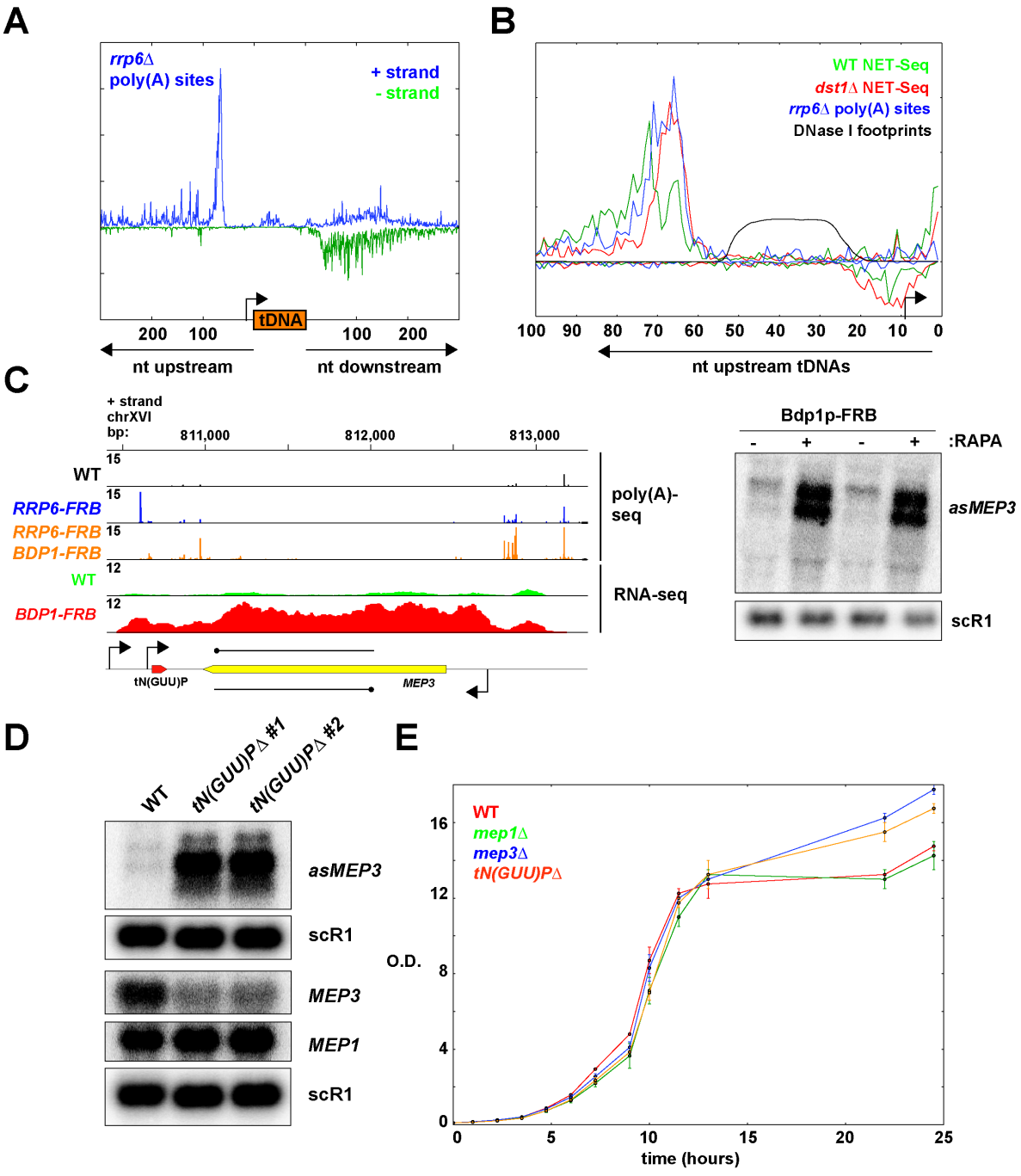

A
rrp6Δ
poly(A) sites
+ strand
- strand
tDNA
200 100
100 200
nt upstream
nt downstream
B
WT NET-Seq
dst1Δ NET-Seq
rrp6Δ poly(A) sites
DNase I footprints
100 90 80 70 60 50 40 30 20 10 0
nt upstream tDNAs
C
+ strand
chrXVI
bp:
811,000
812,000
813,000
WT
RRP6-FRB
RRP6-FRB
BDP1-FRB
WT
BDP1-FRB
poly(A)-seq
RNA-seq
tN(GUU)P
MEP3
Bdp1p-FRB
- + - + :RAPA
asMEP3
scR1
D
WT
tN(GUU)PΔ #1
tN(GUU)PΔ #2
asMEP3
scR1
MEP3
MEP1
scR1
E
WT
mep1Δ
mep3Δ
tN(GUU)PΔ
O.D.
time (hours)

