Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Seth M. Kelly Date

Recognition of Polyadenosine RNA by CCCH Zinc Fingers

By

Seth M. Kelly

Doctor of Philosophy

Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology

_____
Anita H. Corbett, Ph.D.
Advisor

_____
Yue Feng, Ph.D.
Committee Member

_____
T.J. Murphy, Ph.D.
Committee Member

_____
Maureen Powers, Ph.D.
Committee Member

_____
Keith Wilkinson, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D. Dean of the Graduate School

_____
Date

Recognition of Polyadenosine RNA by CCCH Zinc Fingers

By

Seth M. Kelly
B.S., Grove City College, 2003

Advisor: Anita H. Corbett, Ph.D.

An abstract of
a dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology Program
2009

Abstract

Recognition of Polyadenosine RNA by CCCH Zinc Fingers

By Seth M. Kelly

Throughout their lifecycles, mRNA transcripts are coated by a collection of RNA binding proteins. These RNA binding proteins function in a diverse set of processes, but collectively they dictate the fate of the transcripts to which they bind. While some RNA binding proteins recognize sequences found in only a handful of transcripts, others, such as those proteins that bind the poly(A) tail of mRNA transcripts, recognize sequences ubiquitous to all mRNA transcripts. These RNA binding protein binds post-transcriptionally regulate gene expression. Therefore, in order to completely comprehend the regulation of gene expression, it is essential to understand the molecular mechanisms of RNA recognition by RNA binding proteins.

In this dissertation I present data demonstrating that Cys-Cys-Cys-His (CCCH) zinc fingers of the *S. cerevisiae* protein, Nuclear poly (A) Binding protein 2 (Nab2) and its putative human orthologue, ZC3H14, specifically recognize polyadenosine RNA with high affinity. All previously characterized poly(A) RNA binding proteins bind polyadenosine RNA via at least one RNA Recognition Motif. Hence, recognition of polyadenosine RNA via CCCH zinc fingers is a novel mechanism of poly(A) RNA recognition. Genetic and biochemical studies provide compelling evidence that ZnF 5-7 mediate high affinity binding to polyadenosine RNA. To gain further understanding of the mechanism of poly(A) RNA recognition by CCCH zinc fingers, the atomic resolution structure of Nab2 zinc fingers 5 – 7 was solved using NMR. Using this structural data,

we identified several conserved positively charged and aromatic residues that could potentially interact with polyadenosine RNA. Changing these amino acids to alanine resulted in loss of binding to polyadenosine RNA *in vitro* and conferred cold-sensitive growth defects *in vivo*. We have also identified several genes that genetically interact with *NAB2* including components of the mRNA 3'-end processing machinery, a component of the nuclear exosome, and components necessary for the transcriptional termination of the RNA polymerase II. Together, these findings define a novel evolutionarily conserved family of polyadenosine RNA binding proteins. These findings also further demonstrate that polyadenosine RNA binding proteins are key players involved in the post-transcriptional regulation of gene expression.

Recognition of Polyadenosine RNA by CCCH Zinc Fingers

By

Seth M. Kelly
B.S., Grove City College, 2003

Advisor: Anita H. Corbett, Ph.D.

A dissertation submitted to the Faculty of the Graduate School of Emory University
in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Graduate Division of Biological and Biomedical Sciences
Biochemistry, Cell and Developmental Biology Program
2009

Acknowledgements

Herein you will read about (most of) the work I've done in the last six years. There have been ups and downs throughout these last six years, but overall the faculty and students I've encountered at Emory University, and specifically in the Biochemistry, Cell and Developmental Biology Program and the Emory University Biochemistry Department, have made this process an enjoyable one.

To me, a Ph.D. is more than just a collection of experiments that make up a thesis. It's more than the time you spend in the laboratory doing your project. It's more than the time you spend outside of the laboratory thinking about your project. A Ph.D. is the *process* of learning how to think critically and analytically about the data you see in front of you. It's the process of starting with an idea, testing that idea, and then starting all over again and learning how to convey all the things you've learned to both your scientific peers and to your Grandma.

The process of becoming the person (and the scientist) I am today started a long time ago. Along this journey, I have met a collection of amazingly talented people who have challenged me both intellectually and personally. In particular I'd like to acknowledge the following people:

First, I'd like to thank my advisor, Dr. Anita Corbett. When I was trying to decide which lab to join after my first year of graduate school I asked Anita her advice and she told me to choose the lab where I would learn to ask poignant questions that addressed the most important questions in the field. I truly feel that I have had the opportunity to do that in the Corbett lab. Anita has taught me how to *do* science. She's taught me how to write about and present my work in a way that makes sense to people in

and outside the field. The past and present members of the Corbett lab deserve special acknowledgement as well. Specifically, I'd like to thank Michelle Harreman, Milo Fasken, Sara Leung, Laura McLane, Anna Bramley, Shana Kerr, Allison Lange, Kanika Pulliam, Alex Frere, Deanna Greene, Adam Berger, and Maya Kordani. Without their key intellectual and personal encouragements, I would definitely never have finished this process. It's a testament to their commitment to me that this work stands before you. I feel blessed that I've gotten to share 5 years of my life with the people in my lab. More than mere classmates or co-workers, I count them all as dear friends.

This project was also the result of several important collaborations. In the Physics department here at Emory University, Keith Berland and Suzette Pabit worked arduously to develop the fluorescence correlation spectroscopy assay as described herein. Both Keith and Suzette were key contributors to our understanding of the mechanism of polyadenosine recognition by Nab2 I owe much gratitude to them. Chad Kitchen, ChangHui Pak, and Luciano Apponi also contributed immensely to technical aspects of this project. Our ongoing collaboration with Cambridge professor and crystallographer, Murray Stewart, and the members of his laboratory (specifically Christoph Brockmann), allowed us to gain critical insights into the mechanisms by which zinc finger containing proteins recognize poly(A) RNA.

Finally, I'd also like to thank the members of my thesis committee, Yue Feng, T.J. Murphy, Maureen Powers, and Keith Wilkinson for their insightful comments on this project.

Now, on a more personal lever there are several other groups of people I'd like to take the time to thank. First I'd like to acknowledge the numerous science teachers that

have inspired me to learn about the world around (and inside) us. Specifically, I'd like to thank Mrs. Bridget Kennedy, who instilled in me a deep love of the biological sciences and specifically molecular biology and biochemistry. I still remember the thrill of learning, and really understanding, the complicated aspects of cell biology in my senior A.P. Biology class. That memory has kept me going during this long and arduous doctorate process. I'd also like to thank several Grove City College professors including Dr. Durwood Ray, my undergraduate research mentor; Dr. Timothy Homan, whose office door was always open; Dr. Mark Webber and Dr. Kevin Shaw.

Second, I'd next like to specifically thank several students from the Biochemistry, Cell and Developmental Biology (BCDB) program. Dina Greene, Marie Cross, Rob Lyng, Avanti Ghokale, Branch Craig, Emma Delva, and Lori Rowe have all had profound impacts on my viewpoints, both scientifically and personally. I'm eternally grateful for the gift of their friendships.

Third, I'd like to acknowledge the friends who have endured this long journey with me. Dustin Dimit, Brent Nelson, Andrew and Stephanie Bonhaus, Andrea Charbeneau, Pat Blake, Andrew Hicks, Seth and Renee McLaughlin, Phil Hagedorn, Kristin Burgess, Andrew Levine, David and Rachel Smith, Scott Hale, Scott Holland, Jeremy Shankle, Shawn and Sheryl Rummel, Brad Wray, Brian Wallace, Brian Mezey, Matt and Liesel Allgyre, Dan Adamson, Matt Armstrong, Rob and Meeghan Fortson, and Jarrett and Christy Helms, among others, for their continued support.

