Biological and biochemical characterization of mutants of the RNase II family of enzymes

Ana Filipa de Melo Tadeu Pereira dos Reis

Dissertation presented to obtain the Ph.D. degree in Biology

Instituto de Tecnologia Química e Biológica
Universidade Nova de Lisboa

Oeiras, December, 2012

Financial Support from **Fundação para a Ciência e Tecnologia** (**FCT**) – Ph.D. grant SFRH/BD/43207/2008.

Work performed at:

Control of Gene Expression Laboratory

Instituto de Tecnologia Química e Biológica

Av. da República (EAN)

2781-901 Oeiras – Portugal

Tel: +351-21-4469548

Fax: +351-21-4469549

Orientador (Supervisor):

Doutora Cecília Maria Pais de Faria de Andrade Arraiano – Investigadora Coordenadora, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa.

Presidente (President of the Jury):

Doutora Claudina Amélia Marques Rodrigues Pousada - Professora Catedrática Convidada do Instituto de Tecnologia Química e Biológica da Universidade Nova de Lisboa, por delegação.

Vogais (Members of the Jury):

Doutor Phil J. Mitchell – Professor, University of Sheffield, Reino Unido.

Doutora Luísa Maria Ferreira Romão Loison – Investigadora Principal com Habilitação, Instituto Nacional de Saúde Dr. Ricardo Jorge.

Doutor Francisco Javier Enguita Lombardo – Professor Auxiliar da Faculdade de Medicina da Universidade de Lisboa.

Doutora Ana Lúcia Freitas de Mesquita Barbas Sant'Ana de Miranda – Investigadora do Instituto de Biologia Experimental e Tecnológica, Oeiras.

"The noblest pleasure is the joy of understanding."

- **Leonardo da Vinci**

Acknowledgments

I would like to thank all the people that directly and indirectly contributed and supported this work. This Ph.D. dissertation would not have been possible without the support of:

Fundação para a Ciência e Tecnologia (**FCT**) for the financial support crucial of my Ph.D. fellowship (SFRH / BD / 43207 / 2008).

Instituto de Tecnologia Química e Biológica (**ITQB**) for providing me the facilities and resources for the execution of this work.

Professor Cecília Arraiano for giving me the opportunity to start this big adventure, for all the support and endless enthusiasm. I am also thankful for all the opportunities to expose my work in several national and international meetings and for giving me the opportunity to spend some time working in other laboratories, which was very important to expand my scientific knowledge.

Ana Barbas for all the support during my Ph.D. and for giving me the confidence to perform the practical work without being afraid to fail.

Michal Malecki for his interest in my research, the fruitful work discussions and for reading this Dissertation.

To all the past and present colleagues and friends in the lab for making it such an enjoyable place to be and work at: **Ana, Andreia, Cátia, Clémentine, Inês Guinote, Inês Silva, Joana, Margarida Matos, Margarida Saramago, Michal, Mauro, Patrick, Raquel Santos, Ricardo Moreira, Ricardo Santos, Rute, Sandra, Susana Barahona, Susana Domingues, Teresa Pinto, Vânia** and **Zé**.

Professor Ambro van Hoof and his group at Department of Microbiology and Molecular Genetics, University of Texas Health Science Center – Houston, Texas, for kindly receiving me in his laboratory, and for all the guidance and comments since the beginning of this work. A special thanks to **Ale Klauer** and **Daneen Schaeffer** for their help and support during my stay there.

Thank you to our collaborators in Centro de Biologia Molecular "Severo Ochoa" (CSIC-UAM), Campus Universidad Autonoma de Madrid, Spain, **Dr. Paulino Gómez-Puertas** and **Dr. Eduardo López-Viñas**.

To all my colleagues in the other labs in the Institute. Thank you for having participated in this work by discussing science during coffee breaks and lunch. A special thanks to **Dr. Collin McVey**, for helping and advising me in all the questions I could have about proteins.

Thank you to my **Mom** for being unconditionally supportive.
Thank you to my **Dad**, an absence that is always present. We have had some Victories together!
Thank you to my **brother** and my sister in law, **Sara,** for giving me the right words and incentives to follow my path.

Thank you to all my **friends** and **family**, Grandparents, Uncles, Aunts and Cousins. It is impossible to acknowledge you all, one by one. A special thanks to my **Aunt Teresa**, **Uncle Nuno** and **Aunt Gaby**, who have always shown a special interest in my studies.

Thank you all!

Table of Contents

Abstract

The cellular concentration of a given RNA is the result of the balance between its synthesis and degradation. RNA degradation plays a fundamental role in RNA metabolism, since the fast turnover of mRNA is important for the control of gene expression. One of the key components of RNA metabolism are enzymes from the RNase II family. Members of this family are highly processive hydrolytic enzymes that play important roles in RNA decay, processing and quality control pathways. RNase II is the prototype of this family of exoribonucleases and RNase II homologues can be found in all domains of life. Members of this family share sequence identity, a similar domain organization and a highly conserved domain responsible for exonucleolytic activity, the RNB domain. While in *Escherichia coli*, members of this family function as individual proteins, in yeast some of them function in the context of a multiprotein complex called the exosome.

The first objective of this PhD work was to characterize a domain at the N-terminus of Rrp44, whose function was unknown in the beginning of this study. In order to address the specific role of this domain in the activity of the enzyme, Rrp44 truncated mutants were constructed and their ribonucleolytic activity analyzed. A mutant containing only the PIN domain allowed us to demonstrate the presence of an additional endonuclease active site in Rrp44, whose activity is abolished when we mutate a conserved acidic residue important for catalysis. Moreover, we were able to further characterize this activity since we demonstrated that it is sequence independent, but has a preference for 5′ monophosphorylated substrates. Decades after the discovery of the major 3′-5′ exonuclease activity in eukaryotes, this work has demonstrated the presence of an endonuclease active site in Rrp44, and consequently, in the exosome. These findings have serious implications on the mode of action of the exosome, namely the mechanism of recruitment and degradation of RNA, forcing the model of the exosome to be re-evaluated.

In the second part of this Doctoral project, the aim was to study the specific role of the CR3 motif at the N-terminus of Rrp44. This region had been previously identified in *Drosophila melanogaster*, but its function was unknown. Previous data suggest that this region is needed for a large number of exosome functions. Also, previous structural data suggest that it could coordinate a metal ion. To evaluate the role of the N-terminus, a set of mutations in this region were constructed. These Rrp44 mutant proteins were compared to the wild-type protein regarding their endoribonucleolytic activity. These results, coupled with genetic analysis, demonstrated that this region coordinates a metal ion and regulates Rrp44 endonucleolytic activity and its interaction with the exosome. Additionally, these results helped to explain why mutating the CR3 region results in numerous exosome defects, both in the nucleus and the cytoplasm.

In the third part of this work, we focused on the functional characterization of three residues located in the highly conserved RNB catalytic domain: Tyr595, Gln892 and Gly895. To address experimentally their function we carried out point mutations and tested their role in Rrp44 activity. When we mutated residues Gln892 and Tyr595 and tested their activity *in vitro*, the enzymes became 2 fold more active. We also showed that when we mutate Tyr595, the final degradation product of Rrp44 changed from 4 to 6 nucleotides. This result confirms that this residue is responsible for the stacking of the RNA substrate in the catalytic cavity, as was predicted by the structure of Rrp44. Furthermore, we also showed that a strain with a mutation in this residue has a growth defect and affects RNA processing and degradation. This result led us to hypothesize that Tyr595 has an important biological role. This study sheds light on the model previously proposed for Rrp44 activity and mode of action, and gives new insights in RNA processing and degradation by the exosome. Since Rrp44 is the only catalytic subunit of the eukaryotic RNA exosome,

understanding the activity of Rrp44 is essential to understanding the activity of the exosome.

To finalize this study we also studied the prototype of this family of exoribonucleases, *E. coli* RNase II. A thermosensitive allele that was generated *in vitro*, *rnb500*, has been essential for the determination of the cellular function of RNase II. However, little was known about the mutation responsible for such phenotype. During this work we have sequenced the DNA from the *rnb500* allele and identified two substitutions in the sequence of this allele. To determine which mutation is responsible for the thermolability of this allele, we introduced each point mutation individually or both substitutions in the *rnb* gene. RNase II mutant proteins were compared to the wild-type enzyme regarding their exoribonucleolytic activity at the non-permissive temperature. Furthermore, we also performed *in vivo* complementation assays using a temperature-sensitive *E. coli* strain. These assays helped us to identify the mutation responsible for this phenotype - Cys284Tyr. This work also unravelled an unpredicted disulphide bridge in the RNase II structure, which is responsible for a relevant covalent interaction between residues Cys284 and Cys399. The localization of these residues is conserved in other organisms, which led us to propose that the knowledge of these mutations can be an useful tool for studying RNase II enzymes function in other organisms.

The results of this Dissertation allowed for the functional characterization of the N-terminal region of *Saccharomyces cerevisiae* Rrp44 enzyme, namely the PIN domain and the CR3 region. These studies also contribute to a better understanding of Rrp44 activity and mode of action and, consequently, the involvement of the eukaryotic exosome in several processes of RNA metabolism. Moreover, the results obtained from this Dissertation can be applied to better

understand the decay mechanisms in other members of this family with a similar structure and mode of action.

Resumo

A concentração celular de RNA é o resultado do balanço entre a sua síntese e degradação. A degradação de RNA desempenha um papel fundamental no metabolismo de RNA, uma vez que a reciclagem rápida de mRNA é importante para o controlo da expressão génica. Um dos componentes chave do metabolismo de RNA são enzimas da família da RNase II. Os membros desta família são enzimas processivas e hidrolíticas que desempenham papéis importantes na degradação, processamento e controlo de qualidade do RNA. A enzima RNase II é o protótipo desta família de exoribonucleases e enzimas homólogas podem ser encontradas em todos os domínios da vida. Membros desta família têm uma sequência e estrutura semelhantes, e partilham um domínio catalítico característico, o domínio RNB. Enquanto que em *Escherichia coli*, os membros desta família de proteínas funcionam individualmente, em levedura algumas delas funcionam no contexto de um complexo multiproteico para a degradação do RNA – o exossoma.

O primeiro objetivo deste trabalho de Doutoramento foi caracterizar um domínio de função desconhecida na proteína Rrp44. A sua atividade ribonucleolítica foi analisada, a fim de caracterizar o papel deste domínio na atividade da enzima. Mutantes truncados da proteína Rrp44 foram construídos. Um mutante contendo apenas o domínio PIN permitiu-nos demonstrar a presença de um sítio ativo com atividade endonucleolítica na proteína Rrp44. Quando um resíduo conservado importante para a catálise é mutado a atividade é abolida. Adicionalmente, demonstrámos que a atividade deste domínio não depende da sequência do substrato mas tem uma preferência para substratos 5' monofosforilados. Este trabalho demonstrou pela primeira vez a presença de um sítio ativo com atividade endonucleolítica na proteína Rrp44 e, consequentemente, no exossoma. Estes resultados têm implicações importantes

no modo de ação do exossoma, nomeadamente no mecanismo de recrutamento e degradação de RNA, forçando o modelo do exossoma a ser reavaliado.

Na segunda parte do presente projeto de Doutoramento, o objetivo foi estudar o papel específico do motivo CR3 no terminal N da proteína Rrp44. Esta região tinha sido previamente identificada em *Drosophila melanogaster* mas a sua função era desconhecida. Dados anteriores sugeriram que esta região é necessária para um grande número de funções do exossoma. Além disso, dados estruturais da proteína sugerem que esta poderia coordenar um ião metálico. Foram construídas várias mutações nesta região para avaliar o papel do N-terminal no funcionamento da enzima. A atividade endoribonucleolítica das proteínas mutantes foi comparada com a proteína do tipo selvagem. Estes resultados, juntamente com a análise genética, demonstraram que esta região coordena um ião de metal, regula a atividade endonucleolítica da proteína Rrp44 e a sua interação com o exossoma. Além disso, estes resultados ajudam a explicar porque é que uma mutação nesta região ocasiona problemas na degradação pelo exossoma, tanto no núcleo como no citoplasma.

Na terceira parte deste trabalho, debruçámo-nos na caracterização funcional de três resíduos localizados no domínio catalítico altamente conservado, RNB: Tyr595, Gln892 e Gly895. Para abordar experimentalmente a sua função realizámos mutações pontuais e testámos o seu papel na atividade da enzima Rrp44. Quando mutámos os resíduos Gln892 e Tyr595 e testámos a sua atividade *in vitro*, as enzimas tornaram-se 2 vezes mais ativas. Mostrámos também que quando se altera o resíduo Tyr595, o produto da degradação final muda de 4 para 6 nucleótidos. Este resultado confirma que este resíduo é responsável pelo posicionamento do RNA na cavidade catalítica, tal como foi previsto na estrutura da proteína Rrp44. Além disso, também mostrámos que

uma estirpe com uma mutação neste resíduo tem defeitos no crescimento e afeta o processamento e degradação do RNA. Este resultado levou-nos a sugerir que o resíduo Tyr595 tem um papel biológico importante. Estes resultados esclarecem o modelo anteriormente proposto para a atividade e modo de ação da enzima Rrp44, e dá novas perspetivas no processamento e degradação de RNA pelo exossoma. Uma vez que a proteína Rrp44 é a única subunidade catalítica, a compreensão da atividade da enzima Rrp44 é essencial para a compreensão da atividade do exossoma.

Para finalizar este estudo, também estudámos o protótipo desta família de exoribonucleases, a enzima RNase II de *E. coli*. Um alelo termosensível produzido por mutagénese (*rnb500*) tem sido essencial para a determinação da função celular da proteína RNase II. No entanto, pouco se sabe sobre a mutação responsável por tal fenótipo. No âmbito deste trabalho, sequenciámos o DNA deste alelo e identificámos duas substituições na sequência do alelo *rnb500*. Para determinar qual a mutação responsável pela termolabilidade deste alelo, introduzimos cada mutação pontual individualmente ou ambas as substituições no gene *rnb*. A atividade exoribonucleolítica de proteínas mutantes da RNase II foi comparada com a da enzima do tipo selvagem, à temperatura não permissiva. Além disso, também realizámos ensaios de complementação *in vivo* utilizando a estirpe termosensível de *E. coli*. Estes ensaios ajudaram-nos a identificar a mutação responsável por este fenótipo - Cys284Tyr. Este trabalho também desvendou uma ponte dissulfídrica não prevista na estrutura da proteína RNase II, a qual é responsável por uma interação covalente entre os resíduos Cys284 e Cys399. A localização desses resíduos é conservada em outros organismos, o que nos levou a propôr que o conhecimento dessas mutações pode ser uma ferramenta útil para estudar enzimas homólogas da RNase II em outros organismos.

Os resultados desta Dissertação permitiram a caracterização funcional da região N-terminal enzima Rrp44 de *Saccharomyces cerevisiae*, isto é, o domínio PIN e a região CR3. Estes estudos também contribuiram para uma melhor compreensão da atividade da enzima Rrp44 e seu modo de ação. Consequentemente, avançámos na compreensão do papel do exossoma eucariótico em diversos processos de metabolismo do RNA. Além disso, os resultados obtidos a partir desta Dissertação podem ser aplicados para uma melhor compreensão dos mecanismos de degradação em outros membros desta família, com uma estrutura e modo de ação semelhantes.

List of Publications

Scientific articles in international peer reviewed journals

Reis FP, López-Viñas E, Gómez-Puertas P and Arraiano CM. Identification of the temperature-sensitive mutation that impairs the function *in vivo* and *in vitro* of RNase II. *Manuscript in preparation*.

Reis FP, Barbas A, Klauer AA, Schaeffer D, López-Viñas E, Gómez-Puertas P, van Hoof A and Arraiano CM. Modulating the RNA processing and decay by the exosome: altering Rrp44/Dis3 activity and end-product. *Submitted*.

Reis FP, Pobre V, Silva IJ, Malecki M, and Arraiano CM. The RNB Family of Exoribonucleases – Putting the 'Dis' in Disease. *Wiley Interdisciplinary Reviews (WIREs) RNA. Accepted*.

Schaeffer D, **Reis FP**, Johnson SJ, Arraiano CM, van Hoof A. 2012. The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. *Nucleic Acids Res 40*:9298-9307.

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, **Reis FP**, Saramago, M, Silva, IJ and Viegas, S.C. 2010. The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev 34*:883-923.

Schaeffer D, Tsanova B, Barbas A, **Reis FP**, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature Structural & Molecular Biology 16*:56-62.

Domingues S, Matos RG, **Reis FP**, Fialho AM, Barbas A and Arraiano CM. 2009. Biochemical characterization of the hydrolytic exoribonucleases from the human pathogens *Salmonella typhimurium* and *Streptococcus pneumonia*. *Biochemistry 48*:11848-57.

Book chapters

Matos RG, Pobre V, **Reis FP**, Malecki M, Andrade JM and Arraiano CM. 2011. Structure and Mechanisms of 3′ to 5′ Exoribonucleases. In: Nicholson AW, ed. *Ribonucleases*: Springer Berlin Heidelberg. pp 193-222.

Dissertation outline

This Dissertation is divided into six chapters.

Chapter one consists of a general introduction to contextualize the reader in the subject of this Dissertation. It includes a brief overview of the degradation of RNA, with a particular emphasis in *Escherichia coli* RNase II and *Saccharomyces cerevisiae* Rrp44, the main focus of this Dissertation.

The results of this Doctoral work are presented in the following chapters. Each chapter has its own Introduction, Results, Discussion, Materials and Methods, References and Acknowledgments sections.

Chapter two consists of an article published in the *Nature Structural & Molecular Biology* journal in which the author of this Dissertation played an important contribution. In this chapter several mutant proteins of Rrp44, including a truncated protein containing the PIN domain but lacking the RNB domain, were constructed, purified and their ribonucleolytic activity analyzed. This work led to important conclusions regarding the role of the PIN domain.

In **Chapter three** the role of the CR3 region at the N-terminus was studied. Biochemical data and genetic analysis showed that this region of Rrp44 forms a CR3 motif of unpredicted functions and characterized its mechanistic role. This chapter is presented as a manuscript that was published in the *Nucleic Acid Research* journal. The author of this Dissertation played an important contribution.

Chapter four characterizes the role of specific residues in Rrp44 active site. Based on previous studies in the corresponding residues in *E. coli* RNase II, several residues were mutated. The mutations examined were Y595A, G895A, G895E, G895Q and Q892A in yeast Rrp44, and the respective mutants were studied *in*

vitro and *in vivo*. The analysis of these Rrp44 mutants brought new insights and improved the RNA degradation model previously proposed. The author of this Dissertation played a major contribution.

Chapter five focuses on the thermosensitive *E. coli rnb500* allele. Sequence analysis, biochemical *in vitro* assays and *in vivo* complementation assays at the nonpermissive temperature, led to the determination of the mutation responsible for this phenotype. The author of this Dissertation played a major contribution.

Chapter six is an integrated Discussion and the Future Perspectives based on the results presented in the previous chapters.

Abbreviations

5-FOA 5-fluoroorotic acid
A adenine
Å angstrom
Amp ampicillin
ATP adenosine triphosphate
bp base pair
°C degree Celsius
C cytosine
CDS cold shock domain
Cm chloramphenicol
cpm counts per minute
CR3 cysteine-rich with three cysteines
CUTs cryptic unstable transcripts
Δ deletion
Da dalton
DTT dithiothreitol
DNA deoxyribonucleic acid
E. coli *Escherichia coli*
EDTA ethylenediaminetetraacetic acid
eEF eukaryotic translation elongation factor
EF-Tu elongation factor thermo unstable
eRF eukaryotic release factor
ETS external transcribed spacer
G guanine
g relative centrifugal force
Gal galactose
GST glutathione S-transferase
h hour
HEPES hydroxyethyl piperazineethanesulfonic acid
His histidine
IPTG IsoPropyl-β-D-thiogalactopyranoside
Kb kilobase
kDa kilodalton
L liter
LA luria agar
LB luria- bertani broth
M molar/molarity (mol/L)
mg milligram
µg microgram
µl microliter
µM micromolar
Mg magnesium
ml milliliter
min minute
mM milliMolar
mmol millimole
Mn manganese
Mol mole
Mpp6 M-phase phosphoprotein 6
mRNA messenger RNA
MW molecular weight
nM nanoMolar
nmol nanomole
nt nucleotide
OB oligonucleotide binding
OD optical density
oligo oligonucleotide
o.n. over night
^{32}P phosphorus 32 radionucleotide
PAA polyacrylamide
PAGE polyacrylamide gel electrophoresis
PAP I Poly(A) Polymerase I
PCR polymerase chain reaction
PIN PilT N-terminus
pmol picomol
PMSF phenylmethanesulfonylfluoride
PNPase polynucleotide phosphorylase
Poly(A) polyadenylate
rcf relative centrifugal force

RNA ribonucleic acid

RppH RNA pyrophosphohydrolase

rRNA ribosomal RNA

snRNA small nuclear RNA

snoRNA small nucleolar RNA

ssRNA single stranded RNA

dsRNA double stranded RNA

RNase ribonuclease

rpm rotations per minute

rRNA ribosomal RNA

Rrp ribosomal RNA-processing protein

s second

SC synthetic complete

S. cerevisiae *Saccaromyces cerevisiae*

SDS sodium dodecyl sulphate

SRP signal recognition particle

T thymine

Tris trishydroxymethylaminomethane

tRNA transfer RNA

ts thermosensitive

U uracil

UV ultraviolet radiation

vol volume

v/v volume/volume

wt wild-type

w/v weight/volume

YEP yeast-extract peptone

YPD yeast peptone dextrose

Zn zinc

Chapter 1

Introduction

Part of this chapter was based on:
Reis FP, Pobre V, Silva IJ, Malecki M, and Arraiano CM. The RNB Family of Exoribonucleases – Putting the 'Dis' in Disease. *Wiley Interdisciplinary Reviews (WIREs) RNA. Accepted*.

and

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, **Reis FP**, Saramago M, Silva IJ and Viegas SC. 2010. The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev 34*:883-923.

Introduction

The continuous synthesis (transcription) and degradation of RNA allow the metabolic changes that are required for cells to grow, divide and adapt to new environmental conditions. Transcription is important to determine RNA steady-state levels, but the control of RNA stability has been shown to be critical for the regulation of gene expression. RNA degradation is essential for the turnover of mRNA, but is also important for RNA surveillance and processing (Houseley & Tollervey, 2009). Incorrectly synthesized RNAs that would be detrimental for the cell have to be removed. In this way RNA surveillance plays a key role in the fidelity of gene expression (Fig. 1).

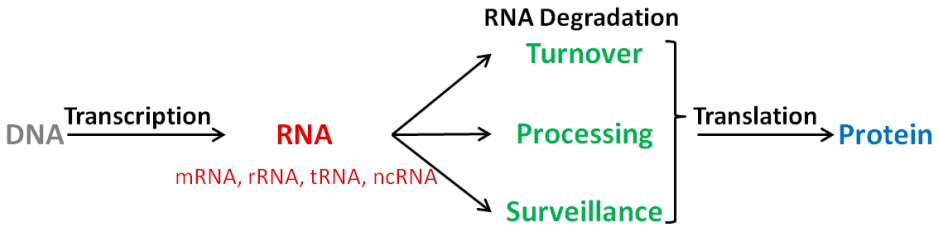


Figure 1. Schematic representation of the involvement of RNA degradation in the control of gene expression. Steady-state levels of RNA in the cells are the result of DNA transcription and RNA turnover. However, the control of gene expression is also dependent on the existence of functional RNAs (RNA processing) and on the fidelity of gene expression (RNA surveillance).

RNA Degradation

Ribonucleases

Ribonucleases (RNases) are the major players of RNA decay processes. However, they are usually assisted by conserved ancillary mRNA-modifying enzymes, namely, RNA helicases and poly(A) polymerases.

RNases can act alone or they can cooperate in RNA degradation complexes. Many of them are essential but others exhibit a functional overlap and are interchangeable, making redundancy a general feature. Many of them are also multifunctional, since they can precisely process some RNA species and completely degrade others.

In terms of mode of action they can be divided into endonucleases, which cleave the RNA internally, or exonucleases, which cleave the RNA from one of the extremities. They can be distributive, releasing the substrate at each catalytic cycle, or processive, remaining bound to the substrate that they degrade. Ribonucleases can act hydrolytically if they use a water molecule to attack the RNA phosphodiester bond releasing nucleotide monophosphates or phosphorolytically, if they use inorganic phosphate to break the RNA backbone releasing mononucleotide 5' diphosphate products.

Model of mRNA Degradation in Prokaryotes

The conventional pathway for RNA degradation in *Escherichia coli* is initiated with an endonucleolytic cleavage at one or more internal sites on the RNA molecule by the action of endoribonuclease (Apirion, 1973), RNase E or RNase III (Fig. 2). While RNase E cleaves single-stranded RNA (ssRNA) (Mackie, 1992), RNase III cleaves double-stranded RNA (dsRNA) (Robertson et al., 1968). RNase E has a significant preference for 5′-monophosphate over the triphosphorylated extremities of primary transcripts (Mackie, 1998). RNA pyrophosphohydrolase (RppH) enzyme facilitates RNase E cleavage by releasing a pyrophosphate from the 5′ end of triphosphorylated transcript, liberating a 5′ monophosphate molecule (Deana et al., 2008). After endoribonucleolytic cleavages, decay intermediates are rapidly degraded by 3′–5′ exoribonucleases, RNase II, RNase R and PNPase (Coburn & Mackie, 1998; Kushner, 2002). The small oligoribonucleotides released by exoribonucleases, two to five nucleotides, are

finally degraded to mononucleotides by oligoribonuclease (Ghosh & Deutscher, 1999).

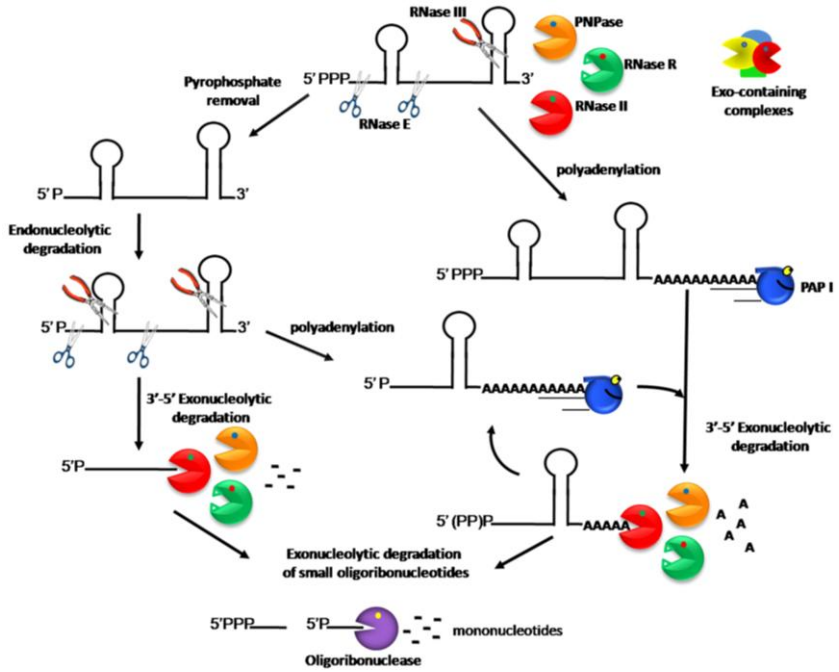


Figure 2. Model of mRNA degradation in *Escherichia coli*. Decay is initiated by one or more endonucleolytic cleavages, followed by the complete degradation of the substrate by exonucleases to oligo- and mononucleotides. The addition of poly(A) tails helps the digestion of structured RNAs. Adapted from Arraiano et al. (2010).

In cases where secondary structures exist, poly(A) polymerase (PAP I) has an important role. The addition of poly(A) tails to the short 3' overhang provides a platform to which exoribonucleases can bind, increasing the susceptibility of RNA to degradation (Coburn & Mackie, 1998). Cycles of polyadenylation and exoribonucleolytic digestion can overcome RNA secondary structures. Also, RNase R, unlike RNase II and PNPase (Spickler & Mackie, 2000), is efficient against highly structured RNAs (Cheng & Deutscher, 2005), provided a 3' toehold is available. Alternatively, the association of PNPase with a DEAD-box RNA helicase RhlB, among other proteins, facilitates the unwinding and

degradation of RNA duplexes (Coburn & Mackie, 1998; Khemici & Carpousis, 2004).

In bacteria, the exoribonucleolytic activity was believed to be just in the 3′ to 5′direction (Zuo & Deutscher, 2001). However, this dogma was broken, since it was observed that there is an enzyme in *Bacillus subtilis* (RNase J1) with both 5′ to 3′ exonucleolytic and endonucleolytic activity (Mathy et al., 2007).

Model of mRNA Degradation in Eukaryotes

Eukaryotes and prokaryotes have evolved distinct RNA degradation pathways. RNA degradation in eukaryotes is much more complex and involves more factors than those in prokaryotes. The eukaryotic cell is divided into two main parts, the nucleus and the cytoplasm, and RNA degradation takes place in both compartments. Compartmentalization causes considerable change in the fate of mRNA: eukaryotic mRNAs have to survive in the cell much longer than prokaryotic messengers, and the molecule synthesized in the nucleus has to be transported to the cytoplasm for protein production.

The predominant cytoplasmic mRNA decay pathway is the deadenylation-dependent RNA degradation (Fig. 3). mRNA decay is initiated with the shortening of the poly(A) tail at the 3′ end by deadenylation enzymes (Muhlrad & Parker, 1992; Decker & Parker, 1993). In yeast, two deadenylase complexes were described, the major Ccr4/NOT complex and Pan2/Pan3 (Brown & Sachs, 1998; Tucker et al., 2001). After deadenylation, there are two possible degradation pathways for the transcripts. One is the removal of the 5′ cap structure of the transcripts by the Dcp1/Dcp2 decapping complex, exposing the RNA molecule to the cytoplasmic Xrn1 5′-3′ exoribonuclease (Hsu & Stevens, 1993; Muhlrad et al., 1994). The other pathway is the 3′-5′ exoribonucleolytic degradation by a large protein complex, termed the exosome (Anderson & Parker, 1998; Wang & Kiledjian, 2001). The 5′ cap from the remaining

oligonucleotide is degraded by the scavenger decapping enzyme DcpS (Liu et al., 2002).

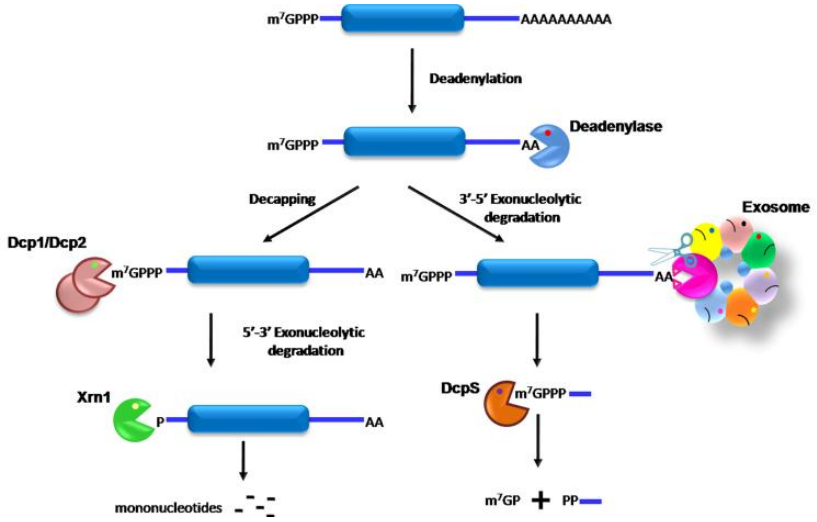


Figure 3. Model of mRNA degradation in eukaryotes. In general, decay is initiated by deadenylation and is followed by one of two pathways: the complete degradation of the substrate by decapping and 5′-3′ exonucleases or 3′-5′ degradation by the exosome. Adapted from Arraiano et al. (2010).

Thus far, we have described the general mRNA decay pathway in eukaryotes. Nevertheless, there are specialized mRNA decay pathways, namely, deadenylation independent decapping (Muhlrad & Parker, 1994), accelerated 3′ to 5′ degradation (van Hoof et al., 2002) and endonuclease cleavage (Doma & Parker, 2006).

RNase II Family of Enzymes

Enzymes from the RNase II family are important in RNA processing, quality control and degradation. Enzymes of this family are sequence-independent hydrolytic exoribonucleases that processively digest RNA in the 3′ to 5′ direction, yielding 5′-nucleoside monophosphates (Zuo & Deutscher, 2001).

This family of enzymes is present in all domains of life and has important roles in the cell. *E. coli* has two members, RNase II and RNase R, which together

with phosphorolytic nuclease PNPase, are responsible for the majority of exonucleolytic RNA degradation (Andrade et al., 2009; Matos et al., 2011). The best studied member in eukaryotes is Rrp44 (also called Dis3), the catalytic component of the exosome (Liu et al., 2006; Dziembowski et al., 2007). Members of this family have been described to be essential for growth (Mitchell et al., 1997), involved in mitotic control (Lim et al., 1997) and cancer (Chapman et al., 2011; Rose et al., 2011; Astuti et al., 2012; Ding et al., 2012; Reis et al., 2012), and required for virulence in several organisms (Cheng et al., 1998; Erova et al., 2008).

Functional Domains

The crystal structures of *E. coli* RNase II and of a catalytically inactive mutant in complex with RNA, revealed the presence of four domains (Frazão et al., 2006; Zuo et al., 2006). Three of these domains have a common oligonucleotide binding-fold (OB fold) characteristic of nucleic-acid-binding proteins (Murzin, 1993), two cold-shock domains (CSD1 and CSD2) and an S1 RNA-binding domain, that are grouped on one side of the 3D structure of the molecule. The well-conserved catalytic domain (RNB domain) is on the other side of the molecule (Frazão et al., 2006). The function of each domain had been previously determined and it was shown that the RNB domain is responsible for the catalytic activity, whereas CSD and S1 domains are responsible for the binding to RNA (Amblar et al., 2006).

In addition to the domains described for *E. coli* RNase II and RNase R (RNase II-like region), Rrp44/Dis3 contains an additional domain at the N-terminus (Fig. 4). Our group and two others have independently shown that the PIN domain of yeast Rrp44 confers endoribonuclease activity to this enzyme (Lebreton & Seraphin, 2008; Schaeffer et al., 2009; Schneider et al., 2009). This domain also plays an important structural role by establishing a link of yeast Rrp44 to the exosome (Bonneau et al., 2009; Schneider et al., 2009).

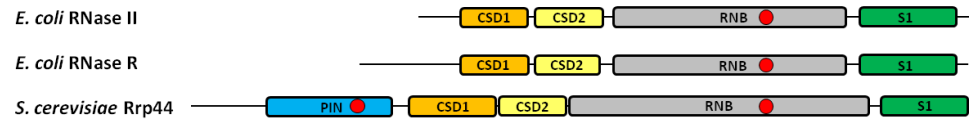


Figure 4. Domain organization of RNase II, RNase R and yeast Rrp44. Members of this family have a similar modular domain organization. The N-terminal region is variable but shares CSD1 (orange) and CSD2 (yellow), followed by a RNB domain (light cyan) and an S1 domain (green). At the N-terminus, yeast Rrp44 also contains a catalytic PIN domain (light blue). Ribonucleolytic active sites are indicated (red circle). Adapted from Chlebowski et al. (2010).

Two crystallographic structures of *Saccharomyces cerevisiae* Rrp44 have been solved: the structure of a catalytically inactive mutant of Rrp44 lacking the PIN domain in complex with RNA (Lorentzen et al., 2008) and the full length Rrp44 in complex with two other subunits of the exosome (Bonneau et al., 2009). The individual domains of Rrp44 are similar in structure to the corresponding domains of RNase II, with 85% of the residues of the RNB domain and more than 70% of the three OB-fold domains superposing in their α-carbon positions. The most striking difference between Rrp44 and RNase II lies in the relative position of the binding domains to the catalytic domain, with implications on the route of RNA access to the catalytic site (Lorentzen et al., 2008).

Exonuclease Activity

The exonucleolytic active site of the RNase II family of enzymes resides in the RNB domain (Amblar & Arraiano, 2005). This domain is only present in the members of this family and contains four highly conserved motifs (I to IV; Mian, 1997). Crystal structures of *E. coli* RNase II (Fig. 5) and *S. cerevisiae* Rrp44 in complex with RNA shed light on the mechanisms of RNA cleavage, translocation and enzyme processivity, as discussed below (Frazão et al., 2006; Lorentzen et al., 2008). The structure of RNase R was not solved yet. Thus, most of the knowledge on RNase R structure and mode of action was inferred from biochemical studies,

RNase II/Rrp44 structures and from models proposed (Barbas et al., 2008; Domingues et al., 2009).

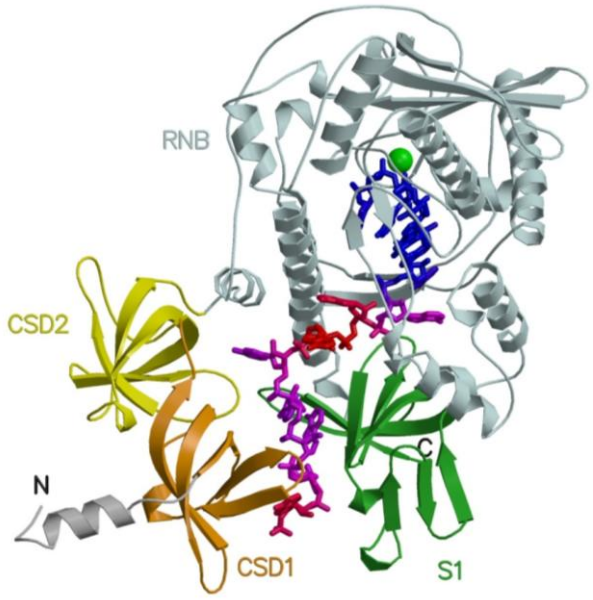


Figure 5. Structure of RNase II in complex with RNA. Representation of RNase II–RNA complex, showing domains CSD1 (orange), CSD2 (yellow), RNB (light cyan) and S1 (green), the Mg^{2+} ion as a sphere (green) and RNA in sticks. Adapted from Frazão et al. (2006).

E. coli RNase II and *S. cerevisiae* Rrp44 active sites lie buried deep within the RNB domain, where four conserved acidic residues coordinate two magnesium ions that facilitate a two-metal ion catalysis mechanism (Steitz & Steitz, 1993; Frazão et al., 2006; Lorentzen et al., 2008). Mutation of one of those conserved residues abolishes the exonucleolytic activity of the enzyme: D209N in RNase II (Amblar & Arraiano, 2005), D280N in RNase R (Matos et al., 2009) and D551N in yeast Rrp44 (Dziembowski et al., 2007).

Single-stranded RNA is threaded to the catalytic cavity through the anchor region, constituted by the RNA-binding domains: CSD2 and S1 domains in RNase II and CSD1 and RNB domains in Rrp44. The first five nucleotides counting from the 3' end of the substrate are stacked and 'clamped' between two

residues of the enzyme that stabilize the RNA in the active site, promoting its translocation between successive cleavage events (Frazão et al., 2006; Lorentzen et al., 2008). When the size of the RNA is not sufficient for the packing of the bases in the catalytic cavity, the translocation of the RNA stops and the final end product is released. In RNase II, residue Glu542 is in close proximity with the leaving nucleotide, and it was proposed that it helps to release the last nucleotide after cleavage (Frazão et al., 2006). However, its substitution by an alanine induced a subtle conformational change in the RNB domain, as predicted by computational modeling, leading to a 110-fold increase in RNase II activity (Barbas et al., 2009). This result shows that Glu542 is involved in the tuning of RNase II activity.

Processivity of RNB family enzymes is achieved by the simultaneous contacts of RNA at two different regions of the enzyme: the catalytic region and the anchor region (Cannistraro & Kennell, 1994; Frazão et al., 2006; Lorentzen et al., 2008). In RNase II, 10 nucleotides is the shortest RNA fragment able to retain contacts with both regions (Frazão et al., 2006). Shorter RNA fragments establish fewer contacts with the protein and the enzyme becomes distributive (Cannistraro & Kennell, 1994). Mutational analysis and biochemical data supported the structural results, since the substitution of Tyr253 in the catalytic site was shown to alter the size of the end-product from 4 to 10 nucleotides (Barbas et al., 2008). This mutation causes the loosening of the RNA substrate at the catalytic region and the RNA fragment is released (Barbas et al., 2008). In RNase R, Tyr324 was also found to be responsible for setting the size of the end-product (Matos et al., 2009).

Endonuclease Activity

Rrp44/Dis3 also has endoribonuclease activity conferred by the PIN domain (Lebreton & Seraphin, 2008; Schaeffer et al., 2009; Schneider et al., 2009). PIN

domains were named for their homology with the N-terminus domain of a bacterial protein involved in the biosynthesis of type IV pili (PilT N terminus; Wall & Kaiser, 1999). The PIN domain family has representatives in all domains of life and are found in proteins with different functions, some of which with nuclease activity (Wall & Kaiser, 1999; Arcus et al., 2004). Members of this family with nuclease activity are found in bacterial toxin-antitoxin systems (Arcus et al., 2005), are involved in eukaryotic ribosomal RNA processing (Fatica et al., 2004; Bleichert et al., 2006), nuclear quality control of mRNPs (Skruzny et al., 2009) and nonsense-mediated decay (Glavan et al., 2006).

The PIN domain consists of ~100 amino acids and shows a similar fold to the catalytic core of T4 RNase H, with four conserved acidic residues involved in the coordination of a metal ion (Bhagwat et al., 1997; Makarova et al., 1999; Clissold & Ponting, 2000; Arcus et al., 2004). In Rrp44, the activity of this domain is highest in the presence of manganese ions and digests linear and circular substrates (Lebreton & Seraphin, 2008; Schaeffer et al., 2009; Schneider et al., 2009). Mutation of acidic residues in the active site impairs its activity, *in vivo* and *in vitro* (Fatica et al., 2004; Bleichert et al., 2006; Glavan et al., 2006; Lebreton & Seraphin, 2008; Schaeffer et al., 2009; Schneider et al., 2009). Interestingly, there are three different versions of Dis3 in some eukaryotes that differ in the conservation of the acidic residues and the presence or absence of the PIN domain. In humans, there are three homologues: hDIS3, hDIS3L1 and hDIS3L2 (Staals & Pruijn, 2010; Tomecki et al., 2010). Two of them, hDIS3 and hDIS3L1, were proved to interact with the exosome ring complex. hDIS3 is mainly nucleoplasmic and has endonucleolytic activity, while hDIS3L1 is strictly cytoplasmic and has no endonucleolytic activity. In humans and *Schizosaccharomyces pombe* there is a third Dis3 homologue - Dis3L2 (Malecki et al., Submitted; Astuti et al., 2012). Although this protein does not have the PIN domain, it was shown to be an active exonuclease in the cytoplasm (Malecki et al.,

Submitted; Astuti et al., 2012). The lack of the PIN domain suggests that hDis3L2 does not interact with the exosome ring, and makes it structurally more similar to bacterial RNase R/II enzymes (Staals et al., 2010).

Substrate Specificity

The activity of the enzymes of RNase II family is sequence independent and they release nucleoside 5′ monophosphates and short oligonucleotides with 2-5 nt in length (Cannistraro & Kennell, 1994; Cheng & Deutscher, 2002; Dziembowski et al., 2007). However, each enzyme has different characteristics.

While yeast Rrp44 is able to digest through dsRNA, provided there is a single-stranded 3′ overhang with at least four nucleotides (Lorentzen et al., 2008), bacterial RNase II stalls around seven nucleotides before it reaches a double-stranded region (Spickler & Mackie, 2000). These differences can be rationalized with the different routes of access to the catalytic active site, proved by the crystallization data (Lorentzen et al., 2008).

A molecular mechanism model was proposed for the unwinding of secondary structures by yeast Rrp44. Accordingly, the unwinding of the RNA is achieved by pulling the 3′ tail into a channel to the catalytic active site that is specific for ssRNA, leading to the separation of duplex regions at the entrance pore (Lorentzen et al., 2008). The driving force would rely on the chemical energy released upon hydrolysis of the RNA chain instead of ATP hydrolysis, as in classical helicases (Lorentzen et al., 2008). Recently it was shown, by single-molecule fluorescence analysis, that this energy is accumulated during multiple single nucleotide step hydrolysis reactions until about four base pairs are unwound in a burst (Lee et al., 2012). Surprisingly, kinetic analysis data shows that RNA unwinding determines the overall RNA degradation rate (Lee et al., 2012).

Similarly to Rrp44, wild-type RNase R is able to digest through dsRNA, provided there is a single-stranded 3' overhang with at least 5 nucleotides (Cheng & Deutscher, 2005). However, a truncated RNase R protein without CSDs and/or the S1 domain is able to degrade double-stranded substrates (Matos et al., 2009). This result shows that the binding domains (CSD1, CSD2 and S1) block the entrance of the substrates into the catalytic channel. Nevertheless, it seems reasonable that the underlying mechanism is similar to Rrp44 (Lorentzen et al., 2008; Lee et al., 2012). Only the solving of the RNase R structure will elucidate its mode of action.

RNase II can bind RNA and DNA, although it is not able to cleave DNA. Furthermore, it is inhibited by DNA oligonucleotides since they bind to its specific binding site (Cannistraro & Kennell, 1994). In the RNase II structure the O2'-ribose oxygens of nucleotides 8, 10 and 12 are engaged in hydrogen bonds with the side chains of conserved residues Glu390, Asp201 and Tyr313 (Frazão et al., 2006). These interactions are important for the proper orientation of RNA and catalysis by RNase II. Instead of three, Rrp44 displays five interactions with 2' OH groups consistent with an higher affinity for ssRNA than for a corresponding ssDNA oligonucleotide (Lorentzen et al., 2008). Similarly, RNase R binds poorly to a DNA oligonucleotide (Cheng & Deutscher, 2002; Vincent & Deutscher, 2006).

RNase II-family members activity does not depend on the RNA sequence, however they show different degrading efficiencies depending on the substrate. For instance, while RNase II and RNase R are more active on the homopolymer poly(A) substrate (Cheng & Deutscher, 2002), Rrp44 seems to be more active on AU-rich sequences and generic substrates (Liu et al., 2006). These preferences may reflect the physiological role of each enzyme in the cell.

Exosome

Rrp44/Dis3 is part of a multiprotein complex known as the eukaryotic exosome. This complex was first identified in *S. cerevisiae* as a complex of 3′ to 5′ exoribonucleases required for 3′ processing of the 5.8S rRNA (Mitchell et al., 1997). Later, it was shown to be evolutionary conserved, to participate in several aspects of cellular RNA metabolism (Allmang et al., 1999a; Allmang et al., 1999b) and to have an endonuclease cleavage site (Lebreton & Seraphin, 2008; Schaeffer et al., 2009; Schneider et al., 2009). Nuclear and cytoplasmic forms of the exosome share 10 essential subunits (Mitchell et al., 1997; Allmang et al., 1999b).

Structure

The eukaryotic core exosome is composed of a six subunits ring structure (Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and Mtr3) that is capped on one side by three other subunits (Csl4, Rrp4 and Rrp40; Fig. 6; Liu et al., 2006).

The subunits that form the ring belong to the PDX family of exoribonucleases, which also includes *E. coli* RNase PH, PNPase and the archaeal exosome (Mian, 1997; Zuo & Deutscher, 2001). Members of this family activate an inorganic phosphate to cleave the substrate in the 3′ to 5′ direction, and release 5' nucleoside diphosphate products (Zuo & Deutscher, 2001). Surprisingly, the eukaryotic exosome lacks the phosphorolytic activity that is characteristic of this family, and in most eukaryotes can induce RNA degradation only when cooperating with the active component Rrp44/Dis3 (Liu et al., 2006; Dziembowski et al., 2007). This suggests that instead of having a catalytic role, the ring acts as a protein scaffold that may have a role in the recruitment of exosome cofactors and substrates.

The proteins that form the cap contain RNA binding domains, namely, S1 domain (ribosomal protein S1; Suryanarayana & Subramanian, 1984) and/or KH

domain (K-homology; Grishin, 2001). The cap is thought to stabilize the exosome and to bind RNA. The core exosome is architecturally similar to the archaeal exosome complex (Buttner et al., 2005; Lorentzen et al., 2005; Lorentzen et al., 2007) and to the bacterial PNPase (Symmons et al., 2000).

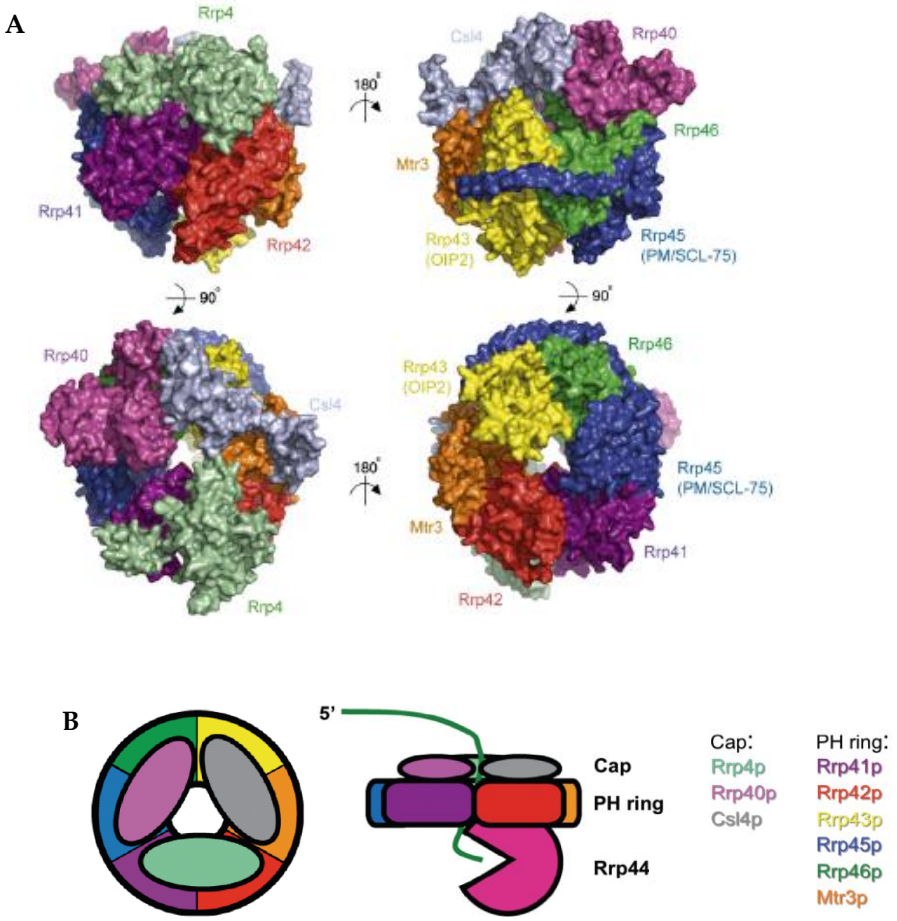


Figure 6. Overall structure of the eukaryotic core exosome. A. Surface representation of the human core exosome, depicting opposing views from sides (on the top), top (bottom left) and bottom (bottom right). Adapted from Liu et al. (2006). **B.** Schematic representation of the eukaryotic exosome, top view (left) and side view of the exosome (right). Adapted from Schaeffer et al. (2009).