Figure 6.8

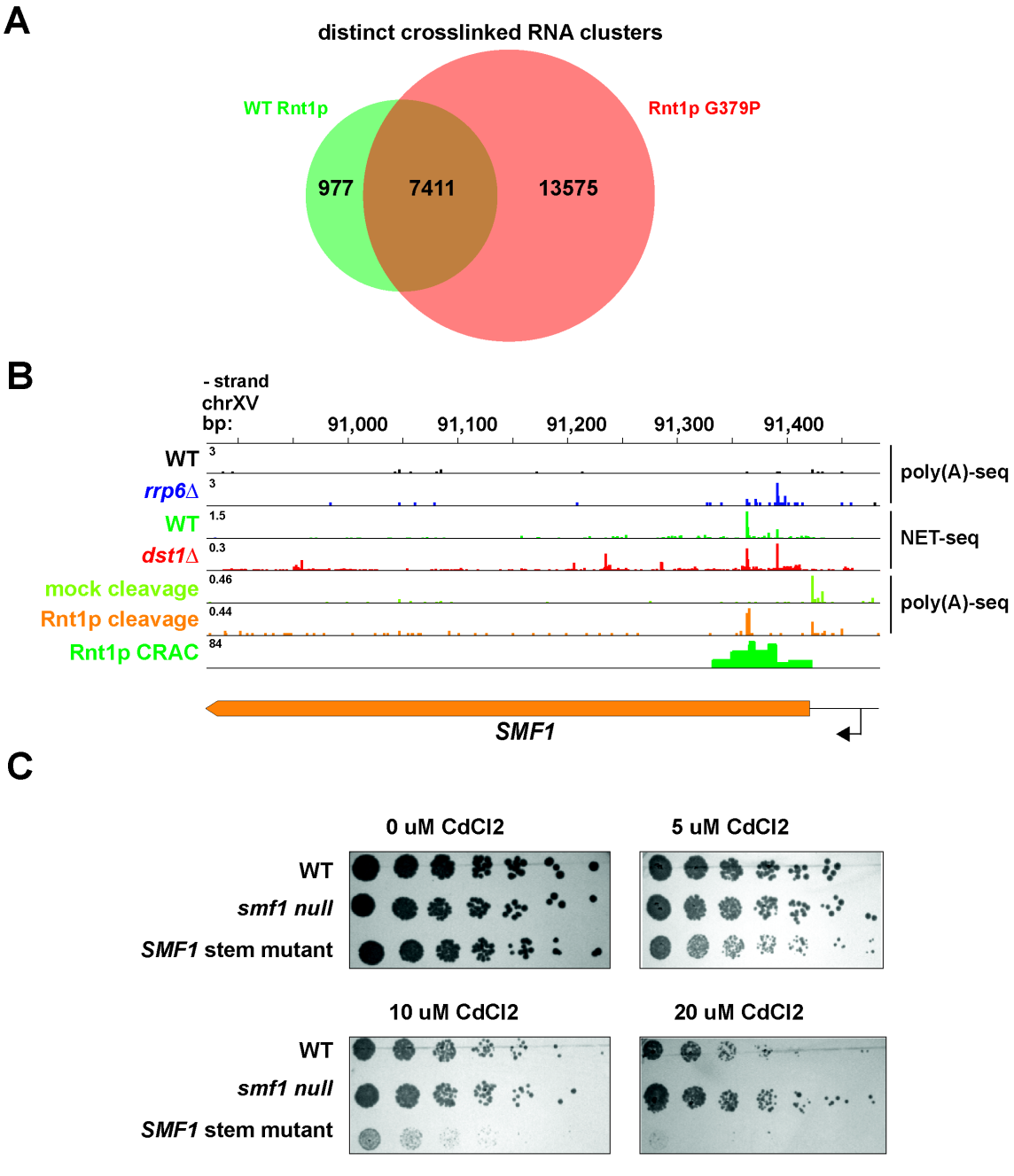