My parents, Timothy and Vicky, as well as my sister, Kara, and her husband, Alex, have been an amazing source of inspiration to me during my Ph.D. Over the years my mother and father have instilled in me a great love of education and have continually

demonstrated that with hard work, anything is possible. Both of my parents have always encouraged me to put 110% of myself into everything I set out to do. Without that mindset, I don't think I would have survived the last 6 years of graduate school. I'd also like to thank my wife's parents, Doug and Linda McKibben as well as my brother-in-law, Craig McKibben.

I feel as if I should share this Ph.D. with my wife, Jill. She could probably recite my thesis from memory at this point. Together we've experienced the entire gamut of emotion during the last six years. From anger and tears to jubilation and beers, I could not have completed this process without her. She has been my constant companion through good times and bad. A mere thank you doesn't really begin to acknowledge the gratuity that is owed her. More than anything, she's taught me what it means to be unconditionally loved. I'd also like to acknowledge my daughter Olivia who has lifted heavy loads from my shoulders with many a toothless smile. I can see great things in her already. I love you both to the depths of my soul.

Finally, praise be to Him from whom all blessings flow and to whom all glory is due.

Table of Contents

Figures

Tables

Abbreviations

3'-UTR	3'-untranslated region
5-FOA	5-fluoroorotic acid
AMP	Ampicillin
BSA	Bovine Serum Albumin
CBC	Cap binding complex
CCCH	Cysteine-Cysteine-Cysteine-Histidine zinc finger
CTD	C-terminal domain of RNA polymerase II
CUT	Cryptic unstable transcript
DAPI	4',6-diamidino-2-phenylindole
DIC	Differential interference microscopy
FCS	Fluorescence correlation spectroscopy
FG	Phenylalanine-Glycine repeat
FISH	Fluorescence *in-situ* hybridization
GFP	Green fluorescent protein
GST	Glutathione S-transferase
IPTG	Isopropyl-thio-galactopyranoside
KH domain	hnRNP K homology domain
LB	Luria broth
mRNP	Messenger RNA ribonucleoprotein complex
NAT	Natamycin
NMD	Nonsense mediate decay
NMR	Nuclear magnetic resonance
NPC	Nuclear pore complex
Pab	Poly(A) RNA binding protein
PCR	Polymerase chain reaction
PLAC	Pepstatin A, leupeptin, aprotinin, chymostatin
PMSF	Phenylmethylsulfonyl fluoride
PVDF	Polyvinylidene fluoride membrane
PWI	Phenyl-tryptophan-isoleucine domain
RGG	Arginine-Arginine-Glycine
RRM	RNA Recognition Motif
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
TRAMP	Trf4/5 – Air1/2 – Mtr4 protein complex
TREX	Transcription and export complex
UBA domain	Ubiquitin associated domain

Chapter 1: A General Introduction

A portion of this chapter is adapted from the following paper:

Kelly, S.M, and Corbett, A.H. *Messenger RNA Export from the Nucleus: A Series of Molecular Wardrobe Changes*. Traffic (2009) Published online: May 27, 2009

Introduction

The central dogma of biology has been well established for several decades: DNA is transcribed into RNA and RNA is then translated into protein. The flow of genetic information from DNA to protein requires, as an absolute necessity, this RNA "blueprint". Like a blueprint sent from factory headquarters to its workers, RNA transcripts dictate the proteins that make up the architecture of every cell. Hence, while the DNA content remains constant from cell to cell in a multi-cellular organism, the type and abundance of these RNA "blueprints" vary widely from one cell type to the next. Essentially, changes in RNA expression between cell types dictate cellular identity and function.

In all eukaryotes, from humans to the single celled budding yeast, *Saccharomyces cerevisiae*, before the translation machinery can effectively translate RNA transcripts into protein, the messenger RNA (mRNA) transcript must first undergo multiple processing steps and subsequently proceed through several quality control checkpoints to ensure these steps were performed correctly. At each of these steps, the RNA is accompanied by a collection of RNA binding proteins. These proteins recognize the transcripts, perform essential processing activities, and package the mRNA into complexes, called mRNA ribonucleoproteins (mRNPs), that regulate transcript stability (1), promote export from the nucleus (2), and modulate translation (3). Hence, RNA binding proteins play a crucial role in the post-transcriptional control of gene expression and have even been likened to post-transcriptional activators and repressors of gene expression (4, 5).

The mechanisms by which many RNA binding proteins recognize specific RNA sequences remain unknown. Many proteins utilize similar RNA binding motifs to

recognize drastically different sequences, while some proteins utilize different motifs to recognize the same sequence. In this dissertation, a novel mechanism of polyadenosine RNA recognition by a family of CCCH zinc finger-containing proteins is identified (Chapters 2 and 3) and analyzed using a structure/function approach (Chapter 3). Furthermore, the *in vivo* effects of the disruption of this interaction are also assessed (Chapters 3 and 4).

The evolutionary conservation of mRNA processing and export factors

Many of the proteins involved in the transcription, processing, and export of mRNA transcripts are highly conserved among eukaryotic species ((6), See also Table 1.1). As demonstrated in Figure 1.1, the overall pathway that produces translation competent mRNA transcripts (7) is present in eukaryotes, from humans to the single celled yeast *Saccharomyces cerevisiae*. Factors ranging from RNA polymerase II and RNA export receptors to RNA binding proteins are highly homologous between different species. Messenger RNA transcripts are initially transcribed by RNA polymerase II in the nucleus of eukaryotic cells. In metazoan cells, typical mRNA transcripts consist of large non-protein-producing (non-coding) sequences called introns (red lines in Figure 1.1) amidst smaller islands of protein coding regions, called exons (gray lines in Figure 1.1) (8, 9). In fact, the average exon in humans is only 150 base-pairs while the average intron is much larger (9). Some introns can be as large as 10,000 base-pairs (9). To assemble the final mRNA transcript, the non-coding intron sequences must be removed and the exons "glued" together in a process called splicing. The budding yeast, *S.*