In yeast, Rrp44 interacts with the nine-protein ring-shaped complex to generate a ribonucleolytically active exosome (Liu et al., 2006; Dziembowski et al.,

2007). Biochemical and structural studies showed that even if the Rrp44 protein by itself is able to degrade RNA, it seems that the substrates that are delivered to this nuclease first have to pass the channel in the exosome ring structure (Bonneau et al., 2009; Malet et al., 2010; Wasmuth & Lima, 2012). We have recently shown that three cysteines at the N-terminus of Rrp44, together with a conserved histidine from the PIN domain, affect the binding to the exosome (Schaeffer et al., 2012). However, in fission yeast, Dis3L2 is able to degrade RNA without binding the exosome (Malecki et al., Submitted).

Cofactors and Functions

The eukaryotic exosome is involved in RNA processing, degradation and surveillance reactions. It is present in two different compartments of the cell, where it has different specialized functions and associates with auxiliary components. Exosome cofactors are thought to direct the exosome to specific RNA substrates and to activate the exosome (Houseley et al., 2006). However, it is not yet known how this activation occurs.

Nuclear Cofactors and Functions

The nuclear exosome has a dual role, the processing and degradation of RNA. It is involved in the processing of different species of RNA precursors, including snRNA, snoRNA and rRNA (Mitchell et al., 1997; Allmang et al., 1999a; Allmang et al., 1999b; van Hoof et al., 2000). In addition, the exosome also participates in the degradation of defective pre-RNAs, which are not correctly processed, spliced or assembled into ribonucleoprotein complexes (Bousquet-Antonelli et al., 2000; Libri et al., 2002; Torchet et al., 2002; Kadaba et al., 2004; LaCava et al., 2005). The nuclear exosome is also implicated in the control of expression levels of mRNA (Kuai et al., 2005), and in the removal of processing byproducts and a myriad of cryptic unstable transcripts (CUTs; Wyers et al., 2005; Davis & Ares, 2006).

The main interactors of the nuclear exosome are Rrp6 (Briggs et al., 1998) and the TRAMP complex (for Trf4/5-Air1/2-Mtr4; LaCava et al., 2005; Vanacova et al., 2005).

Rrp6 is a distributive hydrolytic 3′ to 5′ exoribonuclease that shares sequence similarity to *E. coli* RNase D (Zuo & Deutscher, 2001). Although not essential for cell viability (Briggs et al., 1998), Rrp6 has a critical role in RNA processing and surveillance pathways (Briggs et al., 1998; Allmang et al., 1999a; Allmang et al., 1999b; van Hoof et al., 2000; Thomsen et al., 2003; Kuai et al., 2004). In yeast, Rrp6 is only associated with nuclear exosome complexes (Allmang et al., 1999b), while in humans small fractions were also found in the cytoplasm (Lejeune et al., 2003; van Dijk et al., 2007; Tomecki et al., 2010). It is not clear how the interaction with the core exosome is mediated. Different models place Rrp6 near the cap proteins or the ring proteins (Lehner & Sanderson, 2004; Cristodero et al., 2008), suggesting a distinct recruitment pathway. These data are in agreement with biochemical data showing that the activity of Rrp6 is not affected upon association with the exosome core (Liu et al., 2006). Interestingly, Rrp6 is able to function *in vivo* without interacting with the core exosome (Callahan & Butler, 2008). Additionally, Rrp6 binds to Rrp47 (C1D in humans) and the M-phase phosphoprotein 6, Mpp6 (Mitchell et al., 2003; Stead et al., 2007; Milligan et al., 2008; Butler & Mitchell, 2010; Costello et al., 2011). These nuclear proteins are involved in RNA binding and preferentially interact with structured and pyrimidine-rich sequences, respectively (Stead et al., 2007; Milligan et al., 2008). Both of these proteins are required for 3′ processing of 5.8S rRNA in yeast and depletion of Rrp47 causes similar processing defects on some substrates, to those caused by lack of Rrp6 (Mitchell et al., 2003; Milligan et al., 2008).

The TRAMP complex is responsible for nuclear surveillance of many different RNAs. It is composed of a non-canonical poly(A) polymerase (Trf4 or Trf5), an RNA-binding protein (Air1 or Air2) and a DExH box RNA helicase

(Mtr4; LaCava et al., 2005; Vanacova et al., 2005). The TRAMP complex is thought to activate the exosome by recruiting the exosome to the aberrant RNA substrates through the addition of short poly(A) tails and by stimulating the exonucleolytic activity of Rrp6 and Rrp44 (LaCava et al., 2005; Vanacova et al., 2005; Schneider et al., 2007; Callahan & Butler, 2010). Instead of stabilizing the RNA, poly(A) tails recruit and activate the exosome leading to rapid degradation. The role of polyadenylation in stimulating RNA degradation by 3′ to 5′ exoribonucleases, resembles bacterial mRNA decay (Dreyfus & Regnier, 2002). Mtr4 is thought to unwind secondary structures and to recruit the exosome to TRAMP substrates, since its depletion leads to RNA hyperadenylation *in vivo* (Houseley & Tollervey, 2006). Mtr4 also functions as a cofactor of the exosome independently of the TRAMP complex and is required for both RNA degradation and maturation (de la Cruz et al., 1998).

Cytoplasmic Cofactors and Functions

In the cytoplasm, the exosome is strictly involved in mRNA degradation, where it degrades normal and aberrant transcripts. In this compartment, the transcripts are translated to proteins. Therefore, it is very important to check the translational abilities of RNAs and remove incorrect molecules that can cause the production of aberrant proteins (Doma & Parker, 2007). The types of aberrant mRNAs targeted for degradation by the exosome include transcripts that lack termination codons (nonstop decay surveillance pathway; Frischmeyer et al., 2002; van Hoof et al., 2002), transcripts that stall translation elongation (no-go decay surveillance pathway; Doma & Parker, 2006) or contain premature termination codons (nonsense mediated decay surveillance pathway; Mitchell & Tollervey, 2003).

The cytoplasmic exosome binds the Ski complex (Ski2, Ski3 and Ski8) and Ski7. Ski2 is closely related to Mtr4 and is also a member of the DExH box RNA helicase family (Johnson & Jackson, 2012). Ski7 is an homologue of the eEF1A

(eukaryotic translation elongation factor 1A), eRF3 (eukaryotic release factor 3) and the prokaryotic EF-Tu (elongation factor thermo unstable; Benard et al., 1999) proteins, and is thought to recognize specific RNA substrates and to recruit the exosome to mRNA (Frischmeyer et al., 2002; van Hoof et al., 2002; Mitchell & Tollervey, 2003). Ski7 interacts with the Ski complex and the exosome (Araki et al., 2001), bridging the interactions between both complexes.

Aim of this Dissertation

Enzymes from the RNase II family are involved in RNA metabolism and are conserved both in prokaryotes and eukaryotes. While in *E. coli*, members of this family function as individual proteins, *S. cerevisiae* Rrp44 functions in the context of the exosome. When this study started, not much was known about Rrp44 and its involvement in RNA metabolism. The objectives proposed for this Dissertation were focused on the biochemical characterization of *S. cerevisiae* Rrp44 and *E. coli* RNase II, with the aim of learning more about these enzymes and their implications in RNA metabolism.

Although in the beginning of this work the existence of 5 domains in Rrp44 was already known, only the domains present in RNase II were functionally characterized. Therefore, in the first part of this work we focused on the study of a well conserved PIN domain in yeast Rrp44, which function was unkown. In order to address the specific role of this domain in the activity of the enzyme, we constructed and purified several Rrp44 mutant proteins, and analyzed their ribonucleolytic activity.

In the second part, we wanted to study the role of the CR3 motif at the N-terminus of Rrp44. This region had been previously identified in *Drosophila melanogaster* (Cairrão et al., 2005), but its function was unknown. In order to study

this region we have constructed and purified several truncated proteins, and have studied how this region affects the endonucleolytic activity of the enzyme.

Continuing our study of yeast Rrp44, we have investigated the specific role of certain residues in the evolutionary conserved RNB domain. The residues selected had been previously *in vitro* biochemically characterized in *E. coli* RNase II. We have constructed proteins with point mutations to test the biochemical function of different residues and analyze their physiological implications in yeast.

We were also interested in determining the mutation(s) responsible for the thermosensitive phenotype of *E. coli* RNase II (Donovan & Kushner, 1986), both *in vivo* and *in vitro*. The localization of this(these) mutation(s) can be an useful tool for studying RNase II enzymes and function in other organisms.

This knowledge will contribute to better understand the involvement of these enzymes in RNA metabolism.

References

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, 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, Mitchell P. 1999b. The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development 13*:2148-2158.

Amblar M, Arraiano CM. 2005. A single mutation in Escherichia coli ribonuclease II inactivates the enzyme without affecting RNA binding. *FEBS J 272*:363-374.

Amblar M, Barbas A, Fialho AM, Arraiano CM. 2006. Characterization of the functional domains of Escherichia coli RNase II. *J Mol Biol 360*:921-933.

Anderson JS, Parker RP. 1998. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J 17*:1497-1506.

Andrade JM, Pobre V, Silva IJ, Domingues S, Arraiano CM. 2009. The role of 3'-5' exoribonucleases in RNA degradation. *Prog Mol Biol Transl Sci 85*:187-229.

Apirion D. 1973. Degradation of RNA in Escherichia coli. A hypothesis. *Mol Gen Genet 122*:313-322.

Araki Y, Takahashi S, Kobayashi T, Kajiho H, Hoshino S, Katada T. 2001. Ski7p G protein interacts with the exosome and the Ski complex for 3'-to-5' mRNA decay in yeast. *EMBO J 20*:4684-4693.

Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN. 2004. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem 279*:16471-16478.

Arcus VL, Rainey PB, Turner SJ. 2005. The PIN-domain toxin-antitoxin array in mycobacteria. *Trends Microbiol 13*:360-365.

Arraiano CM, Andrade JM, Domingues S, Guinote IB, Malecki M, Matos RG, Moreira RN, Pobre V, Reis FP, Saramago M, Silva IJ, Viegas SC. 2010. The critical role of RNA processing and degradation in the control of gene expression. *FEMS Microbiol Rev 34*:883-923.

Astuti D, Morris MR, Cooper WN, Staals RH, Wake NC, Fews GA, Gill H, Gentle D, Shuib S, Ricketts CJ, Cole T, van Essen AJ, van Lingen RA, Neri G, Opitz JM, Rump P, Stolte-Dijkstra I, Muller F, Pruijn GJ, Latif F, Maher ER. 2012. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nat Genet 44*:277-284.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2008. New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end-product. *J Biol Chem 283*:13070-13076.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2009. Determination of key residues for catalysis and RNA cleavage specificity: one mutation turns RNase II into a "SUPER-ENZYME". *J Biol Chem 284*:20486-20498.

Benard L, Carroll K, Valle RC, Masison DC, Wickner RB. 1999. The ski7 antiviral protein is an EF1-alpha homolog that blocks expression of non-Poly(A) mRNA in Saccharomyces cerevisiae. *J Virol 73*:2893-2900.

Bhagwat M, Meara D, Nossal NG. 1997. Identification of residues of T4 RNase H required for catalysis and DNA binding. *J Biol Chem 272*:28531-28538.

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

Bonneau F, Basquin J, Ebert J, Lorentzen E, 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, Tollervey D. 2000. Identification of a regulated pathway for nuclear pre-mRNA turnover. *Cell 102*:765-775.

Briggs MW, Burkard KT, Butler JS. 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.

Brown CE, Sachs AB. 1998. Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. *Mol Cell Biol 18*:6548-6559.

Butler JS, Mitchell P. 2010. Rrp6, Rrp47 and cofactors of the nuclear exosome. *Adv Exp Med Biol 702*:91-104.

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

Cairrão F, Arraiano C, Newbury S. 2005. Drosophila gene *tazman,* an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev Dyn 232*:733-737.

Callahan KP, Butler JS. 2008. Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. *Nucleic Acids Res 36*:6645-6655.

Callahan KP, Butler JS. 2010. TRAMP complex enhances RNA degradation by the nuclear exosome component Rrp6. *J Biol Chem 285*:3540-3547.

Cannistraro VJ, Kennell D. 1994. The processive reaction mechanism of ribonuclease II. *J Mol Biol 243*:930-943.

Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, Anderson KC, Ardlie KG, Auclair D, Baker A, Bergsagel PL, Bernstein BE, Drier Y, Fonseca R, Gabriel SB, Hofmeister CC, Jagannath S, Jakubowiak AJ, Krishnan A, Levy J, Liefeld T, Lonial S, Mahan S, Mfuko B, Monti S, Perkins LM, Onofrio R, Pugh TJ, Rajkumar SV, Ramos AH, Siegel DS, Sivachenko A, Stewart AK, Trudel S, Vij R, Voet D, Winckler W, Zimmerman T, Carpten J, Trent J, Hahn WC, Garraway LA, Meyerson M, Lander ES, Getz G, Golub TR. 2011. Initial genome sequencing and analysis of multiple myeloma. *Nature 471*:467-472.

Cheng ZF, Deutscher MP. 2002. Purification and characterization of the Escherichia coli exoribonuclease RNase R. Comparison with RNase II. *J Biol Chem 277*:21624-21629.

Cheng ZF, Deutscher MP. 2005. An important role for RNase R in mRNA decay. *Mol Cell 17*:313-318.

Cheng ZF, Zuo Y, Li Z, Rudd KE, Deutscher MP. 1998. The vacB gene required for virulence in Shigella flexneri and Escherichia coli encodes the exoribonuclease RNase R. *J Biol Chem 273*:14077-14080.

Chlebowski A, Tomecki R, Lopez ME, Seraphin B, Dziembowski A. 2010. Catalytic properties of the eukaryotic exosome. *Adv Exp Med Biol 702*:63-78.

Clissold PM, Ponting CP. 2000. PIN domains in nonsense-mediated mRNA decay and RNAi. *Curr Biol 10*:R888-890.

Coburn GA, Mackie GA. 1998. Reconstitution of the degradation of the mRNA for ribosomal protein S20 with purified enzymes. *J Mol Biol 279*:1061-1074.

Costello JL, Stead JA, Feigenbutz M, Jones RM, Mitchell P. 2011. The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. *J Biol Chem 286*:4535-4543.

Cristodero M, Bottcher B, Diepholz M, Scheffzek K, Clayton C. 2008. The Leishmania tarentolae exosome: purification and structural analysis by electron microscopy. *Mol Biochem Parasitol 159*:24-29.

Davis CA, Ares M, Jr. 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.

de la Cruz J, Kressler D, Tollervey D, Linder P. 1998. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J 17*:1128-1140.

Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature 451*:355-358.

Decker CJ, Parker R. 1993. A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes & development 7*:1632-1643.

Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD, McMichael JF, Wallis JW, Lu C, Shen D, Harris CC, Dooling DJ, Fulton RS, Fulton LL, Chen K, Schmidt H, Kalicki-Veizer J, Magrini VJ, Cook L, McGrath SD, Vickery TL, Wendl MC, Heath S, Watson MA, Link DC, Tomasson MH, Shannon WD, Payton JE, Kulkarni S, Westervelt P, Walter MJ, Graubert TA, Mardis ER, Wilson RK, DiPersio JF. 2012. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature 481*:506-510.

Doma MK, Parker R. 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature 440*:561-564.

Doma MK, Parker R. 2007. RNA quality control in eukaryotes. *Cell 131*:660-668.

Domingues S, Matos RG, Reis FP, Fialho AM, Barbas A, Arraiano CM. 2009. Biochemical characterization of the RNase II family of exoribonucleases from the human pathogens *Salmonella typhimurium* and *Streptococcus pneumoniae*. *Biochemistry 48*:11848-11857.

Donovan WP, Kushner SR. 1986. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proc Natl Acad Sci U S A 83*:120-124.

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

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

Erova TE, Kosykh VG, Fadl AA, Sha J, Horneman AJ, Chopra AK. 2008. Cold shock exoribonuclease R (VacB) is involved in Aeromonas hydrophila pathogenesis. *J Bacteriol 190*:3467-3474.

Fatica A, Tollervey D, Dlakic M. 2004. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *Rna 10*:1698-1701.

Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature 443*:110-114.

Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R, Dietz HC. 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science 295*:2258-2261.

Ghosh S, Deutscher MP. 1999. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proc Natl Acad Sci U S A 96*:4372-4377.

Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *Embo J 25*:5117-5125.

Grishin NV. 2001. KH domain: one motif, two folds. *Nucleic Acids Res 29*:638-643.

Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nat Rev Mol Cell Biol 7*:529-539.

Houseley J, Tollervey D. 2006. Yeast Trf5p is a nuclear poly(A) polymerase. *EMBO reports 7*:205-211.

Houseley J, Tollervey D. 2009. The many pathways of RNA degradation. *Cell 136*:763-776.

Hsu CL, Stevens A. 1993. Yeast cells lacking 5'-->3' exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol Cell Biol 13*:4826-4835.

Johnson SJ, Jackson RN. 2012. Ski2-like RNA helicase structures: Common themes and complex assemblies. *RNA Biology*. Epub 2012 Sep 20 (dx.doi.org/10.4161/rna.22101).

Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. 2004. Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. *Genes & development 18*:1227-1240.

Khemici V, Carpousis AJ. 2004. The RNA degradosome and poly(A) polymerase of Escherichia coli are required in vivo for the degradation of small mRNA decay intermediates containing REP-stabilizers. *Mol Microbiol 51*:777-790.

Kuai L, Das B, Sherman F. 2005. A nuclear degradation pathway controls the abundance of normal mRNAs in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A 102*:13962-13967.

Kuai L, Fang F, Butler JS, Sherman F. 2004. Polyadenylation of rRNA in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A 101*:8581-8586.

Kushner SR. 2002. mRNA decay in Escherichia coli comes of age. *J Bacteriol 184*:4658-4665.

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

Lebreton A, Seraphin B. 2008. Exosome-mediated quality control: substrate recruitment and molecular activity. *Biochim Biophys Acta 1779*:558-565.

Lee G, Bratkowski MA, Ding F, Ke A, Ha T. 2012. Elastic coupling between RNA degradation and unwinding by an exoribonuclease. *Science 336*:1726-1729.

Lehner B, Sanderson CM. 2004. A protein interaction framework for human mRNA degradation. *Genome Res 14*:1315-1323.

Lejeune F, Li X, Maquat LE. 2003. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell 12*:675-687.

Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH. 2002. Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol 22*:8254-8266.

Lim J, Kuroki T, Ozaki K, Kohsaki H, Yamori T, Tsuruo T, Nakamori S, Imaoka S, Endo M, Nakamura Y. 1997. Isolation of murine and human homologues of the fission-yeast dis3+ gene encoding a mitotic-control protein and its overexpression in cancer cells with progressive phenotype. *Cancer Res 57*:921-925.

Liu H, Rodgers ND, Jiao X, Kiledjian M. 2002. The scavenger mRNA decapping enzyme DcpS is a member of the HIT family of pyrophosphatases. *EMBO J 21*:4699-4708.

Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell 127*:1223-1237.

Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell 29*:717-728.

Lorentzen E, Dziembowski A, Lindner D, Seraphin B, Conti E. 2007. RNA channelling by the archaeal exosome. *EMBO reports 8*:470-476.

Lorentzen E, Walter P, Fribourg S, Evguenieva-Hackenberg E, Klug G, Conti E. 2005. The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nature structural & molecular biology 12*:575-581.

Mackie GA. 1992. Secondary structure of the mRNA for ribosomal protein S20. Implications for cleavage by ribonuclease E. *J Biol Chem 267*:1054-1061.

Mackie GA. 1998. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature 395*:720-723.

Makarova KS, Aravind L, Galperin MY, Grishin NV, Tatusov RL, Wolf YI, Koonin EV. 1999. Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res 9*:608-628.

Malecki M, Viegas S, Carneiro T, Golik P, Dressaire C, Ferreira M, Arraiano C. Characterization of Dis3L2 exoribonuclease discloses a novel degradation pathway in eukaryotes. *Submitted*.

Malet H, Topf M, Clare DK, Ebert J, Bonneau F, Basquin J, Drazkowska K, Tomecki R, Dziembowski A, Conti E, Saibil HR, Lorentzen E. 2010. RNA channelling by the eukaryotic exosome. *EMBO reports 11*:936-942.

Mathy N, Benard L, Pellegrini O, Daou R, Wen T, Condon C. 2007. 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell 129*:681-692.

Matos RG, Barbas A, Arraiano CM. 2009. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *Biochem J 423*:291-301.

Matos RG, Pobre V, Reis FP, Malecki M, Andrade JM, Arraiano CM. 2011. Structure and Degradation Mechanisms of 3′ to 5′ Exoribonucleases. In: Nicholson AW, ed. *Ribonucleases*: Springer Berlin Heidelberg. pp 193-222.

Mian IS. 1997. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic Acids Res 25*:3187-3195.

Milligan L, Decourty L, Saveanu C, Rappsilber J, Ceulemans H, Jacquier A, Tollervey D. 2008. A yeast exosome cofactor, Mpp6, functions in RNA surveillance and in the degradation of noncoding RNA transcripts. *Mol Cell Biol 28*:5446-5457.

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

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

Mitchell P, Tollervey D. 2003. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3'-->5' degradation. *Mol Cell 11*:1405-1413.

Muhlrad D, Decker CJ, Parker R. 1994. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. *Genes & development 8*:855-866.

Muhlrad D, Parker R. 1992. Mutations affecting stability and deadenylation of the yeast MFA2 transcript. *Genes & development 6*:2100-2111.

Muhlrad D, Parker R. 1994. Premature translational termination triggers mRNA decapping. *Nature 370*:578-581.

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

Reis FP, Pobre V, Silva IJ, Malecki M, Arraiano CM. 2012. The RNB Family of Exoribonucleases – Putting the 'Dis' in Disease. *Wiley Interdisciplinary Reviews (WIREs) RNA. Accepted.*

Robertson HD, Webster RE, Zinder ND. 1968. Purification and properties of ribonuclease III from Escherichia coli. *J Biol Chem 243*:82-91.

Rose AE, Poliseno L, Wang J, Clark M, Pearlman A, Wang G, Vega YSdMEC, Medicherla R, Christos PJ, Shapiro R, Pavlick A, Darvishian F, Zavadil J, Polsky D, Hernando E, Ostrer H, Osman I. 2011. Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression. *Cancer Res 71*:2561-2571.

Schaeffer D, Reis FP, Johnson SJ, Arraiano CM, van Hoof A. 2012. The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. *Nucleic Acids Res 40*:9298-9307.

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature structural & molecular biology 16*:56-62.

Schneider C, Anderson JT, Tollervey D. 2007. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol Cell 27*:324-331.

Schneider C, Leung E, Brown J, Tollervey D. 2009. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res 37*:1127-1140.

Skruzny M, Schneider C, Racz A, Weng J, Tollervey D, Hurt E. 2009. An endoribonuclease functionally linked to perinuclear mRNP quality control associates with the nuclear pore complexes. *PLoS Biol 7*:e8.

Spickler C, Mackie GA. 2000. Action of RNase II and polynucleotide phosphorylase against RNAs containing stem-loops of defined structure. *J Bacteriol 182*:2422-2427.

Staals RH, Bronkhorst AW, Schilders G, Slomovic S, Schuster G, Heck AJ, Raijmakers R, Pruijn GJ. 2010. Dis3-like 1: a novel exoribonuclease associated with the human exosome. *EMBO J 29*:2358-2367.

Staals RH, Pruijn GJ. 2010. The human exosome and disease. *Adv Exp Med Biol 702*:132-142.

Stead JA, Costello JL, Livingstone MJ, 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.

Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A 90*:6498-6502.

Suryanarayana T, Subramanian AR. 1984. Function of the repeating homologous sequences in nucleic acid binding domain of ribosomal protein S1. *Biochemistry 23*:1047-1051.

Symmons MF, Jones GH, Luisi BF. 2000. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure 8*:1215-1226.

Thomsen R, Libri D, Boulay J, Rosbash M, Jensen TH. 2003. Localization of nuclear retained mRNAs in Saccharomyces cerevisiae. *RNA 9*:1049-1057.

Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, Dziembowski A, Jensen TH. 2010. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J 29*:2342-2357.

Torchet C, Bousquet-Antonelli C, Milligan L, Thompson E, Kufel J, Tollervey D. 2002. Processing of 3'-extended read-through transcripts by the exosome can generate functional mRNAs. *Mol Cell 9*:1285-1296.

Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. 2001. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. *Cell 104*:377-386.

van Dijk EL, Schilders G, Pruijn GJ. 2007. Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways. *RNA 13*:1027-1035.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science 295*:2262-2264.

van Hoof A, Lennertz P, 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.

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

Vincent HA, Deutscher MP. 2006. Substrate recognition and catalysis by the exoribonuclease RNase R. *J Biol Chem 281*:29769-29775.

Wall D, Kaiser D. 1999. Type IV pili and cell motility. *Mol Microbiol 32*:1-10.

Wang Z, Kiledjian M. 2001. Functional link between the mammalian exosome and mRNA decapping. *Cell 107*:751-762.

Wasmuth EV, Lima CD. 2012. Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. *Mol Cell 48*:133-144.

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A. 2005. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell 121*:725-737.

Zuo Y, Deutscher MP. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res 29*:1017-1026.

Zuo Y, Vincent HA, Zhang J, Wang Y, Deutscher MP, Malhotra A. 2006. Structural basis for processivity and single-strand specificity of RNase II. *Mol Cell 24*:149-156.

Chapter 2

Characterization of the PIN domain of yeast Rrp44

This chapter contains data published in:
Schaeffer D, Tsanova B, Barbas A, **Reis FP**, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature Structural & Molecular Biology 16*:56-62.
*Highlighted by Faculty of 1000. Focused in News and Views (*Nature Structural & Molecular Biology*) and Research Highlights (*Nature Reviews Molecular Cell Biology*).

The author of this Dissertation had an important contribution in this article, namely, in the construction of all the clones for the GST-fusion proteins, and was responsible for developing and optimizing the conditions for overexpression and purification of the proteins used in the in vitro nuclease assays.

Abstract

The eukaryotic exosome is a ten-subunit 3′ exoribonucleolytic complex responsible for many RNA-processing and RNA-degradation reactions. How the exosome accomplishes this is unknown. Rrp44 (also known as Dis3), a member of the RNase II family of enzymes, is the catalytic subunit of the exosome. We show that the PIN domain of Rrp44 has endoribonucleolytic activity. The PIN domain is preferentially active toward RNA with a 5′ phosphate, suggesting coordination of 5′ and 3′ processing. We also show that the endonuclease activity is important *in vivo*. The combination of an endoRNase and an exoRNase activity seems to be a widespread feature of RNA-degrading machines.

Introduction

Virtually all RNAs are processed from longer precursors. Although the removal of 3′ extensions seems to be a simple enzymatic reaction, eukaryotic cells contain multiple endoribonucleases (endoRNases) and 3′ exoribonucleases (exoRNases) (Mian, 1997; Moser et al., 1997; Deutscher & Li, 2001; Zuo & Deutscher, 2001). These exoRNases can act redundantly for some reactions, but are specific for other functions (van Hoof et al., 2000a). In addition to these RNA-processing reactions that generate mature RNAs, 3′ exoRNases and endonucleases also carry out mRNA-degradation reactions (Muhlrad et al., 1995; Anderson & Parker, 1998) that can be an important determinant for mRNA quality control (van Hoof et al., 2002; Meaux & Van Hoof, 2006). The same RNase can process some substrates and completely degrade others. It is largely unknown what determines the specificity of different RNases for their particular substrates, or for processing *versus* decay.

The exosome is a major 3′ exoRNase that is present in all eukaryotes, but whose function has been most extensively characterized in *Saccharomyces cerevisiae* (Mitchell et al., 1997; Allmang et al., 1999b). The yeast exosome is responsible for the 3′-end processing of stable RNAs including ribosomal RNA (rRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA), as well as the degradation of RNAs such as the 5′ external transcribed spacer (ETS) of the rRNA precursor and aberrant mRNAs that lack a stop codon or a poly(A) tail (Mitchell et al., 1997; Anderson & Parker, 1998; de la Cruz et al., 1998; Allmang et al., 1999a; Allmang et al., 1999b; van Hoof et al., 2000b; van Hoof et al., 2002; Meaux & Van Hoof, 2006).

The core exosome contains nine subunits and shows extensive structural similarity to bacterial PNPase and archaeal exosomes (Symmons et al., 2000;

Buttner et al., 2005; Lorentzen & Conti, 2005; Lorentzen et al., 2005; Liu et al., 2006; Navarro et al., 2008) (Fig. 1a, below). These nine subunits are all essential for viability. Six subunits have an RNase PH domain and form a ring structure in the exosome (Liu et al., 2006). Although these six subunits show clear sequence and structural similarity to RNases, all six subunits in the yeast and human exosome lack important catalytic residues and are inactive (Liu et al., 2006; Dziembowski et al., 2007). The six PH-ring subunits cannot assemble into a PH-ring *in vitro*, but, when the three cap proteins are added, a stable core exosome can be assembled (Liu et al., 2006). The crystal structure shows that each cap protein contacts two neighboring subunits of the PH ring. These observations suggest that the cap proteins carry out an essential structural role by bridging interactions between PH domains (Liu et al., 2006) (Fig. 1a, below). A second possible function of the cap proteins is to bind exosome substrates, as each contains three putative RNA binding domains (Allmang et al., 1999b; Chekanova et al., 2002; Buttner et al., 2005). Finally, a point mutation in one of the cap proteins (Csl4) disrupts the interaction between the core exosome and one of its cofactors (van Hoof et al., 2002). The exosome requires many cofactors, and a third possible function for the cap proteins is to provide binding sites for these cofactors.

Although the nine subunit exosome core shows remarkable structural similarity to the bacterial PNPase and the archaeal exosome, the yeast exosome contains a tenth essential subunit (Rrp44), which is homologous to bacterial RNase II (Fig. 1a, below) and is responsible for the 3′ exoRNase activity of the exosome (Frazao et al., 2006; Dziembowski et al., 2007; Schneider et al., 2007; Wang et al., 2007; Barbas et al., 2008; Lorentzen et al., 2008). Notably, although Rrp44 is an essential gene, a point mutation that inactivates its exoRNase activity *in vitro* is viable (Dziembowski et al., 2007), suggesting that either the exoRNase activity may not be a central feature of the exosome or the point-mutant protein retains partial exoribonucleolytic activity *in vivo*.

The N terminus of Rrp44 contains a CR3 domain and a PIN domain (Fig. 1a, below). The CR3 domain is a small domain of unknown function that contains three conserved cysteine residues. PIN domains were initially thought to have some regulatory or signaling function, but recent crystal structures of PIN domains revealed a putative active site with similarity to RNase H (Arcus et al., 2004; Levin et al., 2004). Indeed, some PIN domains have endoRNase activity, including the SMG6 protein, which is involved in nonsense-mediated decay (Arcus et al., 2004; Levin et al., 2004; Glavan et al., 2006; Daines et al., 2007; Bunker et al., 2008; Eberle et al., 2009). PIN domains contain four acidic amino acid residues (Asp91, Glu120, Asp171 and Asp198 in Rrp44), and mutations in the third residue abolish the activity of PIN domains *in vitro* and *in vivo* (Fatica et al., 2004; Bleichert et al., 2006; Glavan et al., 2006).

In this paper, we show that the Rrp44 PIN domain is an active endoRNase, revealing that the exosome has endoRNase activity in addition to its known exoRNase activity.

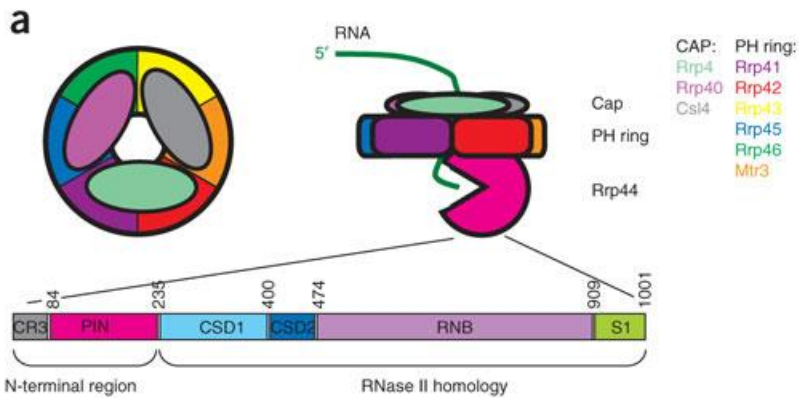
Results

The exosome contains an endoRNase domain

To test whether the Rrp44 PIN domain contains endonuclease activity, we purified a truncated Rrp44 as a glutathione S-transferase (GST) fusion from *Escherichia coli*. This truncated Rrp44 contains the PIN domain, but lacks the RNB domain, and has robust nuclease activity *in vitro* on several RNA substrates (Fig. 1), whereas GST by itself does not (Fig. 1c). The activity level of the truncated Rrp44 was similar to the activity level of the full-length protein (although under different conditions; Supplementary Fig. 1). Mutation of the third conserved acidic residue (Asp171) in the Rrp44 PIN domain abolished the endoRNase

activity (Fig. 1g), as has been shown for some other PIN domains either *in vitro* or *in vivo* (Fatica et al., 2004; Bleichert et al., 2006; Glavan et al., 2006). We conclude that the PIN domain of Rrp44 is a nuclease.

Other PIN domains and the related RNases H act as endonucleases. To test whether the Rrp44 PIN acts as an endonuclease we used 5′- or 3′-labeled substrates. An exonuclease will yield only labeled mononucleotides with one of these substrates. In contrast, the PIN domain yielded a similar pattern of labeled oligonucleotides with each substrate, showing that this domain acts as an endonuclease (Fig. 1d). Notably, the PIN domain was dramatically more active on substrates with a 5′ monophosphate than on substrates with a 5′ hydroxyl group (Fig. 1f), which may have important implications (see Discussion). As expected for an endonuclease, the PIN domain also had some activity on circular U_{30} (Supplementary Fig. 1b). However, degradation of circular RNA was inefficient, probably because it lacks a free 5′ phosphate. All of these results indicate that the Rrp44 PIN domain is an active endoRNase, as has been shown for SMG6 (Levin et al., 2004; Bunker et al., 2008).



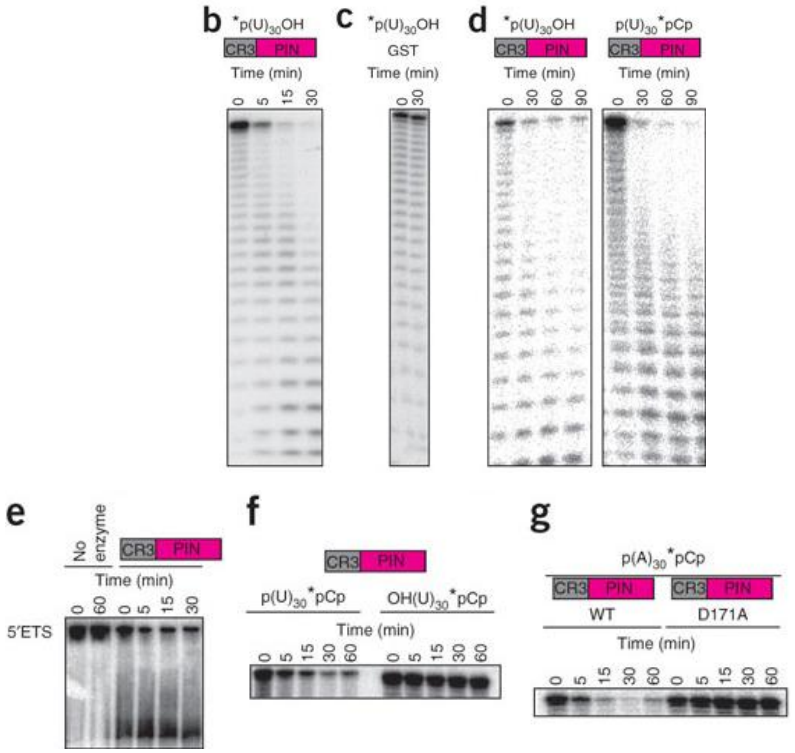


Figure 1. The N-terminal region of Rrp44 has RNase activity. (a) The yeast exosome contains three layers. The exosome contains a PH ring of six proteins with an RNase PH domain. Assembly and/or stability of the PH ring requires three additional proteins that form a cap on one side of the ring. The exoRNase activity of the exosome is provided by a tenth subunit (Rrp44) that associates with the bottom of the ring. RNA substrates are thought to associate with the cap proteins and pass through the central channel in the PH ring, being degraded by Rrp44 after they emerge from the bottom of the ring. Below is a schematic representation of Rrp44's domains. **(b)** The N-terminal region of Rrp44, which contains a CR3 domain and a PIN domain, has RNase activity in vitro. The N-terminal part of Rrp44 was purified as a GST-fusion recombinant protein from *E. coli* and incubated with 5′ end–labeled U_{30} RNA. **(c)** GST by itself, purified and incubated under identical conditions, does not have RNase activity. **(d)** The N-terminal region of Rrp44 degrades both 5′ end–labeled and 3′ end–labeled substrates to smaller labeled oligoribonucleotides, suggesting that this region acts as an endonuclease. **(e)** The N-terminal region of Rrp44 degrades an internally labeled RNA corresponding to the 5′ ETS *in vitro*. **(f)** The N-terminal region of Rrp44 prefers substrates with a 5′ phosphate (left) over those with a 5′ hydroxyl group (right). The apparent slight change in mobility in the right-hand gel is an artifact of this particular gel and does not reflect a slow degradation of the 5′ hydroxyl substrate. **(g)** Mutation of the conserved acidic amino acid residue Asp171 to alanine in the N-terminal region of Rrp44 abolishes RNase activity. The asterisks in **b–d** and **f,g** indicate the position of the P^{32} label.

Either RNase activity can form an active exosome

To address whether this endonuclease activity is important for exosome function, we generated truncations and point mutants of Rrp44. C-terminal truncations that removed the RNase II homology, but left the CR3 and PIN domains, allowed for slow growth (Fig. 2a and Supplementary Fig. 2). Previously it was shown that a point mutation inactivating the catalytic activity of the RNB domain is viable (Amblar & Arraiano, 2005; Dziembowski et al., 2007). The residual growth we observed upon deletion of the RNB domain rules out the possibility that the mutated RNB domain was viable owing to some low-level activity *in vivo*. Instead, our results show that exoRNase activity is not a crucial feature of the exosome core. C-terminal truncations that also removed part of the PIN domain, or truncation of the N terminus, did not support any growth. The smallest viable Rrp44 allele we generated included the CR3 and PIN nuclease domains (amino acids 1–235). We conclude that the PIN domain is essential for exosome function, but we cannot exclude the alternative that the PIN domain (also) has functions separate from the exosome.

To further address the importance of the CR3 and PIN domains, we made point mutants in conserved residues. Mutating the three conserved cysteines of the CR3 domain caused slow growth (Fig. 2b); thus, both N-terminal truncations and point mutants indicate an important role for the CR3 domain. In contrast, the D171A point mutation, which inactivates the PIN domain (or mutations in the other three conserved acidic residues), does not have an obvious growth phenotype (Fig. 2c and Supplementary Fig. 3). However, combining the viable PIN mutations with either the viable D551N or a viable C-terminal truncation resulted in no growth (Fig. 2c,d and Supplementary Fig. 3). These results suggest that either active site is sufficient for viability, but simultaneous inactivation of both catalytic sites results in a nonfunctional exosome.

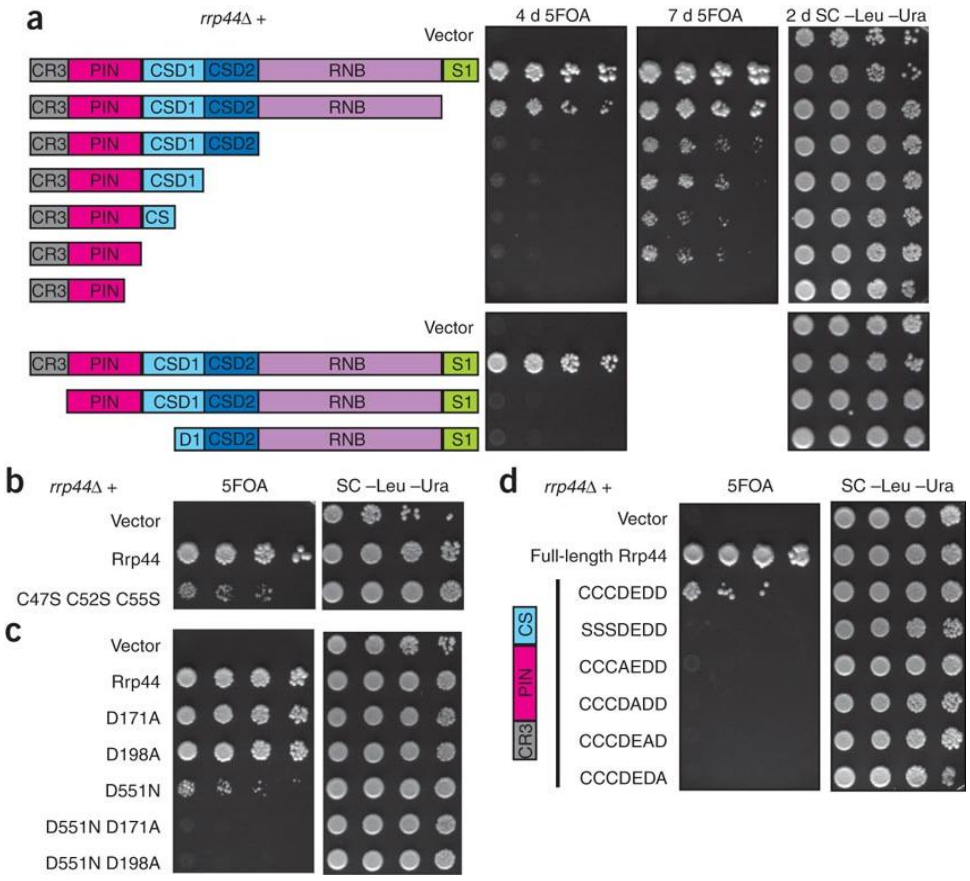


Figure 2. The PIN active site is important for exosome function. (a) The N-terminal region of Rrp44 is required and sufficient for viability. An *rrp44*Δ strain complemented by full-length *RRP44* on a plasmid with a *URA3* marker was transformed with *LEU2* plasmids encoding each of the depicted truncations. Growth on 5-fluoroorotic acid (5FOA) indicates that truncated Rrp44 can carry out the essential function of the exosome. The smallest complementing Rrp44 allele includes amino acids 1–235. Similar results were obtained with a *GAL::rrp44* strain (Supplementary Fig. 2). **(b)** Mutation to serine of the three conserved cysteine residues in the CR3 domain reduces growth. **(c)** Mutations in the PIN domain active site residues of Rrp44 do not have a large effect on growth, but they are lethal in combination with a mutation of the active-site residue of the RNB domain. Similar results were obtained for the D91A and E120A mutations (Supplementary Fig. 3). **(d)** Mutations of the PIN domain active-site residues are lethal in combination with a truncation that removes the RNB domain. CCDEDD indicates Cys47, Cys52, Cys55, Asp91, Glu120, Asp171 and Asp198, which where mutated in the last five rows as indicated. SC –Leu –Ura is synthetic complete media lacking leucine and uracil included as a control.

The exosome carries out both RNA-processing and RNA degradation reactions. Endonuclease activity might be well suited for processing reactions, but exonuclease activity would be more appropriate for degradation reactions. We therefore tested the hypothesis that the PIN domain of Rrp44 is sufficient for exosome-mediated RNA-processing reactions, whereas the RNB domain is required for exosome-mediated RNA decay. Truncations lacking the RNase II homologous domain cause the 5′ ETS and the 7S precursor of the 5.8S rRNA to accumulate (Fig. 3a). Similarly, 7S-processing and 5′ ETS–degradation intermediates accumulated as a result of the D551N mutation in the RNB domain. We did not observe any similar degradation or processing intermediates resulting from the D171A mutation in the PIN domain, and the effect of the combination of D551N with D171A seemed similar to that of the D551N single mutant (Fig. 3b). Thus, both degradation and processing functions of the exosome require the RNB domain. An alternative explanation for the presence of two catalytic sites is that the efficiency of the overall reaction is increased. Under this hypothesis, the endo- and exonuclease domains do not have separate substrates, but both may act together on most or all exosome substrates.

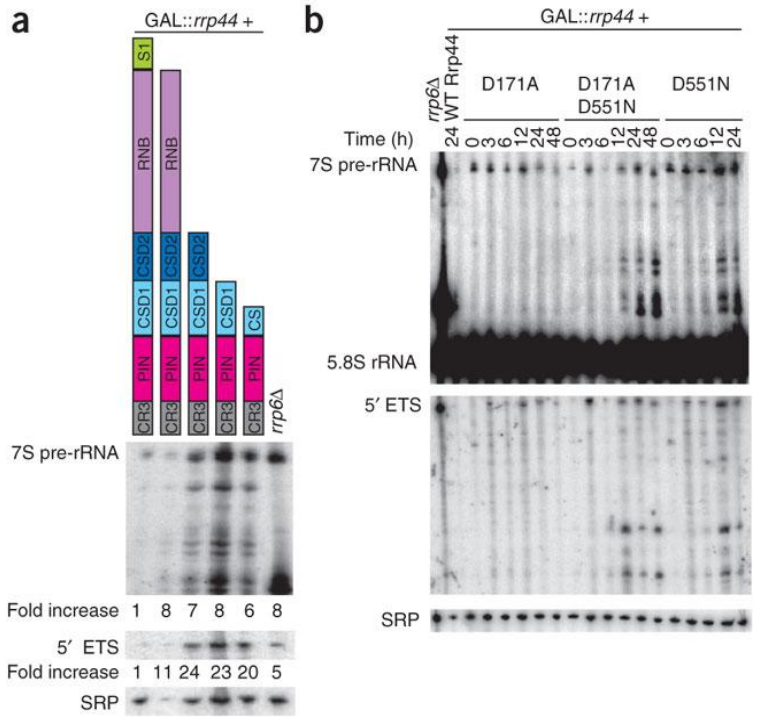


Figure 3. The RNB domain of Rrp44 is required for both RNA-processing and RNA-degradation activities of the exosome. (a) The indicated Rrp44 truncations were expressed in a *GAL::rrp44* yeast strain. RNA was isolated from cultures grown in dextrose (to inhibit expression of the *GAL::rrp44* gene) and analyzed by northern blotting with probes that hybridize to the 7S precursor of 5.8S rRNA (above), the 5′ ETS of the rRNA (middle) and the RNA subunit of the signal recognition particle (SRP; below). The relative levels of the 7S pre-rRNA and the 5′ ETS were normalized for loading using the SRP signal and are indicated under the above and middle blots. Shown is a representative experiment. **(b)** *RRP44* alleles containing the D171A mutation, the D551N mutation or the D171A D551N double mutation were introduced into a *GAL::rrp44* yeast strain. Cultures were first grown in media containing galactose and then incubated in media containing glucose for the indicated time (in hours) and analyzed by northern blotting with probes that hybridize to the 5.8S rRNA (above), the 5′ ETS of the rRNA (middle) and the RNA subunit of the signal recognition particle (below). Shown is a representative experiment.

Discussion

A 5′ phosphate stimulates endoRNase in the exosome

We show for the first time that Rrp44 contains a previously overlooked endoRNase domain—the PIN domain. Two observations explain previous reports that Rrp44 with a mutation in the RNB domain has no nuclease activity *in vitro* (Dziembowski et al., 2007; Schneider et al., 2007). First, the previously reported assays were carried out under conditions favorable for the activity of RNB domains (that is, in Tris buffer in the presence of Mg^{2+}). Under similar conditions, we also failed to detect significant activity (Supplementary Fig. 1). Second, we observed robust endonuclease activity with truncated proteins, but much less activity with full-length protein or any truncations that did not remove the first cold shock domain (CSD1) present in many proteins of the RNase II family of enzymes (data not shown) (Cairrão et al., 2005; Amblar et al., 2006; Frazao et al., 2006; Barbas et al., 2008). This latter observation suggests that CSD1 and perhaps other domains of the exosome regulate the activity of the PIN domain. Nevertheless, the activity level of the truncated Rrp44 under conditions where PIN domains are generally active (that is, in HEPES buffer in the presence of Mn^{2+}) was similar to the activity level of the full-length protein under conditions for the RNB domain (Supplementary Fig. 1). This suggests that, although the activity was not previously detected, it is robust and likely to be important *in vivo* under some circumstances or on some substrates that remain to be determined.

Our observation that Rrp44 needs either its endonuclease or exonuclease activity suggests that the endonuclease is active and important *in vivo*. Furthermore, such functional redundancy seems to be a general theme for a wide variety of RNases and is similar to that found for Rrp44 and Rrp6 in yeast

(Dziembowski et al., 2007), Xrn1 and the exosome in yeast (Johnson & Kolodner, 1995; Anderson & Parker, 1998), several combinations of Rex1, Rex2, Rex3 and Rrp6 in yeast (van Hoof et al., 2000a), and several exoRNase combinations in *E. coli* (Deutscher & Li, 2001; Andrade et al., 2009).