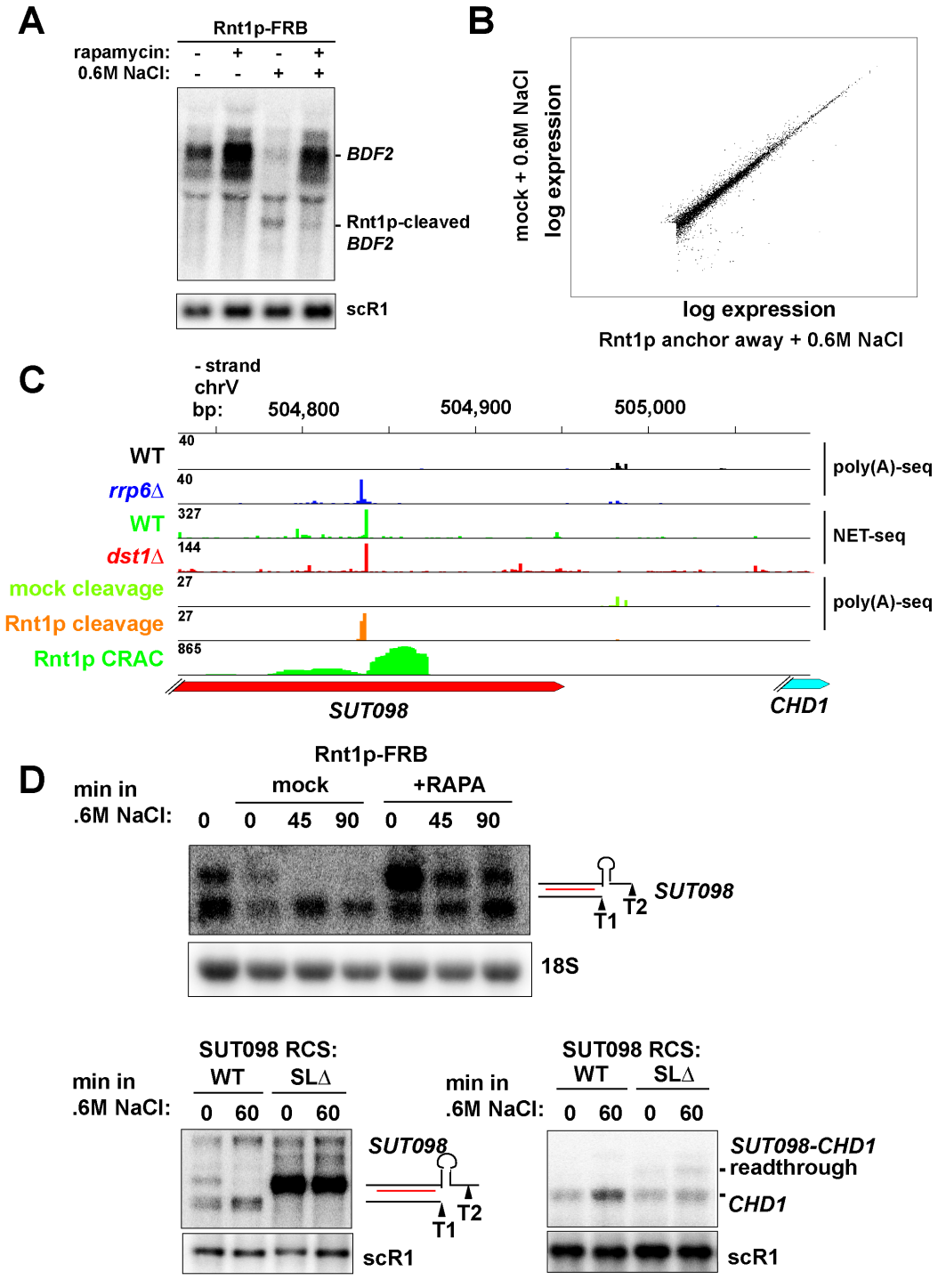