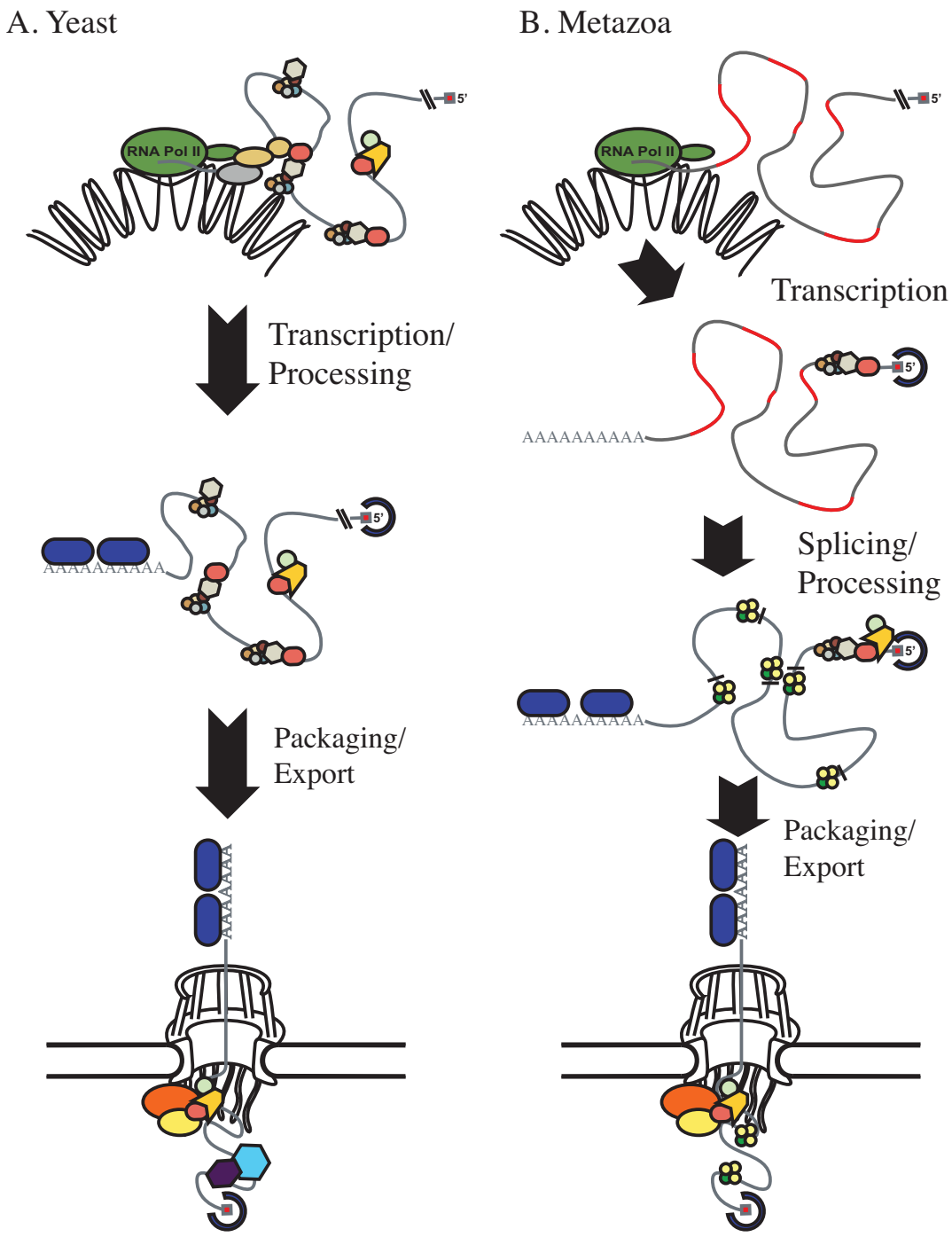

Figure 1.1: General schematic for RNA transcription, processing and export from the nucleus in the budding yeast *Saccharomyces cerevisiae* (A) and in metazoans (B). This figure is adapted from (7) (A) In yeast, mRNA transcripts are initially transcribed by RNA polymerase II. Processing and export factors are recruited to the nascent transcript via interactions with the C-terminal domain of RNA Polymerase II. These factors act to add the 5'-cap, splice introns, and cleave and polyadenylate the 3'-end of the transcript. Export factors, such as Mex67, are recruited during transcription and processing to facilitate the export of mRNA transcripts through the nuclear pore complex and into the

cytoplasm. Following transport through nuclear pore, the conserved RNA helicase, Dbp5, remodels the mRNA ribonucleoprotein (mRNP) complex. (B) In metazoan cells, as in yeast, mRNA transcripts are transcribed by RNA polymerase II. Unlike *S. cerevisiae*, the majority of metazoan transcripts contain both introns (red lines) and exons (gray lines). The splicing machinery removes introns splices together the exons and multi-protein complexes, called exon junction complexes (EJCs) are deposited 20-24 nucleotides away from the exon-exon junction. Following splicing export factors, such as TAP, the metazoan orthologue of Mex67, to export transcripts from the nucleus. As in yeast, following export from the nucleus, Dbp5 remodels mRNP complexes as they exit the nuclear pore complex. In both yeast and metazoans, following export mRNA transcripts are now coated by a new set of proteins and are translated into protein (not shown).

cerevisiae, on the other hand contains a much more compact genome and while the proteins involved in splicing are highly homologous between humans and yeast, the number of yeast genes containing introns is extremely small. As few as 5% of yeast genes contain introns (10, 11).

In both yeast and metazoans, the mRNA transcript undergoes several additional processing steps. These include the addition of a 5'-7-methylguanosine cap on the 5'-end of the transcript and the addition of a polyadenosine tail on the 3'-end of the transcript. Interestingly, these two processing steps help to greatly enhance the stability of mRNA transcripts as the majority of RNA degradation pathways start degrading the RNA from one end or the other (12, 13). Specifically, the 5'-cap serves both a protective role against 5'→3' riboexonucleases (12) and a role in facilitating translation of the transcript in the cytoplasm (14). In all eukaryotes, the cap is bound by a conserved set of proteins called the cap-binding complex (CBC) that helps to facilitate these functions. As mentioned above, mRNA transcripts are also polyadenylated on the 3'-end of the transcript. In order to add the poly(A) tail to the 3'-end of mRNA transcripts, several evolutionarily conserved protein complexes recognize specific sequences in the 3'-end of the mRNA transcript. These complexes then cleave the transcript and add the poly(A) tail (15). The enzymes that recognize and cleave the transcript, as well as add the poly(A) tail, are conserved from yeast to humans. Once added to the transcript, the poly(A) tail is coated by poly(A) binding proteins. Again, these proteins are conserved from yeast to humans (16). Finally, the export factors which actively transport the mRNA out of the nucleus, the components of the nuclear pore complex, through which the mRNA transcript must pass to leave the nucleus, and the proteins that remodel the

RNA:protein complex in the cytoplasm are also well conserved among most eukaryotes. Given the extensive evolutionary conservation of the processes that produce mature mRNA, work from a number of model organisms has contributed to our current understanding of RNA biogenesis. Many interesting studies have exploited the single-celled budding yeast, *Saccharomyces cerevisiae* to define key mechanisms in the lifecycle of an mRNA transcript [reviewed by (17-19)].

The budding yeast *Saccharomyces cerevisiae* as a model system

Beyond the fact that the majority of the proteins involved in post-transcriptional gene expression are conserved, *S. cerevisiae* also provides several other advantages compared to traditional cell culture or animal models for the study of RNA export from the nucleus. Specifically, *S. cerevisiae* cells are very genetically malleable. For example, yeast cells can easily be transformed with extrachromosomal plasmid DNA in order to express a Green Fluorescent Protein (GFP) -tagged protein or in order to test the *in vivo* function of a mutant protein. Furthermore, yeast cells can also exist in either a haploid or diploid state, allowing for the easy combination of genetic mutations by mating. The haploid state of most laboratory yeast strains allows for the expression of phenotypes caused by recessive mutations; these phenotypes would normally not be visible in a diploid organism. In addition, a collection of ~4800 yeast strains exists in which each strain is individually deleted for one non-essential gene. This collection of yeast strains is an invaluable resource when analyzing a network of genetic interactions. Finally, yeast cells can also be grown economically in large quantities for use in large-scale

biochemical experiments. All these points contribute to the utility of *S. cerevisiae* as a model system to study mRNA processing and export from the nucleus.

The mRNA "assembly line"

Each of the mRNA processing events described above, including addition of a 5'-cap, the splicing out of introns, and 3'-end cleavage and polyadenylation, as well as mRNA export itself, is tightly regulated allowing for plasticity in modulating gene expression (4). Particular aspects of transcription, as well as each of these processing events, are coupled to mRNA export, making the entire process one dynamic assembly line from start to finish. This assembly line can produce thousands of copies of a single transcript throughout the lifetime of an organism or only several copies of a transcript once during the development of an organism. The complex regulatory system that dictates when and where a gene is expressed is controlled not only by a complicated series of transcriptional regulators, but also by an equally elaborate network of RNA binding proteins (Table 1.1) that associate with the mRNA transcript during post-transcriptional events. These mRNA binding proteins function in diverse processes ranging from splicing to cytoplasmic RNA localization, but collectively they dictate the fate of each transcript. Like any assembly line, each sequential step constitutes a function that is a necessary prerequisite for the next.

In order for a transcript to leave the transcription "assembly line" in the nucleus and be delivered to ribosomes in the cytoplasm, it must first recruit mRNA export factors.