In the process of testing 5′- or 3′-labeled substrates in *in vitro* degradation reactions, we noticed that substrates with a 5′ phosphate are more rapidly degraded than substrates with a 5′ hydroxyl or circular substrates. This is analogous to what has been described for bacterial RNase E, which prefers a 5′ monophosphate–containing substrate over a circular RNA or a linear RNA with a 5′ hydroxyl or 5′ triphosphate (Mackie, 1998). Despite this functional similarity, we have not detected any sequence similarity between Rrp44 and RNase E. The 5′ end–stimulated activity of Rrp44 might have several important consequences for the substrate specificity of the endonuclease activity. First, PIN domains are part of a large group of enzymes and ribozymes that use two-metal-ion catalysis (Arcus et al., 2004; Levin et al., 2004; Glavan et al., 2006; Bunker et al., 2008). All of these enzymes generate a 5′ cleavage product with a 3′ hydroxyl group and a 3′ cleavage product with a 5′ monophosphate (reviewed in (Yang, 2008)). This cleavage mechanism suggests that the endonuclease activity of Rrp44 might cleave most primary transcripts relatively inefficiently, but the 3′ cleavage product of an initial cleavage will have a 5′ phosphate and be a preferred substrate. Such preferred subsequent cleavage might ensure that RNA degradation is rapidly completed once it has been initiated. Second, many RNA species are processed both at the 5′ and 3′ end, and in most cases it is not clear how these events are coordinated. One possibility offered by our results is that 5′ end processing generates a 5′ monophosphate, which stimulates 3′ end processing by the Rrp44 endonuclease. Testing these hypotheses in detail will require the identification of all, or most, of the endogenous substrates on the endonuclease activity of Rrp44.

The exosome may have multiple ways to interact with RNA

It is not clear how the exosome interacts with RNA substrates. A crystal structure of the archaeal exosome in complex with RNA (Lorentzen & Conti, 2005) suggests that RNA might go through the center of the PH ring, as depicted in Figure 1a, above. On the other hand, it has been proposed that RNA may gain access to Rrp44 without going through the PH ring (Wang et al., 2007). The observation that the endonuclease activity is stimulated by a 5′ monophosphate indicates that Rrp44 can interact with the 5′ end of an exosome substrate. As the 3′ exoRNase active site must interact with the 3′ end, Rrp44 seems to be able to interact with both ends of substrate RNAs. This is unlikely to be compatible with threading the RNA through the PH ring. Thus, our observations seem to be more consistent with the idea that the exosome can interact with RNA substrates in two different ways (compare Fig. 1a, above, to Fig. 4). Moreover, we suggest that substrates that access Rrp44 directly might be identifiable by looking for substrates that are processed *in vivo* in a 5′ end–dependent manner.

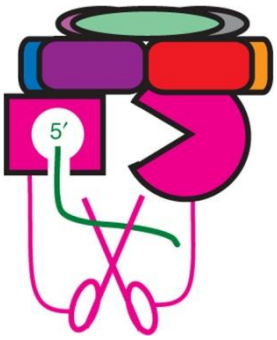


Figure 4. A model of how RNA might access the 5′ end–stimulated endonuclease activity of Rrp44.

Convergence on combining RNases in mRNA decay machines

We show for the first time that the exosome contains both an endonuclease and exonuclease activity. Although this is unexpected for the exosome, it is not

unique in nature. It has been suggested that the archaeal exosome also contains a PIN domain (Koonin et al., 2001), which could provide the archaeal exosome with an endonuclease activity. Similarly, the *E. coli* degradosome combines the RNase E endoRNase with the PNPase 3′ exoRNase (Carpousis et al., 1994; Py et al., 1994), and metazoan P-bodies combine the Argonaute endoRNase, the Xrn1 5′ exoRNase and the Ccr4, Caf1 and Pan2 3′ exoRNases (Bashkirov et al., 1997; Ingelfinger et al., 2002; Cougot et al., 2004; Andrei et al., 2005; Liu et al., 2005; Zheng et al., 2008). Thus, the combination of endoRNases and exoRNases into one RNA-degrading machine is widespread in nature and has evolved multiple times independently, suggesting that such combination offers a fundamental advantage to the cell.

Materials and Methods

Strains

The strains, plasmid and oligonucleotides we used are described in detail in Supplementary Tables 1–3. The heterozygous diploid *RRP44*/*rrp44*Δ has been described and were obtained from Open Biosystems (Giaever et al., 2002). We transformed this strain with a *URA3* plasmid that encoded the corresponding wild-type exosome subunit. We sporulated the transformants and obtained haploid progeny spores by the hydrophobic spore-isolation method essentially as described (Rockmill et al., 1991). Isolated spores were germinated to generate the deletion strain complemented with the *URA3* plasmid. The *GAL::rrp44* strain have been described previously and were obtained from P. Mitchell (University of Sheffield) and D. Tollervey (University of Edinburgh) (Mitchell et al., 1997; Allmang et al., 1999b).

Plasmids

All yeast plasmids used are low-copy plasmids that contain a yeast centromere and the endogenous promoter and 3′-flanking region to control expression of the exosome subunits. All were verified by sequencing.

To generate C-terminal truncations, we initially PCR amplified several hundred base pairs of the 3′-flanking region of the gene of interest and cloned it into pRS415 (Sikorski & Hieter, 1989), which contains a *LEU2* marker. We then PCR amplified the promoter and the desired part of the coding region and cloned it into the plasmid described above. N-terminal truncations were generated similarly, except that the promoter was PCR amplified and cloned first, and the truncated coding region and 3′-flanking region were added in the second step.

We generated GST-fusion plasmids by PCR amplifying the *RRP44* coding region from the plasmids described above, digesting the PCR product with EcoRI and XhoI and ligating it into pGEX-4T-1 (GE Healthcare).

We created site-directed mutants using the QuikChange II kit (Stratagene). Each oligonucleotide also changes a restriction site. We screened putative mutants by restriction digestion and verified them by sequencing. To combine point mutations in the PIN domain of Rrp44 with the D551N mutant in the RNB domain, we digested the plasmids containing each single mutant with SacI and PshAI and ligated them together. PshAI cuts at a unique site within the CSD2 domain.

Overexpression and purification of recombinant Rrp44

Full-length and truncated Rrp44 fused to GST were expressed in the Rosetta(DE3) *E. coli* strain. We grew cultures at 30 °C in 2 liters TB medium supplemented with 200 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol to an optical density at 600 nm (OD_{600}) of 1.2 and then induced expression by addition of 0.2 mM IPTG and

incubation at 18 °C overnight. We harvested cells by centrifugation and froze them. Cells were thawed on ice, resuspended in 30 ml PBS buffer, pH 8, with 10 mM DTT and 1 mM PMSF, and lysed using the French press. We treated the crude extracts with 5 units Benzonase (Sigma) for 30 min at 0 °C, clarified the extract by a 30-min centrifugation at 10,000g, and loaded it onto a GSTrapt FF 1-ml column (GE healthcare) equilibrated in binding buffer. We eluted proteins with a 0–10 mM glutathione gradient in 50 mM Tris-HCl, pH8.5, 10 mM DTT. We pooled fractions containing the purified protein and loaded them onto a gel-filtration column (Superdex 75 10/300 GL or a Superose 12 10/300 GL, GE Healthcare, depending on the protein size) in 20 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$ and 10 mM DTT buffer, pH 8. We concentrated the eluted protein by centrifugation at 15 °C with Vivaspin 500 Centrifugal Concentrators (Vivaspin). We determined protein concentrations by spectrophotometry using a Nanodrop and added 50% (v/v) glycerol to the final fractions before storage at –20 °C.

RNase assays

RNase assays were based on those previously described. Briefly, each assay contained 100–500 nM protein, 200 nM 5′ or 3′ end–labeled U_{30} or A_{30} RNA or an internally labeled 5′ ETS RNA transcribed by T7 RNA polymerase, and 20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM $MnCl_2$, 1 mM DTT.

Growth assays

We performed growth assays essentially as previously described(Meaux & Van Hoof, 2006; Wilson et al., 2007). Briefly, we grew strains in appropriate liquid selection media overnight at 30 °C (or 23 °C for temperature-sensitive mutants). We diluted these cultures in appropriate selection media to a starting OD_{600} of 0.2. We grew the cultures until they reached an OD of 0.8, serially diluted them in 96-well plates by a factor of five, and spotted them onto appropriate media. We

generally incubated these plates for 2–3 d at 23 °C, 30 °C or 37 °C. To monitor the slow growth of Rrp44 truncations, we incubated the plates for up to 20 d.

Northern blotting

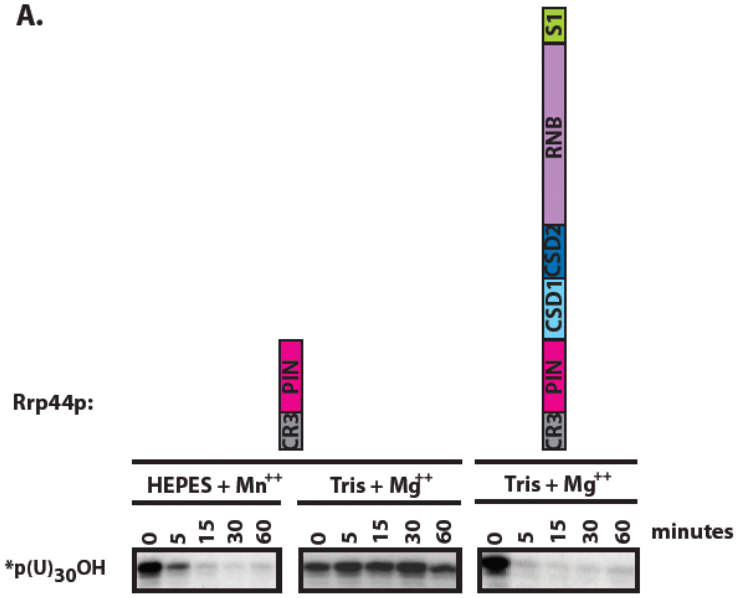
Stability of *PGK1-nonstop* mRNA (van Hoof et al., 2000c), and 5.8S rRNA processing and 5′ ETS degradation (van Hoof et al., 2000c), were assayed as previously described.

Acknowledgments

We are grateful to T. Link and R. Brennan for help with calculating the buried surfaces between the different domains of the cap proteins and the PH ring, and A. Klauer for technical assistance. *GAL::rrp44* strain was kindly provided by P. Mitchell (University of Sheffield) and D. Tollervey (University of Edinburgh). R. Parker, M. Wilkinson, M. Steiger and members of the van Hoof and Arraiano laboratories gave insightful comments on the manuscript. This research was supported by the Pew Scholarship Program in the Biomedical Sciences and by the National Institutes of Health (GM069900) to A.v.H. E.G.D and M.S.-R. were supported by The University of Texas at Houston Medical School-Summer Research Program. The work at the Instituto de Tecnologia Química e Biológica was supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal.

Supplementary Information

Supplementary Figures



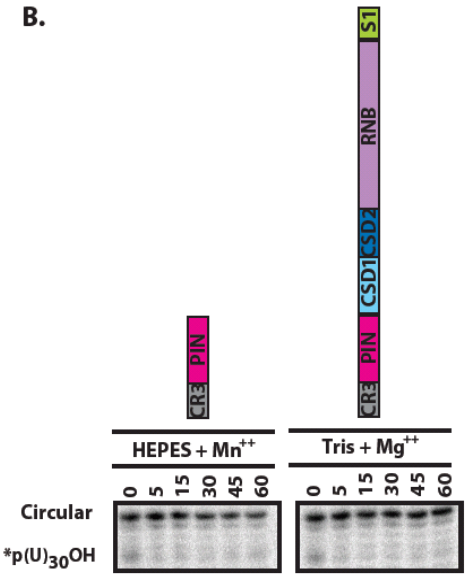


Figure S1. The PIN domain and RNB domain have similar levels of activity *in vitro*. A. The truncated Rrp44 shows robust ribonuclease activity in HEPES buffer containing Mn^{++} ions (left panel), but is much less active in the Tris buffer containing Mg^{++} ions that was previously used to biochemically characterize Rrp44 (middle panel) (Dziembowski et al., 2007). Both the identity of the buffer and the divalent ion contributes to this effect as HEPES + Mg^{++} and Tris + Mn^{++} allow for intermediate activity (data not shown). The activity level of truncated Rrp44 in HEPES/$MnCl_2$ was similar to that of full length Rrp44 in Tris/$MgCl_2$ buffer (compare left to right panels). It has previously been shown that the activity of the full length protein in the Tris/Mg buffer is due to the RNB domain. This strongly suggests that the activity of the PIN domain we detected is robust enough to contribute to overall exosome activity. Furthermore, the observation that the PIN and RNB domains are fully active under different conditions, suggests that one activity or the other might dominate under specific physiological conditions. **B.** The N-terminal region of Rrp44 degrades both linear and circular U_{30} RNA, but prefers the linear substrate (left panel). 40% of the circular substrate was degraded over a 60 minute time course, while 40% degradation of a linear U_{30} substrate under similar conditions typically occurred within 5 minutes. The circular substrate was generated by treating the 5′ end labeled substrate with T4 RNA ligase. Three observations confirm the identification of the circular substrate. First, it is only present after T4 RNA ligase treatment and has the mobility expected of a circular substrate. Second, the P^{32} labeling of the circular form is resistant to treatment with alkaline phosphatase, but the linear form is not (data not shown). Third, the full length Rrp44 degrades a linear substrate under conditions where the RNB exonuclease domain is active, but does not degraded the circular form (right panel).

The indicated recombinant proteins were purified from *E. coli* and incubated with 5′ end labeled U_{30} RNA in either 20 mM HEPES pH 7.5, 150 mM NaCl, 3 mM $MnCl_2$, 1 mM DTT or in 10 mM Tris pH 7.5, 75 mM NaCl, 1mM $MgCl_2$, 1 mM DTT.

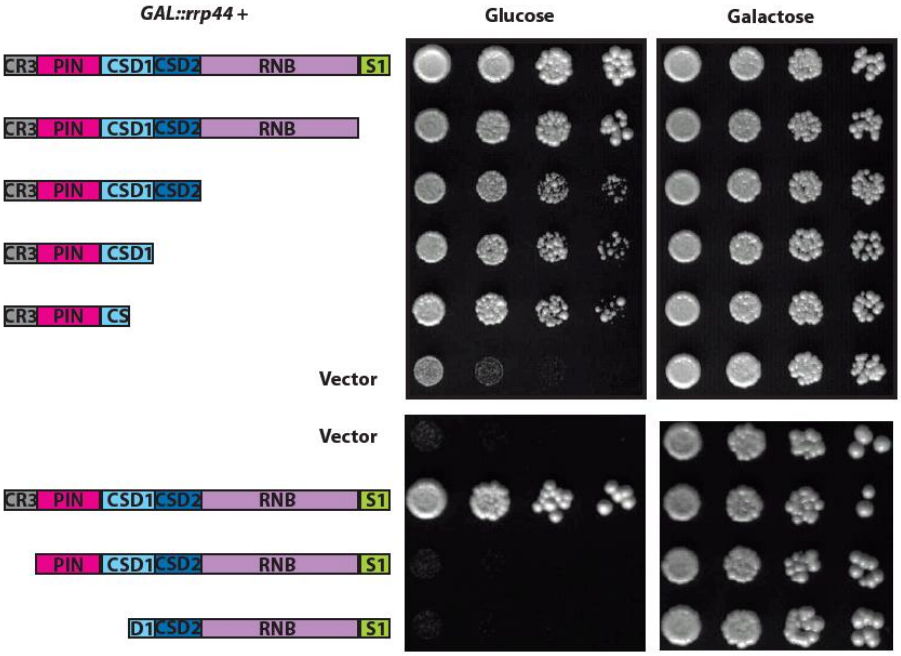


Figure S2. Complementation assay using a *GAL::rrp44* strain confirms that the N-terminal region of Rrp44 is important for exosome function. To confirm the complementation results seen in the *rrp44*Δ strain, the truncations described in the main text were tested in a strain containing *RRP44* under the control of the *GAL10* promoter (*GAL::rrp44*) (Mitchell et al., 1997). In this strain the *RRP44* gene is actively transcribed when the strain is grown on galactose as the sole carbon source. However, growth on glucose represses the *GAL10* promoter, and thus expression of *RRP44*. It has previously been reported that this *GAL::rrp44* strain fails to grow on glucose (Mitchell et al., 1997), but we observed slow growth, suggesting that low levels of Rrp44 remain. Each of the *RRP44* truncations was introduced into this strain. Similar to what was seen in the *rrp44*Δ strain, both full length and the S1 domain deletion restored close-to-wild type growth rates to the *GAL::rrp44* strain. Surprisingly, however, several versions of Rrp44 that are missing the RNB domain that supported very slow growth in the *rrp44*Δ strain supported much faster growth of the *GAL::rrp44* strain. As expected, further C-terminal truncation, or any truncation from the N-terminus failed to support growth. These results confirm that the previously uncharacterized N-terminal region is an important part of the exosome. Furthermore, these results suggest that Rrp44 may have two distinct functions. One function requires the N-terminus and is essential. The second function is important for normal growth rates and requires either the RNase II homologous region, or very low levels of the full length protein.

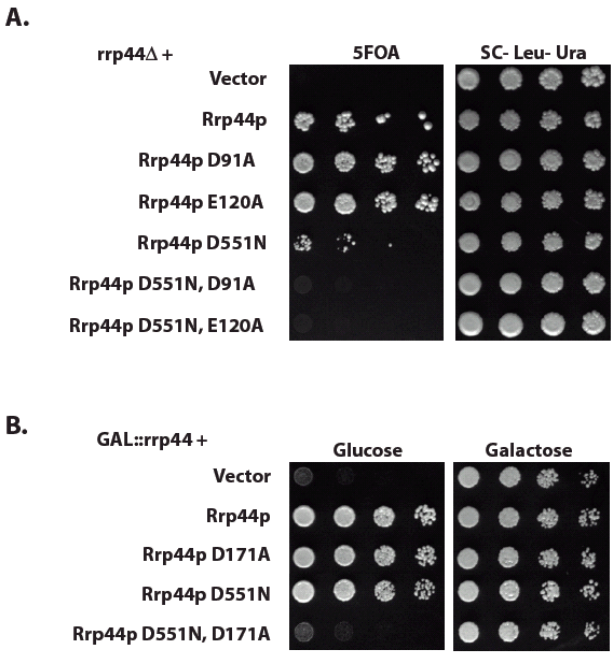


Figure S3. Mutations in the active site of the PIN domain are synthetically lethal with a mutation in the active site of the RNB domain. A. An *rrp44*Δ strain complemented by full length *RRP44* on a plasmid with a *URA3* marker was transformed with *LEU2* plasmids encoding each of the indicated point mutants. Growth on 5FOA indicates that the mutated Rrp44 can carry out the essential function of the exosome. **B.** A *GAL::rrp44* strain was transformed with plasmids encoding each of the indicated point mutants. Growth on glucose containing media indicates that the mutated Rrp44 can carry out the essential function of the exosome.

Supplementary Tables

Table S1: Strains used.

Strain Name	Genotype	Reference
BY4741	*MAT**a**, his3Δ1, leu2Δ0, ura3Δ0, met15Δ0*	Giaever *et al.* 2002
P108	*MAT**a**, his3Δ200, leu2Δ1, trp1, ura3-52, gal2, galΔ108, GAL10::rrp44*	Mitchell *et al.* 1997
yAV1115	*MAT**a**, his3Δ1, leu2Δ0, ura3Δ0, rrp44Δ::NEO [RRP44, URA3]*	This study

Table S2: Plasmids used.

Name	Description	Marker	Insert	Parent plasmid
***RRP44* plasmids**				
pAV317	*RRP44* promoter	*URA3*	oAV185 and oAV188 PCR	pRS415
pAV318	*RRP44* 3' flanking region	*URA3*	oAV189 and oAV186 PCR	pRS415
pAV344	*RRP44* promoter, residues 1-1001, *RRP44* 3' flanking region	*LEU2*	oAV185 and oAV186 PCR	pRS415
pAV363	*RRP44* promoter, residues 792-909, *RRP44* 3' flanking region	*URA3*	oAV354 and oAV452 PCR	pAV318
pAV370	*RRP44* promoter, residues 1-909, *RRP44* 3' flanking region	*LEU2*	Digest pAV363	pAV344
pAV331	*RRP44* promoter, residues 1-440, *RRP44* 3' flanking region	*URA3*	oAV185 and oAV190 PCR	pAV318
pAV343	*RRP44* promoter, residues 1-440, *RRP44* 3' flanking region	*LEU2*	Digest pAV331	pRS415
pAV364	*RRP44* promoter, residues 1-474, *RRP44* 3' flanking region	*LEU2*	oAV185 and oAV356 PCR	pAV343
pAV405	*RRP44* promoter, residues 1-400, *RRP44* 3' flanking region	*URA3*	oAV185 and oAV472 PCR	pAV318
pAV423	*RRP44* promoter, residues 1-400, *RRP44* 3' flanking region	*LEU2*	Digest pAV405	pRS415
pAV404	*RRP44* promoter, residues 1-324, *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV422	*RRP44* promoter, residues 1-324, *RRP44* 3' flanking region	*LEU2*	Digest pAV404	pRS415
pAV414	*RRP44* promoter, residues 1-235, *RRP44* 3' flanking region	*URA3*	oAV185 and oAV468 PCR	pAV318
pAV454	*RRP44* promoter, residues 1-235, *RRP44* 3' flanking region	*LEU2*	Digest pAV414	pRS415
PAV549	*RRP44* promoter, residues 84-1001, *RRP44* 3' flanking region	*LEU2*	oAV541 and oAV186 PCR	pAV341
pAV328	*RRP44* promoter, residues 360-1001, *RRP44* 3' flanking region	*URA3*	oAV187 and oAV186 PCR	pAV317
PAV341	*RRP44* promoter, residues 360-1001, *RRP44* 3' flanking region	*LEU2*	Digest pAV328	pRS415
pAV473	*RRP44* promoter, *RRP44* 3' flanking region	*URA3*	Digest pAV318	pAV317
pAV483	*RRP44* promoter, residues 1-1001, *RRP44* 3' flanking region	-	oAV185 and oAV186 PCR	pBluescript SK-
pAV506	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S), *RRP44* 3' flanking region	-	oAV482 and oAV569 mutagenesis	pAV483
pAV484	*RRP44* promoter, residues 1-1001 (D91A), *RRP44* 3' flanking region	-	oAV570 and oAV571 mutagenesis	pAV483
pAV505	*RRP44* promoter, residues 1-1001 (E120A), *RRP44* 3' flanking region	-	oAV484 and oAV578 mutagenesis	pAV483
pAV485	*RRP44* promoter, residues 1-1001 (D171A), *RRP44* 3' flanking region	-	oAV572 and oAV573 mutagenesis	pAV483
pAV486	*RRP44* promoter, residues 1-1001 (D198A), *RRP44* 3' flanking region	-	oAV574 and oAV575 mutagenesis	pAV483
pAV482	*RRP44* promoter, residues 360-1001 (D551N), *RRP44* 3' flanking region	-	oAV565 and oAV566 mutagenesis	pAV516
pAV523	*RRP44* promoter, residues 1-324 (D91A), *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV525	*RRP44* promoter, residues 1-324 (D171A), *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV526	*RRP44* promoter, residues 1-324 (D198A), *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV524	*RRP44* promoter, residues 1-324 (E120A), *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV522	*RRP44* promoter, residues 1-324 (C47S, C52S, C55S), *RRP44* 3' flanking region	*URA3*	oAV185 and oAV470 PCR	pAV318
pAV528	*RRP44* promoter, residues 1-324 (D91A), *RRP44* 3' flanking region	*LEU2*	Digest pAV523	pRS415
pAV530	*RRP44* promoter, residues 1-324 (D171A), *RRP44* 3' flanking region	*LEU2*	Digest pAV525	pRS415
pAV531	*RRP44* promoter, residues 1-324 (D198A), *RRP44* 3' flanking region	*LEU2*	Digest pAV526	pRS415
pAV529	*RRP44* promoter, residues 1-324 (E120A), *RRP44* 3' flanking region	*LEU2*	Digest pAV524	pRS415
pAV527	*RRP44* promoter, residues 1-324 (C47S, C52S, C55S), *RRP44* 3' flanking region	*LEU2*	Digest pAV522	pRS415
pAV515	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S), *RRP44* 3' flanking region	*LEU2*	Digest pAV506	pRS415
pAV502	*RRP44* promoter, residues 1-1001 (D91A), *RRP44* 3' flanking region	*LEU2*	Digest pAV484	pRS415
pAV514	*RRP44* promoter, residues 1-1001 (E120A), *RRP44* 3' flanking region	*LEU2*	Digest pAV505	pRS415
pAV503	*RRP44* promoter, residues 1-1001 (D171A), *RRP44* 3' flanking region	*LEU2*	Digest pAV485	pRS415
pAV504	*RRP44* promoter, residues 1-1001 (D198A), *RRP44* 3' flanking region	*LEU2*	Digest pAV486	pRS415
pAV501	*RRP44* promoter, residues 1-1001 (D551N), *RRP44* 3' flanking region	*LEU2*	Digest pAV482	pAV344
pAV536	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S, D551N), *RRP44* 3' flanking	*LEU2*	Digest pAV515	pAV501
pAV537	*RRP44* promoter, residues 1-1001 (D91A, D551N), *RRP44* 3' flanking region	*LEU2*	digest pAV502	pAV501
pAV538	*RRP44* promoter, residues 1-1001 (E120A, D551N), *RRP44* 3' flanking region	*LEU2*	Digest pAV514	pAV501
pAV539	*RRP44* promoter, residues 1-1001 (D171A, D551N), *RRP44* 3' flanking region	*LEU2*	Digest pAV503	pAV501
pAV540	*RRP44* promoter, residues 1-1001 (D198A, D551N), *RRP44* 3' flanking region	*LEU2*	Digest pAV504	pAV501
pGEX-4T1	GST	-	-	-
pGEX-Rrp44	GST-RRP44 residues 1-1001	-	GST-44 F and GST-44 R PCR	pGEX4T1
pGEX-454	GST-RRP44 residues 1-235	-	GST-44 F and GST-44 R PCR	pGEX4T1
pAV566	GST-RRP44 residues 1-235 (D171A)	-	GST-44 F and oAV648 PCR	pGEX4T1
pGEX-370	GST-RRP44 residues 1-990	-	GST-44 F and GST-44 R PCR	pGEX4T1

Table S3: Oligonucleotides used.

PCR oligos	Sequence
oAV185	5' ATATAGAGCTCAAAAACTGCCTACGTACCATTTAAC
oAV186	5' ATATACTCGAGCACCACCAAAATGTCAATTTTTTTG
oAV187	5' ATATAACTAGTAACAACATGTCCAATACCACCGTGATTTC
oAV188	5' ATATAACTAGTTTTGGCCTGTATGATGCAAG
oAV189	5' ATATAACTAGTTAGATGGAAGATGCTTCAGTATC
oAV190	5' ATATAACTAGTTTAACGCCTTGTTCTGATTCTGAC
oAV354	5' ATATAACTAGTTTAATTCCTCATTACTTGCCCGAC
oAV356	5' ATATAACTAGTTTAAATTGTACCGAGGTCTCTCACAAAG
oAV365	5' CGCGATCTAGAGGAATTCGAATATGATAGACATCC
oAV366	5' GGTATCGATAAGCTTCGGTGGTATTGTAGG
oAV367	5' CGCGAGGATCCTATGTAACAGCTTATCTACAGTGC
oAV370	5' CGCGAGGATCCTTAAAAAGGTTTGGCACATTTGCGCTT
oAV371	5' CGCGAGGATCCTTAGTCATTCCTTGCAGTGGTCAAA
oAV372	5' CGCGAGGATCCTTAATTAGCATATTTGTTGGTTTTACGAC
oAV404	5' CGATGGCATGCAGGTGTGAGGAAGAAAAAAAAACTGATCAGG
oAV405	5' CGATGGCATGCCTTCCTGGGACAGAAAAGGGTCGTAAAACC
oAV421	5' CGATGGCATGCACTGCAAGGAATGACCTTGGGGTCG
oAV422	5' CGAGGATCCTTAACACCTCACTGTCCCAACTAGTGTAGC
oAV423	5' CGAGGATCCTTATTCTTCCTCTGTAGATTGATCAGTGCC
oAV424	5' CGCGAGGATCCCAAAAGCCGCTGTTTACACATTTATAAAC
oAV425	5' CGCGACTCGAGCTGCTACCTTTGCCAGTGGTAAACC
oAV426	5' CGCGATCTAGAGGAATACTGGCAGTGCGATTTTCATGC
oAV430	5' CGCGAGGATCCATGGTATAATTAAAATATATAAACACATAGAT
oAV431	5' CGCGAAGCTTCGTATTAGATATGTGCACTATAGTTGTCG
oAV432	5' CGCGATCTAGAGCGTTATACTGCGTTATACTGCG
oAV452	5' CGAGCGGCCGCGGTGTGTGGATCCCGAAGATCC
oAV468	5' ATATAACTAGTTTAAATAGAGTCTCTGATGTCGTCTGCA
oAV470	5' ATATAACTAGTTTATGATTGAGGTAGCAGTTCCACGATAAC
oAV472	5' ATATAACTAGTTTACCAGGATCTTCTTTGTATATATACAA
oAV482	5' GATCGGACATCCCAAGTCTTTCTAGAAGTAGTACCAAGAGTCCGCAAATTGTCG
oAV484	5' TTGTCCCCCAGATTGTTCTAGATGCAGTGAGAAACAAGTC
oAV500	5' CGCGAGGATCCTATTAGGCATAACGACCTTTCAATGGAATTACCG
oAV501	5' CGCGAGGATCCTATTACTTTAGAGACCTTGTATGTAAAGAGGC
oAV506	5' CGCGAGGATCCTAGTTTAAATGTATCTCTTACTAGAATAGTCT
oAV507	5' CGCGAGGATCCTAGTTTAAGTTGAGTCAAAACATTCTATTTCGGC
oAV527	5' CGCGATCTAGACGCTTGTGTGACTACGCAAAGATTTAG
oAV528	5' CGCGATCTAGAAACAGTATGGCCCCAGAAACTGGTGATCACGTCGTAGGG
oAV529	5' CGCGATCTAGAAACAGATATGAAGTACGGGAAGTTAAGAAACGGG
oAV530	5' CGCGATCTAGACTCTTGGTTAGTCGAAATGCTGG
oAV531	5' CGCGATCTAGACATACAAGATGATTCCATCTGTAAACGATTTTGTAATCG
oAV532	5' CGCGATCTAGACATACAAGATGACTACAGGACGCGATGCTGGTTTCGG
oAV533	5' CGCGAGAGCTCGGAATACTGGCAGTGCGATTTT
oAV534	5' CGCGAGAGCTCGCGTTATACTGCGTTATACTG
oAV541	5' ATATAACTAGTATGGGTAAGCATTACGTCG
oAV548	5' TACAGTATATCGAGTTACTACCAAATGCAGACGACATCAGAGACTCTATTGGTCGACGGATCCCCGGGTT
oAV549	5' CACCACCAAAATGTCAATTTTTTTGCCATTTTCTAAATAGTTTTCTTCACTTATCGATGAATTCGAGCTCGTT
oAV556	5' CGCCTGCAGGTCGACGGATCCCCGGGTTAATTAATCCA
oAV557	5' CGCAAGCTTTGCCGGTAGAGGTGTGGTCAATAAGAGCGAC
oAV558	5' CGCCTGCAGCGGCATAACGACCTTTCAATGGAATTACCGATAGCA
oAV559	5' CGCCTGCAGCCTTTAGAGACCTTGTATGTAAAGAGGCACT
oAV560	5' CGCCTGCAGCGTTGCCGTTACCTCTCATTTTTTCGGCGGT
oAV561	5' CGCCTGCAGCAATGTATCTCTTACTAGAATAGTCTATATA
oAV562	5' CGCCTGCAGCAGTTGAGTCAAAACATTCTATTTCGGCTTC
oAV563	5' CGCCTGCAGCCTCCTCCTTGACCGTAAGTATTTCTTTAAA
oAV565	5' CTCCAGGATGTGTTGATATCAACGATGCCCTACATGCG
oAV566	5' CGCATGTAGGGCATCGTTGATATCAACACATCCTGGAG
oAV569	5' CGACAATTTGCGGACTCTTGGTACTACTTCTAGAAAGACTTGGGATGTCCGATC
oAV570	5' GCATTACGTCGTCTTGGCCACCAACGTGGTGTTAC
oAV571	5' GTAACACCACGTTGGTGGCCAAGACGACGTAATGC
oAV572	5' CGATTAATGACAGAAACGCGCGCGCTATAAGGAAAACCTGTCAATGG
oAV573	5' CCATGACAGGTTTTCCTTATAGCGCGCGCGTTTCTGTCATTAATCG
oAV574	5' CGTTCTTGTTACCAACGCTCGTTTGAATAGAGAAGC
oAV575	5' GCTTCTCTATTCAAACGAGCGTTGGTAACAAGAACG
oAV578	5' GACTTGTTTCTCACTGCATCTAGAACAATCTGGGGGACAA
oAV604	5' CGCGACTGCAGCTGCCCCAGAAACTGGTGATCACGTCGTAGG
oAV605	5' CGCGACTGCAGCTAAGTACGGGAAGTTAAGAAACGGGATG
oAV606	5' CGCGACTCGAGCTGCTACCTTTGCCAGTGGTAAACC
oAV607	5' CGCGACTGCAGCTATTCCATCTGTAAACGATTTTGTAATCGG
oAV608	5' CGCGACTGCAGCTACTACCGGACGCGATGCTGGTTTCGG
oAV609	5' CGCGCTCGAGCGTATTAGATATGTGCACTATAGTTGTCG
oAV648	5' ATATACTCGAGTTAAATAGAGTCTCTGATGTCGTCTGCA
GST-44 F	5' CGCAGAATTCATGTCAGTTCCCGCTATCG
GST-44 R	5' CCGCTCGAGTATCCTGATACTGAAGCATCTTCC
Northern oligos	
oRP141	5' AATTGATCTATCGAGGAATTCC
oRP100	5' GTCTAGCCGCGAGGAAGG
oJA003	5' TGAGAAGGAAATGACGCT
oRP993	5' CGAACGACAAGCCTACTCG

References

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, 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, Mitchell P. 1999b. The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development 13*:2148-2158.

Amblar M, Arraiano CM. 2005. A single mutation in Escherichia coli ribonuclease II inactivates the enzyme without affecting RNA binding. *FEBS J 272*:363-374.

Amblar M, Barbas A, Fialho AM, Arraiano CM. 2006. Characterization of the functional domains of Escherichia coli RNase II. *J Mol Biol 360*:921-933.

Anderson JS, Parker RP. 1998. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J 17*:1497-1506.

Andrade JM, Pobre V, Silva IJ, Domingues S, Arraiano CM. 2009. The role of 3'-5' exoribonucleases in RNA degradation. *Prog Mol Biol Transl Sci 85*:187-229.

Andrei MA, Ingelfinger D, Heintzmann R, Achsel T, Rivera-Pomar R, Luhrmann R. 2005. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA 11*:717-727.

Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN. 2004. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem 279*:16471-16478.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2008. New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end product. *J Biol Chem 283*:13070-13076.

Bashkirov VI, Scherthan H, Solinger JA, Buerstedde JM, Heyer WD. 1997. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *J Cell Biol 136*:761-773.

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

Bunker RD, McKenzie JL, Baker EN, Arcus VL. 2008. Crystal structure of PAE0151 from Pyrobaculum aerophilum, a PIN-domain (VapC) protein from a toxin-antitoxin operon. *Proteins 72*:510-518.

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

Cairrão F, Arraiano C, Newbury S. 2005. Drosophila gene *tazman*, an orthologue of the yeast exosome component Rrp44/Dis3, is differentially expressed during development. *Dev Dyn 232*:733-737.

Carpousis AJ, Van Houwe G, Ehretsmann C, Krisch HM. 1994. Copurification of E. coli RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell 76*:889-900.

Chekanova JA, Dutko JA, Mian IS, Belostotsky DA. 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.

Cougot N, Babajko S, Seraphin B. 2004. Cytoplasmic foci are sites of mRNA decay in human cells. *J Cell Biol 165*:31-40.

Daines DA, Wu MH, Yuan SY. 2007. VapC-1 of nontypeable Haemophilus influenzae is a ribonuclease. *J Bacteriol 189*:5041-5048.

de la Cruz J, Kressler D, Tollervey D, Linder P. 1998. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J 17*:1128-1140.

Deutscher MP, Li Z. 2001. Exoribonucleases and their multiple roles in RNA metabolism. *Prog Nucleic Acid Res Mol Biol 66*:67-105.

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

Eberle AB, Lykke-Andersen S, Muhlemann O, Jensen TH. 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature structural & molecular biology 16*:49-55.

Fatica A, Tollervey D, Dlakic M. 2004. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *Rna 10*:1698-1701.

Frazao C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature 443*:110-114.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M,

Philippsen P, Davis RW, Johnston M. 2002. Functional profiling of the Saccharomyces cerevisiae genome. *Nature 418*:387-391.

Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *Embo J 25*:5117-5125.

Ingelfinger D, Arndt-Jovin DJ, Luhrmann R, Achsel T. 2002. The human LSm1-7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA 8*:1489-1501.

Johnson AW, Kolodner RD. 1995. Synthetic lethality of sep1 (xrn1) ski2 and sep1 (xrn1) ski3 mutants of Saccharomyces cerevisiae is independent of killer virus and suggests a general role for these genes in translation control. *Mol Cell Biol 15*:2719-2727.

Koonin EV, Wolf YI, Aravind L. 2001. Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. *Genome Res 11*:240-252.

Levin I, Schwarzenbacher R, Page R, Abdubek P, Ambing E, Biorac T, Brinen LS, Campbell J, Canaves JM, Chiu HJ, Dai X, Deacon AM, DiDonato M, Elsliger MA, Floyd R, Godzik A, Grittini C, Grzechnik SK, Hampton E, Jaroszewski L, Karlak C, Klock HE, Koesema E, Kovarik JS, Kreusch A, Kuhn P, Lesley SA, McMullan D, McPhillips TM, Miller MD, Morse A, Moy K, Ouyang J, Quijano K, Reyes R, Rezezadeh F, Robb A, Sims E, Spraggon G, Stevens RC, van den Bedem H, Velasquez J, Vincent J, von Delft F, Wang X, West B, Wolf G, Xu Q, Hodgson KO, Wooley J, Wilson IA. 2004. Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from Archaeoglobus fulgidus at 1.90 A resolution. *Proteins 56*:404-408.

Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R. 2005. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol 7*:719-723.

Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell 127*:1223-1237.

Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell 29*:717-728.

Lorentzen E, Conti E. 2005. Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol Cell 20*:473-481.

Lorentzen E, Walter P, Fribourg S, Evguenieva-Hackenberg E, Klug G, Conti E. 2005. The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nature structural & molecular biology 12*:575-581.

Mackie GA. 1998. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature 395*:720-723.

Meaux S, Van Hoof A. 2006. Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA 12*:1323-1337.

Mian IS. 1997. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic Acids Res 25*:3187-3195.

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

Moser MJ, Holley WR, Chatterjee A, Mian IS. 1997. The proofreading domain of Escherichia coli DNA polymerase I and other DNA and/or RNA exonuclease domains. *Nucleic Acids Res 25*:5110-5118.

Muhlrad D, Decker CJ, Parker R. 1995. Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol Cell Biol 15*:2145-2156.

Navarro MV, Oliveira CC, Zanchin NI, Guimaraes BG. 2008. Insights into the mechanism of progressive RNA degradation by the archaeal exosome. *J Biol Chem 283*:14120-14131.

Py B, Causton H, Mudd EA, Higgins CF. 1994. A protein complex mediating mRNA degradation in Escherichia coli. *Mol Microbiol 14*:717-729.

Rockmill B, Lambie EJ, Roeder GS. 1991. Spore enrichment. *Methods Enzymol 194*:146-149.

Schneider C, Anderson JT, Tollervey D. 2007. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol Cell 27*:324-331.

Sikorski RS, Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics 122*:19-27.

Symmons MF, Jones GH, Luisi BF. 2000. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure 8*:1215-1226.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science 295*:2262-2264.

van Hoof A, Lennertz P, Parker R. 2000a. 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 19*:1357-1365.

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

van Hoof A, Staples RR, Baker RE, Parker R. 2000c. Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol Cell Biol 20*:8230-8243.

Wang HW, Wang j, Ding F, Callahan K, Bratkowski J, Butler JS, Nogles E, Ke A. 2007. Architecture of the yeast Rrp44-exosome complex suggests routes of RNA recruitment for 3′end processing. *Proc Nat Acad Sci USA 104*:16844-16849.

Wilson MA, Meaux S, van Hoof A. 2007. A genomic screen in yeast reveals novel aspects of nonstop mRNA metabolism. *Genetics 177*:773-784.

Yang W. 2008. An equivalent metal ion in one- and two-metal-ion catalysis. *Nature structural & molecular biology 15*:1228-1231.

Zheng D, Ezzeddine N, Chen CY, Zhu W, He X, Shyu AB. 2008. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J Cell Biol 182*:89-101.

Zuo Y, Deutscher MP. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res 29*:1017-1026.

Chapter 3

Demonstrating the importance of the CR3 region of yeast Rrp44

This chapter contains data published in:
Schaeffer D, **Reis FP**, Johnson SJ, Arraiano CM, van Hoof A. 2012. The CR3 motif of Rrp44 is important for interaction with the core exosome and exosome function. *Nucleic Acids Res*. *40*:9298-9307.

The author of this Dissertation was responsible for the construction of the clones for the GST-fusion proteins, protein purification and biochemical nuclease assays experiments described in this chapter.

Abstract

The ten subunit RNA exosome is involved in a large number of diverse RNA processing and degradation events in eukaryotes. These reactions are carried out by the single catalytic subunit, Rrp44/Dis3, which is composed of three parts that are conserved throughout eukaryotes. The exosome is named for the 3' to 5' exoribonuclease activity provided by a large C-terminal region of the Rrp44 subunit that resembles other exoribonucleases. Rrp44 also contains an endoribonuclease domain. Finally, the very N-terminus of Rrp44 contains three Cys residues that are conserved in many eukaryotes but have no known function. These three conserved Cys residues cluster with a previously unrecognized conserved His residue in what resembles a metal-ion-binding site. Genetic and biochemical data show that this CR3 motif affects both endo- and exonuclease activity *in vivo* and both the nuclear and cytoplasmic exosome, as well as the ability of Rrp44 to associate with the other exosome subunits. These data provide the first direct evidence that the exosome-Rrp44 interaction is functionally important and also provides a molecular explanation for the functional defects when the conserved Cys residues are mutated.

Introduction

The RNA exosome is involved in a wide variety of RNA processing and degradation reactions in both the nucleus and the cytoplasm. First, the nuclear exosome processes a subset of RNAs from longer precursors. For example, it processes a 300-nt 7S precursor into the 160-nt 5.8S rRNA (Mitchell et al., 1997). Second, the nuclear exosome completely degrades some RNAs that are byproducts of gene expression, including the 5' external transcribed spacer that is part of the rRNA precursor (de la Cruz et al., 1998). Third, the nuclear exosome degrades aberrant RNAs that fail to complete proper processing, including incompletely modified initiator tRNA (Kadaba et al., 2004). Fourth, the exosome is involved in one of two general pathways of cytoplasmic mRNA decay (Jacobs Anderson & Parker, 1998). Fifth, the cytoplasmic exosome is especially important for degrading aberrant mRNAs, including those that lack a stop codon (nonstop mRNAs) as well as those that are cleaved by a ribozyme (van Hoof et al., 2002; Meaux & van Hoof, 2006).

Although the exosome has a wide variety of substrates, it acts very specifically on those substrates. For example, the exosome degrades initiator tRNA lacking a single methyl group, but not normal initiator tRNA or other tRNAs that lack a modification (Kadaba et al., 2004; Chernyakov et al., 2008). A second example of the exosome's specificity is that the exosome degrades both the poly(A) tail and the body of nonstop mRNAs (van Hoof et al., 2002), while the poly(A) tail of normal mRNAs can only be removed by dedicated deadenylases and not by the exosome (Tucker et al., 2001). What is not yet known is how the exosome carries out these diverse functions while maintaining specificity.

A series of X-ray crystallography and EM studies have resolved the structural organization of the yeast exosome (Buttner et al., 2005; Lorentzen &

Conti, 2005; Lorentzen et al., 2005; Liu et al., 2006; Wang et al., 2007; Lorentzen et al., 2008; Bonneau et al., 2009). The exosome contains a core of ten proteins that are shared between the nuclear and cytoplasmic exosome. At least in fungi and metazoans, only the Rrp44 subunit (also known as Dis3) is catalytically active (Liu et al., 2006; Dziembowski et al., 2007). Rrp44 was initially identified as an exonuclease with similarity to the RNase II family (Mitchell et al., 1997). The similarity to RNase II includes the catalytic RNB domain and three OB-fold RNA binding domains (CSD1, CSD2 and S1; (Frazão et al., 2006). We and others have shown that this exonuclease activity is not essential for viability because Rrp44 contains a second domain in its N-terminus (PIN) with endonuclease activity (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). Mutations that inactivate each of the nuclease activities individually (hereafter referred to as *rrp44-endo*$^{-}$ and *rrp44-exo*$^{-}$) are viable, but simultaneously mutating both (*rrp44-endo*$^{-}$*-exo*$^{-}$) is lethal, suggesting a functional overlap (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). Although the structure containing nine exosome subunits is thought to be conserved in other eukaryotes, whether or not Rrp44 in other eukaryotes stably associates with the other nine subunits is controversial.

The five recognized domains of Rrp44 explain much of the sequence conservation in Rrp44 orthologs. A notable exception is a conserved motif near the N-terminus that has no known function (Fig. 1A). This motif has been named CR3 (Cysteine-Rich with three cysteines; Cairrão et al., 2005). In the process of trying to understand how the exosome structure relates to its diverse functions, we initially constructed a mutant Rrp44 in which all three Cys residues were changed to Ser. This *rrp44-CR3* allele caused slow growth (Schaeffer et al., 2009), and severely reduced the exosome's ability to degrade nonstop mRNAs and ribozyme cleaved mRNAs (Schaeffer & van Hoof, 2011). Interestingly, neither the *rrp44-endo*$^{-}$ nor the *rrp44-exo*$^{-}$ mutation caused a similar defect in mRNA decay,

suggesting that the *rrp44-CR3* mutation somehow affected both catalytic activities. The protein encoded by *rrp44-CR3* accumulated to slightly reduced levels (Schaeffer & van Hoof, 2011). However, this is probably not the cause of the observed phenotypes since overexpression of the *rrp44-CR3* allele from a high copy plasmid did not restore growth. In addition, haplo-insufficiency of *RRP44* did not reduce growth similar to that of *rrp44-CR3* (data not shown). All of these data can be explained if the *rrp44-CR3* mutation reduces both exo- and endonuclease activities of the exosome, thus causing the slow growth and stabilization of nonstop mRNAs. However, the observation that *rrp44-CR3* is viable suggests that it does not completely eliminate catalytic activity.

Recently, the X-ray crystal structure of Rrp44 bound to two other exosome subunits was published, providing the first structural insight into the CR3 motif (Bonneau et al., 2009). We noted that the CR3 motif is close to a His residue and that these four residues are oriented in a tetrahedral configuration that is reminiscent of a mono-metal-binding site. Here, we present genetic, biochemical, and phylogenetic data that offer a coherent explanation of how the CR3 motif alters both the endonuclease and exonuclease activities of Rrp44 *in vivo*: the direct effect of the CR3 mutation is likely a modest perturbation of the local folding and/or structural stability of Rrp44, possibly because of the loss of a metal ion. Even though the CR3 motif is distant from both catalytic sites, this folding defect has two important consequences. First, the CR3 motif is close to residues contacting another exosome subunit, and alterations in the CR3 motif reduce copurification of the other exosome subunits. This, in turn, can explain the defect in exonuclease function *in vivo*. Second, changes in the structure of the CR3 motif appear to cause modest changes in the endonuclease active site, possibly because the two sites are connected by a single α-helix.

Results

Three conserved cysteines in Rrp44 cluster with a conserved histidine

In addition to its five recognizable domains, Rrp44 orthologs contain a conserved motif that is found N-terminal to the PIN domain (Fig. 1A). We have previously shown that deleting as many as 766 residues from the C-terminus of Rrp44 can be tolerated, but that deleting as little as 84 residues from the N-terminus results in lethality, suggesting that this conserved N-terminal region is important (Schaeffer et al., 2009). This region was named CR3 (cysteine-rich domain with three cysteines) because it contains a Cx_4Cx_2C motif that does not resemble any other known domain or motif (Cairrão et al., 2005). A 3.0 Å crystal structure of Rrp44 was recently solved, providing the first structural insight into this motif (Bonneau et al., 2009). The structure reveals that the three conserved Cys residues are physically adjacent to each other and that H_{184} is in close proximity to the three conserved Cys residues (Fig. 1C).

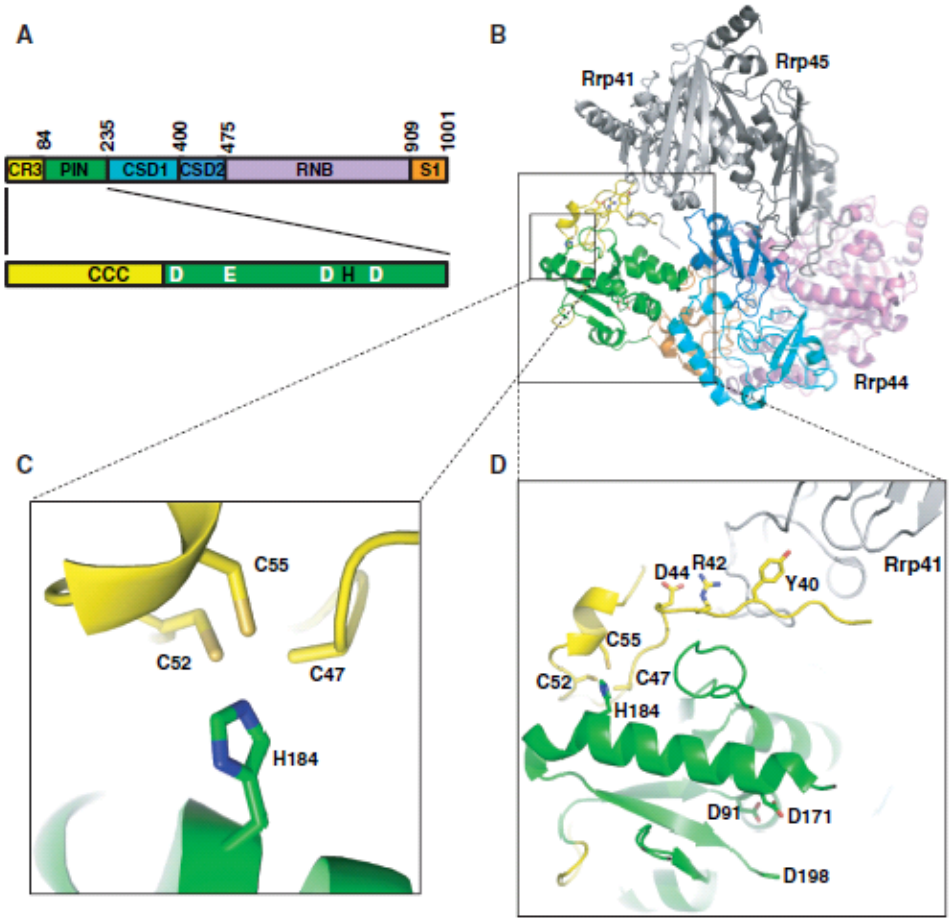


Figure 1. Rrp44 contains a conserved CCCH motif. A. Top: Diagram depicting the five recognized domains in Rrp44 and the CR3 motif. PIN denotes the endonuclease domain, CSD1 and CSD2 denote RNA-binding cold shock domains, RNB denotes the exonuclease domain and S1 denotes an RNA-binding S1 domain. Collectively the CSD1, CSD2, RNB and S1 domains are responsible for exonuclease activity. Bottom: the N-terminus of Rrp44 contains a conserved CCCH motif (black letters) in addition to the catalytic residues of the PIN domain (white letters). **B.** Overview of the Rrp44 structure (colored as in panel A) bound to exosome subunits Rrp41 and Rrp45 (PDB ID 2WP8; Bonneau et al., 2009). **C.** The three conserved Cys residues and the conserved His residue form a tetrahedral cluster in the crystal structure. Note that the sulfur atom of Cys47 was not modeled. **D.** The CR3 motif is physically connected to the exosome binding site (Y40, R42 and D44) and the endonuclease active site of the PIN domain (D91, D171 and D198). Note that the D198 side chain was not modeled. Structures were rendered using PyMOL (The PyMOL Molecular Graphics System. Version 1.5.0.1 Schrödinger, LLC.).