A
distinct crosslinked RNA clusters
WT Rnt1p
Rnt1p G379P
977
7411
13575
B
- strand
chrXV
bp:
91,000
91,100
91,200
91,300
91,400
WT
rrp6Δ
WT
dst1Δ
mock cleavage
Rnt1p cleavage
Rnt1p CRAC
poly(A)-seq
NET-seq
poly(A)-seq
SMF1
C
0 uM CdCl2
5 uM CdCl2
10 uM CdCl2
20 uM CdCl2
WT
smf1 null
SMF1 stem mutant

Figure 6.9


A
Rnt1p-FRB
rapamycin: - + - +
0.6M NaCl: - - + +
BDF2
Rnt1p-cleaved BDF2
scR1
B
mock + 0.6M NaCl log expression
log expression
Rnt1p anchor away + 0.6M NaCl
C
- strand
chrV
bp: 504,800 504,900 505,000
WT
rrp6Δ
WT
dst1Δ
mock cleavage
Rnt1p cleavage
Rnt1p CRAC
40
40
327
144
27
27
865
poly(A)-seq
NET-seq
poly(A)-seq
SUT098
CHD1
D
Rnt1p-FRB
min in .6M NaCl:
mock
+RAPA
0 0 45 90 0 45 90
SUT098
T1
T2
18S
SUT098 RCS:
WT SLΔ
min in .6M NaCl:
0 60 0 60
SUT098
T1
T2
scR1
SUT098 RCS:
WT SLΔ
min in .6M NaCl:
0 60 0 60
SUT098-CHD1 readthrough
CHD1
scR1

References

Allmang, C., Kufel, J., Chanfreau, G., Mitchell, P., Petfalski, E., and Tollervey, D. (1999). Functions of the exosome in rRNA, snoRNA and snRNA synthesis. Embo J. *18*, 5399–5410.

Bailey, T.L. (2011). DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics *27*, 1653–1659.

Berretta, J., and Morillon, A. (2009). Pervasive transcription constitutes a new level of eukaryotic genome regulation. EMBO Rep. *10*, 973–982.

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

Callahan, K.P., and Butler, J.S. (2010). TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. J. Biol. Chem. *285*, 3540–3547.

Carroll, K.L., Ghirlando, R., Ames, J.M., and Corden, J.L. (2007). Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. Rna *13*, 361–373.

Carroll, K.L., Pradhan, D.A., Granek, J.A., Clarke, N.D., and Corden, J.L. (2004). Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol. Cell. Biol. *24*, 6241–6252.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M.P., et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592.

Chanfreau, G., Legrain, P., and Jacquier, A. (1998). Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. J. Mol. Biol. *284*, 975–988.

Churchman, L.S., and Weissman, J.S. (2011). Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature *469*, 368–373.

Clark, M.B., Amaral, P.P., Schlesinger, F.J., Dinger, M.E., Taft, R.J., Rinn, J.L., Ponting, C.P., Stadler, P.F., Morris, K.V., Morillon, A., et al. (2011). The reality of pervasive transcription. PLoS Biol. *9*, e1000625–discussione1001102.

Colin, J., Candelli, T., Porrua, O., Boulay, J., Zhu, C., Lacroute, F., Steinmetz, L.M., and Libri, D. (2014). Roadblock termination by reb1p restricts cryptic and readthrough transcription. Mol. Cell *56*, 667–680.

Davis, C.A., and Ares, M.J. (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc. Natl.

Acad. Sci. U.S.a. *103*, 3262–3267.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Egecioglu, D.E., Henras, A.K., and Chanfreau, G.F. (2006). Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. Rna *12*, 26–32.

Fox, M.J., Gao, H., Smith-Kinnaman, W.R., Liu, Y., and Mosley, A.L. (2015). The exosome component Rrp6 is required for RNA polymerase II termination at specific targets of the Nrd1-Nab3 pathway. PLoS Genet. *11*, e1004999.

Ghazal, G., Gagnon, J., Jacques, P.-E., Landry, J.-R., Robert, F., and Elela, S.A. (2009). Yeast RNase III triggers polyadenylation-independent transcription termination. Mol. Cell *36*, 99–109.

Ghazal, G., Ge, D., Gervais-Bird, J., Gagnon, J., and Abou Elela, S. (2005). Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. Mol. Cell. Biol. *25*, 2981–2994.