Table 1: Protein factors implicated in mRNA export from the nucleus

S. cerevisiae protein	Higher eukaryotic orthologue	RNA accumulation in nucleus?[1]	Function in mRNA export	Other described functions
Hpr1	hHpr1	Poly(A) and heat shock RNA	Component of yeast/human TREX and direct recruiter of Mex67 (yHpr1)	Transcriptional Elongation (yTREX), Genome Stability (yHpr1)
Tho2	hTho2	Poly(A) and heat shock RNA	Component of yeast/human TREX	THO complex member
Mft1	----	Poly(A) and heat shock RNA	Component of yTREX	THO complex member
Thp2	----	Poly(A) and heat shock RNA	Component of yTREX	THO complex member
Tex1	hTex1	N/A[2]	Component of yTREX	THO complex member
----	Thoc5/ fSAP79/ FMIP	Heat shock RNA only[3]	Component of hTREX, required for export of heat shock RNA	Post-transcriptional regulation of genes involved in macrophage/adipocyte differentiation
----	Thoc6/ fSAP35	N/A[2]	Component of hTREX	----
----	Thoc7/ fSAP24	N/A[2]	Component of hTREX	----
Sub2	UAP56	Poly(A) and heat shock RNA	Yra1 adaptor protein	ATP-dependent helicase, function in splicing
Yra1	Aly / REF	Poly(A) and heat shock RNA	Mex67 adaptor protein	RNA annealing activity
Mex67: Mtr2	TAP:p15 / Nxf1:Nxt1	Poly(A) and heat shock RNA	Primary mRNA export factor; contacts mRNA (directly and via adaptors) and nucleoporins	Nuclear export of 60s ribosomes
Pcf11	Pcf11	Poly(A) RNA[4]	Yra1 adaptor protein	mRNA 3'-end formation
Rna14	CstF77	Poly(A) and heat shock RNA	----	mRNA 3'-end formation
Rna15	CstF64	Poly(A) and heat shock RNA	----	mRNA 3'-end formation
Hrp1	----	Heat shock RNA only[3]	----	mRNA 3'-end formation, mRNA transcript stability
Pap1	Pap1	Poly(A) and heat shock RNA	----	Poly(A) polymerase/ mRNA 3'-end formation
Nab2	ZC3H14	Poly(A) RNA[4]	Possible Mex67 adaptor protein	mRNA processing / 3'-end formation
Npl3	----	Poly(A) RNA only[3]	Mex67 adaptor protein	mRNA 3'-end cleavage and polyadenylation site selection, promotes pre-mRNA splicing
Dbp5/Rat8	hDbp5/ DDX19	Poly(A) and heat shock RNA	mRNP remodeling in the cytoplasm following nuclear export	Interacts with transcription and translation-termination factors
Gle1	hGle1	Poly(A) and heat shock RNA	Activator of Dbp5 (in combination with hNup214/yNup159)	Regulation of translation
Thp1	----	Poly(A) and heat shock RNA	Component of yTREX2 complex involved in nuclear export of mRNA	Regulation of genomic integrity
Sac3	GANP/Shd 1	Poly(A) and heat shock RNA	Component of yTREX2 complex involved in nuclear	Regulation of genomic integrity

			export of mRNA	
Sus1	DC6/Eny2	Poly(A) RNA[4]	Component of yTREX2 complex involved in nuclear export of mRNA, possible "bridge" protein between transcription and mRNA export	Component of SAGA histone modification complex, regulation of genomic integrity
Cdc31	CETN3	Poly(A) RNA[4]	Component of yTREX2 complex involved in nuclear export of mRNA	Duplication of microtubule-organizing centers

1 – In either yeast cells expressing mutant proteins or higher eukaryotic cells depleted for the specified factor, does poly(A) or heat shock RNA accumulate in the nucleus as detected by fluorescence *in-situ* hybridization (FISH)?
2 – N/A = information on RNA accumulation is, to the best of our knowledge, not available.
3 – As specified either poly(A) RNA or heat shock RNA accumulates in the nucleus, but not both.
4 – Mutants accumulate poly(A) RNA but, to the best of our knowledge, nuclear accumulation of heat shock RNA has not been tested.

Conventionally, export factors have been defined based on their capacity to bind both the RNA transcript and components of the nuclear pore complex (NPC) and they are, therefore, thought to actively escort mRNA transcripts through the NPC and out of the nucleus. One of the overarching themes in mRNA export is the use of adaptor proteins to recruit these export factors. Although export receptors can bind directly to mRNAs (20-22), their recruitment to mature export-competent mRNA transcripts is greatly enhanced when they are recruited via adaptor proteins (23, 24). These adaptor proteins are hypothesized to specifically recognize RNA sequences to signal that a particular processing step is complete and consequently that the transcript is competent for export to the cytoplasm. Many of these proteins function not only as adaptors but also as important components of other processes, such as splicing or 3'-end processing, hence these processes are "coupled" to mRNA export via these adaptor proteins. While many of the individual proteins involved in mRNA biogenesis are highly conserved from yeast to higher eukaryotes, the particular processes that are coupled to nuclear export are more divergent. In *S. cerevisiae*, the primary mRNA export factor, Mex67 and its heterodimeric partner Mtr2, appear to be recruited through a transcription- and 3'-end processing-dependent mechanism. In higher eukaryotes, however, the recruitment of the Mex67:Mtr2 orthologues, called TAP:p15 (or NXF1:NXT1), to mature mRNA transcripts appears to be dependent upon 5'-cap addition and splicing.

Coupling transcription to mRNA export from the nucleus

During transcription, prior to mRNA export, adaptor proteins are deposited along nascent transcripts. Recent studies have demonstrated that a multi-protein complex,

termed the TREX (transcription and export) complex, is assembled upon the nascent transcript during transcription and is a critical component in determining the efficiency of mRNA export from the nucleus (25-27). Once deposited, components of this complex recruit adaptor proteins to the newly synthesized transcript. The *S. cerevisiae* TREX complex consists of the mRNA export adaptor proteins, Sub2 and Yra1, and components of the THO complex (Hpr1, Mft1, Thp2, and Tho2) (25-28). Sub2, and its mammalian counterpart, UAP56, are putative ATP-dependent helicases that function in splicing and export (29-31). Yra1 (Aly/REF in higher eukaryotes) has also been implicated in pre-mRNA metabolism (32) and mRNA export (23, 33). The THO complex components, several of which are conserved in higher eukaryotes, are required for a wide variety of processes including transcriptional elongation (34, 35) and genome stability (34).

The *S. cerevisiae* THO component, Hpr1, is co-transcriptionally recruited to actively transcribed loci (28, 36) and directly contacts the mRNA export adaptor protein, Sub2 (36). Furthermore, components of the THO complex can be copurified with both Yra1 and Sub2 (28), although THO subunits do not directly interact with Yra1, suggesting that the interaction between THO and Yra1 is bridged by Sub2 (36). These data suggest a model (Figure 1.2) whereby THO components, specifically Hpr1, are recruited to actively transcribed loci and subsequently recruit Sub2 and Yra1. Yra1 can then serve as an adaptor protein for the primary *S. cerevisiae* mRNA export factor, Mex67 (23, 33). In order for this complicated recruitment scheme to function correctly, the transcript must undergo several "molecular wardrobe changes" to properly recruit and subsequently displace these adaptor proteins. For example, Sub2 and Mex67 both interact with the same domain of Yra1 (28), suggesting that these interactions with Yra1

are mutually exclusive. This exclusivity is necessary to displace Sub2 from the transcripts when Yra1 recruits Mex67 (Figure 1.2). Yra1 directly interacts with Mex67 but does not exit the nucleus (23), suggesting that before Mex67 can escort transcripts through the nuclear pore complex, Yra1 must first be removed.