To examine whether H_{184} was conserved and co-occurred with the CR3 motif, we performed a multiple sequence alignment of 20 Rrp44 homologs from very diverse eukaryotes using ClustalW. The alignment showed that both the CR3 motif and H_{184} are conserved in 17 of the 20 included proteins (Supplementary Fig. S1A). Strikingly, two of the Rrp44 orthologs (from *Ciona intestinalis* and *Trichomonas vaginalis*) lack two or all three of the conserved Cys residues of the CR3 motif and also lack H_{184}. The 20^{th} protein, hDis3L1 (and Dis3L1 from other vertebrates), contains His_{184} together with a variation on the CR3 motif with five residues separating the second and third Cys instead of just two. Thus, His_{184} is conserved in 18 Rrp44 homologs that have the CR3 motif, but not in the two homologs that lack the CR3 motif. H_{184} is part of the PIN endonuclease domain, but as shown in Supplementary Fig. S1B, it is not present in PIN domains that are not part of Rrp44 orthologs. Thus, the previously identified CR3 co-occurs and is in close proximity to a conserved histidine. Conserved CCCH motifs in other proteins form binding sites for a metal ion (generally Zn^{2+}). We note, however that the organization in the Rrp44 amino acid sequence and protein fold is distinct from known Zn-binding domains explaining why it previously went unrecognized (See 'Discussion' section).

A triple mutation in the CR3 motif affects growth, but single point mutations do not have the same effect

We have previously shown that mutating the three conserved cysteine residues of the CR3 motif to serine (C47S, C52S, C55S; hereafter referred to as *rrp44-CR3*) resulted in a slow growth phenotype (Schaeffer et al., 2009). One possibility is that two of the conserved Cys residues form a disulfide bond, which is suggested by the proximity of the modeled sulfur atoms in the Rrp44 crystal structure. If this growth defect was caused by the disruption of a disulfide bond between two of the cysteines, then single point mutations in the two residues forming the bond

should have the same effect, whereas the non-interacting Cys would not be expected to affect the disulfide bond, and thus should not affect growth. If, however, the CR3 motif forms a metal-binding site together with H_{184}, any one of the single mutations might decrease affinity for the metal and a triple mutant might further decrease metal affinity. Thus, to distinguish between these two proposed roles of the CR3 motif we created single mutants, C47S, C52S, C55S and H184A, and compared their effect to the *rrp44-CR3* triple mutant. For this assay, we used a *rrp44Δ* strain containing a plasmid with wild-type *RRP44* and *URA3* genes. This strain was transformed with a second plasmid that contained the mutant *RRP44* and *LEU2* genes. If this second plasmid can provide functional Rrp44, then the strain can lose the *URA3* plasmid, resulting in growth on plates containing 5-Fluoroorotic Acid (5-FOA). As shown in Fig. 2A, the single mutants showed a variety of growth phenotypes. The C47S mutant grew worse than the other single mutants, with no obvious growth upon short incubation (2 days; Fig. 2A left panel). C52S grew slightly worse than wild-type, which was obvious at 2 days. Upon longer incubation (6 days; Fig. 2A middle panel) the C47S mutant showed some growth, and the C52S colony size became almost indistinguishable from wild type. The C55S and H184A single mutants grew as well as the wild-type strain after both short and long incubation.

As previously shown, the *rrp44-CR3* mutant has a defect in the rapid degradation of mRNAs that lack a stop codon (Schaeffer & van Hoof, 2011). Because the *rrp44-CR3* mutant stabilizes the *his3*-nonstop mRNA, a *rrp44-CR3 his3-nonstop* strain produces sufficient histidine to support growth in the absence of added histidine, whereas an isogenic *RRP44 his3-nonstop* strain fails to grow under these conditions (Schaeffer & van Hoof, 2011). To determine which residues of the CR3 motif are needed for nonstop mRNA decay, we tested whether the C47S, C52S, C55S or H184A single mutants could suppress a *his3-nonstop* allele. Importantly, the growth results in this assay are counterintuitive:

growth occurs when nonstop mRNA decay is defective, where as functional nonstop mRNA decay limits growth. As shown in Figure 2B, the *his3-nonstop* defects mirrored the growth defects in Figure 2A. The *rrp44-CR3* allele is defective in nonstop mRNA decay, and thus grows. The single C47S mutant shows almost the same defect. C52S showed a less severe defect, and C55S and H184A were similar to wild-type.

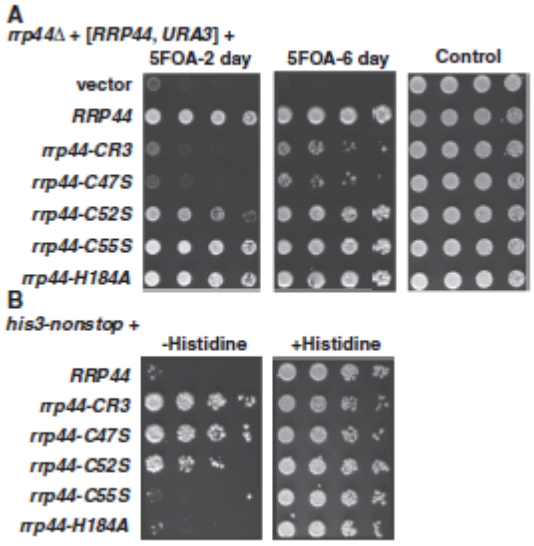


Figure 2. Point mutations in the conserved CCCH motif are inconsistent with a disulfide bond, but are consistent with metal binding. As previously reported, simultaneously mutating the three conserved Cys residues to Ser causes a slow growth phenotype (**A**) and stabilization of *his3-nonstop* mRNA (**B**).

Results similar to ours, where mutating individual metal binding residues have different effects, have been reported for various Zn^{2+} containing proteins (See 'Discussion' section). Presumably this indicates that mutating a single Cys residue to Ser reduced the affinity for Zn-ions, but does not completely eliminate Zn^{2+} binding. Thus, the phenotypes of single amino acid point mutants are consistent with the CCCH motif forming a binding site for a metal ion, with single point mutations decreasing the affinity for the ion and a triple mutation

further decreasing this affinity. In contrast, the phenotypes of the single-point mutants appear inconsistent with a disulfide bond.

Genetic interactions indicate that the CR3 motif affects multiple functions of the exosome *in vivo*

As explained in the Introduction, the known phenotypes of the *rrp44-CR3* mutant can best be explained if this mutation decreases both the endo- and exonuclease activities of Rrp44, but does not completely abolish activity. To better understand the connections between the CR3 motif and nuclease activities of the exosome, we tested for genetic interactions between the *rrp44-CR3* mutation, and mutations inactivating the catalytic activities of the exosome.

The slow growth phenotype of *rrp44-CR3* resembles that of a D551N point mutation that inactivates the exonuclease activity of Rrp44 (hereafter called the *rrp44-exo*$^-$ mutation). One possibility is that the *rrp44-CR3* mutation somehow prevents the exonuclease function of Rrp44, and thereby results in slow growth. This predicts that adding the *rrp44-CR3* mutation to the already inactive *rrp44-exo*$^-$ mutation would have no additional effect. Another possibility is that the reason *rrp44-CR3* grows slowly is distinct from the reason why *rrp44-exo*$^-$ grows slowly. In this case, combining the two mutations might have an additive effect. To test these possibilities, we combined *rrp44-CR3* with *rrp44-exo*$^-$. Fig. 3A shows that an *RRP44* plasmid combining the *rrp44-CR3* and *rrp44-exo*$^-$ mutations (*i.e.* C47S, C52S, C55S, D551N) did not support growth. As previously reported, the plasmids encoding either *rrp44-CR3* or *rrp44-exo*$^-$ allowed for slow growth compared to wild-type *RRP44*. We conclude that the *rrp44-CR3* mutation is synthetically lethal with the *rrp44-exo*$^-$ mutation. Using the same approach, Fig. 3B shows that *rrp44-CR3* is also synthetically lethal with a deletion of the exonuclease domain (*rrp44-exoΔ*). Although *rrp44-CR3* and *rrp44-exo*$^-$ have some of the same molecular effects, disrupting the exonuclease activity further worsens

the growth of *rrp44-CR3*. We conclude that the *rrp44-CR3* mutation does not completely eliminate *in vivo* exonuclease function, and that the slow growth phenotypes of *rrp44-CR3* and *rrp44-exo⁻* are distinct. Furthermore, since combining the *rrp44-endo⁻* and *rrp44-exo⁻* alleles is also synthetic lethal, the *rrp44-CR3-exo⁻* synthetic lethality is consistent with *rrp44-CR3* affecting endonuclease activity.

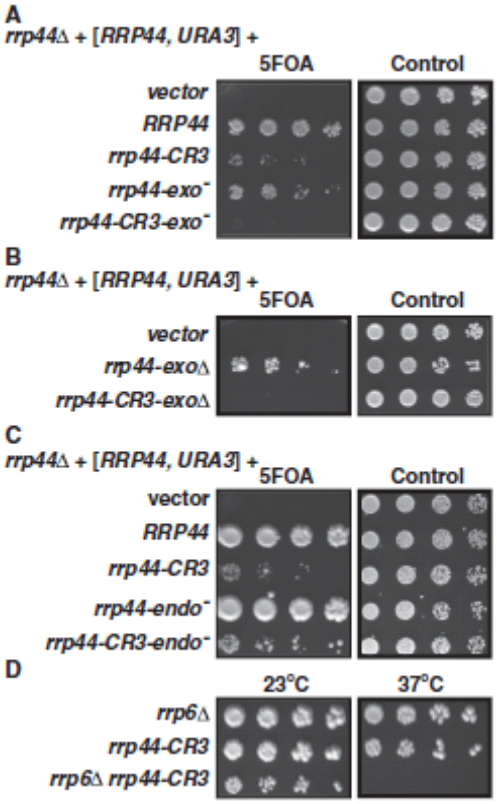


Figure 3. Synthetic lethality of *rrp44-CR3* with mutations in the exosome. Simultaneously mutating the three conserved Cys residues of Rrp44 to Ser is synthetically lethal with point mutations disrupting the Rrp44 exonuclease (**A**), with deletion of the Rrp44 exonuclease domains (**B**) or with deletion of the Rrp6p exonuclease (**D**), but not point mutations disrupting the Rrp44 endonuclease (**C**).

Next, to test the genetic interaction between the CR3 motif and the endonuclease domain of Rrp44, we combined the *rrp44-CR3* allele with *rrp44-*

endo^{-} mutant. Figure 3C shows that this combination has the same growth phenotype as the *rrp44-CR3* single mutant. This is also consistent with the *rrp44-CR3* mutation not completely eliminating exonuclease function, since elimination of exonuclease function (*e.g. rrp44-exo*$^{-}$) is synthetic lethal with *rrp44-endo*$^{-}$.

We next tested for genetic interactions with *rrp6Δ*. Rrp6 is an exoribonuclease that associates with the nuclear form of the exosome, and thus provides it with a third catalytic domain. Figure 3D shows that the *rrp6Δ rrp44-CR3* double mutant grows significantly slower than either single mutant at 23°C and fails to grow at 37°C, indicating that *rrp6Δ* has a synthetic growth defect with *rrp44-CR3*.

Since the nuclear exosome is essential, but the cytoplasmic exosome is not, the above growth phenotypes reflect effects of the mutations on nuclear exosome function. To extend our observations to the cytoplasmic exosome, we tested for effects on mRNA decay. Normal mRNAs are mainly degraded through a decapping pathway. In the absence of decapping activity, normal mRNAs are degraded by the exonuclease activity of the exosome (Jacobs Anderson & Parker, 1998). Thus, mutations that inactivate decapping are synthetically lethal with mutations that inactivate the exonuclease activity of the exosome (Schaeffer & van Hoof, 2011). To test whether the *rrp44-CR3* allele reduces exosome-mediated decay of normal mRNAs, we initially looked for genetic interactions with *dcp1-2*. *dcp1-2* is a temperature sensitive mutation in a subunit of the decapping enzyme. Since decapping is not required for viability, this strain is viable at all temperatures, but requires exosome-mediated mRNA decay at the restrictive temperature. Figure 4A shows that, as expected, the *dcp1-2 rrp44-CR3* double mutant is viable at the permissive temperature (23°C), but inviable at the restrictive temperature for *dcp1-2* (37°C). Therefore, the *rrp44-CR3* mutation likely disrupts exosome-mediated decay of normal mRNAs.

The *dcp1-2 rrp44-CR3* double mutant also allowed us to directly measure the effect of *rrp44-CR3* on the degradation of an endogenous mRNA, *GAL7*. For this experiment we grew *dcp1-2* strains at the permissive temperature in the presence of galactose to mid-log phase. The cultures were then incubated at the restrictive temperature for one hour to inactivate the decapping enzyme, followed by the addition of glucose to inhibit further transcription of the *GAL7* gene. Figure 4B shows that in a *dcp1-2* single mutant, the endogenous *GAL7* mRNA decays with a half-life of about 17 min. In contrast, in the *dcp1-2 rrp44-CR3* double mutant, the *GAL7* mRNA half-life is greater than 60 min, which is similar to the *dcp1-2 ski7Δ* control strain. Both the *dcp1-2* synthetic lethality and *GAL7* mRNA half-lives suggest that the *rrp44-CR3* mutation interferes with exosome-mediated degradation of normal cellular mRNAs.

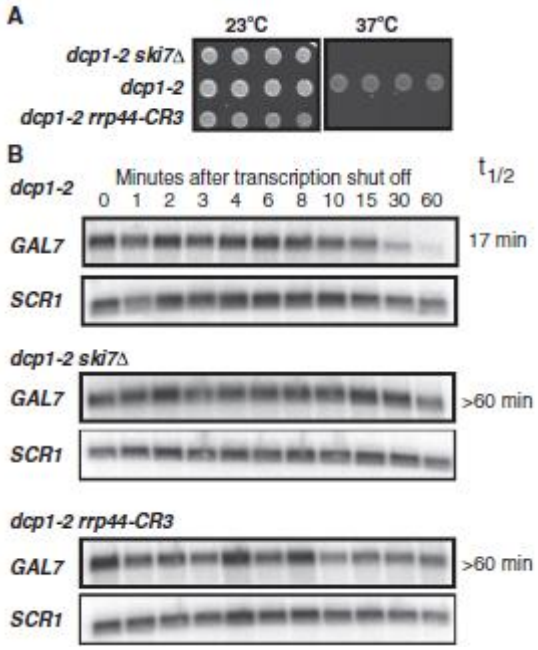


Figure 4. Simultaneously mutating the three conserved Cys residues of Rrp44 to Ser stabilizes cellular mRNAs. (A) The *rrp44-CR3* mutation is synthetically lethal with a temperature-sensitive mutation in the decapping enzyme, *dcp1-2*. (**B**) The *rrp44-CR3* mutation reduces the rate of exosome-mediated degradation of the *GAL7* mRNA.

Overall, these genetic interactions are consistent with the hypothesis that the *rrp44-CR3* mutation reduces, but does not eliminate, the exonuclease activity *in vivo*. Furthermore, the synthetic interactions with *rrp6Δ* and *dcp1-2* suggest that the *rrp44-CR3* allele affects both cytoplasmic and nuclear exosome function.

Mutation of the CR3 motif affects the endoribonuclease activity of Rrp44 *in vitro*

While the genetic analysis above indicates that the CCCH motif is important *in vivo*, it can not address whether these amino acids play a direct role or an indirect role in Rrp44 nuclease activity. In fact, since the CCCH motif is very well separated from the exonuclease domain, we consider it unlikely that these amino acids have a direct role in exonuclease activity. At first glance, the CCCH motif is also far (>25Å) away from the endonuclease active site, however it is important to note that H_{184} of the CCCH motif and D_{171} of the endonuclease active site are part of the same α-helical structure (Fig. 1D and Supplementary Fig. S1A). Because D_{171} is a key residue in the endonuclease active site (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009), we considered the possibility that perturbation to the CCCH site may have an indirect effect on the orientation or flexibility of this α-helix, and thereby affect endonuclease activity. To test this, we purified three truncated Rrp44 proteins as glutathione S-transferase (GST) fusions from *Escherichia coli*. As we and others have previously shown, a truncated Rrp44 protein that contains the PIN domain and CR3 motif (residues 1-235), has endonuclease activity *in vitro* (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). The second recombinant protein purified is the same, except that the three conserved Cys residues were mutated to Ser. This CR3 mutant enzyme was less active in *in vitro* assays containing the same amount of enzyme of similar purity (Fig. 5). Specifically, Figure 5 shows that the wild-type enzyme degraded most of the substrate in the first 5 minutes, with all of the

substrate gone after 15 minutes. In comparison, more substrate was left in the reaction with the same amount of the enzyme with all three Cys residues mutated after 5 or 15 minutes. The third recombinant protein we purified contains only the part of Rrp44 that is similar to other PIN domains (residues 84-235) and lacks the CR3 motif. This PIN domain only protein also exhibited less nuclease activity *in vitro* (Fig. 5). Quantitation of the experiment shown in Figure 5 indicated that the wild-type enzyme was approximately twice as active as either mutant version. Although individual batches of purified protein can have variations in the amount of activity, the differences between the three forms of Rrp44 were reproducible with independent repeat purifications.

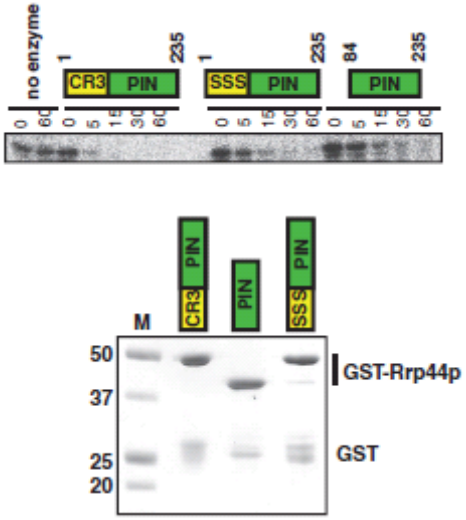


Figure 5. Mutating the CCCH motif affects the *in vitro* endonuclease activity of Rrp44. Recombinant truncated Rrp44 proteins that either contain both an intact CR3 motif and PIN domain (i.e. residues 1–235), a similar protein with the Cys residues in the CR3 motif mutated to Ser, or a protein containing only the PIN domain (i.e. residues 84–235) were used in *in vitro* RNase assays with an A30 substrate as described in materials and methods. Similar results were obtained with a CCCGACACCACCCACUA$_{14}$ substrate. The bottom panel shows a Coomassie stained gel of the purified recombinant proteins.

Mutation of the CR3 motif affects the interaction between Rrp44 and the core exosome

We noted that the CCCH motif is close to three residues that interact with the core exosome. Specifically, in the crystal structure of a trimer of Rrp41, Rrp45 and Rrp44, the conserved Rrp44 Y40, R42 and D44 residues directly interact with Rrp41 (Fig. 1D; Bonneau et al., 2009), and only two residues separate this YxRxD motif from the Cx_4Cx_2C motif (Supplemental Fig. 1A). This led us to hypothesize that mutating the conserved Cys residues to Ser might disorder this part of the protein and reduce the interaction with the core exosome. To test this idea, TAP-tagged wild-type Rrp44 and Rrp44-CR3 proteins were purified using IgG sepharose beads, and exosome interaction was determined using an anti-Rrp4 antibody. Figure 6A shows that the exosome subunit Rrp4 copurifies with the wild-type Rrp44, but no Rrp4 could be detected in the Rrp44-CR3-TAP purification. Thus, even though we suspect that the three conserved Cys residues do not directly contact other exosome subunits, mutating them reduces the ability of Rrp44 to interact with the rest of the exosome, possibly by altering the conformation of Y40, R42 and/or D44.

If the effects of the *rrp44-CR3* mutation are due to an alteration of the interaction of Y40, R42 and/or D44 with Rrp41, then mutating these residues should have a similar effect. We tested this by creating three single-point mutations in these residues (*rrp44-Y40A*, *rrp44-R42A* and *rrp44-D44A*) as well as a triple-point mutant (*rrp44-yrd*) combining these mutations. The single-point mutants did not have an obvious growth defect (data not shown), but the triple *rrp44-yrd* mutant had a slow growth phenotype that closely resembled the growth phenotype of the *rrp44-CR3* allele (Fig. 6B), consistent with the idea that mutating the CCCH motif affects the conformation of Y40, R42 and/or D44, which in turn affects exosome interaction.

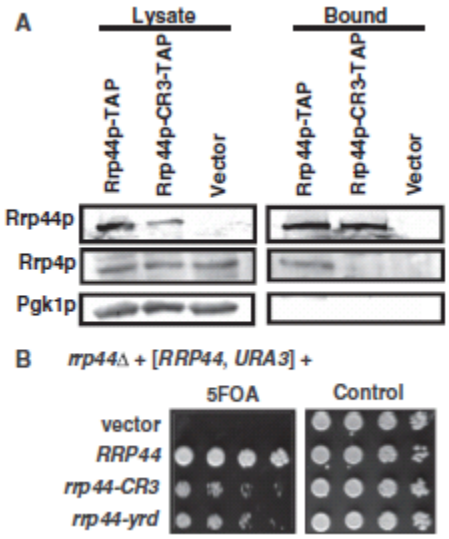


Figure 6. Simultaneously mutating the three conserved Cys residues to Ser reduces interaction with the exosome. A. TAP-tagged Rrp44 containing either a wild-type or mutated CR3 motif were purified from yeast. The lysate (left) and bound (right) fractions were analyzed by Western blot with antibodies for the TAP tag (top), the exosome subunit Rrp4 (middle) or the control protein Pgk1 (bottom). **B.** The *rrp44-yrd* allele that contains Y40A, R42A, and D44A mutations has a slow growth phenotype that resembles that of *rrp44-CR3*.

Discussion

We have previously shown that mutating the CR3 motif leads to a slow growth phenotype, a slightly reduced expression level of Rrp44, and a stabilization of nonstop mRNAs (Schaeffer et al., 2009; Schaeffer & van Hoof, 2011). Although these results hinted at the importance of the conserved motif, it was not clear how changing these Cys residues to Ser affected Rrp44 function. The crystal structure (Bonneau et al., 2009) provided important insight as the published structural model indicated that the cysteines are clustered together. The structure further revealed that, although in the primary sequence His_{184} is far from the CR3 motif, the histidine is in fact positioned in a tetrahedral arrangement with the cysteine residues. While the proximity of Cys residues raises the potential for disulfide bond formation, disulfide bonds are rare in the reduced environment of the

cytoplasm and nucleus. Furthermore, mutating individual Cys residues to Ser should completely disrupt such a disulfide bond, but such mutations did not mimic the *rrp44-CR3* phenotype. This strongly argues against the formation of a disulfide bond in the CR3 motif. Instead the structural clustering of CR3 with His_{184} suggested a metal ion-binding site, since four Cys and/or His residues form Zn^{2+} ion-binding sites in a wide variety of proteins. In such metal-binding sites, mutations of individual residues to serine can have a variety of effects. For example, in the Rpc10 subunit of RNA polymerase III, mutating the first Zn^{2+}-binding Cys residue to Ser is lethal, but changing either the third or fourth Cys (or both) to Ser has no obvious effect on growth (Rubbi et al., 1999). Consistent with the idea that a single Cys to Ser change does not always completely eliminate Zn^{2+} binding, is the natural presence of Ser in some Zn^{2+}-binding sites (Karlin & Zhu, 1997). The suggestion that the CCCH motif forms a metal-binding site is further reinforced by the observation that His_{184} is conserved in most Rrp44 orthologs, but is lacking in two orthologs that also lack the CR3 motif. In addition, although H_{184} is part of the PIN domain, it is not conserved in PIN domains that are not part of Rrp44 homologs. Based on all of the available genetic, biochemical, structural and phylogenetic data we suggest that binding a monoatomic metal ion to this CCCH motif stabilizes folding of this part of Rrp44. Interestingly, Lebreton et al. (2008) have previously shown that the endonuclease activity of Rrp44 *in vitro* is stimulated by addition of either Mn^{2+} or Zn^{2+} ions. One possibility is that either metal ion can bind to the active site. However, an alternative explanation is that one or both of these ions bind to the CCCH motif, and the active site may contain a metal ion that copurifies with the recombinant protein. So far we have not been able to clearly distinguish between these and other possibilities using *in vitro* assays. Moreover, although the CCCH motif most closely resembles a Zn^{2+}-binding site, binding of some other metal ion is also possible, further complicating interpretation of such *in vitro* experiments.

We have previously shown that either the endo- or exonuclease activity of Rrp44 is sufficient for rapid nonstop mRNA degradation, but that the *rrp44-CR3* mutation results in stabilization of nonstop mRNAs. This suggests that the *rrp44-CR3* mutation reduces both nuclease activities. However, since *rrp44-CR3* is viable and Rrp44 nuclease activity is required, the *rrp44-CR3* mutation cannot completely eliminate both functions. Here we test this genetically. If *rrp44-CR3* completely eliminates the exonuclease function, combining it with another mutation that eliminates the exonuclease activity should not have an additional effect. Instead we show that *rrp44-CR3* is synthetically lethal with *rrp44-exo*$^-$. Thus, although *rrp44-CR3* and *rrp44-exo*$^-$ have some of the same molecular defects, we conclude that *rrp44-CR3* does not completely eliminate exonuclease function *in vivo*. Furthermore, since eliminating both endo- and exonuclease activities is lethal, the synthetic lethality of *rrp44-CR3* with *rrp44-exo*$^-$ is consistent with a reduced endonuclease activity. Further synthetic lethality of *rrp44-CR3* with *dcp1-2* and with *rrp6Δ* suggests that both the cytoplasmic and nuclear exosomes are affected, since *dcp1-2* synthetic lethality reflects a role in cytoplasmic mRNA decay, and Rrp6 is exclusively nuclear. Therefore we conclude that the CR3 motif affects many facets of Rrp44 function.

Since the CCCH site is far removed from both catalytic active sites of Rrp44, it is likely not directly involved in either catalytic activity. In many Zn^{2+}-containing proteins, the metal ion stabilizes the overall fold of the protein. The available evidence suggests that the overall structure of Rrp44 can fold in the absence of a metal ion bound to the CCCH motif (Lebreton et al., 2008; Bonneau et al., 2009; Schaeffer et al., 2009; Schneider et al., 2009), but we suggest that binding of a metal ion to the CCCH site has a modest effect on folding and/or flexibility of this part of the protein. This then indirectly affects endo- and exonuclease activities, possibly because the CCCH motif is well-connected to two other important parts of Rrp44 (Fig. 1D). First, His_{184} is part of an α-helix that also

contains the endonuclease active site residue D_{171}. Thus, perturbation of the CCCH motif may slightly alter the position or flexibility of this α-helix, and thereby affect endonuclease activity. Second, we show that the *rrp44-CR3* mutation weakens the interaction between Rrp44 and the remainder of the exosome. Mutations in or depletion of the other nine exosome subunits all cause defects in RNA processing that are similar to those of *rrp44-exo*$^-$. In contrast, the *rrp44-endo*$^-$ allele does not affect growth or any RNA decay or processing reaction, and therefore an endonuclease defect cannot explain the phenotypes of *rrp44-CR3*. Thus, we suggest that the reduced *in vivo* exonuclease activity is a result of the reduced exosome interaction. Although the YxRxD motif is conserved and in the crystal structure interacts with Rrp41, the importance of these residues had not previously been tested. Our data indicate that this motif indeed is a functionally important contact between the core exosome and Rrp44. Strikingly, of the single Cys mutations, C47S would be expected to have the largest effect on the YxRxD motif which may explain why it had the largest effect on exosome function. Interestingly, in many cases Rrp44 orthologs fail to copurify with the exosome (Chen et al., 2001; Estevez et al., 2001; Chekanova et al., 2007; Staals et al., 2010) even though exosome-interacting residues, such as the YxRxD motif, are conserved (Supplementary Fig. S1A and data not shown). Since the YxRxD motif is conserved in Rrp44 orthologs and hDis3L1 orthologs, our results suggest that all of these proteins interact with the core exosome, even though such interaction has proven difficult to demonstrate. Our results further suggest that in some cases the failure to copurify may be because copurification was tested in buffers that included the Zn^{2+} chelator EDTA. Strikingly, two papers reported that hDIS3 does not copurify with the exosome using EDTA containing buffers, while a third paper reports hDIS3-exosome copurification using EDTA-free buffers (Chen et al., 2001; Staals et al., 2010; Tomecki et al., 2010).

Overall we can now offer a coherent model for the importance of the CR3 motif. The primary effect of mutating the CR3 motif is probably modest local misfolding and/or structural instability of Rrp44, possibly due to loss of a metal ion. Because the CR3 motif is structurally connected to both the endonuclease active state and to exosome-interacting residues, this effect on local structure has surprisingly far-reaching effects on exosome function. Furthermore, it opens the possibility that the activity of Rrp44 may be regulated by similar structural alterations.

Materials and Methods

Strains and plasmids

The majority of the experiments used a previously described *rrp44* deletion strain with a *RRP44, URA3* plasmid (yAV1115; Schaeffer et al., 2009). Various alleles of *RRP44* were introduced using the plasmid shuffle as previously described (Schaeffer et al., 2009). All of these plasmids are low copy plasmids, and Rrp44 is expressed under control of its native promoter and 3' flanking sequence. The plasmids used are listed in Supplementary Table S1. The *dcp1-2 rrp44*Δ strain (yAV1143) was previously described (Schaeffer & van Hoof, 2011). To test for synthetic lethality between *rrp6Δ* and *RRP44* alleles, the heterozygous diploid *(RRP44/rrp44Δ::KAN; Giaever et al., 2002)* obtained from Open Biosystems was transformed with the wild-type *RRP44* plasmid, pAV361 (Schaeffer et al., 2009). The diploid was then sporulated to generate a haploid *rrp44Δ::KAN* strain with pAV361. This strain was converted to *rrp44Δ::HYG* by homologous recombination with pAG32 (Goldstein & McCusker, 1999) and selecting for hygromycin resistance. The *rrp44Δ::HYG* strain with pAV361 was then crossed with the *rrp6Δ::KAN* strain from the yeast knock out collection (Giaever et al., 2002). The

resulting diploid was sporulated resulting in yAV1137 (*MATa, leu2-Δ0, ura3-Δ0, his3-Δ1, rrp44Δ::HYG, rrp6Δ::NEO* with a *RRP44, URA3* plasmid).

Growth assays

Synthetic lethality growth assays were performed essentially as described (van Hoof et al., 2000). Briefly, plasmids encoding the *rrp44* point mutations were transformed into the *rrp44Δ* strain, the *dcp1-2 rrp44Δ* strain or the *rrp6Δ rrp44Δ* strain containing a *URA3* plasmid encoding wild-type Rrp44. Transformants were grown in SC–URA–LEU overnight at 23°C. Cultures were diluted in the same selection media to an optical density of 0.8 at 600 nm. Next, cells were serially diluted in 96-well plates by a factor of five and spotted onto media containing 5-Fluoroorotic Acid (5-FOA). 5-FOA selects for cells that have lost the *URA3* plasmid, and therefore growth indicates that the gene on the *LEU2* plasmid is sufficient for viability. Plates were incubated at 23°C, 30°C, and 37°C for 2-5 days. The same effects were seen at all temperatures unless otherwise noted in the results section.

The *his3*-nonstop growth assay was performed essentially as described (van Hoof et al., 2002). Briefly, *rrp44* deletion strains containing a *LEU2* plasmid encoding the *rrp44* point mutations were transformed with a *URA3* plasmid encoding a *his3-nonstop* reporter. This reporter contains a 1-bp deletion which removes the stop codon from the *HIS3* mRNA. Transformants were grown in SC–URA–LEU overnight at 30°C. Cultures were diluted in the same selection media to an optical density of 0.8 at 600 nm. Next, cells were serially diluted in 96-well plates by a factor of five and spotted onto media lacking histidine, SC–HIS–URA–LEU. The plates were incubated at 30°C for 2-6 days.

Stability of *GAL7* mRNA and Northern blotting

To determine the half-life of the *GAL7* mRNA, duplicate cultures of *dcp1-2* yeast strains were grown in YEP + 2% galactose overnight at 23°C. Cultures were diluted in the same media to an optical density of 0.2 at 600 nm and grown to an optical density of 0.8. Cultures were then shifted to 37°C (a non-permissive temperature for *dcp1-2*) for 1 hour. The medium was replaced with 1/10th volume YEP + 4% glucose. At the time points indicated, 2 mL aliquots were pelleted. RNA was isolated and subjected to formaldehyde agarose gel electrophoresis, followed by Northern blot analysis essentially as described (Caponigro et al., 1993; Meaux & van Hoof, 2006). Northern blots were hybridized with ^{32}P 5'-end labeled oligonucleotides specific for *GAL7* (oAV267; 5' CTCTTGCTTCTCTGGAGAGATCGTCAGTC) and the RNA subunit of the signal recognition particle, *SCR1* (oRP100; 5' GTCTAGCCGCGAGGAAGG). Signals were detected and quantitated using a STORM PhosphoImager (Amersham), and corrected for loading by quantitating *SCR1*.

RNase assays on recombinant Rrp44

Truncated Rrp44 fused to GST was expressed in *E. coli*, purified, and assayed as previously described (Schaeffer et al., 2009). Briefly, the Rosetta(DE3) *E. coli* strain with Rrp44 expression plasmids was grown at 30°C to an optical density of 1.2 at 600 nm. Expression was induced with 0.2 mM IPTG at 18°C overnight. Rrp44 was purified over a GSTtrap FF 1-mL column and a HiTrap™Q FF column (GE Healthcare). RNase assays used a 5' end–labeled 30-mer RNA (either A_{30} or $CCCGACACCACCCACUA_{14}$), and 20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM $MnCl_2$, 1 mM DTT.

Rrp44 exosome copurification

The Rrp44-TAP and Rrp44-CR3-TAP plasmids were previously described (Schaeffer & van Hoof, 2011). Cultures were grown in 50 mL media to an optical density of 0.8 at 600 nm, pelleted, and resuspended in 250 μL IP50 buffer [50mM NaCl, 50mM Tris-HCl pH 7.5, 2 mM $MgCl_2$, 0.1% Triton X100, 0.5 mM β-mercaptoethanol, 0.1 mM PMSF, with complete EDTA-free protease inhibitors (Roche)]. 50 μL glass beads (Sigma) were added to the cell suspension and agitated in a vortex mixer for seven minutes at 4°C. Lysates were clarified by centrifugation at 4500g for seven minutes at 4°C.

TAP-tagged proteins were purified by adding 10 μL IgG Sepharose 6 Fast Flow beads (Amersham Biosciences) to 200 μL whole-cell lysate. After incubation for one hour at 4°C, the supernatant was discarded and the beads were washed two times with 200 μL IP50 buffer, and then twice more with 200 μL IP1000 buffer (identical to IP50, but 1M NaCl). Bound proteins were eluted from beads by heating to 95°C for 3 minutes in protein sample buffer and analyzed by western blotting using anti-protein A antibodies (Sigma), anti-Rrp4 antibodies (Mitchell et al., 2003) and anti-Pgk1 antibodies (Invitrogen).

Acknowledgments

Antibodies against Rrp4 were kindly provided by David Tollervey. We thank members of the van Hoof and Arraiano labs and Lisa Berreau for fruitful discussions. This work was funded by grants from National Institutes of Health (GM069900) and the National Science Foundation (MCB1020739) to A.v.H., an EMBO short term fellowship to D.S. and by Fundação para a Ciência e a Tecnologia (FCT) to C.M.A. and F.P.R.

Supplementary Information

Figure S1. Partial sequence alignment of Rrp44 homologs and other yeast PIN domains. **A.** The full length sequences of twenty eukaryotic Rrp44 homologs were aligned using ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) and boxshade (http://www.ch.embnet.org/software/BOX_form.html) with default settings. This figure shows only a small part of the alignment. Red: residues identical in at least half of the sequences. Blue: residues similar in at least half of the sequences. Green: residues forming the putative Zn-binding site. Orange: Two of the endonuclease active site residues. The line below the alignment shows the secondary structure (a=α-helix b=β-sheet) (Bonneau et al., 2009). B. The conserved His184 residue identified in the PIN domain of Rrp44 in part A is not conserved in other PIN domains. Shown is an alignment of part of the PIN domains of 7 other PIN domains in yeast. Note that two of the catalytic site Asp residues are aligned with panel A, but there is no conserved residue corresponding to H184. Residues in panel B are colored to match panel A.

Table S1. Plasmids used.

Name	Short description	Description	Marker
pAV344	*RRP44*	*RRP44* promoter, residues 1-1001, *RRP44* 3' flanking region	*LEU2*
pAV361	*RRP44*	*RRP44* promoter, residues 1-1001, *RRP44* 3' flanking region	*URA3*
pAV515	*rrp44-CR3*	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S), *RRP44* 3' flanking region	*LEU2*
pAV503	*rrp44-endo*$^-$	*RRP44* promoter, residues 1-1001 (D171A), *RRP44* 3' flanking region	*LEU2*
pAV501	*rrp44-exo*$^-$	*RRP44* promoter, residues 1-1001 (D551N), *RRP44* 3' flanking region	*LEU2*
pAV603	*rrp44-C47S*	*RRP44* promoter, residues 1-1001 (C47S), *RRP44* 3' flanking region	*LEU2*
pAV604	*rrp44-C52S*	*RRP44* promoter, residues 1-1001 (C52S), *RRP44* 3' flanking region	*LEU2*
pAV605	*rrp44-C55S*	*RRP44* promoter, residues 1-1001 (C55S), *RRP44* 3' flanking region	*LEU2*
pAV741	*rrp44-H184A*	*RRP44* promoter, residues 1-1001 (H184A), *RRP44* 3' flanking region	*LEU2*
pAV536	*rrp44-CR3-exo*$^-$	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S, D551N), *RRP44* 3' flanking region	*LEU2*
pAV633	*rrp44-CR3-endo*$^-$	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S, D171A), *RRP44* 3' flanking region	*LEU2*
	rrp44-C47S	*RRP44* promoter, residues 1-1001 (Y40A), *RRP44* 3' flanking region	*LEU2*
	rrp44-C47S	*RRP44* promoter, residues 1-1001 (R42A), *RRP44* 3' flanking region	*LEU2*
	rrp44-C47S	*RRP44* promoter, residues 1-1001 (D44A), *RRP44* 3' flanking region	*LEU2*
	rrp44-yrd	*RRP44* promoter, residues 1-1001 (Y40A, R42A, D44A), *RRP44* 3' flanking region	*LEU2*
pAV498	*RRP44-TAP*	*RRP44* promoter, residues 1-1001 TAP, *RRP44* 3' flanking region	*URA3*
pAV622	*rrp44-CR3-TAP*	*RRP44* promoter, residues 1-1001 (C47S, C52S, C55S) TAP, *RRP44* 3' flanking region	*URA3*
pGEX-4T-1		GST	-
pGEX-454		GST-RRP44 residues 1-235	-
pAV575		GST-RRP44 residues 1-235 (C47S, C52S, C55S)	-
pAV576		GST-RRP44 residues 84-235	-
Reporter Plasmids			
pAV188		*his3*-nonstop	*URA3*

References

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

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

Cairrão F, Arraiano C, Newbury S. 2005. Drosophila gene tazman, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev Dyn 232*:733-737.

Caponigro G, Muhlrad D, Parker R. 1993. A small segment of the MAT alpha 1 transcript promotes mRNA decay in Saccharomyces cerevisiae: a stimulatory role for rare codons. *Mol Cell Biol 13*:5141-5148.

Chekanova JA, Gregory BD, Reverdatto SV, Chen H, Kumar R, Hooker T, Yazaki J, Li P, Skiba N, Peng Q, Alonso J, Brukhin V, Grossniklaus U, Ecker JR, Belostotsky DA. 2007. Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell 131*:1340-1353.

Chen CY, Gherzi R, Ong SE, Chan EL, Raijmakers R, Pruijn GJ, Stoecklin G, Moroni C, Mann M, Karin M. 2001. AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell 107*:451-464.

Chernyakov I, Whipple JM, Kotelawala L, Grayhack EJ, Phizicky EM. 2008. Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22 and the 5'-3' exonucleases Rat1 and Xrn1. *Genes Dev 22*:1369-1380.

de la Cruz J, Kressler D, Tollervey D, Linder P. 1998. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J 17*:1128-1140.

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

Estevez AM, Kempf T, Clayton C. 2001. The exosome of Trypanosoma brucei. *EMBO J 20*:3831-3839.

Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature 443*:110-114.

Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian KD, Flaherty P, Foury F, Garfinkel DJ, Gerstein

M, Gotte D, Guldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kotter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang CY, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M. 2002. Functional profiling of the Saccharomyces cerevisiae genome. *Nature 418*:387-391.

Goldstein AL, McCusker JH. 1999. Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast 15*:1541-1553.

Jacobs Anderson JS, Parker R. 1998. The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J 17*:1497-1506.

Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. 2004. Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae. *Genes Dev 18*:1227-1240.

Karlin S, Zhu ZY. 1997. Classification of mononuclear zinc metal sites in protein structures. *Proc Natl Acad Sci U S A 94*:14231-14236.

Lebreton A, Tomecki R, Dziembowski A, Seraphin B. 2008. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature 456*:993-996.

Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell 127*:1223-1237.

Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell 29*:717-728.

Lorentzen E, Conti E. 2005. Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol Cell 20*:473-481.

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

Meaux S, van Hoof A. 2006. Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA 12*:1323-1337.

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

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

Rubbi L, Labarre-Mariotte S, Chedin S, Thuriaux P. 1999. Functional characterization of ABC10alpha, an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *J Biol Chem 274*:31485-31492.

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat Struct Mol Biol 16*:56-62.

Schaeffer D, van Hoof A. 2011. Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. *Proc Natl Acad Sci U S A 108*:2366-2371.

Schneider C, Leung E, Brown J, Tollervey D. 2009. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res 37*:1127-1140.

Staals RH, Bronkhorst AW, Schilders G, Slomovic S, Schuster G, Heck AJ, Raijmakers R, Pruijn GJ. 2010. Dis3-like 1: a novel exoribonuclease associated with the human exosome. *EMBO J 29*:2358-2367.

Tomecki R, Kristiansen MS, Lykke-Andersen S, Chlebowski A, Larsen KM, Szczesny RJ, Drazkowska K, Pastula A, Andersen JS, Stepien PP, Dziembowski A, Jensen TH. 2010. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J 29*:2342-2357.

Tucker M, Valencia-Sanchez MA, Staples RR, Chen J, Denis CL, Parker R. 2001. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. *Cell 104*:377-386.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science 295*:2262-2264.

van Hoof A, Staples RR, Baker RE, Parker R. 2000. Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol Cell Biol 20*:8230-8243.

Wang HW, Wang J, Ding F, Callahan K, Bratkowski MA, Butler JS, Nogales E, Ke A. 2007. Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. *Proc Natl Acad Sci U S A 104*:16844-16849.

Chapter 4

Studying the role of important residues within yeast Rrp44 in RNA metabolism

This chapter contains data from:
Reis FP, Barbas A, Klauer AA, Schaeffer D, López-Viñas E, Gómez-Puertas P, van Hoof A and Arraiano CM. Modulating the RNA processing and decay by the exosome: altering Rrp44/Dis3 activity and end-product. *Submitted*.

The author of this Dissertation performed and planned most of the experiments described in this Chapter, with the exception of the modeling and alignment results that were performed by our collaborators, Dr. Paulino Gómez-Puertas and Dr. Eduardo López-Viñas from Centro de Biologia Molecular "Severo Ochoa" (CSIC-UAM), Campus Universidad Autonoma de Madrid, Spain. The author of this Dissertation is the first author of this paper and was involved in preparing the original manuscript.

Abstract

RNA maturation, turnover and quality control are essential for cell viability. In eukaryotes, the exosome plays a central role in all these processes. The core exosome is composed of nine catalytically inactive subunits constituting a ring structure and the active nuclease Rrp44, also known as Dis3. Rrp44 is a member of the ribonuclease II (RNase II) superfamily of exoribonucleases. In this work we have performed a functional characterization of two residues located in the RNB catalytic domain, to address their precise role in Rrp44 activity. We have constructed Rrp44 mutants and compared their activity to the wild-type Rrp44. When we replaced the Q892 residue with an Ala, the enzyme became two-fold more active *in vitro*. The *Escherichia coli* residue tyrosine 253 (Y253) is critical in determining the size of the end-product in the RNase II enzyme by changing it from 4 to 10 nucleotides. Our results showed that when the equivalent residue in the yeast *Saccharomyces cerevisiae* Rrp44 (Y595) is changed to an alanine (A) the *in vitro* end-product changes from 4 to 6 nucleotides. This suggests that the role of setting the end-product is conserved by this aromatic residue. We also assessed the function of these residues *in vivo*, and showed that although both of them do not affect nonstop decay, a strain containing the Y595A mutation has a growth defect phenotype and affects RNA processing and degradation. These results lead us to hypothesize that this residue has an important biological role. Molecular dynamics modeling of these Rrp44 mutants and the wild-type enzyme showed changes that extended beyond the mutated residues and help explain these results.

Introduction

The exosome is one of the most important protein complexes involved in the maintenance of the correct levels of RNAs in the cell (Houseley et al., 2006; Lykke-Andersen et al., 2009), and it is a highly conserved complex present in both eukaryotes and *Archaea*. The exosome processes some RNAs to shorter forms and also participates in RNA degradation and surveillance. The yeast exosome is a protein complex that consists of a core of nine proteins that are essential for viability, plus a catalytic subunit, and it operates in the nucleus and in the cytoplasm (Allmang et al., 1999b). Previous reports have shown that Dis3/Rrp44 is the only catalytically active nuclease in the yeast core exosome (Liu et al., 2006; Dziembowski et al., 2007) and plays a direct role in RNA surveillance contributing to the recognition and degradation of specific RNA targets (Schneider et al., 2007). Dis3/Rrp44 is a member of the RNase II family of enzymes, since it is closely related to the 3′ hydrolytic exoribonucleases RNase II and RNase R from *E. coli* (Mian, 1997). This family of proteins is ubiquitous and these enzymes can be found in all three domains of life. Members of this family are often essential for growth (Mitchell et al., 1997) and can be developmentally regulated (Cairrão et al., 2005). Mutations in Dis3/Rrp44-like enzymes have been linked with abnormal chloroplast biogenesis (Bollenbach et al., 2005), mitotic control and cancer (Lim et al., 1997). Several diseases, such as multiple myeloma and Perlman syndrome have been associated with defective Dis3 and Dis3-like proteins (Chapman et al., 2011; Astuti et al., 2012).

The determination of the structure of RNase II provided an explanation for the dynamic mechanism of RNA degradation by this family of enzymes (Frazão et al., 2006). Rrp44 presents a similar modular organization to *E. coli* RNase II, consisting of a central nuclease domain flanked by the oligonucleotide

binding domains, two at the N-terminus (cold-shock domain 1 and 2) and one at the C-terminus (S1 domain; Frazão et al., 2006; Lorentzen et al., 2008). In addition to these domains, yeast Rrp44 contains a conserved motif at the N-terminus that contains three conserved cysteine residues - CR3, and a highly conserved PIN domain. We have recently shown that the three cysteines, together with a conserved histidine from the PIN domain, affect the binding to the exosome (Schaeffer et al., 2012). We, and others have shown that in addition to the already known Rrp44 exonucleolytic activity, Rrp44 also has endonucleolytic activity associated with its PIN domain (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009).

In order to ensure proper cellular RNA metabolism, the eukaryotic exosome needs to be tightly regulated. However, it is still unclear how the activity of the exosome is controlled. To address whether the exosome subunits or the exosome cofactors are responsible for the regulation of this macromolecular complex, we focused initially on the biochemical characterization of yeast Dis3/Rrp44, and generated several mutants in conserved residues.

A previous study of RNase II showed that a mutation in E542 resulted in a "super-enzyme" with a significant increase in ribonucleolytic activity (Barbas et al., 2009). Mutations in equivalent Rrp44 residues showed that when the Q892 residue is replaced with an alanine the enzyme became more active *in vitro*.

In other studies it was also shown that Y253 in RNase II and Y324 in RNase R are critical in setting the size of the end-product (Barbas et al., 2008; Matos et al., 2009). In order to investigate if we would have similar results, we studied the role of the corresponding residues in yeast Rrp44. The crystal structure of yeast Rrp44 suggests that Y595 plays a similar role, but its function has not been experimentally addressed (Lorentzen et al., 2008). Our data show that when Y595 residue is changed to an alanine, the *in vitro* end-product changed from the major end-product with 4 nts (Lorentzen et al., 2008) to a 6 nucleotide

RNA fragment. Surprisingly, this substitution also lead to an increase in the ribonucleolytic activity of the mutant protein. Our findings have also shown for the first time that the Y595 residue has an important biological role, since the mutation in this amino acid affects cell growth, processing of the 5.8S rRNA, and degradation of the 5' end external transcribed spacer of the rRNA precursor (5' ETS). These recent findings are important because we show that changing the size of the end-product impairs the activity of the exosome *in vivo*.