Gietz, R.D., and Schiestl, R.H. (2007). High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc *2*, 31–34.

Govind, C.K., Qiu, H., Ginsburg, D.S., Ruan, C., Hofmeyer, K., Hu, C., Swaminathan, V., Workman, J.L., Li, B., and Hinnebusch, A.G. (2010). Phosphorylated Pol II CTD recruits multiple HDACs, including Rpd3C(S), for methylation-dependent deacetylation of ORF nucleosomes. Mol. Cell *39*, 234–246.

Granneman, S., Kudla, G., Petfalski, E., and 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.

Grove, A., Adessa, M.S., Geiduschek, E.P., and Kassavetis, G.A. (2002). Marking the start site of RNA polymerase III transcription: the role of constraint, compaction and continuity of the transcribed DNA strand. Embo J. *21*, 704–714.

Grzechnik, P., and Kufel, J. (2008). Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Mol. Cell *32*, 247–258.

Gudipati, R.K., Neil, H., Feuerbach, F., Malabat, C., and Jacquier, A. (2012). The yeast RPL9B gene is regulated by modulation between two modes of transcription termination. Embo J. *31*, 2427–2437.

Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. Cell *137*, 445–458.

Hartman, E., Wang, Z., Zhang, Q., Roy, K., Chanfreau, G., and Feigon, J. (2013). Intrinsic dynamics of an extended hydrophobic core in the S. cerevisiae RNase III dsRBD contributes to recognition of specific RNA binding sites. J. Mol. Biol. *425*, 546–562.

Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol. Cell *31*, 925–932.

Hesselberth, J.R., Chen, X., Zhang, Z., Sabo, P.J., Sandstrom, R., Reynolds, A.P., Thurman, R.E., Neph, S., Kuehn, M.S., Noble, W.S., et al. (2009). Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nat. Methods *6*, 283–289.

Ju, Q.D., Morrow, B.E., and Warner, J.R. (1990). REB1, a yeast DNA-binding protein with many targets, is essential for growth and bears some resemblance to the oncogene myb. Mol. Cell. Biol. *10*, 5226–5234.

Kim, T., and Buratowski, S. (2009). Dimethylation of H3K4 by Set1 recruits the Set3 histone deacetylase complex to 5' transcribed regions. Cell *137*, 259–272.

Korde, A., Rosselot, J.M., and Donze, D. (2014). Intergenic transcriptional interference is blocked by RNA polymerase III transcription factor TFIIIB in Saccharomyces cerevisiae. Genetics *196*, 427–438.

LaCava, J., Houseley, J., Saveanu, C., Petfalski, E., Thompson, E., Jacquier, A., and Tollervey, D. (2005). RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell *121*, 713–724.

Lapinaite, A., Simon, B., Skjaerven, L., Rakwalska-Bange, M., Gabel, F., and Carlomagno, T. (2013). The structure of the box C/D enzyme reveals regulation of RNA methylation. Nature *502*, 519–523.

Lee, C.Y., Lee, A., and Chanfreau, G. (2003). The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5'-end processing of S. cerevisiae box C/D snoRNAs. Rna *9*, 1362–1370.

Lubas, M., Christensen, M.S., Kristiansen, M.S., Domanski, M., Falkenby, L.G., Lykke-Andersen, S., Andersen, J.S., Dziembowski, A., and Jensen, T.H. (2011). Interaction profiling identifies the human nuclear exosome targeting complex. Mol. Cell *43*, 624–637.

Makino, D.L., Schuch, B., Stegmann, E., Baumgärtner, M., Basquin, C., and Conti, E. (2015). RNA degradation paths in a 12-subunit nuclear exosome complex. Nature *524*, 54–58.

Marini, A.M., Soussi-Boudekou, S., Vissers, S., and Andre, B. (1997). A family of ammonium transporters in Saccharomyces cerevisiae. Mol. Cell. Biol. *17*, 4282–4293.

Nikko, E., Sullivan, J.A., and Pelham, H.R.B. (2008). Arrestin-like proteins mediate ubiquitination and endocytosis of the yeast metal transporter Smf1. EMBO Rep. *9*, 1216–1221.