Although Sub2 directly interacts with both Hpr1 and Yra1 and likely bridges this interaction (36), additional evidence suggests that both Yra1 and Mex67 can be recruited to mRNA transcripts via alternative mechanisms (Figure 1.3) (37). Notably, several studies have demonstrated that the UBiquitin Associated (UBA) domain of Mex67 interacts with the THO component, Hpr1 (38), suggesting that the THO complex can directly recruit Mex67 to mRNA transcripts independent of Yra1 and Sub2. In addition to THO complex-mediated recruitment, other RNA binding proteins can also recruit Mex67 to mRNA transcripts. In particular, the RNA binding protein Npl3 has been implicated in Mex67 recruitment to mRNA transcripts (39). Npl3 is an essential serine-arginine rich (SR) protein that is co-transcriptionally loaded onto nascent transcripts (40, 41) and is required for proper nuclear export of poly(A) RNA (42). Interestingly, Npl3 plays roles in both polyadenylation site choice (40, 43) and early recruitment of spliceosomal proteins to intron-containing transcripts (41), suggesting that these processes could be coupled to Mex67 recruitment.

In addition to the alternative routes for Mex67 recruitment to mRNA transcripts, Yra1, one of the principle Mex67 adaptors, may itself have additional adaptors beyond the commonly accepted Sub2 helicase. Specifically, inactivation of a component of the 3'-end cleavage machinery, Pcf11, causes an ~2-fold reduction in recruitment of Yra1 to

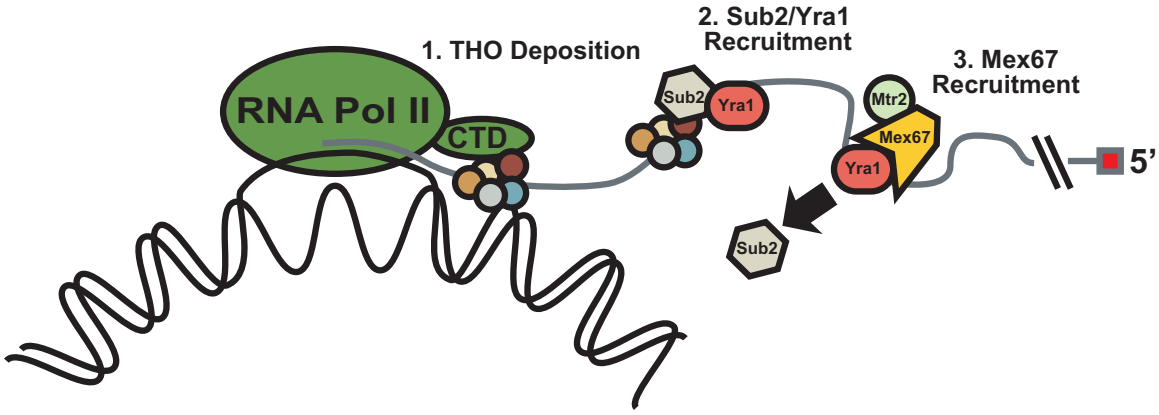

Figure 1.2: "Classical" model of Mex67 recruitment to mRNA transcripts. The principle mRNA export heterodimer, Mex67:Mtr2, is recruited via direct interactions with Yra1 in a Sub2- and THO complex-dependent manner. THO complex members are initially deposited upon nascent transcripts via interactions with the C-terminal domain (CTD) of RNA polymerase II (RNA Pol II). Following THO deposition, Sub2 and Yra1 are recruited to the transcript. Finally, the heterodimeric complex of Mex67:Mtr2 is recruited via interactions between Yra1 and Mex67. Binding of Mex67 to Yra1 displaces Sub2.

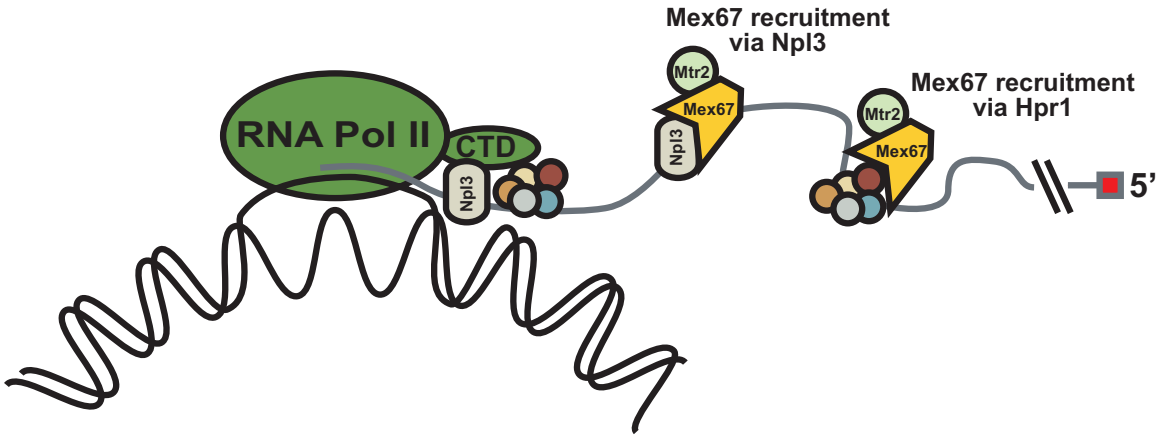

Figure 1.3: Sub2/Yra1- independent recruitment of Mex67 to mRNA transcripts. The heterodimeric export receptor, Mex67:Mtr2, can also be recruited to mRNA transcripts via direct interactions with both a component of the THO complex, Hpr1, and the RNA binding protein, Npl3.

the actively transcribed *PMA1* locus without affecting Sub2 recruitment. Yra1 interacts with components of the 3'-end cleavage machinery, including Pcf11 and Rna15 (37). Furthermore, Pcf11 and other components of the 3'-end processing machinery can be recruited to actively transcribed loci via interaction with the C-terminal domain of RNA polymerase II (44-46). Together, these data suggest a revised model (Figure 1.4) in which Yra1 can be recruited to actively transcribed loci in a Sub2-independent manner via an interaction with the 3'-end processing machinery component, Pcf11. Yra1 is then transferred from Pcf11 to the TREX complex via an interaction with Sub2. Yra1, now bound to the mRNA transcript, recruits the mRNA export heterodimer, Mex67:Mtr2 and the mature mRNA can exit the nucleus.

Even though millions of distinct transcripts are constantly transcribed, most current models for mRNA export suggest that one heterodimeric receptor, Mex67:Mtr2 (TAP:p15 or NXF1:NXT1 in higher eukaryotes), transports all transcripts through the nuclear pore complex into the cytoplasm. These models rely on the initial observation that Mex67:Mtr2 could bind to both RNA and nuclear pore components (20, 47). However, recent genome-wide studies raise the possibility that Mex67-dependent export is not the only route for mRNAs to exit the nucleus. One such study revealed that Mex67 and Yra1 were bound to 1,142 and 1,002 transcripts, respectively. Notably, only 349 transcripts were found in common among the pools of transcripts bound to each protein (48). This finding raises several interesting points. First, Mex67 and Yra1 each associated with distinct classes of RNAs (48), suggesting that expression of functionally related RNAs, such as those that encode proteins required for cell wall biosynthesis, can