Results

Mutation in Q892 enhances Rrp44 activity

The data analyses of both structural and sequence alignments showed that the G895 residue in yeast Rrp44 and the E542 in *E. coli* RNase II occupy the same position (Fig. 1), thus G895 was one of the residues that we chose to mutate. In addition, the close analysis of the sequence alignment and conservation profile shows that in the region where there are two glutamine residues in the human protein (Q795 and Q798 of human Rrp44), there are two homologous residues in yeast Rrp44, namely Q892 and G895 (Fig. 1). This particular region is not fully conserved among taxonomic groups for this family, therefore we hypothesize that a specific role cannot be attributed to a single residue and we believe that the structure of the whole region acts cooperatively in RNA processing. Thus, mutations G895A, G895E, G895Q, and Q892A in yeast Rrp44 could be viable candidates to observe similar effects than those revealed by the RNase II E542A mutant.

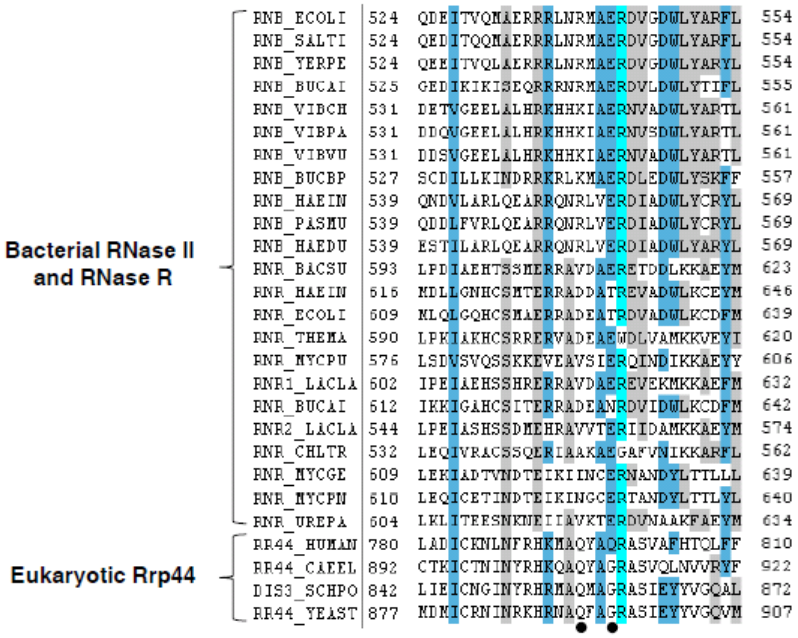


Figure 1. Partial multiple sequence alignment of representative Rrp44, RNase II and RNase R members, in yeast Rrp44 Q892-G895 region. Positions of yeast Q892 and conserved G895 indicated. All the residues positions colored according to sequence conservation.

In order to analyze the role of these residues we constructed these four Rrp44 mutants, purified the proteins containing these mutations, and compared their activities to the purified wild-type Rrp44 protein by performing exoribonuclease assays using a 30-mer single-stranded RNA substrate as described in Methods (Fig. 2; Supplemental Fig. S1A, B). As previously demonstrated, the wild-type Rrp44 enzyme was able to degrade the single-stranded 30-mer RNA substrate to a final major product of 4 nts (Lorentzen et al., 2008). Two of the constructed mutations, G895A and G895Q, showed a slight reduction of the ribonucleolytic activity when compared to the wild-type Rrp44 protein (Supplemental Fig. S1A). The G895E mutant presented activity very similar to that of the wild-type enzyme. However, and most interestingly, the

Rrp44 mutant Q892A presented an increase in catalytic activity (Fig. 2; Supplemental Fig. S1A).

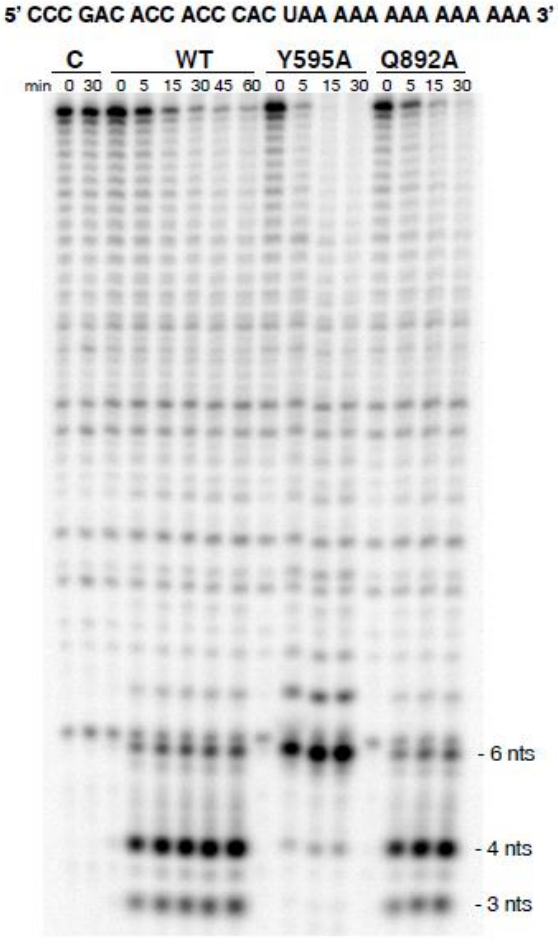


Figure 2. Exoribonuclease activity of Rrp44 and mutant versions of Rrp44 – Activity assays were performed as described in Materials and Methods using 20 nM of enzyme and 4 nM of 30 mer RNA substrate (5′-CCC GAC ACC ACC CAC UAA AAA AAA AAA AAA-3′) and samples were taken at the time-points indicated. Control reactions (C) were incubated with no enzyme added. Length of substrates and degradation products are indicated in the figure.

We determined the exoribonucleolytic activity of the wild-type Rrp44 and Q892A proteins by measuring disappearance of the substrate. Our data showed that the Q892A mutant enzyme has an increase in ribonucleolytic activity of more

than 2 fold when compared to its wild-type counterpart (17.4 pmol.min^{-1}.nmol^{-1} *versus* 6.8 pmol.min^{-1}.nmol^{-1}; Table 1).

Table 1. Exoribonucleolytic activity of wild-type Rrp44 and point mutant enzymes – The exoribonucleolytic activity was assayed using a 30-mer oligoribonucleotide (5′-CCCGACACCAACCACUAAAAAAAAAAAAAA-3′) as substrate. Each value represents the mean of at least three activity assays as described in Methods section and respective standard deviation.

Enzyme	Enzyme activity (pmol.min^{-1}.nmol^{-1})	Relative activity (%)
Wild-type Rrp44	6.78 ± 0.76	100
Q892A mutant	17.35 ± 4.28	256
Y595A mutant	14.35 ± 2.90	212

A mutation in Y595 alters the Rrp44 end-product

Previous reports in *E. coli* showed that residue Y253 in RNase II and Y324 in RNase R are critical in setting the size of the end-product (Barbas et al., 2008; Matos et al., 2009). This residue in RNase II was determined to be responsible for the stabilization of the 3′-end of the RNA molecule in the catalytic site. One of the aims of our work was to determine the specific role of the equivalent residue in yeast Rrp44. This residue is highly conserved in RNase II family of enzymes and by analysis of the sequence alignment of RNase II, RNase R, and Dis3/Rrp44 proteins we and others determined that this specific residue is equivalent to Y595 in yeast Rrp44 (Lorentzen et al., 2008; Matos et al., 2009). We mutated the aromatic residue Y595 in Rrp44 into an alanine, purified the mutant protein, and assessed its enzymatic activity (Fig. 2; Supplemental Fig. S1A). Our results demonstrated that wild-type Rrp44 degraded a generic substrate to a final product with smaller size than the Y595A mutant. As shown in Fig. 2 the major end-product of Rrp44 is 4 nts in length, while the Y595A mutant produced a major final product of 6 nts. Our data also suggest that this mutation affects the protein activity, since the Y595A substitution resulted in an increase of more than

2 fold in catalytic activity, by comparing with wild-type (6.8 $pmol.min^{-1}nmol^{-1}$ and 14.3 $pmol.min^{-1}nmol^{-1}$ for the wild-type and Y595A mutant, respectively; Table 1).

Theoretical structural insights of Dis3/Rrp44 interaction to RNA at the RNB domain

In order to explain the results obtained, we have performed theoretical structural models, based on the x-ray structure of Rrp44 D551N-RNA bound complex (Lorentzen et al., 2008) and the Rrp44 D551N mutant *apo* form (Bonneau et al., 2009). After six nanoseconds of free molecular dynamics, representative Dis3/Rrp44 models for the wild-type, Y595A and Q892A mutants showed no relevant differences in their secondary and tertiary structures. According to RMSD data, only the single stranded RNA (ssRNA)-bound RNB Q892A mutant has significant flexibility in its peptide-backbone (Supplemental Fig. S2). However, a detailed inspection of the different RNB domains within the catalytic cleft provided a set of interesting local variations that modify ssRNA binding patterns in the three models.

Wild-type model

The theoretical structure of Rrp44/Dis3 wild-type suggests that the general pattern of binding interactions between its RNB domain and the five most 3' ssRNA nucleotides (RA1-RA5) could be essentially similar to the one in the RNA-bound D551N crystal structure (Lorentzen et al., 2008). Nevertheless, upon *in silico* reversion of the D551N mutation, simulation reveals that Tyr654 could also contribute to binding of the outgoing nucleotide (RA1) by its phosphate group, as well as holding nucleotide RA2 by its 2'-OH moiety (Fig. 3A). The putative role of Arg847, as a tight binder of the phosphate group in RA1, seems to be reinforced in the model. Surprisingly, a new possible interaction could also participate in substrates' sugar-phosphate anchoring: our model proposes that residue Asn724

would establish a double hydrogen-bonding union to nucleotide RA2 by the 2'-OH group and the oxygen atom in its ribose ring, contributing to stability of the preceding nucleotide. In comparison to the experimentally determined structure of yeast Rrp44/Dis3 D551N mutant, our theoretical model of the wild-type enzyme proposes that a strengthening of the interactions that mediate the binding to Rrp44/Dis3 of nucleotides RA1 and RA2 could be feasible. However it is still important to clarify the precise structure of the complex and the dynamics of the outgoing nucleotide during catalysis in the wild-type.

Q892A mutant model

The most conspicuous structural effect of Q892A mutation is the disruption of the canonical hydrogen-bond formed by the amide group in the glutamine side chain to nucleotide RA5 (Fig. 3B). Inspection of the whole set of interactions at the catalytic cleft of this RNB domain, suggests weaker and more labile substrate binding than in the wild-type model and D551N structure (Lorentzen et al., 2008). For example, the side chain of residue Tyr654 could be too far away (beyond 4.0 Angstroms) from the ssRNA molecule to form strong stable hydrogen-bonds. In addition, the electrostatic interaction of Arg847 with the phosphate-backbone at RA1 could be relaxed in terms of spatial proximity. However, despite that the Q892A mutation could soften some direct interactions, the model also predicts that the hydrogen-bond between Glu700 and nucleotide RA3 would remain preserved. Such contact would help to maintain the continuous self-stacking organization of nucleotides RA1-RA5 deployed by Tyr595. Although introduction of the short non-polar side chain of alanine in this position could be responsible for the easing of referred interactions, the model of the Q892A protein is consistent with its near wild-type ability to bind ssRNA conformation in a catalytically active form. The increased lability may explain the enhanced activity of this mutant.

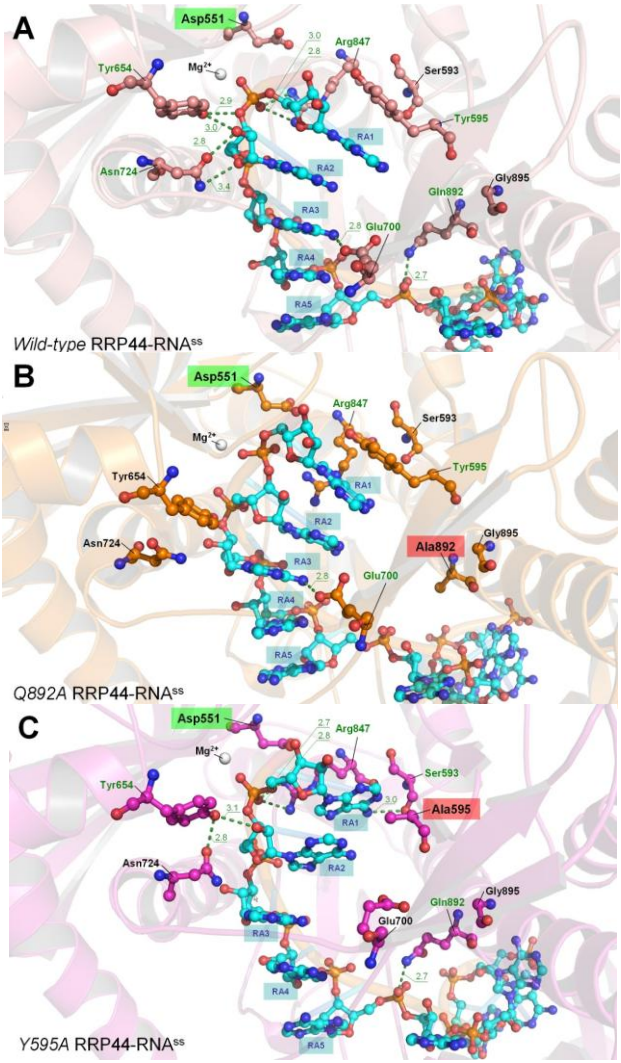


Figure 3. **Theoretical catalytic site of Rrp44/Dis3 RNB domain in *S. cerevisiae*** - Wild-type **(A)** and mutants Q892A **(B)** and Y595A **(C)**, in complex with one single strand poly(A) RNA molecule. Relative locations of some functionally relevant residues described in the text, depicted over every modeled protein's 3D structure. Also layed out inside color text boxes, *in silico* reverted mutation Asp551 (green), Q892A and Y595A mutations (red), and nucleotide bases (blue) starting from RA1 (outgoing). Hypothetical hydrogen-bond mediated interactions, residue tags and annotations of corresponding point distances (in Angstroms, measured over every final representative model) depicted in green colour.

Other important residues, hypothetically invariant among modeled phenotypes, not shown for clarity.

Y595A mutant model

Analyses of Y595A theoretical model in the catalytic cleft region support the hypothesis that this mutation distorts self-stacking interactions among the five most 3' ssRNA nucleotides (RA1-RA5; Fig. 3C). Residues RA3-RA5 are stacked similarly to the wild-type enzyme. Stacking between RA1 and RA2 is also maintained, but stacking between RA2 and RA3 is disrupted. In addition, other hypothetical modifications at the catalytic cleft suggest that a general easing of binding interactions of nucleotides RA2-RA5 might take place. For example, in contrast to the wild-type model and the D551N mutant structure (Lorentzen et al., 2008), the side chain of Glu700 in Y595A mutant would be exposed to solvent, unable to stabilize the adenine base in nucleotide RA3 directly. At the same time, Asn724 does not seem to exert any relevant influence to ssRNA, but it would determine the position of Tyr654 side chain to bind 2'-OH moiety in nucleotide RA2. So far, the model proposes that (a) alterations in continuity of π-stacked interactions and (b) the easing of strong binding interactions to nucleotides RA2 and RA3 could result in an increase in the lability of nucleotides RA2-RA5, induce unstable bound conformations in ssRNA substrates smaller than 6 nucleotides, and consequently hinder catalysis with such fragments. However, other relevant changes in the local environment of the outgoing nucleotide (RA1) indicate that its tight binding would be preserved during the catalytic events. For example, rather than being exposed to solvent, the 6-amine moiety in adenine base of RA1 could be interacting to peptide carbonylic oxygen of Ser593, and a stronger electrostatic interaction (in terms of distance to phosphate group) of Arg847 to phosphate in RA1 could be also possible.

Assessing the functions *in vivo* of the Q892 and Y595 residues

In order to test whether the previous results affect the function of the exosome *in vivo*, we performed growth assays with both *rrp44-Q892A* and *rrp44-Y595A* mutants. An *rrp44Δ* strain complemented with *RRP44*-wild-type on a plasmid with a *URA3* selectable marker was transformed with *LEU2* plasmids encoding each of the two point mutations, Q892A and Y595A. The empty vector and plasmid with wild-type *RRP44* gene were used as negative and positive controls, respectively. The transformants were spotted onto plates containing 5-fluoro-orotic acid (5-FOA), which kills cells that contain the *URA3* plasmid. Growth on 5-FOA indicates that the *URA3-RRP44* plasmid was lost and that the mutated Rrp44 encoded in the *LEU2* plasmid can carry out the essential function of the exosome.

As compared to the wild-type *RRP44*, the mutant *rrp44-Q892A* strain does not have an obvious growth defect, suggesting that this mutation does not affect the essential function of the exosome *in vivo* (Fig. 4).

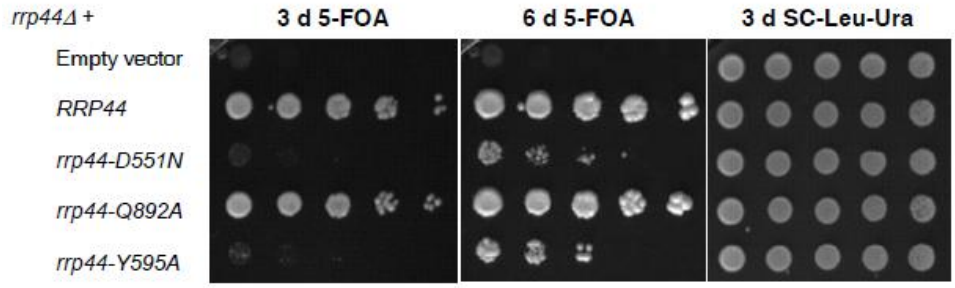


Figure 4. The Y595 residue is important for exosome function - A *rrp44Δ* strain complemented by full-length *RRP44* on a plasmid with a *URA3* marker was transformed with *LEU2* plasmids encoding each of the mutations. Growth on 5-fluoroorotic acid (5-FOA) indicates that *rrp44-Q892A* can carry out the essential function of the exosome. Mutation to alanine of the tyrosine 595 residue reduces growth.

The strain containing the *rrp44-Y595A* version is viable, but grows significantly slower than the wild-type strain. This slow growth is not as strong as

the phenotype caused by the D551N mutation that impairs the catalytic activity of the RNB domain (Dziembowski et al., 2007). Previously, it was shown that r*rp44-D551N* mutant has a growth defect but is still viable (Dziembowski et al., 2007; Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). The similar growth defect of the strains carrying *rrp44-Y595A* with the already described *rrp44-D551N* mutant suggests an important biological role for the Y595 residue. Furthermore, this growth defect may be related to the alteration of the end-product induced by this mutation.

Q892A and Y595A mutations do not affect nonstop mRNA decay

The inactivation of the cytoplasmic exosome results in the stabilization of transcripts that lack stop codons (Frischmeyer et al., 2002; van Hoof et al., 2002). To investigate whether these mutations are required for exosome-mediated mRNA decay, we analyzed the expression of a *his3-nonstop* reporter mRNA. The *his3-nonstop* growth assay exploits a mutant *HIS3* gene that lacks in-frame stop codons (van Hoof et al., 2002). The resulting aberrant reporter mRNA is rapidly degraded by the cytoplasmic exosome. This results in low levels of His3 protein, which results in a failure to grow in the absence of histidine. A defect in nonstop mRNA decay results in growth on media lacking histidine due to stabilization of the *his3-nonstop* mRNA. As previously described, deletion of the gene encoding the cytoplasmic exosome cofactor Ski7 allowed for growth on media lacking histidine; an important control showing that the *his3-nonstop* mRNA transcript was stabilized in this strain (van Hoof et al., 2002; Fig. 5). In contrast, the wild-type strain and both mutants of *RRP44* (*rrp44-Q892A* and *rrp44-Y595A*) failed to grow on the plate without histidine. Thus, this result indicates that the *his3-nonstop* mRNA was unstable and that none of the mutations in *rrp44* affect *nonstop* mRNA degradation.

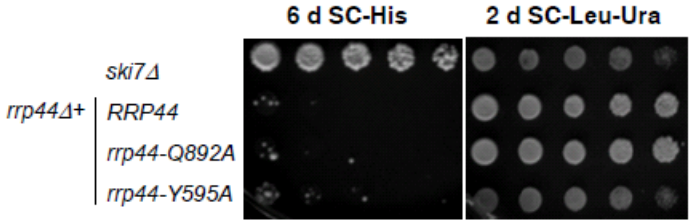


Figure 5. *rrp44-Y595* and *rrp44-Q892* do not have a defect in *nonstop* mRNA decay - A *ski7Δ* strain [Vector, *LEU2*], *rrp44Δ* [*RRP44, LEU2*], *rrp44Δ* [*rrp44-Q892A, LEU2*] and *rrp44Δ* [*rrp44-Q892A, LEU2*] were transformed with a mutant *his3* gene on a plasmid with a *URA3* marker. The transformants were selected on SC-Leu-Ura. Growth on SC-His indicates that the *nonstop* mRNA is stabilized while failure to grow indicates normal nonstop mRNA decay. SC-Leu-Ura was used as control.

Y595 is important for RNA-processing and degradation

The exosome carries out both RNA-processing and RNA degradation reactions. In order to determine if Rrp44 residues Q892 and Y595 are required for any of these reactions in the nuclear exosome, we isolated RNA from each mutant strain and analyzed it by Northern blotting for aberrant processing of 5.8S rRNA and 7S pre-RNA, and degradation of the 5′-external transcribed spacer (5′-ETS). These RNA species are products of the 35S polycistronic precursor rRNA.

The 5.8S rRNA is generated from a 7S precursor by the exosome and Mtr4 (Briggs et al., 1998; de la Cruz et al., 1998; Allmang et al., 1999a). It was already reported that 7S-processing intermediates accumulated as a result of the D551N mutation in the RNB domain but not from the D171A mutation in the PIN domain (Schaeffer et al., 2009; Schneider et al., 2009; Fig. 6A). In contrast, the *rrp6*Δ strain does not lead to the accumulation of these species, but instead causes the accumulation of a distinct RNA species that is 30 nucleotides longer than the normal 5.8S rRNA (Briggs et al., 1998). Similar to the D551N mutant, the *rrp44-Y595A* strain causes the accumulation of heterogeneous species that range in length between 7S pre-RNA and 5.8S rRNA. This indicates that residue Y595 is contributes to the processing of the 5.8S rRNA (Fig. 6A).

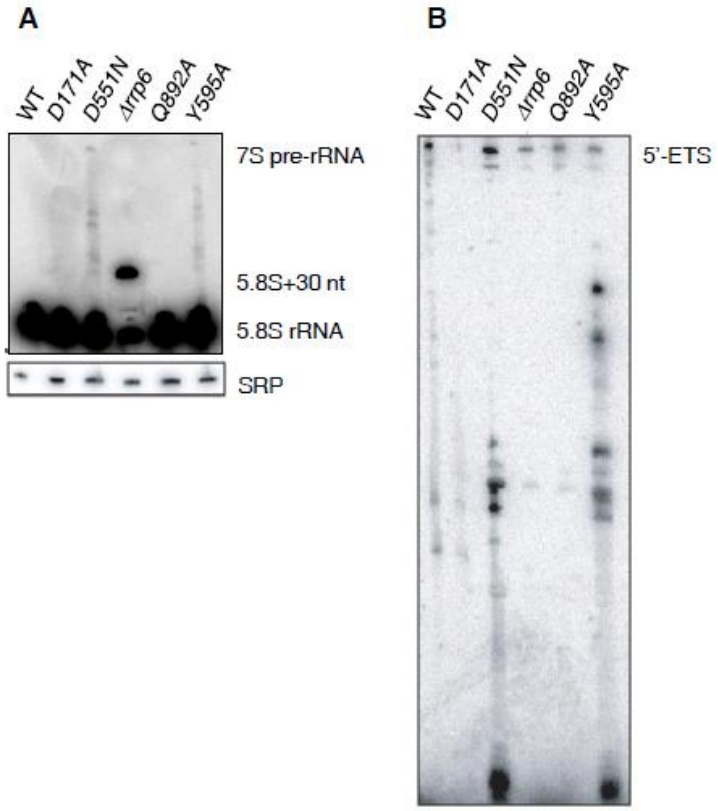


Figure 6. Y595 residue is required for both RNA-processing and degradation - The indicated Rrp44 mutations were expressed in a *rrp44*Δ yeast strain. RNA was isolated from cultures grown in rich medium and analyzed by northern blotting with probes that hybridize to the 5′ ETS region, 5.8S rRNA and the RNA subunit of the signal recognition particle (SRP).

The 5′-ETS is a by-product of rRNA processing which is degraded by the exosome after cleavage from 35S pre-RNA (Allmang et al., 1999a). Northern blotting using a probe for the 5′-ETS revealed the accumulation of intermediates smaller than 5′ETS in *rrp44-Y595A* and *rrp44-D551N*, but not in *rrp44-D171A* and wild-type *RRP44* (Fig. 6B). This result is in agreement with previous findings obtained for the *rrp44-D551N* and *rrp44-D171A* strains (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). However, the size of the intermediates accumulated in *rrp44-Y595A* strain is different from the ones that

accumulate in *rrp44-D551N* strain. In addition to the involvement in the maturation of 7S pre-rRNA, the Y595 residue appears to play an important role in the degradation of the excised 5'-ETS pre-rRNA region. However, its function may be different from the Rrp44 D551N residue.

The *rrp44-Q892A* mutant strain does not show any specific RNA degradation or processing defects on 7S pre-rRNA, 5.8S rRNA and 5' ETS, since it always showed the same pattern of degradation as wild-type *RRP44* and *rrp44-D171A* (Fig. 6A, B). This result is not surprising, since this mutation did not cause a growth defect phenotype.

Discussion

Mutations in Rrp44 change exosome activity and the size of the end-product

We performed multiple sequence alignments of RNase II homologues in order to find which Rrp44 residues corresponded to *E. coli* RNase II residues that were shown to be important for activity and size of the end-product. The data analysis of both structural and sequence alignments showed that G895 in yeast Rrp44 is the residue that occupies the exact same position in sequence as E542 in *E. coli* RNase II. Previous studies reported that the carboxylic group of E542 is in close proximity to the nitrogen atoms of the leaving nucleotide, and the establishment of one or more hydrogen-bonds between them could facilitate the elimination of this nucleotide after cleavage of the RNA substrate by RNase II (Barbas et al., 2009). However, our results demonstrate that mutations in the G895 residue do not affect the protein catalytic activity, indicating that this amino acid is functionally different from the E542 residue of RNase II. Moreover, this amino acid is not fully conserved among taxonomic groups of this family of proteins.

This is the case of the eukaryotic homologues, which do not have a glutamic acid in this position. Instead, all of the homologues contain a glutamine residue in the vicinity (Q892 in the Rrp44 from yeast) that only differs from the glutamic acid in the absence of the negative charge. Substitution of Q892 by alanine in yeast Rrp44 (Q892A mutation) led to an increase in ribonucleolytic activity, similar to what happens with the *E. coli* RNase II E542A mutant. However, the increase in activity of the Rrp44 Q892A mutant was much more moderate.

Here we report that when Y595 residue is substituted by an alanine, the *in vitro* end-product changed from 4 nts to 6 nts. This data suggest that the role of this aromatic residue in setting the protein end-product is highly conserved. Moreover, this replacement also resulted in a significant increase in the protein catalytic activity, suggesting that the release of the product is rate-limiting, and residue Y595 seems to slow down the release of the cleaved mononucleotide. Consequently, mutant Y595A increases the rate of nucleotide release and RNA turnover.

Theoretical structural responses and ribonucleolytic activity

It is intriguing that two distant mutations, as Q892A and Y595A, increase the catalytic activity of yeast Dis3/Rrp44 to a similar extent. Both substitutions seem to modify the relative organization of nucleotides RA2-RA5, making it more labile than what is expected in the wild-type model presented here. Although further structural modifications induced by Q892A and Y595A mutations are possible, a faster processing capacity could explain the experimentally determined two-fold higher activity in both cases. Nevertheless, this work cannot dismiss the possibility that other changes in the catalytic complex could in turn modify the energetic barriers, facilitating the reaction and processing. From an evolutionary perspective, it is remarkable that the Dis3/Rrp44 RNB domain has a wide capacity to accommodate mutations in the catalytic cleft. Interestingly, by

performing undetermined global adaptive conformational changes, as it happens in the Q892A model (Supplemental Fig. S2), the whole protein is able to remain functionally competent despite solvent displacement effects or changes in direct binding interactions. In relation to the latter hypothesis, experimental results assessed in mutations G895A, G895Q, and G895E are also extraordinarily meaningful, as they show that this position is minimally related or irrelevant to the structure of protein-RNA-solvent complex. Nevertheless, more extensive studies involving the other domains of Dis3/Rrp44 are still required to generate deeper insights in the fine structure and dynamics of the RNB-RNA complex during catalysis. In the Y595A model, the lability of nucleotides RA2-RA5 also seem to induce unsteady bound conformations in single-stranded RNA substrates smaller than 6 nucleotides, hindering catalysis.

Y595 residue has an important biological role

We showed for the first time that Y595 residue seems to be a critical residue in setting the smallest product generated by Rrp44. Moreover, without reducing the catalytic activity of the enzyme, an alteration in this residue was shown to have dramatic implications for the cell, since the strain carrying the Y595A mutation has an effect in growth phenotype. This result lead us to hypothesize that the alteration of the size of the end-product leads to the accumulation of degradation fragments in the cell. In *E. coli*, oligoribonuclease (*orn* gene) is required for the complete degradation of mRNA to mononucleotides and this process is required for cell viability (Niyogi & Datta, 1975; Yu & Deutscher, 1995). Oligoribonuclease has been described as a single-stranded specific RNase that has strong affinity for small RNA fragments (2-5 nts), with a 5 mer oligoribonucleotide being its preferred substrate (Niyogi & Datta, 1975; Yu & Deutscher, 1995). The 6 nucleotide end-product left by the Rrp44-Y595A mutant may not be a suitable substrate for degradation by the yeast oligoribonuclease, and subsequently, the

accumulation of small RNA fragments might have toxic effects for the cell and impair cellular growth. Close homologues of the *orn* gene were found in a wide range of eukaryotes, raising the possibility that oligoribonuclease also participates in mRNA degradation in higher organisms (Ghosh & Deutscher, 1999). In yeast, an oligoribonuclease homologue was identified (*YNT20/REX2* gene) that has 27.5% sequence identity and 32% sequence similarity with the *E. coli* homologue (Hanekamp & Thorsness, 1999). Although the yeast homologue Ynt20/Rex2 was initially implicated in mitochondrial functions, it was later reported to function in the nucleus, namely in the processing of stable and small RNAs (van Hoof et al., 2000), and a Ynt20/Rex2-GFP fusion protein is localized to both the nucleus and mitochondria (Huh et al., 2003). Alternatively, the accumulation of these fragments could inhibit essential cellular functions.

This study sheds light on the model previously proposed for Rrp44 activity and mode of action, and allows for an important step forward in the understanding of the mechanism of RNA processing and degradation by the exosome.

Materials and Methods

Materials

Restriction enzymes, T4 DNA ligase and T4 polynucleotide kinase were purchased from Fermentas. Phusion DNA polymerase is from Finnzymes. Unlabeled oligonucleotide primers were synthesized by STAB Vida, Portugal.

Strains

The majority of the experiments used a previously described *rrp44* deletion strain The *E. coli* strains used were XL1-Blue (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB* lac$I^{q}ZΔM15$ Tn*10* (Tet^{r})]) *E. coli* strain (Stratagene) for cloning

experiments, and Rosetta(DE3) [pLysS (F^- ompT $hsdS_B$(r_B^- m_B^-) gal dcm (DE3) pLysSRARE (Cam^r)] *E. coli* strain (Novagen) for expression and purification of enzymes.

Construction of *rrp44* mutants by site-directed mutagenesis

The point mutations Q892A and Y595A were introduced into *pGEX-RRP44* and *pRS415- RRP44* (Schaeffer et al., 2009) by site-directed mutagenesis. The oligonucleotides used in this study were the following (base changes are indicated in lower case letters): forward oligonucleotide Q892Aa, 5′-CAGAAACGCCgcATTCGCCGGTAGGGCTaGCATAG-3′ and reverse oligonucleotide Q892Ab, 5′-ATGCtAGCCCTACCGGCGAATgcGGCGTTTCTGTG-3′; forward oligonucleotide Y595Aa, 5′-GGTACTTCTGTAgcTTTGGTcGAC-3′ and reverse oligonucleotide Y595Ab, 5′-GTCgACCAAAcgTACAGAAGTACC-3′; forward oligonucleotide G895Aa, 5′-CGCCCAATTCGCCGcTAGGGCtAGC-3′ and reverse oligonucleotide G895Ab, 5′-GCTaGCCCTAgCGGCGAATTGGGCG-3′; forward oligonucleotide G895Ea, 5′-CGCCCAATTCGCCGaaAGGGCtAGC-3′ and reverse oligonucleotide G895Eb, 5′- GCTaGCCCTttCGGCGAATTGGGCG-3′; forward oligonucleotide G895Qa, 5′- CGCCCAATTCGCCcaaAGGGCtAGC-3′ and reverse oligonucleotide G895Qb, 5′-GCTaGCCCTttgGGCGAATTGGGCG-3′. Each oligonucleotide also changes a restriction site. We screened putative mutants by restriction digestion and verified them by sequencing at STAB Vida, Portugal.

Overexpression and purification of wild-type Rrp44 and mutants

Wild-type and mutated Rrp44 fused to GST were expressed in the Rosetta(DE3)pLysS *E. coli* strain. Cultures were grown in 2 liters TB medium supplemented with 200 mg ml^{-1} ampicillin and 34 mg ml^{-1} chloramphenicol at 30 ºC to an optical density at 600 nm (OD_{600}) of 1.2 and then induced expression by

addition of 0.2 mM IPTG and finally incubated overnight at 18 ºC. Cells were harvested by centrifugation and subsequently the pellets were frozen. Cells were thawed on ice, resuspended in 30 ml PBS buffer, pH 7.3, with 10 mM DTT and 1 mM PMSF, and lysed using the French Press. The crude extracts were treated with 5 units Benzonase (Sigma) for 30 min, and the extract was clarified by a 30-min centrifugation at 10,000 g. After this the protein extract was loaded onto a GSTrap™ FF 1-ml column (GE healthcare), previously equilibrated in binding buffer. Proteins were eluted with 50 mM Tris-HCl, pH 8.2, 10 mM reduced glutathione. The fractions containing the purified protein were pooled and loaded onto a HiTrap™ Q FF column (GE Healthcare), previously equilibrated in 50 mM Tris-HCl, pH 8.2. Proteins were then eluted with a 0–1 M NaCl gradient in 50 mM Tris-HCl, pH 8.2, 1 M NaCl, 1 mM DTT. The eluted protein was concentrated by centrifugation at 4 ºC with Vivaspin 500 Centrifugal Concentrators (Vivaspin). The protein concentrations were determined by spectrophotometry using a NanoDrop 1000 instrument (Alfagene) and a 50% (v/v) glycerol solution and 1 mM DTT was added to the final fractions before storage at –20 ºC.

RNase assays

Exoribonucleolytic activity was assayed using a generic 30-mer RNA oligoribonucleotide (5′-CCCGACACCAACCACUAAAAAAAAAAAAAA-3′) labeled at the 5′-end with [γ-^{32}ATP] and T4 polynucleotide kinase. Exoribonuclease activities were performed in a final volume of 12,5 µl containing 62,500 cpm of substrate, 20 nM of enzyme, and 10 mM, Tris pH 8, 75 mM KCl, 400 µM $MgCl_2$ and 1 mM β-mercaptoethanol (Dziembowski et al., 2007). Reactions were started by the addition of the enzyme and incubated at 37 °C. Samples were withdrawn at the time points indicated in the figures, and the reaction was stopped by adding formamide-containing dye supplemented with 10 mM EDTA. Reaction products were resolved in a 20% polyacrylamide, 7M urea. Gels were

imaged using STORM 860 Molecular Imager scanner (GE Heathcare). The exoribonucleolytic activity of the enzymes was determined by quantifying the disappearance of the substrate in several distinct experiments, and each value obtained represents the mean of at least three independent assays. Quantification was carried out using ImageJ (NIH).

Comparative modeling of Rrp44 wild-type and mutants Q892A and Y595A proteins

Preliminary 3D-structural models of *S. cerevisiae* Rrp44 wild-type and mutant proteins Q892A and Y595A, were built by using standard comparative methods, the software DeepView (Guex & Peitsch, 1997) and the crystallographic structure of yeast Rrp44 D551N mutant protein bound to a 10-nucleotide poly(A) RNA strand (Protein Data Bank code 2VNU; Lorentzen et al., 2008). In order to obtain uninterrupted theoretical structures in the Rrp44 region comprised between residues 252-1001, template set was partially enriched with segments from Rrp44 D551N mutant *apo* form (Protein Data Bank code 2WP8; Bonneau et al., 2009). Only fragments not affected by conformational changes between *apo* and *bound* states were used for this purpose (Lorentzen et al., 2008; Bonneau et al., 2009). Overall structural quality of all rendered models was checked using the analyses programs (Anolea, Gromos and Verify3D) provided by SWISS-MODEL server (http://swissmodel.expasy.org; Peitsch, 1995; Arnold et al., 2006; Kiefer et al., 2009).

Modeling of Rrp44 molecular interactions to single-stranded RNA

In order to model 3D-structural interaction of yeast RRP44/DIS3 wild-type, and mutants Q892A and Y595A, to one single-stranded poly(A) RNA molecule (ssRNA, RA1-RA10), preliminary structures were set up by translating the RNA ligand coordinates from the PDB structure 2VNU over each comparative model.

To obtain a set of refined models, a standard molecular dynamics protocol of 6 ns was performed over every Rrp44-ssRNA system, essentially as described elsewhere (Barbas et al., 2008, 2009). All the energy minimizations and molecular dynamics simulations were performed with the PMEMD module of the AMBER11 package (Case et al., 2010), using molecular mechanics parameters in *parm99* (Case et al., 2010) plus *parm99SBild* modifications (Wang, 2000) and *ff99bsc0* (Lindorff-Larsen et al., 2010) sets. Finally and for every RRP44/DIS3-ssRNA system, one minimized average structure was taken from the final most stable segment of simulation, as representative of the whole ensemble. Root-mean-square-deviation (RMSD) trace analyses of the peptide-backbone in the RNB domain (residues 475-911), and the phosphate-backbone in nucleotides RA1-RA5 were determined with the program PTRAJ from the same Amber suite. All RMSD point values were calculated from the initial conformation of the three modeled Rrp44-ssRNA systems. All structures were manipulated and visually rendered using Pymol.

Yeast complementation assay

An *rrp44Δ* strain complemented with by *RRP44*-wild-type on a plasmid with a *URA3* selectable marker was transformed with LEU2 plasmids encoding each of the point mutations. To test whether these mutations were essential for growth, the transformants were serially diluted and spotted onto plates containing 5-fluoro-orotic acid (5-FOA) or control plates (SC-LEU-URA). Growth on 5-FOA indicates that the mutated Rrp44 can carry out the essential function of the exosome.

his3-nonstop growth assay

To investigate whether the mutations of Rrp44 are required for exosome-mediated mRNA decay we used the two viable Rrp44 mutations in a *his3-nonstop*

growth assay. A *ski7Δ* strain [Vector, *LEU2*], *rrp44Δ* [*RRP44, LEU2*], *rrp44Δ* [*rrp44-Q892A, LEU2*] and *rrp44Δ* [*rrp44-Q892A, LEU2*] were transformed with a mutant *his3* gene, lacking a stop codon, on a plasmid with a *URA3* marker. The transformants were selected on SC-Ura-Leu. SC-Ura-Leu was used as control. A ski7Δ strain was included because it is known to stabilize the his3-nonstop mRNA. Single transformants were serially diluted and spotted onto SC-HIS plates to assess for growth, which indicates nonstop mRNA defects.

Northern blot analysis

The *RRP44*, *rrp44-D171A*, *rrp44-D551N*, *rrp6Δ*, *rrp44-Q892A*, *rrp44-Y595A* strains were grown in YPD, total RNA was isolated, resolved on polyacrylamide gels, and probed with 5′ ^{32}P-labelled oligonucleotides for 5.8S processing defects (5′-TTTCGCTGCGTTCTTCATC-3′), 5′ ETS degradation defects (5′-CGAACGACAAGCCTACTCG-3′), and as a loading control the RNA subunit of the signal recognition particle (SPR - 5′-GTCTAGCCGCGAGGAAGG-3′).

Acknowledgments

We thank Mónica Amblar for critical reading and Michal Malecki for fruitful discussion.

This work was supported by grants from Fundação para a Ciência e Tecnologia (FCT) to [C.M.A.], a Ph.D. fellowship from FCT to [F.P.R.], and a PostDoc fellowship from FCT to [A.B.].

Work in the A.v.H. laboratory was funded by grants from National Institutes of Health [R01GM099790], the Welch foundation [AU-1773] and an EMBO short term fellowship to [D.S.].

Work at the P.G-P. laboratory was supported by: the Spanish Ministerio de Ciencia e Innovación through [grant SAF2007-61926, IPT2011-0964-900000 and

SAF2011-13156-E] and the European Commission through [grant FP7 HEALTH-F3-2009-223431 (EU project "Divinocell")] and [FP7 HEALTH-2011-278603 (EU project "Dorian")]. Support from the "Fundación Ramón Areces" and the Centro de Computación Científica CCC-UAM" for computational support is also acknowledged. Work at Biomol-Informatics was partially financed by the European Social Fund.

Supplementary Information

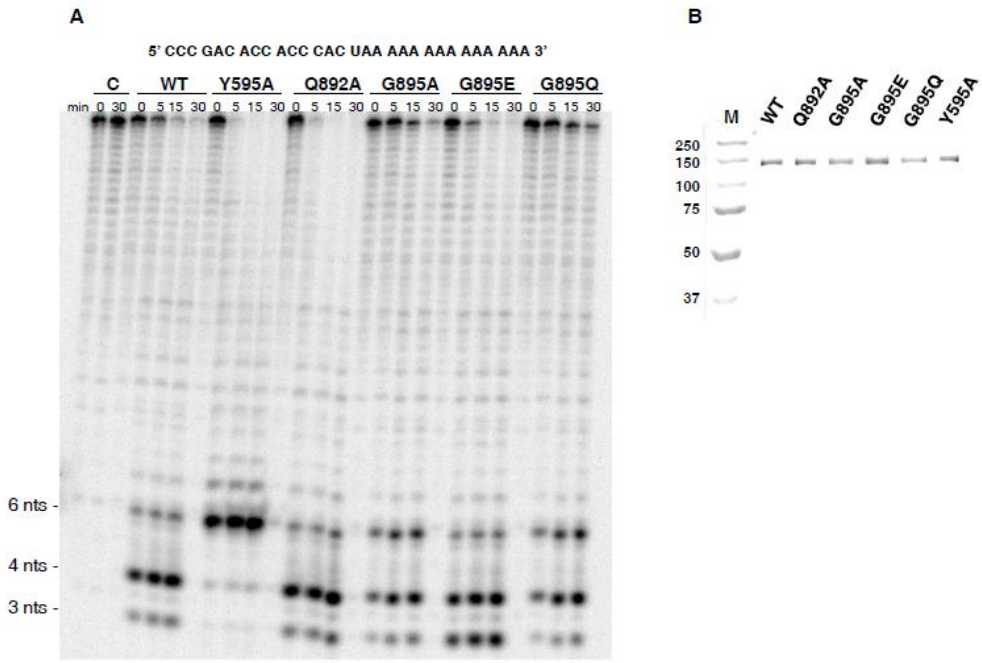


Figure S1. Exoribonuclease activity of Rrp44 and the different mutants. A. Activity assays were performed as described in Materials and Methods using 20 nM of enzyme and 4 nM of 30 mer RNA substrate (5′-CCC GAC ACC ACC CAC UAA AAA AAA AAA AAA-3′) and samples were taken at the time-points indicated. Control reactions (C) were incubated with no enzyme added. Length of substrates and degradation products are indicated in the figure. **B.** SDS-PAGE analysis of wild type and Rrp44 mutants. Equal amounts of mutant proteins were loaded into each lane; the gel was stained with Coomassie blue for detection.

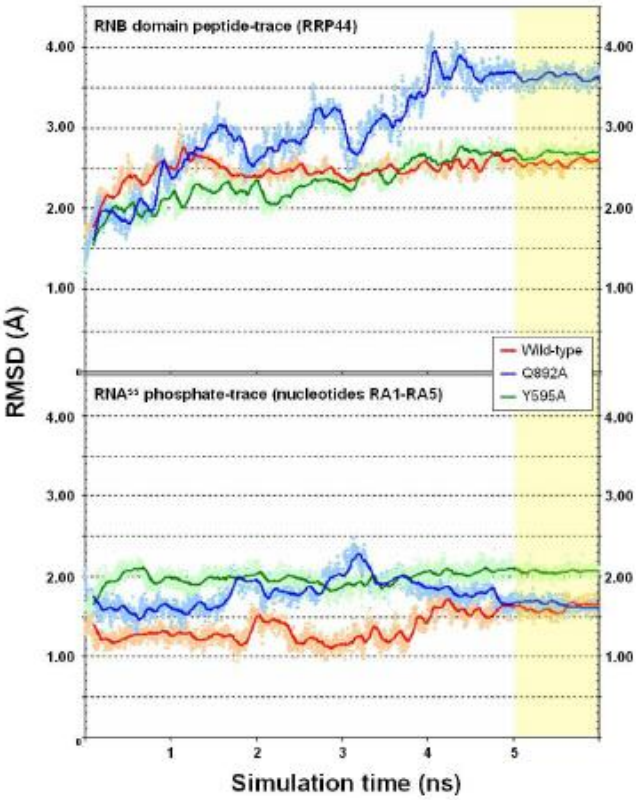


Figure S2. General structural derive of RNB-ssRNA theoretical models from yeast Rrp44/Dis3 wild-type, and mutants Y595A and Q892A. RMSD temporal profiles of RNB domain peptide-backbones (upper panel) and five most 3' ssRNA nucleotides (RA1-RA5) phosphate-backbone (lower panel) traces. RMSD point values and running averages (over 20) plotted in colours red (wild-type), blue (Q892A) and green (Y595A). RMSD measurements to initial conformations, revealed some interesting differential responses during the last segment of the simulations (Table S1), highlighted in soft yellow colour. Data from Rrp44/Dis3 Q892A suggest that this mutation might be able to induce significant variability in the peptide-backbone of the RNB domain, in comparison to the wild type and Y595A mutant. On the other hand, little variations seem to be affecting the phosphate-backbone structure of ssRNA within the catalytic cleft (nucleotides RA1-RA5) in the three phenotypes, although subtle differences could be altering it when bound to Y595A mutant.

Table S1. Average RMSD values obtained from the last nanosecond of free molecular dynamics trajectories of the three yeast Rrp44/Dis3-ssRNA interaction models calculated. Data correspond to peptide-backbone trace of RNB domains and nucleotides RA1-RA5 in RNA bound molecules respectively.

Enzyme	Rrp44/Dis3 RNB domain (Å)	poly(A)-RNA RA1-RA5 (Å)
Wild-type Rrp44	2.58±0.08	1.59±0.07
Y595A mutant	2.68±0.08	1.69±0.08
Q892A mutant	3.63±0.10	2.07±0.06

References

Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. 1999a. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *The EMBO journal 18*:5399-5410.

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P. 1999b. The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development 13*:2148-2158.

Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics 22*:195-201.

Astuti D, Morris MR, Cooper WN, Staals RH, Wake NC, Fews GA, Gill H, Gentle D, Shuib S, Ricketts CJ, Cole T, van Essen AJ, van Lingen RA, Neri G, Opitz JM, Rump P, Stolte-Dijkstra I, Muller F, Pruijn GJ, Latif F, Maher ER. 2012. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nat Genet 44*:277-284.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2008. New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end product. *The Journal of biological chemistry 283*:13070-13076.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2009. Determination of key residues for catalysis and RNA cleavage specificity: one mutation turns RNase II into a "SUPER-ENZYME". *The Journal of biological chemistry 284*:20486-20498.

Bollenbach TJ, Lange H, Gutierrez R, Erhardt M, Stern DB, Gagliardi D. 2005. RNR1, a 3'-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in Arabidopsis thaliana. *Nucleic acids research 33*:2751-2763.

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

Briggs MW, Burkard KT, Butler JS. 1998. Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. *The Journal of biological chemistry 273*:13255-13263.

Cairrão F, Arraiano C, Newbury S. 2005. Drosophila gene *tazman*, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev Dyn 232*:733-737.

Case DA, Darden TA, Cheatham TE, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts B, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani, F., Vanicek, J., Wu, X., Brozell,

S.R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, M.-J., Cui, G., Roe, D.R., Mathews, D.H., Seetin, M.G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A. and Kollman, P.A. AMBER 11, University of California, San Francisco. . 2010.

Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, Anderson KC, Ardlie KG, Auclair D, Baker A, Bergsagel PL, Bernstein BE, Drier Y, Fonseca R, Gabriel SB, Hofmeister CC, Jagannath S, Jakubowiak AJ, Krishnan A, Levy J, Liefeld T, Lonial S, Mahan S, Mfuko B, Monti S, Perkins LM, Onofrio R, Pugh TJ, Rajkumar SV, Ramos AH, Siegel DS, Sivachenko A, Stewart AK, Trudel S, Vij R, Voet D, Winckler W, Zimmerman T, Carpten J, Trent J, Hahn WC, Garraway LA, Meyerson M, Lander ES, Getz G, Golub TR. 2011. Initial genome sequencing and analysis of multiple myeloma. *Nature 471*:467-472.

de la Cruz J, Kressler D, Tollervey D, Linder P. 1998. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *The EMBO journal 17*:1128-1140.

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

Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature 443*:110-114.

Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R, Dietz HC. 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science 295*:2258-2261.

Ghosh S, Deutscher MP. 1999. Oligoribonuclease is an essential component of the mRNA decay pathway. *Proceedings of the National Academy of Sciences of the United States of America 96*:4372-4377.

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis 18*:2714-2723.