Ozsolak, F., and Milos, P.M. (2011). Transcriptome profiling using single-molecule direct RNA sequencing. Methods Mol. Biol. *733*, 51–61.

Pelechano, V., Wei, W., and Steinmetz, L.M. (2013). Extensive transcriptional heterogeneity revealed by isoform profiling. Nature *497*, 127–131.

Porrua, O., and Libri, D. (2015). Transcription termination and the control of the transcriptome: why, where and how to stop. Nat. Rev. Mol. Cell Biol.

Powers, R.W., Kaeberlein, M., Caldwell, S.D., Kennedy, B.K., and Fields, S. (2006). Extension of chronological life span in yeast by decreased TOR pathway signaling. Genes Dev. *20*, 174–184.

Rhee, H.S., and Pugh, B.F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. Cell *147*, 1408–1419.

Rondón, A.G., Mischo, H.E., Kawauchi, J., and Proudfoot, N.J. (2009). Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae. Mol. Cell *36*, 88–98.

Roy, K., and Chanfreau, G. (2014). Stress-Induced Nuclear RNA Degradation Pathways Regulate Yeast Bromodomain Factor 2 to Promote Cell Survival. PLoS Genet. *10*, e1004661.

Santos, J., Sousa, M.J., and Leão, C. (2012). Ammonium is toxic for aging yeast cells, inducing death and shortening of the chronological lifespan. PLoS ONE *7*, e37090.

Schuch, B., Feigenbutz, M., Makino, D.L., Falk, S., Basquin, C., Mitchell, P., and Conti, E. (2014). The exosome-binding factors Rrp6 and Rrp47 form a composite surface for recruiting the Mtr4 helicase. Embo J. *33*, 2829–2846.

Steinmetz, E.J., Conrad, N.K., Brow, D.A., and Corden, J.L. (2001). RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature *413*, 327–331.

Stuckey, S., and Storici, F. (2013). Gene knockouts, in vivo site-directed mutagenesis and other modifications using the delitto perfetto system in Saccharomyces cerevisiae. Meth. Enzymol. *533*, 103–131.

Supek, F., Supekova, L., Nelson, H., and Nelson, N. (1996). A yeast manganese transporter related to the macrophage protein involved in conferring resistance to mycobacteria. Proc. Natl. Acad. Sci. U.S.a. *93*, 5105–5110.

Tisseur, M., Kwapisz, M., and Morillon, A. (2011). Pervasive transcription - Lessons from yeast. Biochimie *93*, 1889–1896.

van Dijk, E.L., Chen, C.L., d'Aubenton-Carafa, Y., Gourvennec, S., Kwapisz, M., Roche, V., Bertrand, C., Silvain, M., Legoix-Né, P., Loeillet, S., et al. (2011). XUTs are a class of Xrn1-

sensitive antisense regulatory non-coding RNA in yeast. Nature *475*, 114–117.

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

Vanácová, S., Wolf, J., Martin, G., Blank, D., Dettwiler, S., Friedlein, A., Langen, H., Keith, G., and Keller, W. (2005). A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol. *3*, e189.

Wang, Q., Nowak, C.M., Korde, A., Oh, D.-H., Dassanayake, M., and Donze, D. (2014). Compromised RNA polymerase III complex assembly leads to local alterations of intergenic RNA polymerase II transcription in Saccharomyces cerevisiae. BMC Biol. *12*, 89.

Webb, S., Hector, R.D., Kudla, G., and Granneman, S. (2014). PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast. Genome Biol. *15*, R8.

Wyers, F., Rougemaille, M., Badis, G., Rousselle, J.-C., Dufour, M.-E., Boulay, J., Régnault, B., Devaux, F., Namane, A., Séraphin, B., et al. (2005). Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell *121*, 725–737.

Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. Nature *457*, 1033–1037.

Xue, Y., Van, C., Pradhan, S.K., Su, T., Gehrke, J., Kuryan, B.G., Kitada, T., Vashisht, A., Tran, N., Wohlschlegel, J., et al. (2015). The Ino80 complex prevents invasion of euchromatin into silent chromatin. Genes Dev. *29*, 350–355.

Zhou, J., Liang, B., and Li, H. (2011). Structural and functional evidence of high specificity of Cbf5 for ACA trinucleotide. Rna *17*, 244–250.