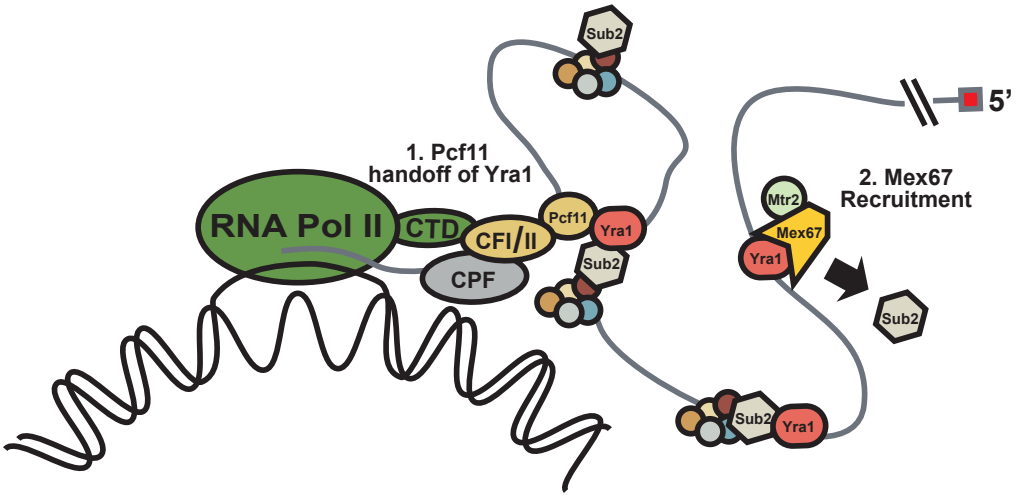

Figure 1.4. Revised "classical" model of Mex67 recruitment to mRNA transcripts. Mex67:Mtr2 is recruited to mRNA transcripts via interactions with 3'-end processing components. While initial studies demonstrated Yra1 was recruited to transcripts directly via THO components and Sub2, more recent data suggests a revised model where Yra1 is initially recruited to the mRNA transcript via interactions with the 3'-end processing factor, Pcf11 and subsequently transferred to the TREX complex via an interaction with Sub2. Yra1 then recruits Mex67 and transcripts are exported.

be post-transcriptionally coordinated by regulation of these RNA export factors (48). Interestingly, a separate genome-wide study investigating the transcripts associated with the *S. cerevisiae* RNA binding proteins, Npl3, Nab2, and Hrp1, demonstrated that these proteins preferentially associate with functionally distinct classes of RNAs (49), providing further evidence that particular RNA binding proteins associate with specific transcripts. Second, these genome-wide studies indicate that the Yra1 and Mex67-centric mRNA export model may not be applicable for all mRNA transcripts and suggest that other RNA binding proteins and protein complexes may actively facilitate mRNA export from the nucleus by interacting with both mRNA transcripts and nucleoporins.

Coupling 3'-end formation and mRNA export

Although there are some discrepancies between yeast and higher eukaryotes in 3'-end formation, the overall processes of cleavage and polyadenylation are remarkably conserved [reviewed by (50)]. The initial step in 3'-end formation involves the recognition of specific sequences within the pre-mRNA 3'- untranslated region (3'-UTR) and cleavage of the transcript. Once the transcript has been cleaved, poly(A) polymerase (PAP) adds the 200 to 250 adenosines (70 to 90 adenosines in *S. cerevisiae*) that comprise the poly(A) tail (15). In both budding yeast and mammalian cells, the poly(A) tail is bound by a complement of poly(A) binding proteins (Pabs) (51). Although coupling of mRNA export and 3'-end processing is somewhat controversial in metazoans, in *S. cerevisiae*, 3'-end processing has been more directly implicated in mRNA export from the nucleus. Notably, mutations within several yeast 3'-end processing factors, including Rna14, Rna15, Pcf11, and Pap1 cause accumulation of bulk

poly(A) RNA in the nucleus (52-54). In addition, mutations in numerous *S. cerevisiae* mRNA export factors, including Mex67, Yra1, and the cytoplasmic NPC-associated helicase Dbp5 (see below), cause hyperpolyadenylation of transcripts (52, 55). Components of the 3'-end processing machinery also genetically interact with mRNA export factors (56, 57).

Several studies in *S. cerevisiae* have investigated the link between 3'-end processing and mRNA export using reporter transcripts truncated by a hammerhead ribozyme rather than processed by the normal 3'-cleavage and polyadenylation machinery (58, 59). Hammerhead ribozymes are self-cleaving RNA sequences which were originally isolated from plant viruses (60). RNA transcripts synthesized by RNA polymerase II that contain a ribozyme sequence in lieu of a standard 3'-UTR are not polyadenylated efficiently and are also not efficiently exported from the nucleus (58). Export of these ribozyme truncated reporter RNAs is not entirely blocked however, as deletion of the gene encoding the cytoplasmic riboexonuclease, *XRN1*, results in an increase in cytoplasmic reporter RNA, suggesting that a fraction of these reporter transcripts exits the nucleus but is rapidly degraded due to the absence of a poly(A) tail (58). Interestingly, export of these reporters is rescued by an encoded stretch of adenosines immediately upstream of the self-cleaving ribozyme sequence that mimics a poly(A) tail. This result suggests that while the presence of a poly(A) tail helps facilitate mRNA export from the nucleus, the poly(A) tail is not an absolute requirement. Indeed, *TRP4* transcripts terminated at their 3'-ends by a hammerhead ribozyme can partially complement a *trp4* deletion mutant, indicating that a fraction of these non-polyadenylated transcripts can exit the nucleus and be translated (61). Whether these ribozyme-

terminated transcripts are exported by Mex67 and the canonical mRNA export machinery or by components of another RNA export pathway remains unclear.

A link between 3'-end formation and nuclear export: Poly(A) RNA binding proteins

Although 3'-end formation appears to be coupled to mRNA export from the nucleus in *S. cerevisiae* (52, 55, 57, 58), little information exists as to the actual physical link between components of the 3'-end processing machinery and mRNA export factors. One candidate class of proteins consists of the poly(A) binding proteins (Pabs). Pabs are conserved from yeast to higher eukaryotes and are important in the regulation of transcript polyadenylation, stability, translation, and nuclear export (62). The most well characterized *S. cerevisiae* Pab, Pab1, localizes to the cytoplasm at steady state and regulates both translation and mRNA stability (62). In addition, Pab1 shuttles into the nucleus (63, 64) and regulates poly(A) tail length (65, 66). This collection of observations, along with the fact that mutations within Pab1 only show limited effects on poly(A) RNA export (64), suggests that while Pab1 may enter the nucleus, its principle role is likely in the cytoplasm and not in coupling mRNA 3'-end formation to nuclear export.

More likely candidate proteins that couple 3'-end processing and mRNA export are nuclear poly(A) binding proteins, such as *S. pombe* Pab2 [PABPN1 in higher eukaryotes (62)] or *S. cerevisiae* Nab2 [ZC3H14 in higher eukaryotes (67)]. Although Pab2 and its orthologue PABPN1 bind specifically to polyadenosine RNA and modulate polyadenylation (68, 69), neither protein has been linked to mRNA export from the nucleus. A more likely candidate Pab that couples mRNA 3'-end processing to nuclear

export may be Nab2, which binds specifically to polyadenosine RNA *in vitro* (67, 70, 71) and also regulates poly(A) tail length (71, 72). Nab2 mutants also show nuclear accumulation of bulk poly(A) RNA (71, 73) and genetically interact with both Mex67 (57, 74) and Yra1 (75). Therefore, Nab2 could function as a factor involved in 3'-end formation that serves as an adaptor for Mex67 recruitment to export competent transcripts.

Coupling splicing and mRNA export from the nucleus

While both yeast and higher eukaryotes employ a conserved set of factors to facilitate the nuclear export of mRNA transcripts, the mechanisms that recruit these factors are somewhat divergent. In higher eukaryotes, where most transcripts are subject to splicing, mRNA export receptors seem to be recruited to the 5'-end of transcripts in a splicing- and 5'-cap dependent manner.