Hanekamp T, Thorsness PE. 1999. YNT20, a bypass suppressor of yme1 yme2, encodes a putative 3'-5' exonuclease localized in mitochondria of Saccharomyces cerevisiae. *Curr Genet 34*:438-448.

Houseley J, LaCava J, Tollervey D. 2006. RNA-quality control by the exosome. *Nature reviews 7*:529-539.

Huh WK, Falvo JV, Gerke LC, Carroll AS, Howson RW, Weissman JS, O'Shea EK. 2003. Global analysis of protein localization in budding yeast. *Nature 425*:686-691.

Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. 2009. The SWISS-MODEL Repository and associated resources. *Nucleic acids research 37*:D387-392.

Lebreton A, Tomecki R, Dziembowski A, Seraphin B. 2008. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature 456*:993-996.

Lim J, Kuroki T, Ozaki K, Kohsaki H, Yamori T, Tsuruo T, Nakamori S, Imaoka S, Endo M, Nakamura Y. 1997. Isolation of murine and human homologues of the fission-yeast dis3+ gene encoding a mitotic-control protein and its overexpression in cancer cells with progressive phenotype. *Cancer Res 57*:921-925.

Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins 78*:1950-1958.

Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell 127*:1223-1237.

Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Molecular cell 29*:717-728.

Lykke-Andersen S, Brodersen DE, Jensen TH. 2009. Origins and activities of the eukaryotic exosome. *Journal of cell science 122*:1487-1494.

Matos RG, Barbas A, Arraiano CM. 2009. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *The Biochemical journal 423*:291-301.

Mian IS. 1997. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic acids research 25*:3187-3195.

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

Niyogi SK, Datta AK. 1975. A novel oligoribonuclease of Escherichia coli. I. Isolation and properties. *The Journal of biological chemistry 250*:7307-7312.

Peitsch MC. 1995. Protein modelling by E-mail. . *Bio/Technology 13*:658-660.

Schaeffer D, Reis FP, Johnson SJ, Arraiano CM, van Hoof A. 2012. The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. *Nucleic acids research 40*:9298-9307.

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature structural & molecular biology 16*:56-62.

Schneider C, Anderson JT, Tollervey D. 2007. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol Cell 27*:324-331.

Schneider C, Leung E, Brown J, Tollervey D. 2009. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic acids research*.

van Hoof A, Frischmeyer PA, Dietz HC, Parker R. 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science 295*:2262-2264.

van Hoof A, Lennertz P, Parker R. 2000. 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 19*:1357-1365.

Wang JC, P.; Kollman, P.A. . 2000 How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem 21*:1049–1074.

Yu D, Deutscher MP. 1995. Oligoribonuclease is distinct from the other known exoribonucleases of Escherichia coli. *J Bacteriol 177*:4137-4139.

Chapter 5

Identification of the temperature-sensitive mutation that impairs the function *in vivo* and *in vitro* of RNase II

This chapter contains data from:
Reis FP, López-Viñas E, Gómez-Puertas P and Arraiano CM. A single point mutation is responsible for the temperature-sensitive phenotype of *E. coli* RNase II. *Manuscript in preparation*.

The author of this Dissertation performed and planned all the experiments described in this Chapter, with the exception of the modelling and alignment results that were performed by our collaborators, Professor Paulino Gómez-Puertas and Dr. Eduardo López-Viñas from Centro de Biologia Molecular "Severo Ochoa" (CSIC-UAM), Campus Universidad Autonoma de Madrid, Spain.

Abstract

A thermosensititve ribonuclease II (RNase II) allele, *rnb500*, which encodes RNase II with thermolabile activity, has been widely used for many years in the study of RNase II (Donovan & Kushner, 1986; Arraiano et al., 1988; Yancey & Kushner, 1990). Strains harbouring the *rnb500* allele produce an RNase II having thermolabile activity at 44 °C, *in vitro* and *in vivo* (Donovan & Kushner, 1986). Although this allele has been extensively used, the mutation responsible for the synthesis of an inactive RNase II at the nonpermissive temperature is still unknown. Here we describe for the first time the residues responsible for this temperature-sensitive phenotype.

Introduction

Escherichia coli ribonuclease II (RNase II, encoded by the *rnb* gene) is the prototype of RNase II family of enzymes (Mian, 1997), which are widespread among the three domains of life. *E. coli* RNase II is responsible for 90% of the exoribonucleolytic activity in crude extracts (Deutscher & Reuven, 1991). Therefore, RNase II has been the subject of extensive genetic, biochemical and structural studies aimed at understanding its precise biological role and mechanism of action.

A thermosensitive (ts) *rnb500* allele has been essential for the determination of the cellular function of RNase II, and also of other ribonucleases (Donovan & Kushner, 1986; Arraiano et al., 1988; Cruz et al., 1996; Mohanty & Kushner, 2000). This mutant allele was generated *in vitro* by hydroxylamine mutagenesis of the cloned wild-type *rnb* gene, and is thermolabile *in vitro* and *in vivo* at 44 °C (Donovan & Kushner, 1986). A strain with this temperature-sensitive mutant demonstrated that either PNPase or RNase II is required for cell viability in *E coli*, suggesting the involvement of both ribonucleases in a common essential process in the cell, more specifically in RNA decay (Donovan & Kushner, 1986). This allele was also important in the identification and characterization of 3' polyadenylation in *E. coli* (Cohen, 1995). Despite the wide use of *rnb500* during the last 25 years, nothing is known about the mutation responsible for such phenotype. The work presented here describes the identification of the mutation responsible for the temperature-sensitive *rnb500* allele.

Results and Discussion

In order to identify the mutation(s) responsible for the ts phenotype, we sequenced the gene encoding the thermolabile RNase II protein. The

chromosomal DNA of *E. coli* SK5689 strain (Arraiano et al., 1988) was used to amplify the *rnb500* gene by PCR, and the sequence of the gene was determined. To our surprise, two missense mutations were observed. One is the result of a G to A substitution at nucleotide 376 that resulted in an Asp to an Asn substitution at amino acid 126 (D126N). The other point mutation is a G to A transition at nucleotide 851, resulting in a Tyr residue in place of a Cys residue at amino acid 284 (C284Y). The RNB family exhibit the same modular organization: a central nuclease domain (known as RNB) flanked by oligonucleotide-binding family (OB) folds at the N-terminus (cold shock domains (CSD) 1 and 2) and at the C-terminus (S1 domain; (Frazão et al., 2006). The RNB domain is well conserved and is exclusive to RNase II proteins (Mian, 1997). Accordingly to the crystal structure of RNase II, D126 lies in CSD2 and C284 lies in the RNB catalytic domain. Although RNase II C284 has not been shown to be directly involved in substrate binding or catalysis, the conservation of the RNB domain within the multiple organisms that contains RNB-like enzymes suggests that this residue is important; perhaps it is involved in maintaining the structure and consequently the function of RNase II.

Strains deficient for both PNPase and RNase II are not viable (Donovan & Kushner, 1983). Therefore, to determine which mutant is responsible for the ts phenotype we decided to use a double exonuclease mutant (strain CMA401) harbouring the *pnp200* allele encoding a thermosensitive PNPase (Yancey & Kushner, 1990) in conjunction with a *rnb* deletion allele (Piedade et al., 1995). We chose pCMA01 (Zilhão et al., 1993) as the vector for genetic complementation because it contains the natural *rnb* promoter. To determine whether a single mutation or both are responsible for the thermolability of *rnb500* allele, we introduced each point mutation individually (single mutant) or the two substitutions (double mutant) in pCMA01, by site-directed mutagenesis. Transformation of the temperature-sensitive *pnp200* Δ*rnb E. coli* strain with a

plasmid bearing the wild-type *rnb* gene can abolish the ts phenotype at 44 °C (Fig. 1). Consequently, it is easy to score which mutant supports RNase II function *in vivo*.

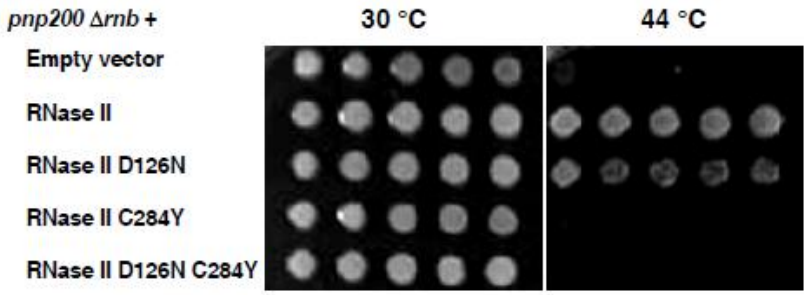


Figure 1. RNase II C284Y mutation is enough to confer thermosensitivity to RNase II. Growth at the permissive (30 °C) and nonpermissive temperature (44 °C) of a thermosensitive *E. coli* strain, harbouring empty vector or vectors encoding either wt RNase II or enzymes with the indicated mutations. Bacterial cultures were grown on LA medium, supplemented with ampicillin (100 μg/ml) and thymine (50 μg/ml) for one day.

Comparing to the wild-type, RNase II D126N exhibited no major phenotype, indicating that this individual mutation was not detrimental for the function of RNase II *in vivo*. Strikingly, the expression of RNase II C284Y and the double mutant (D126NC284Y) did not allow growth at the nonpermissive temperature (44 °C; Fig. 1). This result demonstrates that a single amino acid change (C284Y) is sufficient to generate the ts phenotype.

To see whether these substitutions altered RNase II activity, *in vitro* enzyme activity assays were carried. Therefore, the single mutations were cloned by site-directed mutagenesis into the previously described inducible plasmid pFCT6.9 (Cairrão et al., 2001) and the new plasmids were transferred to *E. coli* BL21(DE3), to overproduce the corresponding six-histidine-tagged RNase II variant proteins. The *in vitro* thermolability of RNase II was measured by preincubating the purified proteins at 44 °C for 30 min in the appropriate buffer assay, and then RNase II activity was measured at 30 °C (Donovan & Kushner,

1986). In the activity assay wild-type RNase II behaved as expected (Fig. 2), catalyzing the complete degradation of a 30 ss substrate to a final product of 4 to 6 nt, in less than 5 min with 1 nM of enzyme (Amblar et al., 2006). Not surprisingly, D126N exhibited almost the same activity as wild-type, confirming that this individual mutation is not responsible for the ts phenotype of *rnb500*. In contrast, the C284Y mutant enzyme was unable to degrade the substrate, even after 30 min of incubation. The decreased activity relative to wild-type, after the shift to the nonpermissive in the *in vitro* experiment, is in good agreement with what has been seen previously for enzyme activity in experiments with a strain carrying the *rnb500* allele (Donovan & Kushner, 1986). When lysates of a RNase II-deficient (*rnb296*) strain that harboured a *rnb500* plasmid (pDK39) were incubated at 44 °C for 30 min, the lysates lost more than 95% of activity (Donovan & Kushner, 1986). These results demonstrate that the thermolability of the *rnb500* proteins is due to a single amino acid substitution in the RNB domain.

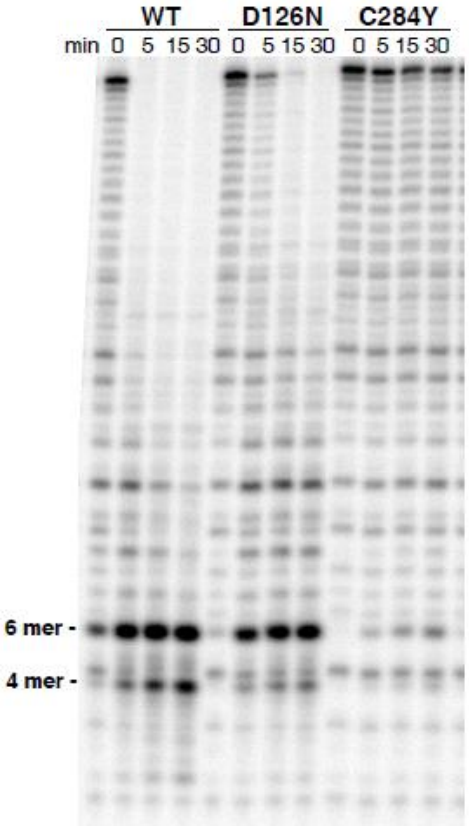


Figure 2. Exoribonucleolytic activity of RNase II and mutant versions of RNase II. Activity assays were performed as described in Materials and Methods using 1 nM of enzymes. Samples were taken at the time-points indicated. Length of substrates and degradation products are indicated in the figure.

The analyses of RNase II structure revealed that there is a second Cys residue in the close vicinity of C284 (C399), raising the possibility that a disulfide bridge can be formed between these two residues, increasing the thermostability of the protein. In order to confirm this hypothesis, we substituted a G for an A at nucleotide 1196 in pCMA01, by site-directed mutagenesis, and tested *in vivo* if mutant C399Y is also responsible for the ts phenotype. After transformation of the temperature-sensitive *pnp200 Δrnb E. coli* strain, the cells were not viable at 44 °C, indicating that like C284Y mutant, C399Y mutant is detrimental for the function of RNase II at the non-permissive temperature (Fig. 3). This result is consistent

with the hypothesis that the ts phenotype of *rnb500* allele is the result of a disruption of a disulfide bond between residues C284 and C399, essential for the proper function of RNase II.

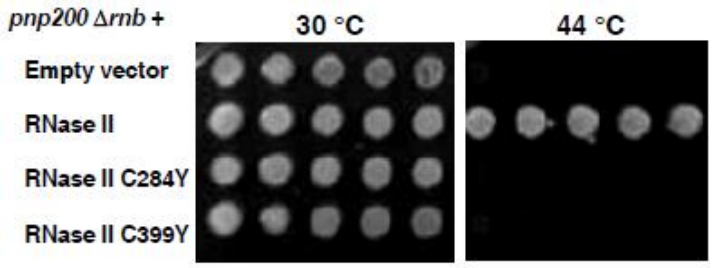


Figure 3. RNase II C399Y mutation cause very severe growth defect. Growth at the permissive (30 °C) and nonpermissive temperature (44 °C) of a thermosensitive *E. coli* strain, harbouring empty vector or vectors encoding either wt RNase II or enzymes with the indicated mutations. Bacterial cultures were grown on LA medium, supplemented with ampicillin (100 μg/ml) and thymine (50 μg/ml) for one day.

After confirming that changes in residues C284 or C399 are responsible for the ts phenotype, the level of conservation of these residues was analyzed in other RNase II-like enzymes. Sequence alignments revealed that most of the RNase II enzymes have conserved cysteine residues, with chances of forming a possible covalent interaction (Fig. 4). Regarding other members of the family, namely RNase R, Rrp44 and Dis3, the sequences and structures of the proteins do not appear to exhibit conserved Cys residues in the respective positions.

```
RNB_ECOLI      272 SLRANEVRPVLACRMTLSADGTIED 296     388 IVEEAMIAANICAARVLRDKLG.F 410
RNB_SALTI      272 SLRANEVRPALACRMIIAADGTIDD 296     388 IVEESMIAANLCAARVLRDKLG.F 410
RNB_YERPE      272 SLRPNERRPVLVCRVTITEEGTLSN 296     388 IVEECMIAANVCAALALREHLG.F 410
RNB_VIBCH      276 SLMENQVRPALCCSVTIRKDGVIGD 300     393 LVEESMITANICAGKTLQTTFG.F 415
RNR_VIBCH      350 SLNPQVDRLCMVCEMTVSETGKLSG 374     465 LIEECMILANIASASLVEKA.KEA 487
RNR_ECOLI      343 SLNPQVDRLCMVCEMTVSSKGRLTG 367     458 LIEECMILANISAARFVEKA.KEP 480
RNR_BACSU      331 SLNPKVDRLTLSCEMTINSQGQVTE 355     447 LIEEFMLVANETVAEHFHWM.NVP 469
RNR_MYCPN      347 SLNPNEKRYVVVCELNFDHEARLNF 371     462 MIENLMVVANEAVAWTLTNH.KVH 484
RRP44_HUMAN    550 SLKCDVDRLAFSCIWEMNHN.AEIL 573     662 MVEEFMLLANISVAKKIHEEFSEH 685
RRP44_MOUSE    550 SLRSNVDRLAFSCIWEMNHN.AEIL 573     662 MVEEFMLLANISVAKKIHEEFSEH 685
RRP44_YEAST    614 SLKPYVDRFAFSVIWELDDS.ANIV 637     726 LVEEFMLLANISVARKIYDAFPQT 749
DI3L1_HUMAN    549 SLLGGVDRYAVSIMWELDKASYEIK 573     680 TVAECMILANHWVAKKIWESFPHQ 703
DI3L1_MOUSE    549 SLLGGVDRYAVSVMWELDKTSYEIK 573     680 TVAECMILANHWVAKKIWESFPHQ 703
DI3L1_DANRE    544 SLLGGVDRYAMSVLWELDAETLEVC 568     672 TVAECMIYANHWVARKIQESFPHH 695
DI3L2_HUMAN    454 SLNPMSDKLTFSVIWTLTPE.GKIL 477     581 LVEEFMLLANMAVAHKIHRAFPEQ 604
DI3L2_MOUSE    452 SLNPMTDKLTFSVIWKLTPE.GKIL 475     579 LVEEFMLLANMAVAHKIFRTFPEQ 602
Q9RYD0_DEIRA   163 GLGLHEVSPALSICLDLDPDGNAEA 187     265 VVQECMTLAGWGTAIFADDNE.IP 287
D4H014_HALVD   152 SLVPEEDRLAHTVEMEIKKDTLSFE 176     243 IIEECMLKANKAVTHELMWNRGVE 266
```

Figure 4. Partial multiple sequence alignment of representative RNase II, RNase R, Rrp44 and Dis3. Position of *E. coli's* C284 and C399 is pointed out with black circles. All the residues positions colored according to sequence conservation.

Conclusion

In this report, we have determined for the first time the residues in the *rnb500* allele responsible for the thermolabile phenotype of RNase II at 44 °C. The *in vivo* and *in vitro* results pointed out the importance of a single residue in *rnb500* phenotype. A single substitution of a cysteine for a tyrosine (C284Y) leads to the disruption of a disulfide bridge that is linked to a loss in enzyme thermostability and the subsequent ts phenotype. This work also unraveled an unpredicted disulfide bridge in RNase II structure, responsible for a relevant covalent interaction between residues C284 and C399. The localization of these mutations can be an useful tool for studying RNase II enzymes function in other organisms, especially the ones that do not have a temperature sensitive protein and which function is still unknown.

Materials and Methods

Materials

Restriction enzymes, T4 DNA ligase, and T4 polynucleotide kinase were purchased from Fermentas. Phusion DNA polymerase is from Finnzymes. Oligonucleotides were synthesized by STABVida, Portugal. The primers used for the construction of RNase II mutants are given in Table 1.

Table 1 – Mutagenic primers used in this study. The mismatched nucleotide in each mutagenic primer is underlined and the restriction site is in lowercase letter.

Primer	Sequence	Amino acid change
oFRA3	5'-GAAGGCAACTGGGCGGTTGCCGAAATGCGaCGTCATCC-3'	D126N
oFRA4	5'-ATGACGtCGCATTTCGGCAACCGCCCAGTTGCCTTC-3'	
oFRA5	5'-TGGCATACCGCATGACGCTCTCaGCTGATG-3'	C284Y
oFRA6	5'-CAGCtGAGAGCGTCATGCGGTATGCCAGTA-3'	
oFRA30	5'-AAGCGATGATTGCCGCTAgCATTTATGCGGCCCGCGTACTGC-3'	C399Y

oFRA31	5'-GCATAAATGcTAGCGGCAATCATCGCTTCTTCGACGATACG-3'

E. coli strains

The following *E. coli* strains were used: DH5α [fhuA2 Δ(argF-lacZ)U169 phoAglnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17 (Taylor et al., 1993)] for cloning experiments, strain CMA401, an isogenic derivative of *E. coli* MG1693 strain [*thy*A715, *pnp*200 *rnb*Δ::tet^R], was used for the genetic complementation assay, and BL21(DE3) [F$^-$ r$_B$$^-m_B$$^-$ gal ompT (int::P$_{lacUV5}$ T7 gen1 imm21 nin5) (Studier & Moffatt, 1986)] was used for expression and purification of enzymes. The *rnb500* allele was obtained from SK5689 *E. coli* strain (Arraiano et al., 1988).

Identification and cloning of *rnb500* mutations

The *rnb*500 allele was amplified from the chromosomal DNA of *E. coli* SK5689 (Arraiano et al., 1988) using a standard PCR reaction with Pfu DNA polymerase. The primers used for the amplification matched perfectly 37 nucleotides upstream of the initiation codon (5'-CTGTCAGCCGCTCTAATGGCC-3') and 33 nucleotides downstream of the stop codon (5'-GCCCATCCATGAGGAATGGGC-3') of RNase II, respectively. The DNA sequence of the 2005 bp PCR product was determined (STABvida, Oeiras, Portugal). Two point mutations were detected: a G → A substitution at position 376 (residue substitution D126N) and a G → A substitution at position 851 (residue substitution C284Y) of the *rnb* gene. These mutations were further inserted into the previously described pCMA01 (Zilhão et al., 1993) and pFCT6.9 plasmid (Cairrão et al., 2001), by site-directed mutagenesis. The nucleotide sequences of the mutagenic primers used to generate point mutations are shown in Table 1. We screened all putative mutant constructs by restriction digestion and verified them by sequencing at STAB Vida, Portugal.

Genetic complementation assays

E. coli MG1693 derivative cells (strain CMA401) [*thy*A715, *pnp*200 *rnb*Δ::tet^R] were transformed with pBluescript SK+ (Stratagene) or pCMA01 (Zilhão et al., 1993) encoding either wild-type RNase II protein or its mutant derivatives, and grown on agar plates supplemented with 100 μg/ml ampicillin and 50 μg/ml thymine. The genetic complementation ability of the various RNase II mutant derivatives were assessed by examining growth of colonies after one day incubation at the non-permissive temperature (44 ºC). In the complementation experiments, cells transformed with pBluescript SK+ (vector alone) and pCMA01 (i.e., wild-type RNase II protein) served as the negative and positive controls, respectively.

Overexpression and purification of RNase II proteins

The plasmid used for expression of wild-type *E. coli* histidine-tagged RNase II protein was pFCT6.9 plasmid (Cairrão et al., 2001). This plasmid contains the *rnb* gene cloned into pET-15b vector (Novagen) under the control of ϕ10 promoter, allowing the expression of the $(His)_6$-tagged RNase II fusion protein or its mutant derivatives.

E. coli BL21(DE3) containing the recombinant plasmid of interest was grown at 30 ºC in 200 mL of TB medium supplemented with 200 μg/mL ampicillin to an optical density at 600 nm (OD_{600}) of 1.5. Protein expression was then induced by addition of 0.5 mM IPTG for 20 h at 26 ºC. Cells were harvested by centrifugation and stored at -80 ºC.

Purification was performed by histidine affinity chromatography using HisTrap Chelating HP columns (GE Healthcare) and an ÄKTA HPLC system (GE Healthcare) following the protocol previously described (Amblar et al., 2006; Arraiano et al., 2008). Briefly, frozen cells were thawed and resuspended on 6 mL of buffer A [20 mM Tris-HCl, 0.5 M NaCl and 20 mM imidazole, pH 8]. Cell

suspensions were lysed using a French press at 9000 psi in the presence of 1 mM PMSF. The crude extracts were treated with Benzonase (Sigma) to degrade the nucleic acids and clarified by a 60 min centrifugation at 10,000 g. The clarified extracts were then loaded into a HisTrap Chelating Sepharose 1 mL column equilibrated in buffer A. Protein elution was achieved with a continuous imidazole gradient (from 20 to 500 mM) in buffer A. Eluted protein was buffer exchanged with buffer C [20 mM Tris-HCl, 250 mM NaCl, pH 8]. Protein concentrations were determined by spectrophotometry using a NanoDrop 1000 instrument (Alfagene), and 50% (v/v) glycerol and 5 mM DTT was added to the final fractions prior to storage at -20 ºC.

Activity assays

Exoribonucleolytic activity was assayed using a 30-mer oligoribonucleotide (5'-CCCGACACCAACCACUAAAAAAAAAAAAAA-3') as substrate. The oligoribonucleotide was labeled at 5'-end with [γ-^{32}P]ATP with T4 polynucleotide kinase. Before the reaction, the RNA substrate (62,500 cpm per reaction) were heated for 5 min and cooled to 30 ºC temperature prior to determining activities. The exoribonucleolytic reactions were performed in a final volume of 12.5 μL containing 20 mM of Tris/HCl, pH 8, 1 mM of DTT, 100 mM of KCl and 1 mM of $MgCl_2$ (Arraiano et al., 2008). The *in vitro* thermolability of RNase II was measured by preincubating 1 nM of each protein (unless otherwise specified) in the reaction buffer at 44 ºC for 30 min, to inactivate the enzymes, and cooled to 30 ºC for 20 min (Donovan & Kushner, 1986). Reactions were started by the addition of the substrate and mixtures were incubated at 30 ºC. Samples (2 μl) were withdrawn at the times indicated in the figures, and the reaction was stopped via addition of formamide-containing dye supplemented with 10 mM EDTA. Samples were loaded onto a 20% polyacrylamide/7 M urea gel for electrophoresis. Gels were imaged using a STORM 860 scanner.

Acknowledgments

We thank José Andrade, Michal Malecki and Susana Domingues for helpful advice and discussions. Strain CMA401 was kindly provided by José Andrade. We also thank Andreia Aires for technical assistance. This work was funded by Fundação para a Ciência e a Tecnologia (FCT), Spanish Ministerio de Ciencia, Innovación and the European Commission, "Fundación Ramón Areces", Centro de Computación Científica CCC-UAM" and by the European Social Fund.

References

Amblar M, Barbas A, Fialho AM, Arraiano CM. 2006. Characterization of the functional domains of Escherichia coli RNase II. *Journal of molecular biology 360*:921-933.

Arraiano CM, Barbas A, Amblar M. 2008. Characterizing ribonucleases in vitro examples of synergies between biochemical and structural analysis. *Methods in enzymology 447*:131-160.

Arraiano CM, Yancey SD, Kushner SR. 1988. Stabilization of discrete mRNA breakdown products in ams pnp rnb multiple mutants of Escherichia coli K-12. *J Bacteriol 170*:4625-4633.

Cairrão F, Chora A, Zilhão R, Carpousis AJ, Arraiano CM. 2001. RNase II levels change according to the growth conditions: characterization of gmr, a new Escherichia coli gene involved in the modulation of RNase II. *Molecular microbiology 39*:1550-1561.

Cohen SN. 1995. Surprises at the 3' end of prokaryotic RNA. *Cell 80*:829-832.

Cruz AA, Marujo PE, Newbury SF, Arraiano CM. 1996. A new role for RNase II in mRNA decay: striking differences between RNase II mutants and similarities with a strain deficient in RNase E. *FEMS Microbiol Lett 145*:315-324.

Deutscher MP, Reuven NB. 1991. Enzymatic basis for hydrolytic versus phosphorolytic mRNA degradation in Escherichia coli and Bacillus subtilis. *Proceedings of the National Academy of Sciences of the United States of America 88*:3277-3280.

Donovan WP, Kushner SR. 1983. Amplification of ribonuclease II (rnb) activity in Escherichia coli K-12. *Nucleic acids research 11*:265-275.

Donovan WP, Kushner SR. 1986. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proceedings of the National Academy of Sciences of the United States of America 83*:120-124.

Frazão C, McVey CE, Amblar M, Barbas A, Vonrhein C, Arraiano CM, Carrondo MA. 2006. Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature 443*:110-114.

Mian IS. 1997. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic acids research 25*:3187-3195.

Mohanty BK, Kushner SR. 2000. Polynucleotide phosphorylase, RNase II and RNase E play different roles in the in vivo modulation of polyadenylation in Escherichia coli. *Molecular microbiology 36*:982-994.

Piedade J, Zilhao R, Arraiano CM. 1995. Construction and characterization of an absolute deletion mutant of Escherichia coli ribonuclease II. *FEMS Microbiol Lett 127*:187-193.

Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology 189*:113-130.

Taylor RG, Walker DC, McInnes RR. 1993. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic acids research 21*:1677-1678.

Yancey SD, Kushner SR. 1990. Isolation and characterization of a new temperature-sensitive polynucleotide phosphorylase mutation in Escherichia coli K-12. *Biochimie 72*:835-843.

Zilhão R, Camelo L, Arraiano CM. 1993. DNA sequencing and expression of the gene rnb encoding Escherichia coli ribonuclease II. *Molecular microbiology 8*:43-51.

Chapter 6

Discussion and Future Perspectives

RNA is degraded by a large variety of nucleases, and ribonucleases from the RNase II family are key players in the processing and degradation of RNA in organisms from all kingdoms of life. The precise definition of the biochemical properties and mechanism of action of enzymes from the RNase II family was the main objective of this Doctoral work.

The PIN domain is an endonuclease

The eukaryotic exosome is a versatile ribonucleolytic protein complex, with a major function in the 3'–5'processing and degradation of RNA. All ten subunits of the yeast exosome are essential for viability (Mitchell et al., 1997; Allmang et al., 1999), underscoring the importance of this complex in RNA metabolism. Previously, biochemical studies demonstrated that all the subunits of the exosome are inactive, except Rrp44/Dis3, which displays a processive and hydrolytic activity dependent on the intact RNB domain (Liu et al., 2006; Dziembowski et al., 2007). Rrp44 displays a modular domain organization, with a region similar to *Escherichia coli* RNase II (RNase II-like region), that includes a central RNB domain and three OB-fold RNA binding domains (CSD1, CSD2 and S1). Additionally, at the N-terminus, Rrp44 contains a PIN domain whose function was unknown until the beginning of this study. The first objective of this Doctoral work was to determine the role of the PIN domain (**Chapter 2**).

Two assumptions made us hypothesize that this domain was an active nuclease. First, the PIN domain family has representatives in all domains of life, some of which with nuclease activity (Wall & Kaiser, 1999; Arcus et al., 2004). Second, although Rrp44 is an essential gene, a point mutation that inactivates its exoribonuclease activity is still viable suggesting the existence of another nuclease activity. In order to assess its activity, we cloned, overexpressed and purified in *E. coli* different versions of yeast Rrp44 protein, and analyzed their ribonucleolytic

activity. Expression of these recombinant proteins was a major bottleneck of this work. A number of strategies had to be designed to improve expression and the yield of soluble protein. After testing different tags, vectors, strains and cell growth parameters, we were able to increase the yield of the soluble fusion proteins and perform *in vitro* activity assays. We, and three other groups, showed that PIN domain has a robust endonucleolytic activity that is abolished by mutations in acidic residues coordinating manganese (Lebreton et al., 2008; Schaeffer et al., 2009; Schneider et al., 2009). It is not surprising that two previous studies did not detect this endonucleolytic activity, and had reported that Rrp44 with a mutation in the RNB domain was inactive *in vitro* (Dziembowski et al., 2007; Schneider et al., 2007). The endonucleolytic activity had not been noticed because Rrp44 enzymatic activities have different buffer preferences. While Mg^{2+} is the most favorable divalent cation for the RNB domain, the PIN domain is more active in the presence of Mn^{2+}. This is in good agreement with another eukaryotic PIN domain protein (Glavan et al., 2006).

These results were a major breakthrough in the RNA field because they uncovered a novel mechanism of degradation of RNA, and the model of the exosome had to be revisited. The aforementioned results also created much excitement and discussions about the conservation of the degradation mechanisms and complexes in prokaryotes and eukaryotes.

Regulation of the catalytic activities of Rrp44

Excessive or uncontrolled RNA degradation might cause deleterious effects to the cell. Given the wide impact of the exosome in RNA processing and degradation pathways, the precise tuning of Rrp44 activity is therefore of particular importance. Consequently, mechanisms that control the activities of this enzyme are essential for the cells. Biochemical and structural data revealed that Rrp44

binds to the exosome ring mainly due to interactions with the PIN domain (Wang et al., 2007; Bonneau et al., 2009). These observations were supported by genetic analysis demonstrating that the PIN domain deletion results in the dissociation of Rrp44 from the yeast exosome ring (Schneider et al., 2009). Threading of RNA substrates through the ring (Bonneau et al., 2009; Malet et al., 2010) constitutes a powerful means to prevent uncontrolled RNA degradation, by limiting the access to the exoribonucleolytic site. Therefore, the association of Rrp44 with the exosome core modulates its exonuclease activity on structured substrates (Liu et al., 2006; Lorentzen et al., 2008). The association with the ring does not inhibit the degradation of circular substrates by the PIN domain that seems to face the solvent (Bonneau et al., 2009). However, recent data suggest that its activity is modulated by the exosome (Wasmuth & Lima, 2012). Results from this Doctoral work (results from **Chapter 2** and **Chapter 3**), contributed to unravel different modes of regulation of exo- and endonuclease activities.

First, substrates with a 5′ monophosphate are more rapidly degraded by the PIN domain than substrates with a 5′ hydroxyl or circular substrates (results from **Chapter 2**). This phenomenon is known as "5′-sensing" and is analogous to *E. coli* RNase E (Mackie, 1998), a catalytic component of the degradosome. This endoribonuclease is also more active on 5′-monophosphorylated substrates and it was shown that for an efficient cleavage, RNA pyrophosphohydrolase (RppH) should convert the 5′ terminus of primary transcripts from a triphosphate to a monophosphate (Deana et al., 2008). This result raises another parallel between prokaryotes and eukaryotes, and suggests that maybe the PIN domain is more active on substrates that are decapped, restricting its action on capped substrates. However, this mechanism needs to be clarified (see **Future Perspectives**).

Second, we and others noticed that the truncated protein, containing only the PIN domain, is much more active than the full-length protein (results from **Chapter 2**; Lebreton et al., 2008). A similar result was previously reported for

RNase II, where was shown that CSD domains seem to be involved in the control of mRNA degradation (Amblar et al., 2006). This observation suggests that other Rrp44 domains somehow regulate the activity of the PIN domain.

Lastly, Rrp44 contains at the N-terminus a CR3 (cysteine-rich with three cysteines) region (Cairrão et al., 2005). This region contains a Cx_4Cx_2C motif with no homology to any other domain or motif, and unknown function. Previously, we had shown that mutation of the three Cys of this region results in a slow growth phenotype (Schaeffer et al., 2009), suggesting an important function. Later, a crystal structure of Rrp44 provided the first insight into this motif (Bonneau et al., 2009) but this structure did not contain a complete N-terminus. The spatial position of three Cys and His184 resembled a metal binding site. Through genetic analysis we showed that the N-terminal CR3 region in conjunction with H184 residue has metal binding capacity (results from **Chapter 3**). Moreover, we suggest that the binding of the metal could stabilize the folding of this part of the protein, affecting interactions with the exosome and consequently regulating exonuclease activity. Moreover, since H184 residue is part of the same α-helix of D171 residue (one of the four conserved residues essential for metal coordination in the endonuclease catalytic site), we tested how this motif could affect endonuclease activity *in vitro*. Our results showed that the wild-type protein is two-fold more active than the protein lacking the CR3 motif, suggesting that this motif also affects endonuclease activity. As a result, the CR3 metal binding capacity regulates endonuclease activity directly, and exonuclease activity through the interaction of the protein with the exosome. This study gave a molecular explanation for the fact that the mutation of the Cys residues leads to a decrease in enzymatic activity and uncovered the mechanistic role of CR3.

Studying the role of important residues in RNA degradation

Previous reports have shown that Rrp44 is responsible for the activity of the core exosome (Liu et al., 2006; Dziembowski et al., 2007). In addition, the individual domains of Rrp44 are similar in sequence and structure to the corresponding domains of RNase II. The most similar is the catalytic domain RNB, with the highest conservation at the active site. Moreover, most of the residues that interact with the 3′ end of RNA are conserved through RNase II superfamily, suggesting that the two nucleases have a similar hydrolytic mechanism (Lorentzen et al., 2008; Matos et al., 2009). In this part of the Doctoral work, the aim was to characterize the role of some residues from the evolutionary conserved RNB domain in catalysis and analyze the physiological importance of these residues (**Chapter 4**).

Previous studies in our laboratory have shown that when *E. coli* RNase II Glu542 is changed by an Ala, the enzyme becomes 120 fold more active (Barbas et al., 2009). This residue is in close proximity with the leaving nucleotide and it is thought to prevent nucleotide degradation by slowing down the activity of RNase II. We have performed mutations in equivalent Rrp44 residues and our results showed that when we replaced Gln892 residue with an alanine the enzyme became twice more active *in vitro*. However, this increase in activity does not seem to affect the function of the exosome *in vivo*.

In *E. coli* RNase II, Tyr253 was determined to be responsible for the stabilization of the 3′-end of the RNA molecule in the catalytic site and to be critical in setting the size of the end-product (Barbas et al., 2008). This residue is highly conserved in RNase II family of enzymes being equivalent to Y595 in yeast Rrp44 (Lorentzen et al., 2008; Matos et al., 2009). The crystal structure of yeast Rrp44 suggests that Y595 plays a similar role, but its function had not been

experimentally addressed (Lorentzen, 2008). Similarly to what was reported for *E. coli* Y253, when Y595 residue is changed to an alanine, the *in vitro* end-product changed. In addition, an alteration in this residue has an important biological role since the mutation in this amino acid affected cell growth, RNA processing and degradation. These results lead us to hypothesize that the alteration of the size of the end-product leads to the accumulation of degradation fragments in the cell, inhibiting the action of oligoribonuclease homologues in yeast.

This study brought new insights and improved the proposed model for RNA degradation by Rrp44, defining the precise role and function of some important residues present in the active site of the enzyme. These findings are quite important since the determination of the role of the conserved residues in Rrp44 is essential to understand how the exosome ensures proper cellular RNA metabolism.

An *in vitro* tool to study RNA metabolism

E. coli RNase II is the prototype of the RNase II superfamily of exoribonucleases (Mian, 1997), whose homologues are present in all three domains of life. RNase II has been the subject of extensive genetic, biochemical and structural studies to understand its biological role and mechanism of action. A thermosensitive (ts) *E. coli rnb500* allele generated *in vitro*, has been essential for the determination of the cellular function of RNase II (Donovan & Kushner, 1986). Despite the wide use of *rnb500* during the past 25 years, the mutation that impairs the function of RNase II *in vivo* and *in vitro* at the nonpermissive temperature is still unknown. In the final part of this Doctoral work, we intended to identify the mutation(s) responsible for this phenotype (**Chapter 5**). Sequencing of the DNA and sequence alignments of wild-type *rnb* gene and *rnb500* allele allowed us to identify two changes in the sequence. Through *in vivo* complementation assays and *in vitro*

enzymatic activity assays performed at the nonpermissive temperature, the mutation responsible for the ts phenotype was identified. This mutation was the result of a substitution of a cysteine to a tyrosine at residue 284. The analysis of the structure of RNase II, by our collaborators Professor Paulino Gómez-Puertas and Dr. Eduardo López-Viñas from Centro de Biologia Molecular "Severo Ochoa" (CSIC-UAM), Campus Universidad Autonoma de Madrid, confirmed this result and revealed the existence of a second Cys residue (C399) in the close vicinity of C284, unravelling an unpredicted disulphide bridge in RNase II structure. This bridge is responsible for a relevant covalent interaction between residues C284 and C399.

In summary, we identified the mutations responsible for the ts phenotype. These residues are conserved in other RNase II proteins. As such, the homologous ts enzyme can be used as a powerful tool in the study of RNase II enzymes function in other organisms, especially the ones that do not have a ts protein or those like *Synechocystis sp*. in which RNase II is the only exoribonuclease.

Future Perspectives

In vivo genetic data indicated that either the endo- or exonuclease active sites of Rrp44 can be sufficient for viability, but the simultaneous inactivation of both catalytic sites impairs cell growth. This suggests that both domains are active and important for the cell. We showed that the RNB domain is required for some functions of the exosome, namely, 5.8S rRNA processing and 5'ETS degradation (Schaeffer et al., 2009). However, inactivation of the PIN domain does not cause large defects in the processing and degradation of the same transcripts. The presence of two catalytic sites can be important to increase the efficiency of the overall reaction or to facilitate the degradation of RNA secondary structures that

cause the exosome to stall in the 3′ to 5′ direction. However, it remains to be investigated how these activities are coordinated and the relative contribution of each activity in the processing and degradation of RNA in the cell. As a future perspective of this study, it would be interesting to determine the substrates of each catalytic site using microarray analysis and/or deep sequencing.

We identified a regulatory mechanism of the PIN domain, since our results showed that the activity of this domain is more pronounced on 5′ monophosphorylated substrates than an hydroxyl or circular substrate. Thus, it would be interesting to characterize the 5′-end dependency of this domain. The 5′-end dependency of Rrp44 endonuclease can be tested by activity assays using the purified protein lacking the RNase II-like region and a biological substrate with different 5′ ends, including a 5′ monophophate, triphosphate and cap.

In the first and second part of this Doctoral work we noticed that other regions or domains of Rrp44 can regulate the activity of the PIN and RNB domains. Therefore, it would be appealing to study the involvement of other domains in the regulation of these two activities. For this study, different truncated Rrp44 mutants should be constructed, overexpressed and purified from *E. coli*. The comparison of their endo- and exoribonucleolytic with the wild type Rrp44 enzyme, could help to elucidate their involvement in the regulation of RNA decay.

In the second part of this work we have identified an unpredicted metal binding site in the CR3 motif. Although a zinc ion is most probably the metal coordinated in this site, other metals could also be considered. To determine whether Rrp44 coordinates a zinc ion, a full length Rrp44 should be purified from *E. coli* lysates, and be subjected to atomic absorption spectroscopy. Mn^{2+}, which is coordinated by the endonuclease active site, and Mg^{2+}, which is coordinated by the exonuclease active site, should also be detected.

In the fourth part of this work, we were able to identify the mutation responsible for the thermolability of a RNase II allele, *rnb500*. It will be interesting to introduce this mutation in other RNase II enzymes from other organisms and test if causes the same phenotype. The localization of these mutations can be an useful tool for studying RNase II enzymes function in other organisms.

The findings that we present in this Dissertation were very important to understand the mechanism of action of yeast Rrp44. Since Rrp44 is the only active nuclease of the exosome, understanding the activity of Rrp44 is essential to understand the activity of the exosome core. In addition, most of the results can be extrapolated to other members of the RNase II family of enzymes.

References

Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, Mitchell P. 1999. The yeast exosome and human PM-Scl are related complexes of 3' --> 5' exonucleases. *Genes & development 13*:2148-2158.

Amblar M, Barbas A, Fialho AM, Arraiano CM. 2006. Characterization of the functional domains of Escherichia coli RNase II. *J Mol Biol 360*:921-933.

Arcus VL, Backbro K, Roos A, Daniel EL, Baker EN. 2004. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J Biol Chem 279*:16471-16478.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2008. New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end product. *J Biol Chem 283*:13070-13076.

Barbas A, Matos RG, Amblar M, Lopez-Vinas E, Gomez-Puertas P, Arraiano CM. 2009. Determination of key residues for catalysis and RNA cleavage specificity: one mutation turns RNase II into a "SUPER-ENZYME". *J Biol Chem 284*:20486-20498.

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

Cairrão F, Arraiano C, Newbury S. 2005. Drosophila gene *tazman*, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev Dyn 232*:733-737.

Deana A, Celesnik H, Belasco JG. 2008. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature 451*:355-358.

Donovan WP, Kushner SR. 1986. Polynucleotide phosphorylase and ribonuclease II are required for cell viability and mRNA turnover in Escherichia coli K-12. *Proc Natl Acad Sci U S A 83*:120-124.

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

Glavan F, Behm-Ansmant I, Izaurralde E, Conti E. 2006. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *Embo J 25*:5117-5125.

Lebreton A, Tomecki R, Dziembowski A, Seraphin B. 2008. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature 456*:993-996.

Liu Q, Greimann JC, Lima CD. 2006. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell 127*:1223-1237.

Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. 2008. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell 29*:717-728.

Mackie GA. 1998. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature 395*:720-723.

Malet H, Topf M, Clare DK, Ebert J, Bonneau F, Basquin J, Drazkowska K, Tomecki R, Dziembowski A, Conti E, Saibil HR, Lorentzen E. 2010. RNA channelling by the eukaryotic exosome. *EMBO reports 11*:936-942.

Matos RG, Barbas A, Arraiano CM. 2009. RNase R mutants elucidate the catalysis of structured RNA: RNA-binding domains select the RNAs targeted for degradation. *The Biochemical journal 423*:291-301.

Mian IS. 1997. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic Acids Res 25*:3187-3195.

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

Schaeffer D, Tsanova B, Barbas A, Reis FP, Dastidar EG, Sanchez-Rotunno M, Arraiano CM, van Hoof A. 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature structural & molecular biology 16*:56-62.

Schneider C, Anderson JT, Tollervey D. 2007. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol Cell 27*:324-331.

Schneider C, Leung E, Brown J, Tollervey D. 2009. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res 37*:1127-1140.

Wall D, Kaiser D. 1999. Type IV pili and cell motility. *Mol Microbiol 32*:1-10.

Wang HW, Wang j, Ding F, Callahan K, Bratkowski J, Butler JS, Nogles E, Ke A. 2007. Architecture of the yeast Rrp44-exosome complex suggests routes of RNA recruitment for 3′end processing. *Proc Nat Acad Sci USA 104*:16844-16849.

Wasmuth EV, Lima CD. 2012. Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. *Mol Cell 48*:133-144.

Appendix

Publications

nature structural & molecular biology

The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities

Daneen Schaeffer[1], Borislava Tsanova[1], Ana Barbas[2], Filipa Pereira Reis[2], Eeshita Ghosh Dastidar[1], Maya Sanchez-Rotunno[1], Cecília Maria Arraiano[2] & Ambro van Hoof[1,3]

The eukaryotic exosome is a ten-subunit 3′ exoribonucleolytic complex responsible for many RNA-processing and RNA-degradation reactions. How the exosome accomplishes this is unknown. Rrp44 (also known as Dis3), a member of the RNase II family of enzymes, is the catalytic subunit of the exosome. We show that the PIN domain of Rrp44 has endoribonucleolytic activity. The PIN domain is preferentially active toward RNA with a 5′ phosphate, suggesting coordination of 5′ and 3′ processing. We also show that the endonuclease activity is important *in vivo*. Furthermore, the essential exosome subunit Csl4 does not contain any domains that are required for viability, but its zinc-ribbon domain is required for exosome-mediated mRNA decay. These results suggest that specific exosome domains contribute to specific functions, and that different RNAs probably interact with the exosome differently. The combination of an endoRNase and an exoRNase activity seems to be a widespread feature of RNA-degrading machines.

Virtually all RNAs are processed from longer precursors. Although the removal of 3′ extensions seems to be a simple enzymatic reaction, eukaryotic cells contain multiple endoribonucleases (endoRNases) and 3′ exoribonucleases (exoRNases)[1–4]. These exoRNases can act redundantly for some reactions, but are specific for other functions[5]. In addition to these RNA-processing reactions that generate mature RNAs, 3′ exoRNases and endonucleases also carry out mRNA-degradation reactions[6,7] that can be an important determinant for mRNA quality control[8,9]. The same RNase can process some substrates and completely degrade others. It is largely unknown what determines the specificity of different RNases for their particular substrates, or for processing versus decay.

The exosome is a major 3′ exoRNase that is present in all eukaryotes, but whose function has been most extensively characterized in *Saccharomyces cerevisiae*[10,11]. The yeast exosome is responsible for the 3′-end processing of stable RNAs including ribosomal RNA (rRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA), as well as the degradation of RNAs such as the 5′ external transcribed spacer (ETS) of the rRNA precursor and aberrant mRNAs that lack a stop codon or a poly(A) tail[7–14].

The core exosome contains nine subunits and shows extensive structural similarity to bacterial PNPase and archaeal exosomes[15–20] (**Fig. 1a**, above). These nine subunits are all essential for viability. Six subunits have an RNase PH domain and form a ring structure in the exosome[15]. Although these six subunits show clear sequence and structural similarity to RNases, all six subunits in the yeast and human exosome lack important catalytic residues and are inactive[15,21]. The six PH-ring subunits cannot assemble into a PH-ring *in vitro*, but, when the three cap proteins are added, a stable core exosome can be assembled[15]. The crystal structure shows that each cap protein contacts two neighboring subunits of the PH ring. These observations suggest that the cap proteins carry out an essential structural role by bridging interactions between PH domains[15] (**Fig. 1a**, above). A second possible function of the cap proteins is to bind exosome substrates, as each contains three putative RNA binding domains[11,20,22]. Finally, a point mutation in one of the cap proteins (Csl4) disrupts the interaction between the core exosome and one of its cofactors[8]. The exosome requires many cofactors, and a third possible function for the cap proteins is to provide binding sites for these cofactors.

Although the nine subunit exosome core shows remarkable structural similarity to the bacterial PNPase and the archaeal exosome, the yeast exosome contains a tenth essential subunit (Rrp44), which is homologous to bacterial RNase II (**Fig. 1a**) and is responsible for the 3′ exoRNase activity of the exosome[21,23–27]. Notably, although Rrp44 is an essential gene, a point mutation that inactivates its exoRNase activity *in vitro* is viable[21], suggesting that either the exoRNase activity may not be a central feature of the exosome or the point-mutant protein retains partial exoribonucleolytic activity *in vivo*.

[1]Department of Microbiology and Molecular Genetics, University of Texas Health Science Center-Houston, 6431 Fannin Street, MSB 1.212 Houston, Texas 77030, USA. [2]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Apartado 127, 2781-901 Oeiras, Portugal. [3]Correspondence should be addressed to A.v.H. (ambro.van.hoof@uth.tmc.edu).