The human TREX complex contains the adaptor proteins UAP56 and Aly/REF, orthologues of budding yeast Sub2 and Yra1, respectively, as well as the human THO complex members, hHpr1, hTho2, Thoc5/fSAP79, Thoc6/fSAP35, and Thoc7/fSAP24 (76). Unlike their *S. cerevisiae* counterparts, human TREX (hTREX) constituents (including UAP56, Aly/REF, and THO members) have not been directly linked to transcription but instead have more directly been linked to the addition of the 5'-7-methyl guanosine cap and the splicing out of introns (7, 76-78). Early studies suggested that hTREX recruitment may be coupled to splicing since UAP56 copurified with spliceosomal proteins, specifically $U2AF^{65}$ (79, 80), and Aly/REF associates with the exon junction complex (EJC), a multi-protein complex deposited 20 to 24 nucleotides

upstream of exon-exon junctions (81). hTREX components also preferentially associate with mRNA transcripts that have undergone splicing rather than artificial transcripts manufactured from cDNA constructs (77). Furthermore, Aly and Thoc5 colocalize with nuclear speckles, which are sub-nuclear domains thought to store processing factors and components of the spliceosome (76, 78, 82). Together, these results suggest that hTREX is associated with spliceosomes and potentially is recruited as a part of the EJC.

Interestingly, more recent studies have shed some doubt on the idea that TREX is recruited to nascent transcripts as part of the spliceosome or the EJC in higher eukaryotes. In particular, both hTREX and Aly are recruited to the 5'-end of mRNA transcripts (76, 77). Several hTREX components, including Aly, UAP56, and hTho2, interact with the cap-binding complex, which specifically recognizes the 7-methyl guanosine cap on the 5'-end of mRNA transcripts (77, 83). In addition, recruitment of Aly and a component of the hTHO complex, hTho2, to decapped reporter transcripts is dramatically decreased compared to properly capped transcripts (77). Moreover, eIF4A3, a component of the EJC, is recruited to capped and decapped transcripts equally well (77), suggesting that TREX recruitment is dependent upon the 5'-cap, whereas recruitment of the EJC is not. As the addition of a 5'-cap dramatically increases the efficiency of nuclear export of spliced transcripts (77), recruitment of hTREX, and thus mRNA export receptors, is highly likely to be dependent upon both capping and splicing.

Translocation through the nuclear pore complex

Once an mRNA transcript has been properly processed, packaged, and has recruited the correct export receptor(s), the resulting mRNA ribonucleoprotein (mRNP) complex is

translocated through nuclear pore complexes to the cytoplasm. The nuclear pore complex consists of several classes of nucleoporins (Nups), including structural Nups and Nups containing domains rich with phenylalanine-glycine (FG) repeats. FG-Nups line the interior cavity of the NPC and allow for regulated macromolecular transport into and out of the nucleus (84). Multiple hypotheses exist (85) as to the exact mechanism by which nuclear pores maintain cargo selectivity while still retaining the capacity to transport cargoes efficiently and rapidly. Generally, these FG repeats are thought to extend into the central cavity of the NPC and form multiple low affinity interactions with soluble transport factors, such as Mex67/TAP (20, 47, 84) as they transit the NPC. Interestingly, recent work has demonstrated that different transport receptors (i.e. mRNA export vs. different pathways for protein import) may require different subsets of FG-Nups (21, 84, 86), suggesting different transport receptors may take different routes through nuclear pore complexes.

A molecular wardrobe change completes nuclear export

Throughout the assembly line of mRNA processing that culminates in export from the nucleus, a multitude of different proteins associate with the mRNA transcript (Figure 1.5). Initially, mRNA processing proteins are recruited to the nascent transcript during transcription via interactions with the C-terminal domain of RNA polymerase II. Many of these processing factors are displaced following completion of processing or prior to export from the nucleus. Export factors then recognize the mature transcripts and convey them through the nuclear pore complex to the cytoplasm. The export factors are

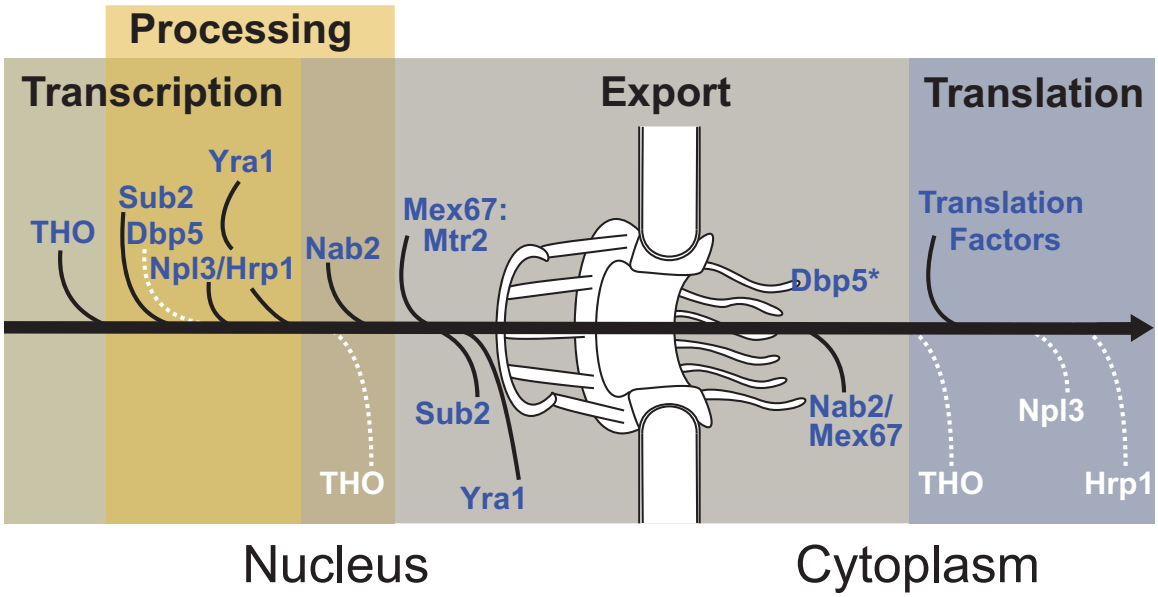

Figure 1.5. A timeline for the molecular displacements that occur in the course of mRNA export. In *S. cerevisiae*, adaptor proteins responsible for the recruitment of mRNA export receptors are deposited upon transcripts (denoted by black curved lines above the large directional arrow) during transcription and processing, coupling these processes to mRNA export. During transcription, components of the THO complex are initially deposited on the nascent transcript. Other processing factors and RNA binding proteins, such as Hrp1, Npl3, Sub2, Yra1, and Nab2 are subsequently recruited to the maturing transcript through a combination of interactions with THO components, the C-terminal domain of RNA polymerase II, and other mechanisms. The principle yeast mRNA export receptor heterodimer, Mex67:Mtr2, is subsequently recruited via interactions with adaptor proteins. Recruitment of Mex67:Mtr2 displaces Sub2 (denoted by blacked curved lines below the large directional arrow), eventually Yra1 is also displaced from the mRNA transcript, and the mRNP exits the nucleus through the nuclear pore complex. Once in the cytoplasm, the Dbp5 helicase remodels the mRNP, displacing export factors, such as Mex67 and Nab2, and subsequently allowing translation factors to bind to the transcript. As both Npl3 and Hrp1 associate with polyribosomes, the mechanism by which these proteins dissociate from the transcript is unclear (represented by white dashed lines below the large directional arrow). The mechanism and compartment of THO displacement also remains unclear, as the hTHO components shuttle between the nucleus and the cytoplasm but shuttling of *S. cerevisiae* THO components has not been reported.

subsequently displaced and factors that regulate the cytoplasmic destiny of the transcript bind. These cycles of protein displacement occur continually throughout the lifecycle of an mRNA transcript and help to functionally coordinate mRNA biogenesis. One of the best-characterized examples of this cycle of molecular displacement occurs immediately following translocation of the mRNP through the nuclear pore complex. Upon reaching the cytoplasmic side of the nuclear pore complex, the mRNP must undergo a significant remodeling event to replace nuclear export factors with a new complement of proteins that regulate the cytoplasmic fate of the transcript. For example, in *S. cerevisiae* the nuclear poly(A) binding protein Nab2 is not detected in polyribosomes (87) in the cytoplasm, suggesting that it most likely is removed and replaced by Pab1, the principle cytoplasmic poly(A) binding protein important for mRNA stability and translation efficacy (62).