Received 6 November; accepted 13 November; published online 7 December 2008; doi:10.1038/nsmb.1528

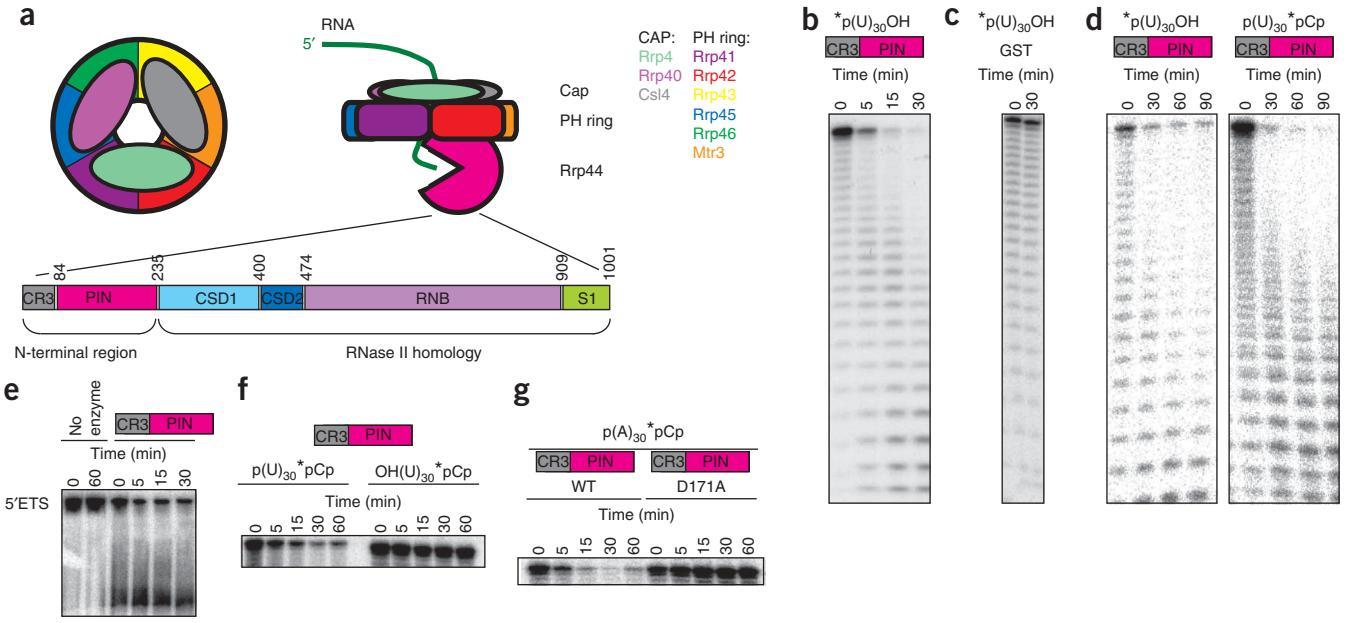


Figure 1 The N-terminal region of Rrp44 has RNase activity. (**a**) The yeast exosome contains three layers. The exosome contains a PH ring of six proteins with an RNase PH domain. Assembly and/or stability of the PH ring requires three additional proteins that form a cap on one side of the ring. The exoRNase activity of the exosome is provided by a tenth subunit (Rrp44) that associates with the bottom of the ring. RNA substrates are thought to associate with the cap proteins and pass through the central channel in the PH ring, being degraded by Rrp44 after they emerge from the bottom of the ring. Below is a schematic representation of Rrp44's domains. (**b**) The N-terminal region of Rrp44, which contains a CR3 domain and a PIN domain, has RNase activity *in vitro*. The N-terminal part of Rrp44 was purified as a GST-fusion recombinant protein from *E. coli* and incubated with 5′ end–labeled U_{30} RNA. (**c**) GST by itself, purified and incubated under identical conditions, does not have RNase activity. (**d**) The N-terminal region of Rrp44 degrades both 5′ end–labeled and 3′ end–labeled substrates to smaller labeled oligoribonucleotides, suggesting that this region acts as an endonuclease. (**e**) The N-terminal region of Rrp44 degrades an internally labeled RNA corresponding to the 5′ ETS *in vitro*. (**f**) The N-terminal region of Rrp44 prefers substrates with a 5′ phosphate (left) over those with a 5′ hydroxyl group (right). The apparent slight change in mobility in the right-hand gel is an artifact of this particular gel and does not reflect a slow degradation of the 5′ hydroxyl substrate. (**g**) Mutation of the conserved acidic amino acid residue Asp171 to alanine in the N-terminal region of Rrp44 abolishes RNase activity. The asterisks in **b**–**d** and **f**,**g** indicate the position of the P^{32} label.

The N terminus of Rrp44 contains a CR3 domain and a PIN domain (**Fig. 1a**, below). The CR3 domain is a small domain of unknown function that contains three conserved cysteine residues. PIN domains were initially thought to have some regulatory or signaling function, but recent crystal structures of PIN domains revealed a putative active site with similarity to RNase H[28,29]. Indeed, some PIN domains have endoRNase activity, including the SMG6 protein, which is involved in nonsense-mediated decay[28–33]. PIN domains contain four acidic amino acid residues (Asp91, Glu120, Asp171 and Asp198 in Rrp44), and mutations in the third residue abolish the activity of PIN domains *in vitro* and *in vivo*[30,34,35].

In this paper, we show that the Rrp44 PIN domain is an active endoRNase, revealing that the exosome has endoRNase activity in addition to its known exoRNase activity. Furthermore, we characterize the functions of the cap proteins of the yeast exosome and show that, whereas Rrp4 and Rrp40 might have a largely structural role, Csl4 does not seem to have an important role in exosome structure but, instead, seems to have specific domains that are required for specific functions.

RESULTS

The exosome contains an endoRNase domain

To test whether the Rrp44 PIN domain contains endonuclease activity, we purified a truncated Rrp44 as a glutathione *S*-transferase (GST) fusion from *Escherichia coli*. This truncated Rrp44 contains the PIN domain, but lacks the RNB domain, and has robust nuclease activity *in vitro* on several RNA substrates (**Fig. 1**), whereas GST by itself does not (**Fig. 1c**). The activity level of the truncated Rrp44 was similar to the activity level of the full-length protein (although under different conditions; **Supplementary Fig. 1** online). Mutation of the third conserved acidic residue (Asp171) in the Rrp44 PIN domain abolished the endoRNase activity (**Fig. 1g**), as has been shown for some other PIN domains either *in vitro* or *in vivo*[30,34,35]. We conclude that the PIN domain of Rrp44 is a nuclease.

Other PIN domains and the related RNases H act as endonucleases. To test whether the Rrp44 PIN acts as an endonuclease we used 5′- or 3′-labeled substrates. An exonuclease will yield only labeled mononucleotides with one of these substrates. In contrast, the PIN domain yielded a similar pattern of labeled oligonucleotides with each substrate, showing that this domain acts as an endonuclease (**Fig. 1d**). Notably, the PIN domain was dramatically more active on substrates with a 5′ monophosphate than on substrates with a 5′ hydroxyl group (**Fig. 1f**), which may have important implications (see Discussion). As expected for an endonuclease, the PIN domain also had some activity on circular U_{30} (**Supplementary Fig. 1b**). However, degradation of circular RNA was inefficient, probably because it lacks a free 5′ phosphate. All of these results indicate that the Rrp44 PIN domain is an active endoRNase, as has been shown for SMG6 (refs. 29,32).

Either RNase activity can form an active exosome

To address whether this endonuclease activity is important for exosome function, we generated truncations and point mutants of Rrp44. C-terminal truncations that removed the RNase II homology, but left

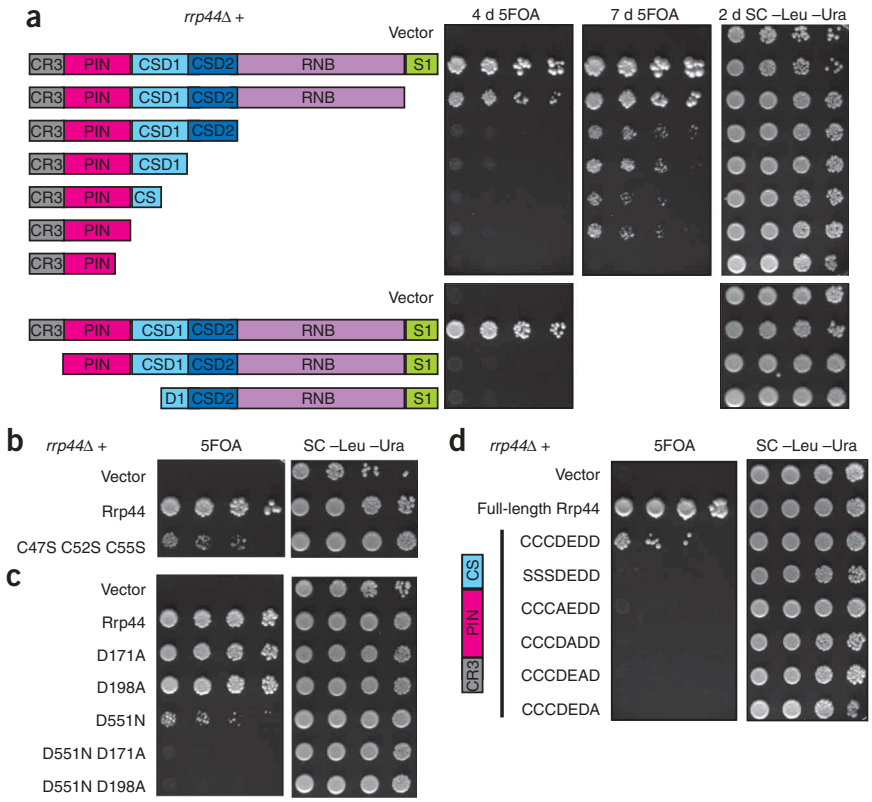


Figure 2 The PIN active site is important for exosome function. (**a**) The N-terminal region of Rrp44 is required and sufficient for viability. An *rrp44*Δ strain complemented by full-length *RRP44* on a plasmid with a *URA3* marker was transformed with *LEU2* plasmids encoding each of the depicted truncations. Growth on 5-fluoroorotic acid (5FOA) indicates that truncated Rrp44 can carry out the essential function of the exosome. The smallest complementing Rrp44 allele includes amino acids 1–235. Similar results were obtained with a *GAL::rrp44* strain (**Supplementary Fig. 2**). (**b**) Mutation to serine of the three conserved cysteine residues in the CR3 domain reduces growth. (**c**) Mutations in the PIN domain active site residues of Rrp44 do not have a large effect on growth, but they are lethal in combination with a mutation of the active-site residue of the RNB domain. Similar results were obtained for the D91A and E120A mutations (**Supplementary Fig. 3**). (**d**) Mutations of the PIN domain active-site residues are lethal in combination with a truncation that removes the RNB domain. CCDEDD indicates Cys47, Cys52, Cys55, Asp91, Glu120, Asp171 and Asp198, which where mutated in the last five rows as indicated. SC –Leu –Ura is synthetic complete media lacking leucine and uracil included as a control.

the CR3 and PIN domains, allowed for slow growth (**Fig. 2a** and **Supplementary Fig. 2** online). Previously it was shown that a point mutation inactivating the catalytic activity of the RNB domain is viable[21,36]. The residual growth we observed upon deletion of the RNB domain rules out the possibility that the mutated RNB domain was viable owing to some low-level activity *in vivo*. Instead, our results show that exoRNase activity is not a crucial feature of the exosome core. C-terminal truncations that also removed part of the PIN domain, or truncation of the N terminus, did not support any growth. The smallest viable Rrp44 allele we generated included the CR3 and PIN nuclease domains (amino acids 1–235). We conclude that the PIN domain is essential for exosome function, but we cannot exclude the alternative that the PIN domain (also) has functions separate from the exosome.

To further address the importance of the CR3 and PIN domains, we made point mutants in conserved residues. Mutating the three conserved cysteines of the CR3 domain caused slow growth (**Fig. 2b**); thus, both N-terminal truncations and point mutants indicate an important role for the CR3 domain. In contrast, the D171A point mutation, which inactivates the PIN domain (or mutations in the other three conserved acidic residues), does not have an obvious growth phenotype (**Fig. 2c** and **Supplementary Fig. 3** online). However, combining the viable PIN mutations with either the viable D551N or a viable C-terminal truncation resulted in no

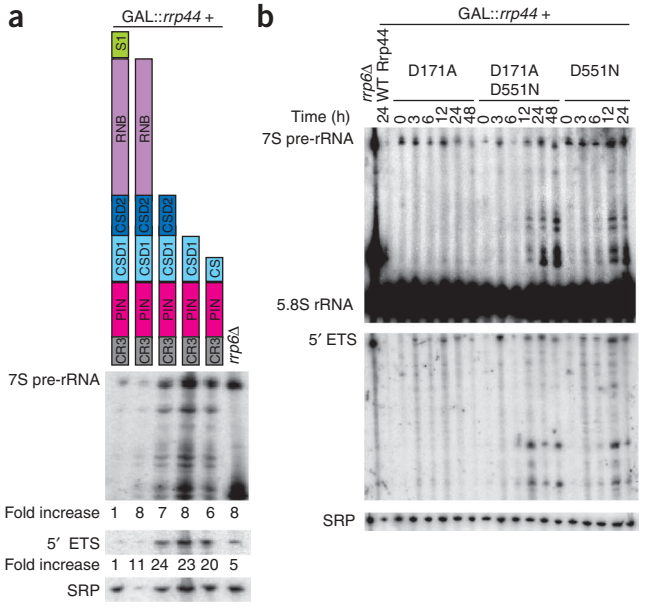


Figure 3 The RNB domain of Rrp44 is required for both RNA-processing and RNA-degradation activities of the exosome. (**a**) The indicated Rrp44 truncations were expressed in a *GAL::rrp44* yeast strain. RNA was isolated from cultures grown in dextrose (to inhibit expression of the *GAL::rrp44* gene) and analyzed by northern blotting with probes that hybridize to the 7S precursor of 5.8S rRNA (above), the 5′ ETS of the rRNA (middle) and the RNA subunit of the signal recognition particle (SRP; below). The relative levels of the 7S pre-rRNA and the 5′ ETS were normalized for loading using the SRP signal and are indicated under the above and middle blots. Shown is a representative experiment. (**b**) *RRP44* alleles containing the D171A mutation, the D551N mutation or the D171A D551N double mutation were introduced into a *GAL::rrp44* yeast strain. Cultures were first grown in media containing galactose and then incubated in media containing glucose for the indicated time (in hours) and analyzed by northern blotting with probes that hybridize to the 5.8S rRNA (above), the 5′ ETS of the rRNA (middle) and the RNA subunit of the signal recognition particle (below). Shown is a representative experiment.

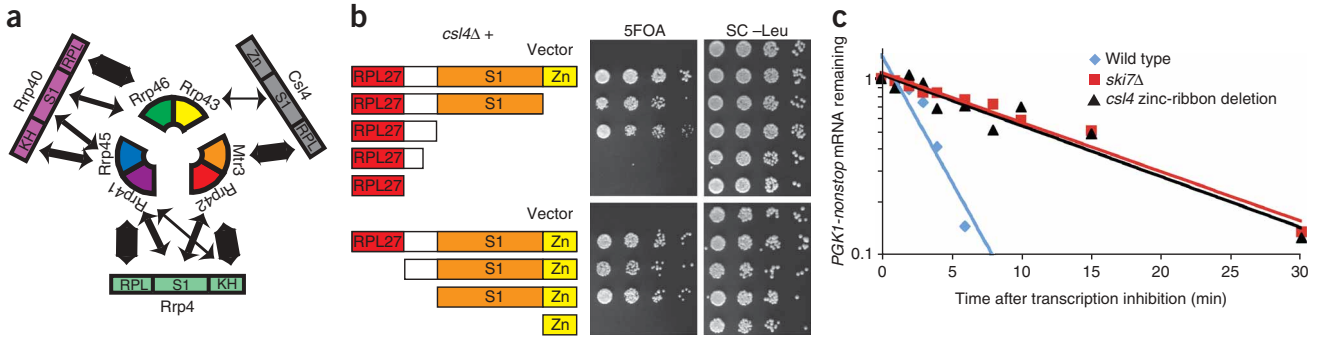


Figure 4 The essential Csl4 does not contain any essential domains, but its zinc-ribbon domain is required for cytoplasmic exosome-mediated mRNA decay. (**a**) Rrp4 and Rrp40 make extensive contacts with two neighboring subunits of the PH ring. Csl4 makes extensive contacts with Mtr3 but relatively small contacts with Rrp43. The width of the double-headed arrows is proportional to the extent of the buried surface calculated using PyMol (http://pymol.sourceforge.net). (**b**) The essential Csl4 does not contain any essential domains. Csl4 contains a RPL27-like domain (red), a 38-residue linker not present in other eukaryotes (white), an S1 domain (orange) and a zinc ribbon–like domain that is missing the zinc-coordinating residues (yellow). A *csl4*Δ strain complemented by full-length *CSL4* on a plasmid with a *URA3* marker was transformed with *LEU2* plasmids encoding the depicted truncations. Growth on 5-fluoroorotic acid (5FOA) indicates that truncated Csl4 can carry out the essential function of the exosome. Similar results were obtained with a *GAL::csl4* strain (**Supplementary Fig. 4**). (**c**) The zinc-ribbon domain of Csl4 is required for exosome-mediated mRNA decay. The decay rate of *PGK1pG-nonstop* mRNA was measured in a wild-type strain (half-life = 2 min), a strain lacking the zinc-ribbon domain of Csl4 (half-life = 10 min) and a *ski7*Δ strain (half-life = 10 min). Shown is the average value from two independent experiments. The conclusion that the zinc-ribbon domain of Csl4 is required for exosome-mediated mRNA decay was confirmed using two other approaches (**Supplementary Fig. 7**).

growth (**Fig. 2c,d** and **Supplementary Fig. 3**). These results suggest that either active site is sufficient for viability, but simultaneous inactivation of both catalytic sites results in a nonfunctional exosome.

The exosome carries out both RNA-processing and RNA-degradation reactions. Endonuclease activity might be well suited for processing reactions, but exonuclease activity would be more appropriate for degradation reactions. We therefore tested the hypothesis that the PIN domain of Rrp44 is sufficient for exosome-mediated RNA-processing reactions, whereas the RNB domain is required for exosome-mediated RNA decay. Truncations lacking the RNase II homologous domain cause the 5′ ETS and the 7S precursor of the 5.8S rRNA to accumulate (**Fig. 3a**). Similarly, 7S-processing and 5′ ETS–degradation intermediates accumulated as a result of the D551N mutation in the RNB domain. We did not observe any similar degradation or processing intermediates resulting from the D171A mutation in the PIN domain, and the effect of the combination of D551N with D171A seemed similar to that of the D551N single mutant (**Fig. 3b**). Thus, both degradation and processing functions of the exosome require the RNB domain. An alternative explanation for the presence of two catalytic sites is that the efficiency of the overall reaction is increased. Under this hypothesis, the endo- and exonuclease domains do not have separate substrates, but both may act together on most or all exosome substrates.

The essential Csl4 does not contain an essential domain

To gain further insight into the structure-function relationship of the exosome, we next focused on the function of the three cap proteins. The Csl4 subunit contains an N-terminal domain that resembles RPL27, a central S1 domain and a C-terminal domain that shows structural similarity to zinc ribbons, although the zinc-coordinating residues have been mutated. In addition, the yeast Csl4 contains a 38-residue linker that is absent in Csl4 from humans and most other eukaryotes. The first 22 residues of this inserted sequence include ten glutamate and two aspartate residues. To test which domains of Csl4 are essential, we made several truncations. Markedly, each of the domains could be removed, and either the RPL27-like or S1 plus zinc-ribbon domains are sufficient for the essential function of the exosome (**Fig. 4** and **Supplementary Fig. 4** online). Thus, the essential Csl4 subunit does not seem to have an essential domain.

The other two cap proteins, Rrp4 and Rrp40, each contain an N-terminal RPL27-like domain, a central S1 domain and a C-terminal KH domain. In contrast to Csl4, eight truncations of Rrp4 or Rrp40 each failed to complement *rrp4*Δ and *rrp40*Δ (**Supplementary Fig. 5** online). Thus, the observation that the essential Csl4 does not contain any essential domains is specific for Csl4. Furthermore, the function of Rrp4 or Rrp40 could not be provided by chimeric proteins where one domain of Rrp4 was replaced by the paralogous Rrp40 domain (or vice versa), or by coexpressing two parts of Rrp4 (**Supplementary Fig. 6** online). Overall, these results are consistent with Rrp4 and Rrp40 providing important bridging contacts for the PH-ring[15]. However, Csl4 does not seem to share this function.

A specific domain for exosome-mediated mRNA decay

Although the zinc-ribbon domain of Csl4 is not required for the essential exosome function(s), three assays each indicate that it is required for cytoplasmic exosome-mediated mRNA decay. First, *PGK1-nonstop* mRNA, which is normally rapidly degraded, is stabilized upon inactivation of the cytoplasmic exosome (for example, *ski7*Δ). The Csl4 truncation that deletes the zinc-ribbon domain results in a defect similar to *ski7*Δ (**Fig. 4c**). Second, the zinc ribbon is also required for the rapid degradation of the *his3-nonstop* mRNA (**Supplementary Fig. 7** online). Third, a role of the exosome in general mRNA decay can be monitored through synthetic lethality with the *dcp1-2* mutation, and this assay also indicates that the zinc ribbon of Csl4 is required for exosome-mediated mRNA decay (**Supplementary Fig. 7**). Therefore, the zinc-ribbon domain of Csl4 is not required for exosome structure or the essential exosome function(s), but it is required for cytoplasmic, exosome-mediated mRNA decay.

DISCUSSION

A 5′ phosphate stimulates endoRNase in the exosome

We show for the first time that Rrp44 contains a previously overlooked endoRNase domain—the PIN domain. Two observations explain

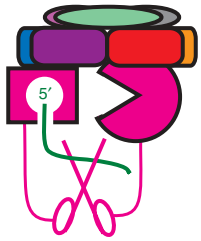


Figure 5 A model of how RNA might access the 5′ end–stimulated endonuclease activity of Rrp44.

previous reports that Rrp44 with a mutation in the RNB domain has no nuclease activity *in vitro*[21,24]. First, the previously reported assays were carried out under conditions favorable for the activity of RNB domains (that is, in Tris buffer in the presence of Mg^{2+}). Under similar conditions, we also failed to detect significant activity (**Supplementary Fig. 1**). Second, we observed robust endonuclease activity with truncated proteins, but much less activity with full-length protein or any truncations that did not remove the first cold shock domain (CSD1) present in many proteins of the RNase II family of enzymes (data not shown)[26,37–39]. This latter observation suggests that CSD1 and perhaps other domains of the exosome regulate the activity of the PIN domain. Nevertheless, the activity level of the truncated Rrp44 under conditions where PIN domains are generally active (that is, in HEPES buffer in the presence of Mn^{2+}) was similar to the activity level of the full-length protein under conditions for the RNB domain (**Supplementary Fig. 1**). This suggests that, although the activity was not previously detected, it is robust and likely to be important *in vivo* under some circumstances or on some substrates that remain to be determined.

Our observation that Rrp44 needs either its endonuclease or exonuclease activity suggests that the endonuclease is active and important *in vivo*. Furthermore, such functional redundancy seems to be a general theme for a wide variety of RNases and is similar to that found for Rrp44 and Rrp6 in yeast[21], Xrn1 and the exosome in yeast[7,40], several combinations of Rex1, Rex2, Rex3 and Rrp6 in yeast[5], and several exoRNase combinations in *E. coli*[3,41].

In the process of testing 5′- or 3′-labeled substrates in *in vitro* degradation reactions, we noticed that substrates with a 5′ phosphate are more rapidly degraded than substrates with a 5′ hydroxyl or circular substrates. This is analogous to what has been described for bacterial RNase E, which prefers a 5′ monophosphate–containing substrate over a circular RNA or a linear RNA with a 5′ hydroxyl or 5′ triphosphate[42]. Despite this functional similarity, we have not detected any sequence similarity between Rrp44 and RNase E. The 5′ end–stimulated activity of Rrp44 might have several important consequences for the substrate specificity of the endonuclease activity. First, PIN domains are part of a large group of enzymes and ribozymes that use two-metal-ion catalysis[28–30,32]. All of these enzymes generate a 5′ cleavage product with a 3′ hydroxyl group and a 3′ cleavage product with a 5′ monophosphate (reviewed in ref. 43). This cleavage mechanism suggests that the endonuclease activity of Rrp44 might cleave most primary transcripts relatively inefficiently, but the 3′ cleavage product of an initial cleavage will have a 5′ phosphate and be a preferred substrate. Such preferred subsequent cleavage might ensure that RNA degradation is rapidly completed once it has been initiated. Second, many RNA species are processed both at the 5′ and 3′ end, and in most cases it is not clear how these events are coordinated. One possibility offered by our results is that 5′ end processing generates a 5′ monophosphate, which stimulates 3′ end processing by the Rrp44 endonuclease. Testing these hypotheses in detail will require the identification of all, or most, of the endogenous substrates on the endonuclease activity of Rrp44.

Functional specialization of an exosome domain

The exosome contains ten different essential subunits containing 21 domains. How this structural complexity correlates with the multiple functions of the exosome is not clear. To begin to address these questions, we generated many deletions of domains, both in Rrp44 and in the three cap proteins. The cap proteins have been suggested to be important for exosome structure, by simultaneously interacting with two different PH-ring subunits. The finding that the RPL27-like domain and S1 domain of Csl4 are not both essential is inconsistent with Csl4 providing important bridging interactions between subunits of the PH-ring, as deletion of the S1 and zinc-ribbon domains results in Csl4 interacting with only Mtr3 (**Fig. 4a**). Conversely, deletion of the RPL27-like domain results in Csl4 interacting with only Rrp43.

A high-throughput screen previously identified transposon insertions in the essential *CSL4* gene[44]. Unexpectedly, these insertions are viable, a result that is explained by our observation that a severely truncated Csl4 protein is still functional. Similarly, T-DNA insertions into the *Arabidopsis thaliana CSL4* gene are viable[45]. This observation was interpreted to mean that Csl4 is not an essential subunit of the *A. thaliana* exosome, but our results suggest the alternative explanation that the T-DNA mutants produce a truncated Csl4 that retains some function.

A corollary of the conclusion that Csl4 is not important for exosome structure is that it must be functionally important. One possible model is that Csl4 provides important contacts with either an essential substrate RNA or an essential cofactor. If such an RNA or protein molecule makes contacts with both the RPL27-like domain and the S1 or zinc-ribbon domains, it might still be able to interact if either binding site is deleted, but not if both are deleted. Notably, some surface-exposed conserved amino acid residues of the Csl4 zinc-ribbon domain that are not essential become essential if the RPL27-like domain is deleted (**Supplementary Fig. 4**). This is consistent with these residues forming one binding site and the RPL27-like domain containing a second binding site for an exosome cofactor. Csl4 is also thought to provide the binding sites for Ski7 (ref. 8); thus, Csl4 and perhaps the other cap proteins may provide important contacts for exosome cofactors, and may thus be important for specific exosome functions.

The exosome may have multiple ways to interact with RNA

It is not clear how the exosome interacts with RNA substrates. A crystal structure of the archaeal exosome in complex with RNA[19] suggests that RNA might go through the center of the PH ring, as depicted in **Figure 1a**, above. Our conclusion that Csl4 provides important contacts for interaction with cytoplasmic cofactors seems most consistent with the suggestion that cytoplasmic mRNA follows this route. On the other hand, it has been proposed that RNA may gain access to Rrp44 without going through the PH ring[25]. The observation that the endonuclease activity is stimulated by a 5′ monophosphate indicates that Rrp44 can interact with the 5′ end of an exosome substrate. As the 3′ exoRNase active site must interact with the 3′ end, Rrp44 seems to be able to interact with both ends of substrate RNAs. This is unlikely to be compatible with threading the RNA through the PH ring. Thus, our observations seem to be more consistent with the idea that the exosome can interact with RNA

substrates in two different ways (compare **Fig. 1a**, above, to **Fig. 5**). Moreover, we suggest that substrates that access Rrp44 directly might be identifiable by looking for substrates that are processed *in vivo* in a 5′ end–dependent manner.

Convergence on combining RNases in mRNA decay machines

We show for the first time that the exosome contains both an endonuclease and exonuclease activity. Although this is unexpected for the exosome, it is not unique in nature. It has been suggested that the archaeal exosome also contains a PIN domain[46], which could provide the archaeal exosome with an endonuclease activity. Similarly, the *E. coli* degradosome combines the RNase E endoRNase with the PNPase 3′ exoRNase[47,48], and metazoan P-bodies combine the Argonaute endoRNase, the Xrn1 5′ exoRNase and the Ccr4, Caf1 and Pan2 3′ exoRNases[49–54]. Thus, the combination of endoRNases and exoRNases into one RNA-degrading machine is widespread in nature and has evolved multiple times independently, suggesting that such combination offers a fundamental advantage to the cell.

Note added in proof: After this manuscript was submitted, endonuclease activity of the Rrp44 PIN domain was also reported by others[60].

METHODS

Strains. The strains, plasmid and oligonucleotides we used are described in detail in **Supplementary Tables 1–3** online. The heterozygous diploids *RRP44*/*rrp44*Δ, *CSL4*/*csl4*Δ, *RRP4*/*rrp4*Δ and *RRP40*/*rrp40*Δ have been described and were obtained from Open Biosystems[55]. We transformed each of these strains with a *URA3* plasmid that encoded the corresponding wild-type exosome subunit. We sporulated the transformants and obtained haploid progeny spores by the hydrophobic spore-isolation method essentially as described[56]. Isolated spores were germinated to generate the deletion strains complemented with the *URA3* plasmid. The *GAL::rrp44*, *GAL::csl4*, *GAL::rrp4* and *GAL::rrp40* strains have been described previously and were obtained from P. Mitchell (University of Sheffield) and D. Tollervey (University of Edinburgh)[10,11].

Plasmids. All yeast plasmids used are low-copy plasmids that contain a yeast centromere and the endogenous promoter and 3′-flanking region to control expression of the exosome subunits. All were verified by sequencing.

To generate C-terminal truncations, we initially PCR amplified several hundred base pairs of the 3′-flanking region of the gene of interest and cloned it into pRS415 (ref. 57), which contains a *LEU2* marker. We then PCR amplified the promoter and the desired part of the coding region and cloned it into the plasmid described above. N-terminal truncations were generated similarly, except that the promoter was PCR amplified and cloned first, and the truncated coding region and 3′-flanking region were added in the second step.

We generated GST-fusion plasmids by PCR amplifying the *RRP44* coding region from the plasmids described above, digesting the PCR product with EcoRI and XhoI and ligating it into pGEX-4T-1 (GE Healthcare).

We created site-directed mutants using the QuikChange II kit (Stratagene). Each oligonucleotide also changes a restriction site. We screened putative mutants by restriction digestion and verified them by sequencing. To combine point mutations in the PIN domain of Rrp44 with the D551N mutant in the RNB domain, we digested the plasmids containing each single mutant with SacI and PshAI and ligated them together. PshAI cuts at a unique site within the CSD2 domain.

Overexpression and purification of recombinant Rrp44. Full-length and truncated Rrp44 fused to GST were expressed in the Rosetta(DE3) *E. coli* strain. We grew cultures at 30 °C in 2 liters TB medium supplemented with 200 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol to an optical density at 600 nm (OD_{600}) of 1.2 and then induced expression by addition of 0.2 mM IPTG and incubation at 18 °C overnight. We harvested cells by centrifugation and froze them. Cells were thawed on ice, resuspended in 30 ml PBS buffer, pH 8, with 10 mM DTT and 1 mM PMSF, and lysed using the French press. We treated the crude extracts with 5 units Benzonase (Sigma) for 30 min at 0 °C, clarified the extract by a 30-min centrifugation at 10,000*g*, and loaded it onto a GSTrap™ FF 1-ml column (GE healthcare) equilibrated in binding buffer. We eluted proteins with a 0–10 mM glutathione gradient in 50 mM Tris-HCl, pH8.5, 10 mM DTT. We pooled fractions containing the purified protein and loaded them onto a gel-filtration column (Superdex 75 10/300 GL or a Superose 12 10/300 GL, GE Healthcare, depending on the protein size) in 20 mM Tris-HCl, 150 mM NaCl, 5 mM $MgCl_2$ and 10 mM DTT buffer, pH 8. We concentrated the eluted protein by centrifugation at 15 °C with Vivaspin 500 Centrifugal Concentrators (Vivaspin). We determined protein concentrations by spectrophotometry using a Nanodrop and added 50% (v/v) glycerol to the final fractions before storage at –20 °C.

RNase assays. RNase assays were based on those previously described. Briefly, each assay contained 100–500 nM protein, 200 nM 5′ or 3′ end–labeled U_{30} or A_{30} RNA or an internally labeled 5′ ETS RNA transcribed by T7 RNA polymerase, and 20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM $MnCl_2$, 1 mM DTT.

Growth assays. We performed growth assays essentially as previously described[9,58]. Briefly, we grew strains in appropriate liquid selection media overnight at 30 °C (or 23 °C for temperature-sensitive mutants). We diluted these cultures in appropriate selection media to a starting OD_{600} of 0.2. We grew the cultures until they reached an OD of 0.8, serially diluted them in 96-well plates by a factor of five, and spotted them onto appropriate media. We generally incubated these plates for 2–3 d at 23 °C, 30 °C or 37 °C. To monitor the slow growth of Rrp44 truncations, we incubated the plates for up to 20 d.

Northern blotting. Stability of *PGK1-nonstop* mRNA[59], and 5.8S rRNA processing and 5′ ETS degradation[59], were assayed as previously described.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We are grateful to T. Link and R. Brennan for help with calculating the buried surfaces between the different domains of the cap proteins and the PH ring, and A. Klauer for technical assistance. *GAL::rrp44*, *GAL::csl4*, *GAL::rrp4* and *GAL::rrp40* strains were kindly provided by P. Mitchell (University of Sheffield) and D. Tollervey (University of Edinburgh). R. Parker, M. Wilkinson, M. Steiger and members of the van Hoof and Arraiano laboratories gave insightful comments on the manuscript. This research was supported by the Pew Scholarship Program in the Biomedical Sciences and by the National Institutes of Health (GM069900) to A.v.H. E.G.D and M.S.-R. were supported by The University of Texas at Houston Medical School-Summer Research Program. The work at the Instituto de Tecnologia Química e Biológica was supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal. A.B. was a recipient of a post-doctoral fellowship from FCT, Portugal.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Moser, M.J., Holley, W.R., Chatterjee, A. & Mian, I.S. The proofreading domain of *Escherichia coli* DNA polymerase I and other DNA and/or RNA exonuclease domains. *Nucleic Acids Res.* **25**, 5110–5118 (1997).
2. Mian, I.S. Comparative sequence analysis of ribonucleases HII, III, II PH and D. *Nucleic Acids Res.* **25**, 3187–3195 (1997).
3. Deutscher, M.P. & Li, Z. Exoribonucleases and their multiple roles in RNA metabolism. *Prog. Nucleic Acid Res. Mol. Biol.* **66**, 67–105 (2001).
4. Zuo, Y. & Deutscher, M.P. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res.* **29**, 1017–1026 (2001).
5. 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.* **19**, 1357–1365 (2000).
6. Muhlrad, D., Decker, C.J. & Parker, R. Turnover mechanisms of the stable yeast PGK1 mRNA. *Mol. Cell. Biol.* **15**, 2145–2156 (1995).
7. Jacobs Anderson, J.S. & Parker, R. The 3′ to 5′ degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3′ to 5′ exonucleases of the exosome complex. *EMBO J.* **17**, 1497–1506 (1998).
8. van Hoof, A., Frischmeyer, P.A., Dietz, H.C. & Parker, R. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* **295**, 2262–2264 (2002).

9. Meaux, S. & Van Hoof, A. Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA* **12**, 1323–1337 (2006).
10. Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M. & Tollervey, D. The exosome: a conserved eukaryotic RNA processing complex containing multiple 3′→5′ exoribonucleases. *Cell* **91**, 457–466 (1997).
11. Allmang, C. *et al.* The yeast exosome and human PM-Scl are related complexes of 3′ → 5′ exonucleases. *Genes Dev.* **13**, 2148–2158 (1999).
12. Allmang, C. *et al.* Functions of the exosome in rRNA, snoRNA and snRNA synthesis. *EMBO J.* **18**, 5399–5410 (1999).
13. van Hoof, A., Lennertz, P. & Parker, R. Yeast exosome mutants accumulate 3′-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs. *Mol. Cell. Biol.* **20**, 441–452 (2000).
14. de la Cruz, J., Kressler, D., Tollervey, D. & Linder, P. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3′ end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J.* **17**, 1128–1140 (1998).
15. Liu, Q., Greimann, J.C. & Lima, C.D. Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**, 1223–1237 (2006).
16. Lorentzen, E. *et al.* The archaeal exosome core is a hexameric ring structure with three catalytic subunits. *Nat. Struct. Mol. Biol.* **12**, 575–581 (2005).
17. Symmons, M.F., Jones, G.H. & Luisi, B.F. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. *Structure* **8**, 1215–1226 (2000).
18. Navarro, M.V., Oliveira, C.C., Zanchin, N.I. & Guimaraes, B.G. Insights into the mechanism of progressive RNA degradation by the archaeal exosome. *J. Biol. Chem.* **283**, 14120–14131 (2008).
19. Lorentzen, E. & Conti, E. Structural basis of 3′ end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol. Cell* **20**, 473–481 (2005).
20. Buttner, K., Wenig, K. & Hopfner, K.P. Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol. Cell* **20**, 461–471 (2005).
21. Dziembowski, A., Lorentzen, E., Conti, E. & Seraphin, B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.* **14**, 15–22 (2007).
22. Chekanova, J.A., Dutko, J.A., Mian, I.S. & Belostotsky, D.A. *Arabidopsis thaliana* exosome subunit AtRrp4p is a hydrolytic 3′→5′ exonuclease containing S1 and KH RNA-binding domains. *Nucleic Acids Res.* **30**, 695–700 (2002).
23. Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A. & Conti, E. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell* **29**, 717–728 (2008).
24. Schneider, C., Anderson, J.T. & Tollervey, D. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol. Cell* **27**, 324–331 (2007).
25. Wang, H.W. *et al.* Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3′ end processing. *Proc. Natl. Acad. Sci. USA* **104**, 16844–16849 (2007).
26. Frazão, C. *et al.* Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* **443**, 110–114 (2006).
27. Barbas, A. *et al.* New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end product. *J. Biol. Chem.* **283**, 13070–13076 (2008).
28. Arcus, V.L., Backbro, K., Roos, A., Daniel, E.L. & Baker, E.N. Distant structural homology leads to the functional characterization of an archaeal PIN domain as an exonuclease. *J. Biol. Chem.* **279**, 16471–16478 (2004).
29. Levin, I. *et al.* Crystal structure of a PIN (PilT N-terminus) domain (AF0591) from *Archaeoglobus fulgidus* at 1.90 resolution. *Proteins* **56**, 404–408 (2004).
30. Glavan, F., Behm-Ansmant, I., Izaurralde, E. & Conti, E. Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J.* **25**, 5117–5125 (2006).
31. Daines, D.A., Wu, M.H. & Yuan, S.Y. VapC-1 of nontypeable *Haemophilus influenzae* is a ribonuclease. *J. Bacteriol.* **189**, 5041–5048 (2007).
32. Bunker, R.D., McKenzie, J.L., Baker, E.N. & Arcus, V.L. Crystal structure of PAE0151 from *Pyrobaculum aerophilum*, a PIN-domain (VapC) protein from a toxin-antitoxin operon. *Proteins* **72**, 510–518 (2008).
33. Eberle, A.B., Lykke-Andersen, S., Muhlemann, O. & Jensen, T.H. SMG6 promoted endonucleoytic cleavage of nonsense mRNA in human cells. *Nat. Struct. Mol. Biol.* advance online publication, doi:10.1038/nsmb.1530 (07 December 2008).
34. Fatica, A., Tollervey, D. & Dlakic, M. PIN domain of Nob1p is required for D-site cleavage in 20S pre-rRNA. *RNA* **10**, 1698–1701 (2004).
35. Bleichert, F., Granneman, S., Osheim, Y.N., Beyer, A.L. & Baserga, S.J. The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci. USA* **103**, 9464–9469 (2006).
36. Amblar, M. & Arraiano, C.M. A single mutation in *Escherichia coli* ribonuclease II inactivates the enzyme without affecting RNA binding. *FEBS J.* **272**, 363–374 (2005).
37. Amblar, M., Barbas, A., Fialho, A.M. & Arraiano, C.M. Characterization of the functional domains of *Escherichia coli* RNase II. *J. Mol. Biol.* **360**, 921–933 (2006).
38. Barbas, A. *et al.* New insights into the mechanism of RNA degradation by ribonuclease II: identification of the residue responsible for setting the RNase II end product. *J. Biol. Chem.* **283**, 13070–13076 (2008).
39. Cairrao, F., Arraiano, C. & Newbury, S. *Drosophila* gene *tazman*, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev. Dyn.* **232**, 733–737 (2005).
40. Johnson, A.W. & Kolodner, R.D. Synthetic lethality of *sep1* (*xrn1*) *ski2* and *sep1* (*xrn1*) *ski3* mutants of *Saccharomyces cerevisiae* is independent of killer virus and suggests a general role for these genes in translation control. *Mol. Cell. Biol.* **15**, 2719–2727 (1995).
41. Andrade, J.M., Pobre, V., Silva, I.J., Domingues, S. & Arraiano, C.M. The role of 3′ to 5′ exonucleases in RNA degradation. *Prog. Nucleic Acid Res. Mol. Biol.* (in the press).
42. Mackie, G.A. Ribonuclease E is a 5′-end-dependent endonuclease. *Nature* **395**, 720–723 (1998).
43. Yang, W. An equivalent metal ion in one- and two-metal-ion catalysis. *Nat. Struct. Mol. Biol.* **15**, 1228–1231 (2008).
44. Ross-Macdonald, P. *et al.* Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* **402**, 413–418 (1999).
45. Chekanova, J.A. *et al.* Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the *Arabidopsis* transcriptome. *Cell* **131**, 1340–1353 (2007).
46. Koonin, E.V., Wolf, Y.I. & Aravind, L. Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. *Genome Res.* **11**, 240–252 (2001).
47. Carpousis, A.J., Van Houwe, G., Ehretsmann, C. & Krisch, H.M. Copurification of *E. coli* RNAase E and PNPase: evidence for a specific association between two enzymes important in RNA processing and degradation. *Cell* **76**, 889–900 (1994).
48. Py, B., Causton, H., Mudd, E.A. & Higgins, C.F. A protein complex mediating mRNA degradation in *Escherichia coli*. *Mol. Microbiol.* **14**, 717–729 (1994).
49. Liu, J., Valencia-Sanchez, M.A., Hannon, G.J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat. Cell Biol.* **7**, 719–723 (2005).
50. Ingelfinger, D., Arndt-Jovin, D.J., Luhrmann, R. & Achsel, T. The human LSm1–7 proteins colocalize with the mRNA-degrading enzymes Dcp1/2 and Xrnl in distinct cytoplasmic foci. *RNA* **8**, 1489–1501 (2002).
51. Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M. & Heyer, W.D. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *J. Cell Biol.* **136**, 761–773 (1997).
52. Cougot, N., Babajko, S. & Seraphin, B. Cytoplasmic foci are sites of mRNA decay in human cells. *J. Cell Biol.* **165**, 31–40 (2004).
53. Zheng, D. *et al.* Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *J. Cell Biol.* **182**, 89–101 (2008).
54. Andrei, M.A. *et al.* A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* **11**, 717–727 (2005).
55. Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391 (2002).
56. Rockmill, B., Lambie, E.J. & Roeder, G.S. Spore enrichment. *Methods Enzymol.* **194**, 146–149 (1991).
57. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
58. Wilson, M.A., Meaux, S. & van Hoof, A. A genomic screen in yeast reveals novel aspects of nonstop mRNA metabolism. *Genetics* **177**, 773–784 (2007).
59. van Hoof, A., Staples, R.R., Baker, R.E. & Parker, R. Function of the ski4p (Csl4p) and Ski7p proteins in 3′-to-5′ degradation of mRNA. *Mol. Cell. Biol.* **20**, 8230–8243 (2000).
60. Lebreton, A., Tomecki, R., Dziembowski, A. & Séraphin, B. Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* advance online publication, doi:10.1038/nature07480 (7 December 2008).

9298–9307 *Nucleic Acids Research, 2012, Vol. 40, No. 18* *Published online 24 July 2012*
doi:10.1093/nar/gks693

The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function

Daneen Schaeffer[1], Filipa Pereira Reis[2], Sean J. Johnson[3], Cecília Maria Arraiano[2] and Ambro van Hoof[1,*]

[1]Department of Microbiology and Molecular Genetics, University of Texas Health Science Center-Houston, Houston, TX 77030, USA, [2]Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República, 2780-157 Oeiras, Portugal and [3]Department of Chemistry and Biochemistry, Utah State University, Logan, UT, 84322 USA

Received May 15, 2012; Revised June 19, 2012; Accepted June 25, 2012

ABSTRACT

The 10-subunit RNA exosome is involved in a large number of diverse RNA processing and degradation events in eukaryotes. These reactions are carried out by the single catalytic subunit, Rrp44p/Dis3p, which is composed of three parts that are conserved throughout eukaryotes. The exosome is named for the 3′ to 5′ exoribonuclease activity provided by a large C-terminal region of the Rrp44p subunit that resembles other exoribonucleases. Rrp44p also contains an endoribonuclease domain. Finally, the very N-terminus of Rrp44p contains three Cys residues (CR3 motif) that are conserved in many eukaryotes but have no known function. These three conserved Cys residues cluster with a previously unrecognized conserved His residue in what resembles a metal-ion-binding site. Genetic and biochemical data show that this CR3 motif affects both endo- and exonuclease activity *in vivo* and both the nuclear and cytoplasmic exosome, as well as the ability of Rrp44p to associate with the other exosome subunits. These data provide the first direct evidence that the exosome-Rrp44p interaction is functionally important and also provides a molecular explanation for the functional defects when the conserved Cys residues are mutated.

INTRODUCTION

The RNA exosome is involved in a wide variety of RNA processing and degradation reactions in both the nucleus and the cytoplasm. First, the nuclear exosome processes a subset of RNAs from longer precursors. For example, it processes a 300-nt 7S precursor into the 160-nt 5.8S rRNA (1). Second, the nuclear exosome completely degrades some RNAs that are byproducts of gene expression, including the 5′-external transcribed spacer that is part of the rRNA precursor (2). Third, the nuclear exosome degrades aberrant RNAs that fail to complete proper processing, including incompletely modified initiator tRNA (3). Fourth, the exosome is involved in one of two general pathways of cytoplasmic mRNA decay (4). Fifth, the cytoplasmic exosome is especially important for degrading aberrant mRNAs, including those that lack a stop codon (nonstop mRNAs) as well as those that are cleaved by a ribozyme (5,6).

Although the exosome has a wide variety of substrates, it acts very specifically on those substrates. For example, the exosome degrades initiator tRNA lacking a single methyl group, but not normal initiator tRNA or other tRNAs that lack a modification (3,7). A second example of the exosome's specificity is that the exosome degrades both the poly(A) tail and the body of nonstop mRNAs (6), while the poly(A) tail of normal mRNAs can only be removed by dedicated deadenylases and not by the exosome (8). What is not yet known is how the exosome carries out these diverse functions while maintaining specificity.

A series of X-ray crystallography and EM studies have resolved the structural organization of the yeast exosome (9–15). The exosome contains a core of ten proteins that are shared between the nuclear and cytoplasmic exosome. At least in fungi and metazoans, only the Rrp44p subunit (also known as Dis3p) is catalytically active (9,16). Rrp44p was initially identified as an exonuclease with similarity to the RNase II family (1). The similarity to RNase II includes the catalytic RNB domain and three OB-fold RNA-binding domains (CSD1, CSD2 and S1 (17)). We and others have shown that this exonuclease activity is not essential for viability because Rrp44p

*To whom correspondence should be addressed. Tel: +1 713 500 5234; Fax: +1 713 500 5499; Email: ambro.van.hoof@uth.tmc.edu
Present address:
Daneen Schaeffer, Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA.

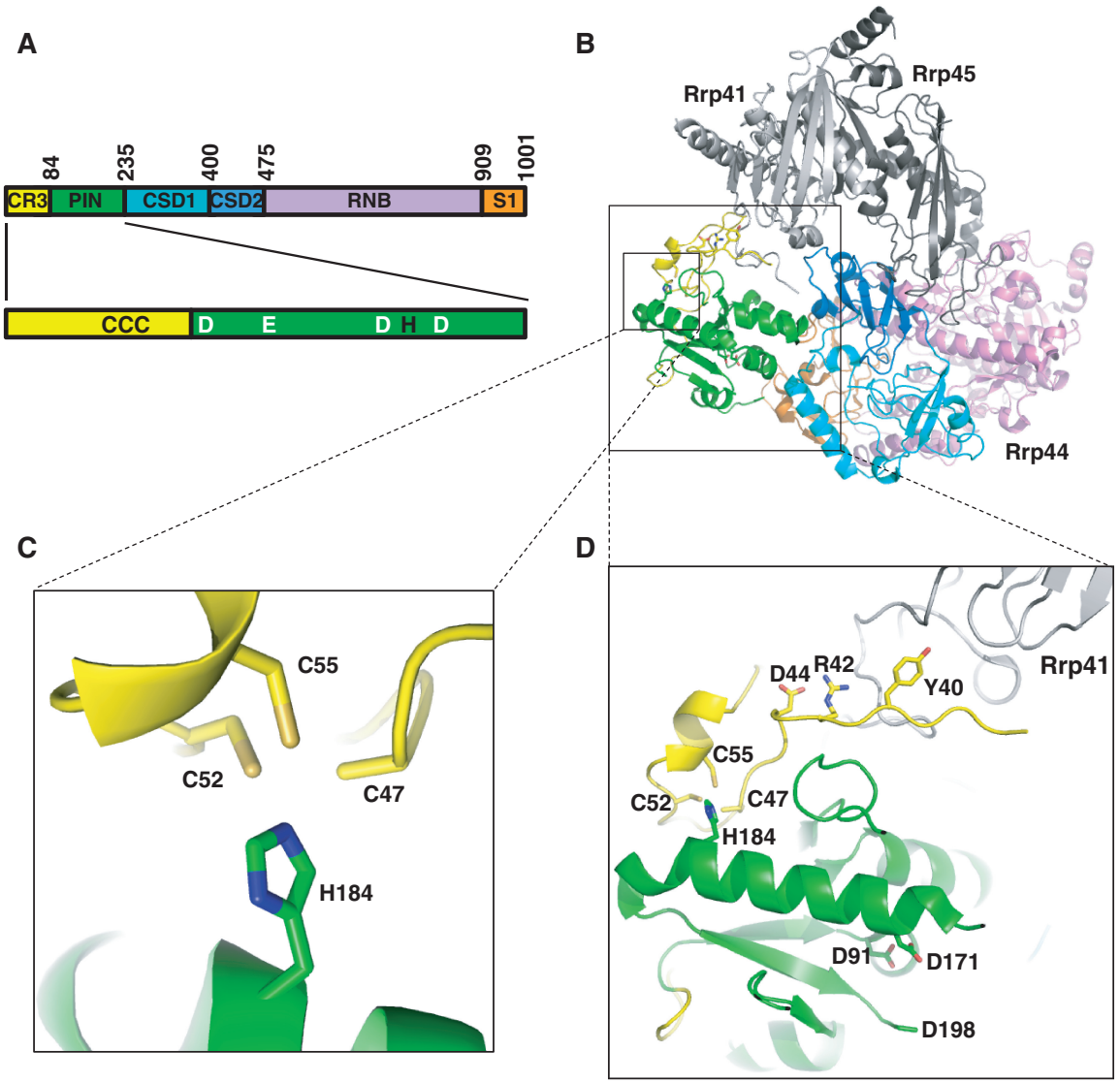


Figure 1. Rrp44 contains a conserved CCCH motif. (**A**) Top: Diagram depicting the five recognized domains in Rrp44 and the CR3 motif. PIN denotes the endonuclease domain, CSD1 and CSD2 denote RNA-binding cold shock domains, RNB denotes the exonuclease domain and S1 denotes an RNA-binding S1 domain. Collectively the CSD1, CSD2, RNB and S1 domains are responsible for exonuclease activity. Bottom: the N-terminus of Rrp44p contains a conserved CCCH motif (black letters) in addition to the catalytic residues of the PIN domain (white letters). (**B**) Overview of the Rrp44 structure (colored as in panel A) bound to exosome subunits Rrp41 and Rrp45 (PDB ID 2WP8; 13). (**C**) The three conserved Cys residues and the conserved His residue form a tetrahedral cluster in the crystal structure. Note that the sulfur atom of Cys47 was not modeled. (**D**) The CR3 motif is physically connected to the exosome-binding site (Y40, R42 and D44) and the endonuclease active site of the PIN domain (D91, D171 and D198). Note that the D198 side chain was not modeled. Structures were rendered using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC.).

contains a second domain in its N-terminus (PIN) with endonuclease activity (18–20). Mutations that inactivate each of the nuclease activities individually (hereafter referred to as *rrp44–endo*$^-$ and *rrp44–exo*$^-$) are viable, but simultaneously mutating both (*rrp44–endo*$^-$*exo*$^-$) is lethal, suggesting a functional overlap (18–20). Although the structure containing nine exosome subunits is thought to be conserved in other eukaryotes, whether or not Rrp44p in other eukaryotes stably associates with the other nine subunits is controversial.

The five recognized domains of Rrp44p explain much of the sequence conservation in Rrp44p orthologs. A notable exception is a conserved motif near the N-terminus that has no known function (Figure 1A). This motif has been named CR3 (Cysteine-Rich with three cysteines (21)). In the process of trying to understand how the exosome structure relates to its diverse functions, we initially constructed a mutant Rrp44p in which all three Cys residues were changed to Ser. This *rrp44–CR3* allele caused slow growth (19), and severely reduced the exosome's ability to degrade nonstop mRNAs and ribozyme cleaved mRNAs (22). Interestingly, neither the *rrp44–endo*$^-$ nor the *rrp44–exo*$^-$ mutation caused a similar defect in mRNA decay, suggesting that the *rrp44–CR3* mutation somehow

affected both catalytic activities. The protein encoded by *rrp44–CR3* accumulated to slightly reduced levels (22). However, this is probably not the cause of the observed phenotypes since overexpression of the *rrp44–CR3* allele from a high copy plasmid did not restore growth. In addition, haplo-insufficiency of *RRP44* did not reduce growth similar to that of *rrp44–CR3* (data not shown). All of these data can be explained if the *rrp44–CR3* mutation reduces both exo- and endonuclease activities of the exosome, thus causing the slow growth and stabilization of nonstop mRNAs. However, the observation that *rrp44–CR3* is viable suggests that it does not completely eliminate catalytic activity.

Recently, the x-ray crystal structure of Rrp44p bound to two other exosome subunits was published, providing the first structural insight into the CR3 motif (13). We noted that the CR3 motif is close to a His residue and that these four residues are oriented in a tetrahedral configuration that is reminiscent of a mono-metal-binding site. Here, we present genetic, biochemical and phylogenetic data that offer a coherent explanation of how the CR3 motif alters both the endonuclease and exonuclease activities of Rrp44p *in vivo*: the direct effect of the CR3 mutation is likely a modest perturbation of the local folding and/or structural stability of Rrp44p, possibly because of the loss of a metal ion. Even though the CR3 motif is distant from both catalytic sites, this folding defect has two important consequences. First, the CR3 motif is close to residues contacting another exosome subunit, and alterations in the CR3 motif reduce copurification of the other exosome subunits. This, in turn, can explain the defect in exonuclease function *in vivo*. Second, changes in the structure of the CR3 motif appear to cause modest changes in the endonuclease active site, possibly because the two sites are connected by a single α-helix.

MATERIALS AND METHODS

Strains and plasmids

The majority of the experiments used a previously described *rrp44* deletion strain with a *RRP44, URA3* plasmid (yAV1115; (19)). Various alleles of *RRP44* were introduced using the plasmid shuffle as previously described (19). All of these plasmids are low copy plasmids, and Rrp44p is expressed under control of its native promoter and 3′-flanking sequence. The plasmids used are listed in Supplementary Table S1. The *dcp1–2 rrp44*Δ strain (yAV1143) was previously described (22). To test for synthetic lethality between *rrp6Δ* and *RRP44* alleles, the heterozygous diploid (*RRP44/rrp44Δ::KAN;* 23) obtained from Open Biosystems was transformed with the wild-type *RRP44* plasmid, pAV361 (19). The diploid was then sporulated to generate a haploid *rrp44Δ::KAN* strain with pAV361. This strain was converted to *rrp44Δ::HYG* by homologous recombination with pAG32 (24) and selecting for hygromycin resistance. The *rrp44Δ::HYG* strain with pAV361 was then crossed with the *rrp6Δ::KAN* strain from the yeast knock out collection (23). The resulting diploid was sporulated resulting in yAV1137 (*MATa, leu2-Δ0, ura3-Δ0, his3-Δ1, rrp44Δ::HYG, rrp6Δ::NEO* with a *RRP44, URA3* plasmid).

Growth assays

Synthetic lethality growth assays were performed essentially as described (25). Briefly, plasmids encoding the *rrp44* point mutations were transformed into the *rrp44Δ* strain, the *dcp1–2 rrp44Δ* strain or the *rrp6Δ rrp44Δ* strain containing a *URA3* plasmid encoding wild-type Rrp44p. Transformants were grown in SC–URA–LEU overnight at 23°C. Cultures were diluted in the same selection media to an optical density of 0.8 at 600 nm. Next, cells were serially diluted in 96-well plates by a factor of five and spotted onto media containing 5-Fluoroorotic Acid (5-FOA). 5-FOA selects for cells that have lost the URA3 plasmid, and therefore growth indicates that the gene on the LEU2 plasmid is sufficient for viability. Plates were incubated at 23°C, 30°C and 37°C for 2–5 days. The same effects were seen at all temperatures unless otherwise noted in the results section.

The *his3*-nonstop growth assay was performed essentially as described (6). Briefly, *rrp44* deletion strains containing a *LEU2* plasmid encoding the *rrp44* point mutations were transformed with a *URA3* plasmid encoding a *his3-nonstop* reporter. This reporter contains a 1-bp deletion which removes the stop codon from the *HIS3* mRNA. Transformants were grown in SC–URA–LEU overnight at 30°C. Cultures were diluted in the same selection media to an optical density of 0.8 at 600 nm. Next, cells were serially diluted in 96-well plates by a factor of five and spotted onto media lacking histidine, SC–HIS–URA–LEU. The plates were incubated at 30°C for 2–6 days.

Stability of *GAL7* mRNA and northern blotting

To determine the half-life of the *GAL7* mRNA, duplicate cultures of the *dcp1–2* yeast strains were grown in YEP + 2% galactose overnight at 23°C. Cultures were diluted in the same media to an optical density of 0.2 at 600 nm and grown to an optical density of 0.8. Cultures were then shifted to 37°C (a non-permissive temperature for *dcp1–2*) for 1 h. The medium was replaced with 1/10th volume YEP + 4% glucose. At the time points indicated, 2 ml aliquots were pelleted. RNA was isolated and subjected to formaldehyde agarose gel electrophoresis, followed by northern blot analysis essentially as described (5,26). Northern blots were hybridized with ^{32}P 5′-end-labeled oligonucleotides specific for *GAL7* (oAV267; 5′-CTCTTGCTTCTCTGGAGAGATCGTCAGTC) and the RNA subunit of the signal recognition particle, *SCR1* (oRP100; 5′-GTCTAGCCGCGAGGAAGG). Signals were detected and quantitated using a STORM PhosphoImager (Amersham), and corrected for loading by quantitating *SCR1*.

RNase assays on recombinant Rrp44p

Truncated Rrp44p fused to GST was expressed in *Escherichia coli*, purified, and assayed as previously described (19). Briefly, the Rosetta(DE3) *E. coli* strain

with Rrp44 expression plasmids was grown at 30°C to an optical density of 1.2 at 600 nm. Expression was induced with 0.2 mM IPTG at 18°C overnight. Rrp44 was purified over a GSTtrap FF 1-ml column and a HiTrap™Q FF column (GE Healthcare). RNase assays used a 5′-end-labeled 30-mer RNA (either A_{30} or CCCGACACCACC $CACUA_{14}$), and 20 mM HEPES, pH 7.5, 150 mM NaCl, 3 mM $MnCl_2$, 1 mM DTT.

Rrp44p exosome copurification

The Rrp44p–TAP and Rrp44p–CR3–TAP plasmids were previously described (22). Cultures were grown in 50 ml media to an optical density of 0.8 at 600 nm, pelleted, and resuspended in 250 µl IP50 buffer [50 mM NaCl, 50 mM Tris–HCl pH 7.5, 2 mM $MgCl_2$, 0.1% Triton X100, 0.5 mM β-mercaptoethanol, 0.1 mM PMSF, with complete EDTA-free protease inhibitors (Roche)]. 50 µl glass beads (Sigma) were added to the cell suspension and agitated in a vortex mixer for 7 min at 4°C. Lysates were clarified by centrifugation at 4500*g* for 7 min at 4°C.

TAP-tagged proteins were purified by adding 10 µl IgG Sepharose 6 Fast Flow beads (Amersham Biosciences) to 200 µl whole-cell lysate. After incubation for 1 h at 4°C, the supernatant was discarded and the beads were washed two times with 200 µl IP50 buffer, and then twice more with 200 µl IP1000 buffer (identical to IP50, but 1 M NaCl). Bound proteins were eluted from beads by heating to 95°C for 3 min in protein sample buffer and analyzed by western blotting using anti-protein A antibodies (Sigma), anti-Rrp4p antibodies (27) and anti-Pgk1p antibodies (Invitrogen).

RESULTS

Three conserved cysteines in Rrp44p cluster with a conserved histidine

In addition to its five recognizable domains, Rrp44p orthologs contain a conserved motif that is found N-terminal to the PIN domain (Figure 1A). We have previously shown that deleting as many as 766 residues from the C-terminus of Rrp44p can be tolerated, but that deleting as little as 84 residues from the N-terminus results in lethality, suggesting that this conserved N-terminal region is important (19). This region was named CR3 (cysteine-rich domain with three cysteines) because it contains a Cx_4Cx_2C motif that does not resemble any other known domain or motif (21). A 3.0 Å crystal structure of Rrp44p was recently solved, providing the first structural insight into this motif (13). The structure reveals that the three conserved Cys residues are physically adjacent to each other and that H_{184} is in close proximity to the three conserved Cys residues (Figure 1C).

To examine whether H_{184} was conserved and co-occurred with the CR3 motif, we performed a multiple sequence alignment of 20 Rrp44p homologs from very diverse eukaryotes using ClustalW. The alignment showed that both the CR3 motif and H_{184} are conserved in 17 of the 20 included proteins (Supplementary Figure S1A). Strikingly, two of the Rrp44p orthologs (from *Ciona intestinalis* and *Trichomonas vaginalis*) lack two or all three of the conserved Cys residues of the CR3 motif and also lack H_{184}. The 20th protein, hDis3L1 (and Dis3L1 from other vertebrates), contains His_{184} together with a variation on the CR3 motif with five residues separating the second and third Cys instead of just two. Thus, His_{184} is conserved in 18 Rrp44p homologs that have the CR3 motif, but not in the two homologs that lack the CR3 motif. H_{184} is part of the PIN endonuclease domain, but as shown in Supplementary Figure S1B, it is not present in PIN domains that are not part of Rrp44p orthologs. Thus, the previously identified CR3 co-occurs and is in close proximity to a conserved histidine. Conserved CCCH motifs in other proteins form binding sites for a metal ion (generally Zn^{2+}). We note, however that the organization in the Rrp44p amino acid sequence and protein fold is distinct from known Zn-binding domains explaining why it previously went unrecognized (See 'Discussion' section).

A triple mutation in the CR3 motif affects growth, but single point mutations do not have the same effect

We have previously shown that mutating the three conserved cysteine residues of the CR3 motif to serine (C47S, C52S, C55S; hereafter referred to as *rrp44–CR3*) resulted in a slow growth phenotype (19). One possibility is that two of the conserved Cys residues form a disulfide bond, which is suggested by the proximity of the modeled sulfur atoms in the Rrp44p crystal structure. If this growth defect was caused by the disruption of a disulfide bond between two of the cysteines, then single point mutations in the two residues forming the bond should have the same effect, whereas the non-interacting Cys would not be expected to affect the disulfide bond, and thus should not affect growth. If, however, the CR3 motif forms a metal-binding site together with H_{184}, any one of the single mutations might decrease affinity for the metal and a triple mutant might further decrease metal affinity. Thus, to distinguish between these two proposed roles of the CR3 motif we created single mutants, C47S, C52S, C55S and H184A, and compared their effect to the *rrp44–CR3* triple mutant. For this assay, we used a *rrp44Δ* strain containing a plasmid with wild-type *RRP44* and *URA3* genes. This strain was transformed with a second plasmid that contained the mutant *RRP44* and *LEU2* genes. If this second plasmid can provide functional Rrp44p, then the strain can lose the *URA3* plasmid, resulting in growth on plates containing 5-FOA. As shown in Figure 2A, the single mutants showed a variety of growth phenotypes. The C47S mutant grew worse than the other single mutants, with no obvious growth upon short incubation (2 days; Figure 2A left panel). C52S grew slightly worse than wild-type, which was obvious at 2 days. Upon longer incubation (6 days; Figure 2A middle panel) the C47S mutant showed some growth, and the C52S colony size became almost indistinguishable from wild-type. The C55S and H184A single mutants grew as well as the wild-type strain after both short and long incubation.

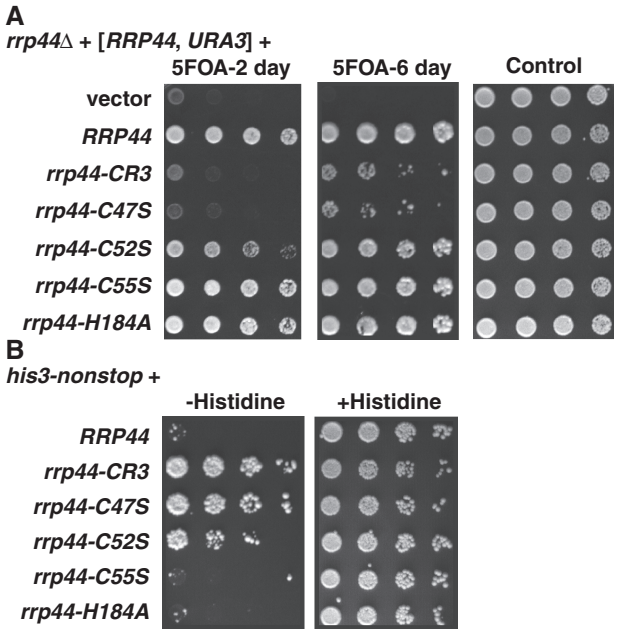


Figure 2. Point mutations in the conserved CCCH motif are inconsistent with a disulfide bond, but are consistent with metal binding. As previously reported, simultaneously mutating the three conserved Cys residues to Ser causes a slow growth phenotype (**A**) and stabilization of *his3-nonstop* mRNA (**B**). The single C47S mutation has a similar, but slightly less severe defect, which can be seen after 6 days on 5-FOA plates in the middle panel. Shorter incubations (e.g. 2 days; left panel) show that the *rrp44–C52S* allele also has a significant but less severe defect. For panel A a *rrp44Δ* strain with a low copy plasmid with the wild-type *RRP44* and *URA3* genes was transformed with a second plasmid with a *LEU2* gene and the indicated *RRP44* gene (none, wild-type, *rrp44–CR3*, *rrp44–C47S*, *rrp44–C52S*, *rrp44–C55S*, or *rrp44–H184A*). Transformants were selected and then serially diluted and spotted on 5FOA containing plates that select for cells that have lost the URA3 plasmid or SC–URA–LEU plates. For Panel B, the colonies that grew on 5FOA were then transformed with a reporter plasmid that contains the his3-nonstop and URA3 genes. Transformants were selected and then serially diluted and spotted on SC–URA–LEU–HIS plates or SC–URA–LEU plates.

We have also shown previously that the *rrp44–CR3* mutant has a defect in the rapid degradation of mRNAs that lack a stop codon (22). Because the *rrp44–CR3* mutant stabilizes the *his3*-nonstop mRNA, a *rrp44–CR3 his3-nonstop* strain produces sufficient histidine to support growth in the absence of added histidine, whereas an isogenic *RRP44 his3-nonstop* strain fails to grow under these conditions (22). To determine which residues of the CR3 motif are needed for nonstop mRNA decay, we tested whether the C47S, C52S, C55S or H184A single mutants could suppress a *his3-nonstop* allele. Importantly, the growth results in this assay are counterintuitive: Growth occurs when nonstop mRNA decay is defective, where as functional nonstop mRNA decay limits growth. As shown in Figure 2B, the *his3-nonstop* defects mirrored the growth defects in Figure 2A. The *rrp44–CR3* allele is defective in nonstop mRNA decay, and thus grows. The single C47S mutant shows almost the same defect. C52S showed a less severe defect, and C55S and H184A were similar to wild type.

Results similar to ours, where mutating individual metal-binding residues have different effects, have been reported for various Zn^{2+}-containing proteins (see 'Discussion' section). Presumably this indicates that mutating a single Cys residue to Ser reduced the affinity for Zn-ions, but does not completely eliminate Zn^{2+} binding. Thus, the phenotypes of single amino acid point mutants are consistent with the CCCH motif forming a binding site for a metal ion, with single point mutations decreasing the affinity for the ion and a triple mutation further decreasing this affinity. In contrast, the phenotypes of the single-point mutants appear inconsistent with a disulfide bond.

Genetic interactions indicate that the CR3 motif affects multiple functions of the exosome *in vivo*

As explained in the Introduction, the known phenotypes of the *rrp44–CR3* mutant can best be explained if this mutation decreases both the endo- and exonuclease activities of Rrp44, but does not completely abolish activity. To better understand the connections between the CR3 motif and nuclease activities of the exosome, we tested for genetic interactions between the *rrp44–CR3* mutation, and mutations inactivating the catalytic activities of the exosome.

The slow growth phenotype of *rrp44–CR3* resembles that of a D551N point mutation that inactivates the exonuclease activity of Rrp44p (hereafter called the *rrp44–exo*$^-$ mutation). One possibility is that the *rrp44–CR3* mutation somehow prevents the exonuclease function of Rrp44p, and thereby results in slow growth. This predicts that adding the *rrp44–CR3* mutation to the already inactive *rrp44–exo*$^-$ mutation would have no additional effect. Another possibility is that the reason *rrp44–CR3* grows slowly is distinct from the reason why *rrp44–exo*$^-$ grows slowly. In this case, combining the two mutations might have an additive effect. To test these possibilities, we combined *rrp44–CR3* with *rrp44–exo*$^-$. Figure 3A shows that an *RRP44* plasmid combining the *rrp44–CR3* and *rrp44–exo*$^-$ mutations (i.e. C47S, C52S, C55S, D551N) did not support growth. As previously reported, the plasmids encoding either *rrp44–CR3* or *rrp44–exo*$^-$ allowed for slow growth compared to wild-type *RRP44*. We conclude that the *rrp44–CR3* mutation is synthetically lethal with the *rrp44–exo*$^-$ mutation. Using the same approach, Figure 3B shows that *rrp44–CR3* is also synthetically lethal with a deletion of the exonuclease domain (*rrp44–exoΔ*). Although *rrp44–CR3* and *rrp44–exo*$^-$ have some of the same molecular defects disrupting the exonuclease activity further worsens the growth of *rrp44–CR3*. We conclude that the *rrp44–CR3* mutation does not completely eliminate *in vivo* exonuclease function, and that the slow growth phenotypes of *rrp44–CR3* and *rrp44–exo*$^-$ are distinct. Furthermore, since combing the *rrp44–endo*$^-$ and *rrp44–exo*$^-$ alleles is also synthetic lethal, the *rrp44–CR3-exo*$^-$ synthetic lethality is consistent with *rrp44–CR3* affecting endonuclease activity.

Next, to test the genetic interaction between the CR3 motif and the endonuclease domain of Rrp44p, we combined the *rrp44–CR3* allele with *rrp44–endo*$^-$

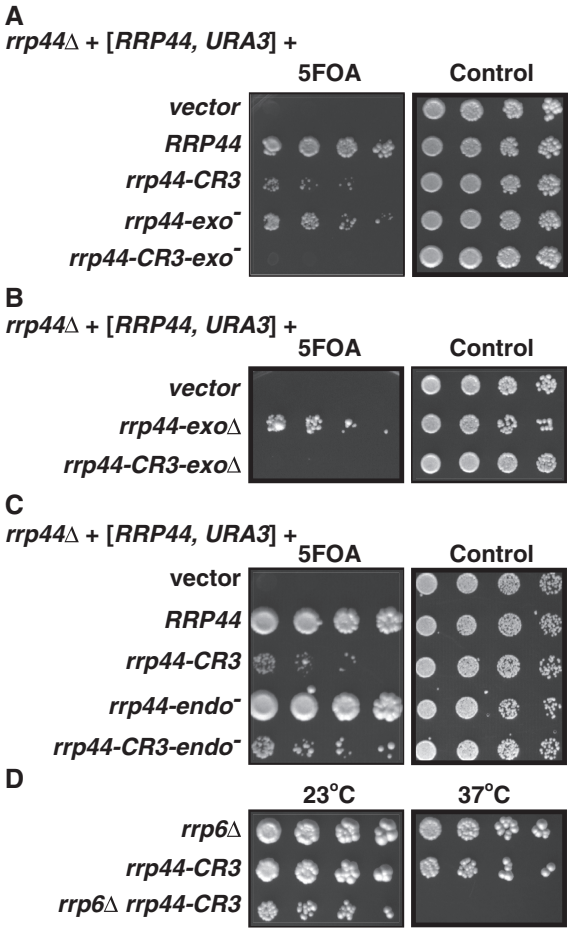


Figure 3. Synthetic lethality of *rrp44–CR3* with mutations in the exosome. Simultaneously mutating the three conserved Cys residues of Rrp44p to Ser is synthetically lethal with point mutations disrupting the Rrp44p exonuclease (**A**), with deletion of the Rrp44p exonuclease domains (**B**) or with deletion of the Rrp6p exonuclease (**D**), but not point mutations disrupting the Rrp44p endonuclease (**C**) For panels A, B and C synthetic lethality was tested using a plasmid shuffle assay as in Figure 2. Failure to grow on 5-FOA plates in panels A to C indicates that the tested *RRP44* allele is inviable. For panel D, for the top row and bottom rows an *rrp6Δ*, *rrp44Δ* double mutant with a wild-type *RRP44*, *URA3* plasmid was transformed with either a wild-type *RRP44*, *LEU2* plasmid (top row), or a *rrp44–CR3*, *LEU2* plasmid (bottom row). For the middle row an *rrp44Δ* mutant with a wild-type *RRP44*, *URA3* plasmid was transformed with a *rrp44–CR3*, *LEU2* plasmid. All three double transformants were then serially diluted and spotted on 5FOA containing plates that were then incubated at the indicated temperatures.

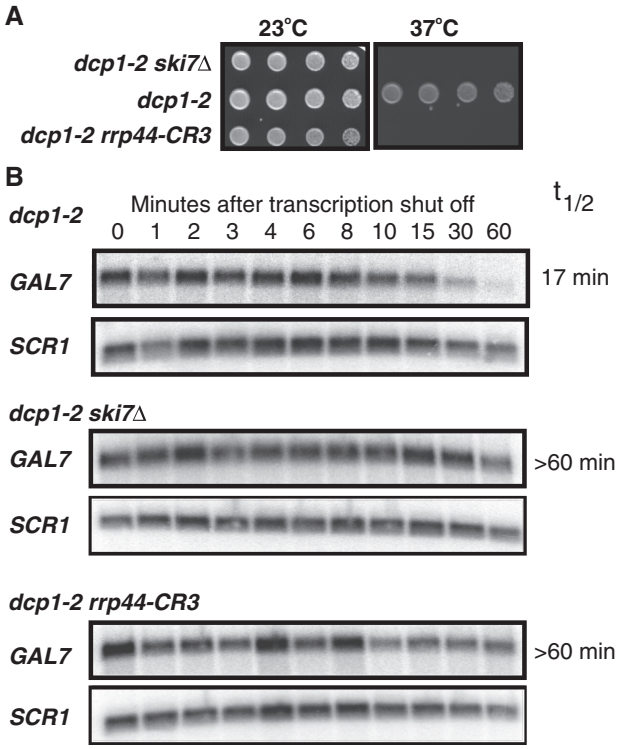


Figure 4. Simultaneously mutating the three conserved Cys residues of Rrp44p to Ser stabilizes cellular mRNAs. (**A**) The *rrp44–CR3* mutation is synthetically lethal with a temperature-sensitive mutation in the decapping enzyme, *dcp1–2*. The experiment was performed as described for Figures 2 and 3. (**B**) The *rrp44–CR3* mutation reduces the rate of exosome-mediated degradation of the *GAL7* mRNA. duplicate cultures of the indicated strains were grown in media containing galactose, and then incubated for 1 h at 37°C to inactivate the decapping enzyme. The media was then replaced with media containing glucose to shut off transcription of the GAL7 gene and RNA was isolated at the indicated time points and analyzed by northern blotting. The RNA subunit of the signal recognition particle (encoded by the SCR1 gene) was used as a loading control. RNA levels were quantitated using a STORM PhosphoImager and half-lives determined. Both half-life measurements for the dcp1–2 strain were 17 min. In the *dcp1–2 ski7Δ* and *dcp1–2 rrp44–CR3* strains the half-lives were more than 60 min.

mutant. Figure 3C shows that this combination has the same growth phenotype as the *rrp44–CR3* single mutant. This is also consistent with the *rrp44–CR3* mutation not completely eliminating exonuclease function, since elimination of exonuclease function (e.g. *rrp44–exo*$^-$) is synthetic lethal with *rrp44–endo*$^-$.

We next tested for genetic interactions with *rrp6Δ*. Rrp6p is an exoribonuclease that associates with the nuclear form of the exosome, and thus provides it with a third catalytic domain. Figure 3D shows that the *rrp6Δ rrp44–CR3* double mutant grows significantly slower than either single mutant at 23°C and fails to grow at 37°C, indicating that *rrp6Δ* has a synthetic growth defect with *rrp44–CR3*.

Since the nuclear exosome is essential, but the cytoplasmic exosome is not, the above growth phenotypes reflect effects of the mutations on nuclear exosome function. To extend our observations to the cytoplasmic exosome, we tested for effects on mRNA decay. Normal mRNAs are mainly degraded through a decapping pathway. In the absence of decapping activity, normal mRNAs are degraded by the exonuclease activity of the exosome (4). Thus, mutations that inactivate decapping are synthetically lethal with mutations that inactivate the exonuclease activity of the exosome (22). To test whether the *rrp44–CR3* allele reduces exosome-mediated decay of normal mRNAs, we initially looked for genetic interactions with *dcp1–2*. *dcp1–2* is a temperature sensitive mutation in a subunit of the decapping enzyme. Since decapping is not required for viability, this strain is viable at all temperatures, but requires exosome-mediated mRNA decay at the restrictive temperature. Figure 4A shows that, as expected,

the *dcp1–2 rrp44–CR3* double mutant is viable at the permissive temperature (23°C), but inviable at the restrictive temperature for *dcp1–2* (37°C). Therefore, the *rrp44–CR3* mutation likely disrupts exosome-mediated decay of normal mRNAs.

The *dcp1–2 rrp44–CR3* double mutant also allowed us to directly measure the effect of *rrp44–CR3* on the degradation of an endogenous mRNA, *GAL7*. For this experiment we grew *dcp1–2* strains at the permissive temperature in the presence of galactose to mid-log phase. The cultures were then incubated at the restrictive temperature for 1 h to inactivate the decapping enzyme, followed by the addition of glucose to inhibit further transcription of the *GAL7* gene. Figure 4B shows that in a *dcp1–2* single mutant, the endogenous *GAL7* mRNA decays with a half-life of about 17 min. In contrast, in the *dcp1–2 rrp44–CR3* double mutant, the *GAL7* mRNA half-life is greater than 60 min, which is similar to the *dcp1–2 ski7Δ* control strain. Both the *dcp1–2* synthetic lethality and *GAL7* mRNA half-lives suggest that the *rrp44–CR3* mutation interferes with exosome-mediated degradation of normal cellular mRNAs.

Overall, these genetic interactions are consistent with the hypothesis that the *rrp44–CR3* mutation reduces, but does not eliminate, the exonuclease activity *in vivo*. Furthermore, the synthetic interactions with *rrp6Δ* and *dcp1–2* suggest that the *rrp44–CR3* allele affects both cytoplasmic and nuclear exosome function.

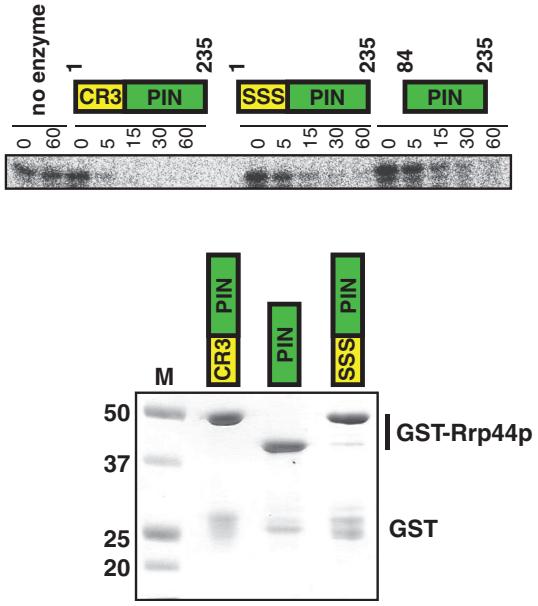


Figure 5. Mutating the CCCH motif affects the *in vitro* endonuclease activity of Rrp44p. Recombinant truncated Rrp44 proteins that either contain both an intact CR3 motif and PIN domain (i.e. residues 1–235), a similar protein with the Cys residues in the CR3 motif mutated to Ser, or a protein containing only the PIN domain (i.e. residues 84–235) were used in *in vitro* RNase assays with an A_{30} substrate as described in materials and methods. Similar results were obtained with a $CCCGACACCACCCACUA_{14}$ substrate. The bottom panel show a Coomassie stained gel of the purified recombinant proteins.

Mutation of the CR3 motif affects the endoribonuclease activity of Rrp44p *in vitro*

While the genetic analysis above indicates that the CCCH motif is important *in vivo*, it can not address whether these amino acids play a direct role or an indirect role in Rrp44p nuclease activity. In fact, since the CCCH motif is very well separated from the exonuclease domain, we consider it unlikely that these amino acids have a direct role in exonuclease activity. At first glance, the CCCH motif is also far (>25 Å) away from the endonuclease active site, however it is important to note that H_{184} of the CCCH motif and D_{171} of the endonuclease active site are part of the same α-helical structure (Figure 1D and Supplementary Figure S1A). Because D_{171} is a key residue in the endonuclease active site (18–20), we considered the possibility that perturbation to the CCCH site may have an indirect effect on the orientation or flexibility of this α-helix, and thereby affect endonuclease activity. To test this, we purified three truncated Rrp44 proteins as glutathione S-transferase (GST) fusions from *E. coli*. As we and others have previously shown, a truncated Rrp44 protein that contains the PIN domain and CR3 motif (residues 1–235), has endonuclease activity *in vitro* (18–20). The second recombinant protein purified is the same, except that the three conserved Cys residues were mutated to Ser. This CR3 mutant enzyme was less active in *in vitro* assays containing the same amount of enzyme of similar purity (Figure 5). Specifically, Figure 5 shows that the wild-type enzyme degraded most of the substrate in the first 5 min, with all of the substrate gone after 15 min. In comparison, more substrate was left in the reaction with the same amount of the enzyme with all three Cys residues mutated after 5 or 15 min. The third recombinant protein we purified contains only the part of Rrp44p that is similar to other PIN domains (residues 84–235) and lacks the CR3 motif. This PIN domain only protein also exhibited less nuclease activity *in vitro* (Figure 5). Quantitation of the experiment shown in Figure 5 indicated that the wild-type enzyme was approximately twice as active as either mutant version. Although individual batches of purified protein can have variations in the amount of activity, the differences between the three forms of Rrp44p were reproducible with independent repeat purifications.

Mutation of the CR3 motif affects the interaction between Rrp44p and the core exosome

We noted that the CCCH motif is close to three residues that interact with the core exosome. Specifically, in the crystal structure of a trimer of Rrp41p, Rrp45p and Rrp44p, the conserved Rrp44p Y_{40}, R_{42} and D_{44} residues directly interact with Rrp41p (Figure 1D; (13)) and only two residues separate this YxRxD motif from the Cx_4Cx_2C motif (Supplementary Figure S1A). This led us to hypothesize that mutating the conserved Cys residues to Ser might disorder this part of the protein and reduce the interaction with the core exosome. To test this idea, TAP-tagged wild-type Rrp44p and Rrp44p–CR3 proteins were purified using IgG sepharose beads, and exosome interaction was determined using an anti-Rrp4p antibody. Figure 6A shows that the exosome

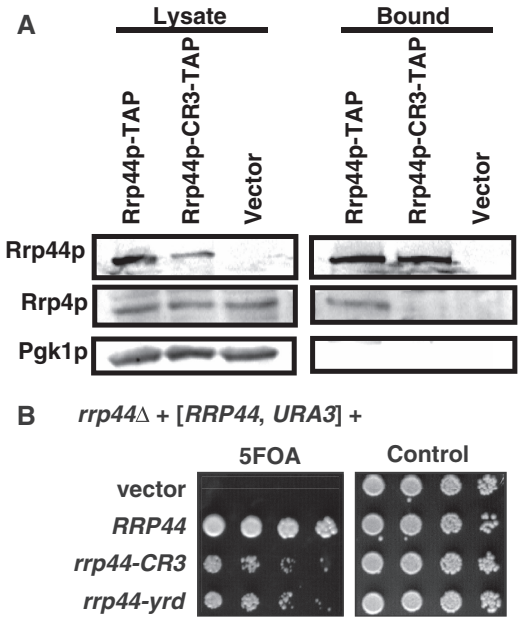


Figure 6. Simultaneously mutating the three conserved Cys residues to Ser reduces interaction with the exosome. (**A**) TAP-tagged Rrp44 containing either a wild-type or mutated CR3 motif were purified from yeast. The lysate (left) and bound (right) fractions were analyzed by western blot with antibodies for the TAP tag (top), the exosome subunit Rrp4p (middle) or the control protein Pgk1p (bottom). (**B**) The *rrp44–yrd* allele that contains Y40A, R42A and D44A mutations has a slow growth phenotype that resembles that of *rrp44–CR3*. The experiment was performed using the plasmid shuffle as described in Figure 2.

subunit Rrp4p copurifies with the wild-type Rrp44p, but no Rrp4p could be detected in the Rrp44p–CR3–TAP purification. Thus, even though we suspect that the three conserved Cys residues do not directly contact other exosome subunits, mutating them reduces the ability of Rrp44p to interact with the rest of the exosome, possibly by altering the conformation of Y_{40}, R_{42} and/or D_{44}.

If the effects of the *rrp44–CR3* mutation are due to an alteration of the interaction of Y_{40}, R_{42} and/or D_{44} with Rrp41p, then mutating these residues should have a similar effect. We tested this by creating three single point mutations in these residues (*rrp44–Y40A*, *rrp44–R42A* and *rrp44–D44A*) as well as a triple-point mutant (*rrp44–yrd*) combining these mutations. The single-point mutants did not have an obvious growth defect (data not shown), but the triple *rrp44–yrd* mutant had a slow growth phenotype that closely resembled the growth phenotype of the *rrp44–CR3* allele (Figure 6B), consistent with the idea that mutating the CCCH motif affects the conformation of Y40, R42 and/or D44, which in turn affects exosome interaction.

DISCUSSION

We have previously shown that mutating the CR3 motif leads to a slow growth phenotype, a slightly reduced expression level of Rrp44p, and a stabilization of nonstop mRNAs (19,22). Although these results hinted at the importance of the conserved motif, it was not clear how changing these Cys residues to Ser affected Rrp44p function. The crystal structure (13) provided important insight as the published structural model indicated that the cysteines are clustered together. The structure further revealed that, although in the primary sequence His_{184} is far from the CR3 motif, the histidine is in fact positioned in a tetrahedral arrangement with the cysteine residues. While the proximity of Cys residues raises the potential for disulfide bond formation, disulfide bonds are rare in the reduced environment of the cytoplasm and nucleus. Furthermore, mutating individual Cys residues to Ser should completely disrupt such a disulfide bond, but such mutations did not mimic the *rrp44–CR3* phenotype. This strongly argues against the formation of a disulfide bond in the CR3 motif. Instead the structural clustering of CR3 with His_{184} suggested a metal ion-binding site, since four Cys and/or His residues form Zn^{2+} ion-binding sites in a wide variety of proteins. In such metal-binding sites, mutations of individual residues to serine can have a variety of effects. For example, in the Rpc10p subunit of RNA polymerase III, mutating the first Zn^{2+}-binding Cys residue to Ser is lethal, but changing either the third or fourth Cys (or both) to Ser has no obvious effect on growth (28). Consistent with the idea that a single Cys to Ser change does not always completely eliminate Zn^{2+} binding, is the natural presence of Ser in some Zn^{2+}-binding sites (29). The suggestion that the CCCH motif forms a metal-binding site is further reinforced by the observation that His_{184} is conserved in most Rrp44p orthologs, but is lacking in two orthologs that also lack the CR3 motif. In addition, although H_{184} is part of the PIN domain, it is not conserved in PIN domains that are not part of Rrp44p homologs. Based on all of the available genetic, biochemical, structural and phylogenetic data we suggest that binding a monoatomic metal ion to this CCCH motif stabilizes folding of this part of Rrp44p. Interestingly, Lebreton *et al.* (18) have previously shown that the endonuclease activity of Rrp44p *in vitro* is stimulated by addition of either Mn^{2+} or Zn^{2+} ions. One possibility is that either metal ion can bind to the active site. However, an alternative explanation is that one or both of these ions bind to the CCCH motif, and the active site may contain a metal ion that copurifies with the recombinant protein. So far we have not been able to clearly distinguish between these and other possibilities using *in vitro* assays. Moreover, although the CCCH motif most closely resembles a Zn^{2+}-binding site, binding of some other metal ion is also possible, further complicating interpretation of such *in vitro* experiments.

We have previously shown that either the endo- or exonuclease activity of Rrp44p is sufficient for rapid nonstop mRNA degradation, but that the *rrp44–CR3* mutation results in stabilization of nonstop mRNAs. This suggests that the *rrp44–CR3* mutation reduces both nuclease activities. However, since *rrp44–CR3* is viable and Rrp44p nuclease activity is required, the *rrp44–CR3* mutation cannot completely eliminate both functions. Here we test this genetically. If *rrp44–CR3* completely eliminates the exonuclease function, combining it with another mutation that eliminates the exonuclease activity should not have an additional effect. Instead we show that

rrp44–CR3 is synthetically lethal with *rrp44–exo*$^-$. Thus, although *rrp44–CR3* and *rrp44–exo*$^-$ have some of the same molecular defects, we conclude that *rrp44–CR3* does not completely eliminate exonuclease function *in vivo*. Furthermore, since eliminating both endo- and exonuclease activities is lethal, the synthetic lethality of *rrp44–CR3* with *rrp44–exo*$^-$ is consistent with a reduced endonuclease activity. Further synthetic lethality of *rrp44–CR3* with *dcp1–2* and with *rrp6Δ* suggests that both the cytoplasmic and nuclear exosomes are affected, since *dcp1–2* synthetic lethality reflects a role in cytoplasmic mRNA decay, and Rrp6p is exclusively nuclear. Therefore we conclude that the CR3 motif affects many facets of Rrp44p function.

Since the CCCH site is far removed from both catalytic active sites of Rrp44p, it is likely not directly involved in either catalytic activity. In many Zn^{2+}-containing proteins, the metal ion stabilizes the overall fold of the protein. The available evidence suggests that the overall structure of Rrp44p can fold in the absence of a metal ion bound to the CCCH motif (13,18–20), but we suggest that binding of a metal ion to the CCCH site has a modest effect on folding and/or flexibility of this part of the protein. This then indirectly affects endo- and exonuclease activities, possibly because the CCCH motif is well-connected to two other important parts of Rrp44p (Figure 1D). First, His_{184} is part of an α-helix that also contains the endonuclease active site residue D_{171}. Thus, perturbation of the CCCH motif may slightly alter the position or flexibility of this α-helix, and thereby affect endonuclease activity. Second, we show that the *rrp44–CR3* mutation weakens the interaction between Rrp44p and the remainder of the exosome. Mutations in or depletion of the other nine exosome subunits all cause defects in RNA processing that are similar to those of *rrp44–exo*$^-$. In contrast, the *rrp44–endo*$^-$ allele does not affect growth or any RNA decay or processing reaction, and therefore an endonuclease defect cannot explain the phenotypes of *rrp44–CR3*. Thus, we suggest that the reduced *in vivo* exonuclease activity is a result of the reduced exosome interaction. Although the YxRxD motif is conserved and in the crystal structure interacts with Rrp41p, the importance of these residues had not previously been tested. Our data indicate that this motif indeed is a functionally important contact between the core exosome and Rrp44p. Strikingly, of the single Cys mutations, C47S would be expected to have the largest effect on the YxRxD motif which may explain why it had the largest effect on exosome function. Interestingly, in many cases Rrp44p orthologs fail to copurify with the exosome (30–33) even though exosome-interacting residues, such as the YxRxD motif, are conserved (Supplementary Figure S1A and data not shown). Since the YxRxD motif is conserved in Rrp44p orthologs and hDis3L1 orthologs, our results suggest that all of these proteins interact with the core exosome, even though such interaction has proven difficult to demonstrate. Our results further suggest that in some cases the failure to copurify may be because copurification was tested in buffers that included the Zn^{2+} chelator EDTA. Strikingly, two papers reported that hDIS3 does not copurify with the exosome using EDTA containing buffers, while a third paper reports hDIS3-exosome copurification using EDTA-free buffers (32–34).

Overall we can now offer a coherent model for the importance of the CR3 motif: The primary effect of mutating the CR3 motif is probably modest local misfolding and/or structural instability of Rrp44p, possibly due to loss of a metal ion. Because the CR3 motif is structurally connected to both the endonuclease active state and to exosome-interacting residues, this effect on local structure has surprisingly far-reaching effects on exosome function. Furthermore, it opens the possibility that the activity of Rrp44p may be regulated by similar structural alterations.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figure 1 and Supplementary Table 1.

ACKNOWLEDGEMENTS

Antibodies against Rrp4p were kindly provided by David Tollervey. The authors thank members of the van Hoof and Arraiano labs and Lisa Berreau for fruitful discussions.

FUNDING

The National Institutes of Health (NIH) [GM069900 and GM099790 to A.v.H.]; the National Science Foundation [MCB1020739 to A.v.H.]; an EMBO short term fellowship [to D.S.]; grants from Fundação para a Ciência e Tecnologia [FCT; PEst-OE/EQB/LA0004/2011 to C.M.A.]; Ph.D. fellowship from FCT (to F.P.R.). Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. 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.
2. de la Cruz,J., Kressler,D., Tollervey,D. and Linder,P. (1998) Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in Saccharomyces cerevisiae. *EMBO J.*, **17**, 1128–1140.
3. 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.
4. Jacobs Anderson,J.S. and Parker,R. (1998) The 3' to 5' degradation of yeast mRNAs is a general mechanism for mRNA turnover that requires the SKI2 DEVH box protein and 3' to 5' exonucleases of the exosome complex. *EMBO J.*, **17**, 1497–1506.
5. Meaux,S. and van Hoof,A. (2006) Yeast transcripts cleaved by an internal ribozyme provide new insight into the role of the cap and poly(A) tail in translation and mRNA decay. *RNA*, **12**, 1323–1337.
6. van Hoof,A., Frischmeyer,P.A., Dietz,H.C. and Parker,R. (2002) Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science*, **295**, 2262–2264.
7. Chernyakov,I., Whipple,J.M., Kotelawala,L., Grayhack,E.J. and Phizicky,E.M. (2008) Degradation of several hypomodified mature tRNA species in Saccharomyces cerevisiae is mediated by Met22

and the 5'-3' exonucleases Rat1 and Xrn1. *Genes Dev.*, **22**, 1369–1380.
8. Tucker,M., Valencia-Sanchez,M.A., Staples,R.R., Chen,J., Denis,C.L. and Parker,R. (2001) The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in Saccharomyces cerevisiae. *Cell*, **104**, 377–386.
9. Liu,Q., Greimann,J.C. and Lima,C.D. (2006) Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell*, **127**, 1223–1237.
10. Lorentzen,E. and Conti,E. (2005) Structural basis of 3' end RNA recognition and exoribonucleolytic cleavage by an exosome RNase PH core. *Mol. Cell*, **20**, 473–481.
11. 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.
12. Lorentzen,E., Basquin,J., Tomecki,R., Dziembowski,A. and Conti,E. (2008) Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol. Cell*, **29**, 717–728.
13. 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.
14. Buttner,K., Wenig,K. and Hopfner,K.P. (2005) Structural framework for the mechanism of archaeal exosomes in RNA processing. *Mol. Cell*, **20**, 461–471.
15. Wang,H.W., Wang,J., Ding,F., Callahan,K., Bratkowski,M.A., Butler,J.S., Nogales,E. and Ke,A. (2007) Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. *Proc. Natl Acad. Sci. USA*, **104**, 16844–16849.
16. Dziembowski,A., Lorentzen,E., Conti,E. and Seraphin,B. (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat. Struct. Mol. Biol.*, **14**, 15–22.
17. Frazao,C., McVey,C.E., Amblar,M., Barbas,A., Vonrhein,C., Arraiano,C.M. and Carrondo,M.A. (2006) Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature*, **443**, 110–114.
18. Lebreton,A., Tomecki,R., Dziembowski,A. and Seraphin,B. (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature*, **456**, 993–996.
19. Schaeffer,D., Tsanova,B., Barbas,A., Reis,F.P., Dastidar,E.G., Sanchez-Rotunno,M., Arraiano,C.M. and van Hoof,A. (2009) The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nat. Struct. Mol. Biol.*, **16**, 56–62.
20. Schneider,C., Leung,E., Brown,J. and Tollervey,D. (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res.*, **37**, 1127–1140.
21. Cairrao,F., Arraiano,C. and Newbury,S. (2005) Drosophila gene tazman, an orthologue of the yeast exosome component Rrp44p/Dis3, is differentially expressed during development. *Dev. Dy./*, **232**, 733–737.
22. Schaeffer,D. and van Hoof,A. (2011) Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. *Proc. Natl Acad. Sci. USA*, **108**, 2366–2371.
23. Giaever,G., Chu,A.M., Ni,L., Connelly,C., Riles,L., Veronneau,S., Dow,S., Lucau-Danila,A., Anderson,K., Andre,B. *et al.* (2002) Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, **418**, 387–391.
24. Goldstein,A.L. and McCusker,J.H. (1999) Three new dominant drug resistance cassettes for gene disruption in Saccharomyces cerevisiae. *Yeast*, **15**, 1541–1553.
25. van Hoof,A., Staples,R.R., Baker,R.E. and Parker,R. (2000) Function of the ski4p (Csl4p) and Ski7p proteins in 3'-to-5' degradation of mRNA. *Mol. Cell Biol.*, **20**, 8230–8243.
26. Caponigro,G., Muhlrad,D. and Parker,R. (1993) A small segment of the MAT alpha 1 transcript promotes mRNA decay in Saccharomyces cerevisiae: a stimulatory role for rare codons. *Mol. Cell Biol.*, **13**, 5141–5148.
27. Mitchell,P., Petfalski,E., Houalla,R., Podtelejnikov,A., Mann,M. and Tollervey,D. (2003) Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. *Mol. Cell Biol.*, **23**, 6982–6992.
28. Rubbi,L., Labarre-Mariotte,S., Chedin,S. and Thuriaux,P. (1999) Functional characterization of ABC10alpha, an essential polypeptide shared by all three forms of eukaryotic DNA-dependent RNA polymerases. *J. Biol. Chem.*, **274**, 31485–31492.
29. Karlin,S. and Zhu,Z.Y. (1997) Classification of mononuclear zinc metal sites in protein structures. *Proc. Natl Acad. Sci. USA*, **94**, 14231–14236.
30. Chekanova,J.A., Gregory,B.D., Reverdatto,S.V., Chen,H., Kumar,R., Hooker,T., Yazaki,J., Li,P., Skiba,N., Peng,Q. *et al.* (2007) Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell*, **131**, 1340–1353.
31. Estevez,A.M., Kempf,T. and Clayton,C. (2001) The exosome of Trypanosoma brucei. *EMBO J.*, **20**, 3831–3839.
32. Staals,R.H., Bronkhorst,A.W., Schilders,G., Slomovic,S., Schuster,G., Heck,A.J., Raijmakers,R. and Pruijn,G.J. (2010) Dis3-like 1: a novel exoribonuclease associated with the human exosome. *EMBO J.*, **29**, 2358–2367.
33. Chen,C.Y., Gherzi,R., Ong,S.E., Chan,E.L., Raijmakers,R., Pruijn,G.J., Stoecklin,G., Moroni,C., Mann,M. and Karin,M. (2001) AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell*, **107**, 451–464.
34. Tomecki,R., Kristiansen,M.S., Lykke-Andersen,S., Chlebowski,A., Larsen,K.M., Szczesny,R.J., Drazkowska,K., Pastula,A., Andersen,J.S., Stepien,P.P. *et al.* (2010) The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.*, **29**, 2342–2357.