One component of the machinery in *S. cerevisiae* responsible for mRNP reorganization upon entry into the cytoplasm is the RNA helicase, Dbp5 (88). Dbp5 (also known as Rat8) is conserved from yeast to higher eukaryotes (88, 89) and belongs to the family of DEAD-box RNA helicases, which unwind short stretches of double-stranded RNA or remodel RNA-protein interactions (90). Early work demonstrated that Dbp5 is localized to the cytoplasmic fibrils of the nuclear pore complex at steady-state (89) and is required for proper nuclear export of poly(A) RNA (88), hinting at a role for Dbp5 in the terminal stages of poly(A) RNA export as mRNPs exit the nuclear pore complex. More recent studies have corroborated that idea and provided new insight into the role of Dbp5 in poly(A) RNA export. During the final stage of nuclear export Dbp5 contacts its activator, the NPC-associated Gle1, as well as the small co-activator molecule, inositol

hexakisphosphate ($InsP_6$) (89, 91), leading to activation of Dbp5 at the cytoplasmic face of the nuclear pore complex. Once activated, Dbp5 facilitates the removal of mRNA export factors, including Nab2 and Mex67 (91, 92). Whether it is Dbp5 removing proteins from transcripts as they exit the NPC or other RNA helicases remodeling complexes during the splicing out of introns or other processing events (90), a collection of RNA helicases play critical roles in remodeling mRNP complexes throughout mRNA biogenesis.

The exosome contains both quality control and processing functions

Errors in many of the processing steps between the initial transcription of an mRNA and its final destination at the translation machinery in the cytoplasm can cause the production of erroneous transcripts, which, if translated, could be deleterious to the cell (18, 19). Importantly, the cell has evolved several quality control checkpoints that monitor the correctness of processing and prevent the accumulation of error-containing transcripts. Error-containing transcripts are normally targeted for degradation via the exosome, a multi-protein complex containing two active 3'→5' riboexonucleases, Rrp6 and Dis3/Rrp44 (93-96). Recent evidence has also demonstrated that a non-canonical poly(A) polymerase containing complex, called the TRAMP complex helps to facilitate the degradation of transcripts containing processing errors (18, 93, 97). The TRAMP complex contains several proteins, including one of two non-canonical poly(A) polymerases, Trf4 or Trf5, one of two RNA binding proteins, Air1 or Air2, and a putative RNA helicase, Mtr4 (18, 94, 97). The different combination of the TRAMP complex members most likely dictate the specificity of the complex to different RNA transcripts.

Together the proteins of the TRAMP complex recognize error-containing transcripts and add short poly(A) tails to their 3'-ends (18, 94, 97). These oligoadenylated transcripts are then recognized by the nuclear exosome and rapidly degraded (18, 94, 97).

Beyond the degradation of faulty transcripts, the exosome also trims the 3'-end of rRNAs, snoRNAs, and tRNAs (98, 99). These trimming activities are essential for the production of fully mature RNA transcripts. Interestingly, a recent study demonstrated that trimming was also facilitated by Trf4/5 and poly(A) polymerase dependent oligo-adenylation of these short RNA transcripts (100). Although at this time it is unclear, the cell must also have a mechanism for distinguishing between transcripts containing short poly(A) tails and destined for degradation and those containing short poly(A) tails and destined for mere trimming. Most likely, other RNA binding proteins bound to these transcripts dictate the eventual fate of these short RNAs.

Molecular recognition of RNA via multiple conserved domains

Throughout their lifecycles, mRNA transcripts are coated with a variety of RNA binding proteins. As mentioned above, the complement of RNA binding proteins associated with a transcript dictates the eventual fate of that transcript (4). At the heart of all these interactions lies specific molecular recognition of RNA sequences by RNA binding proteins. Only recently, however, have several of these RNA binding domains been structurally analyzed to reveal the mechanisms of RNA recognition. One challenge in the field is to structurally analyze not only the RNA binding protein but also the bound target sequence. The structural analysis of proteins bound to their specific RNA targets reveals important details about these interactions.

One of the best-characterized single-stranded RNA (ssRNA) binding domains is the RNA Recognition Motif (RRM). RRM domains are typically 80-100 residues in length and fold into a four-stranded antiparallel β-sheet packed against two α-helices (101). The three-dimensional structure of the β-sheets forms the RNA binding platform from which amino acid side chains extend to form specific protein:RNA interactions (101). These motifs are among the most abundant RNA binding domains, being present in about 0.5-1% of all human genes (102). Although all RRMs form a similar structural architecture to recognize ssRNA, these domains are present in many different proteins that recognize many different sequences (102). For example, the RRM of the human spliceosomal protein U1A binds to an exposed 7-nucleotide hairpin in the U1 snRNA (101) while the 4 tandem RRMs of the yeast Poly(A) binding protein, Pab1, specifically recognize the poly(A) tail of mRNA transcripts (103). The question remains as to how the same motif can recognize drastically different sequences. The answer lies in the fact that although the three-dimensional fold of the domain is very similar, the amino acids that comprise the domain are completely different. For example, the RRMs of human poly(A) binding protein, PABPC1, uses a combination of base-stacking interactions and hydrogen bonding to interact with adenines and 2'-OH groups of poly(A) RNA (101, 104). Several of these hydrogen-bonding interactions are highlighted in Figure 1.6 [adapted from (104), Protein data bank file ID = 1CVJ]. For example, in the crystal structure of PABPC1, the N6 nitrogen of adenine 7 (A7) forms a hydrogen bond with the carbonyl group of Asp 45. In addition, the double bonded oxygen in the side chain of Asn 105 forms a hydrogen bond with the hydrogens of the N6 nitrogen of adenine 2 (A2). Alternatively, the *Drosophila melanogaster* protein, sex-lethal (SXL), encodes an

RRM domain that specifically recognizes a sequence of pyrimidines by forming hydrogen bonds with the functional groups of the bases as well as with the sugar-phosphate backbone (101, 105). These differences illustrate the diverse nature and plasticity of RNA Recognition Motifs.

Although typically thought of as DNA binding motifs, zinc fingers are another prevalent RNA binding domain (106, 107). Zinc finger domains consist of cysteine (C) and histidine (H) residues, typically found in CCHH, CCCH, or CCHC arrangements, which chelate a zinc ion (106, 107). The amino acids between the structurally important cysteines and histidines are then correctly positioned in flexible loops to form specific interactions with RNA. The diversity of RNA targets recognized by this class of RNA binding proteins is astounding. Sequences and motifs ranging from AU-rich elements (AREs) in the 3'-UTR of certain mRNA transcripts (108) to specific hairpin loops within the 5s ribosomal RNA (rRNA) (109) are recognized by different varieties of zinc fingers. However, due to their flexible nature, zinc fingers are notoriously difficult to crystallize. Therefore, very little is known about the exact mechanisms of RNA recognition by zinc fingers. Importantly, several recent structural studies using NMR have shed new light on this mode of molecular recognition. One of these studies investigated the molecular recognition of AREs by the tandem CCCH-zinc fingers of human TIS11d (108). TIS11d is a member of a family of proteins that post-transcriptionally modulates gene expression by binding AU-rich elements in the 3'-UTR of specific transcripts and thereby controlling the stability of those transcripts. The NMR solution structure of TIS11d in complex with an AU-rich RNA oligonucleotide (Figure 1.7) revealed that ARE

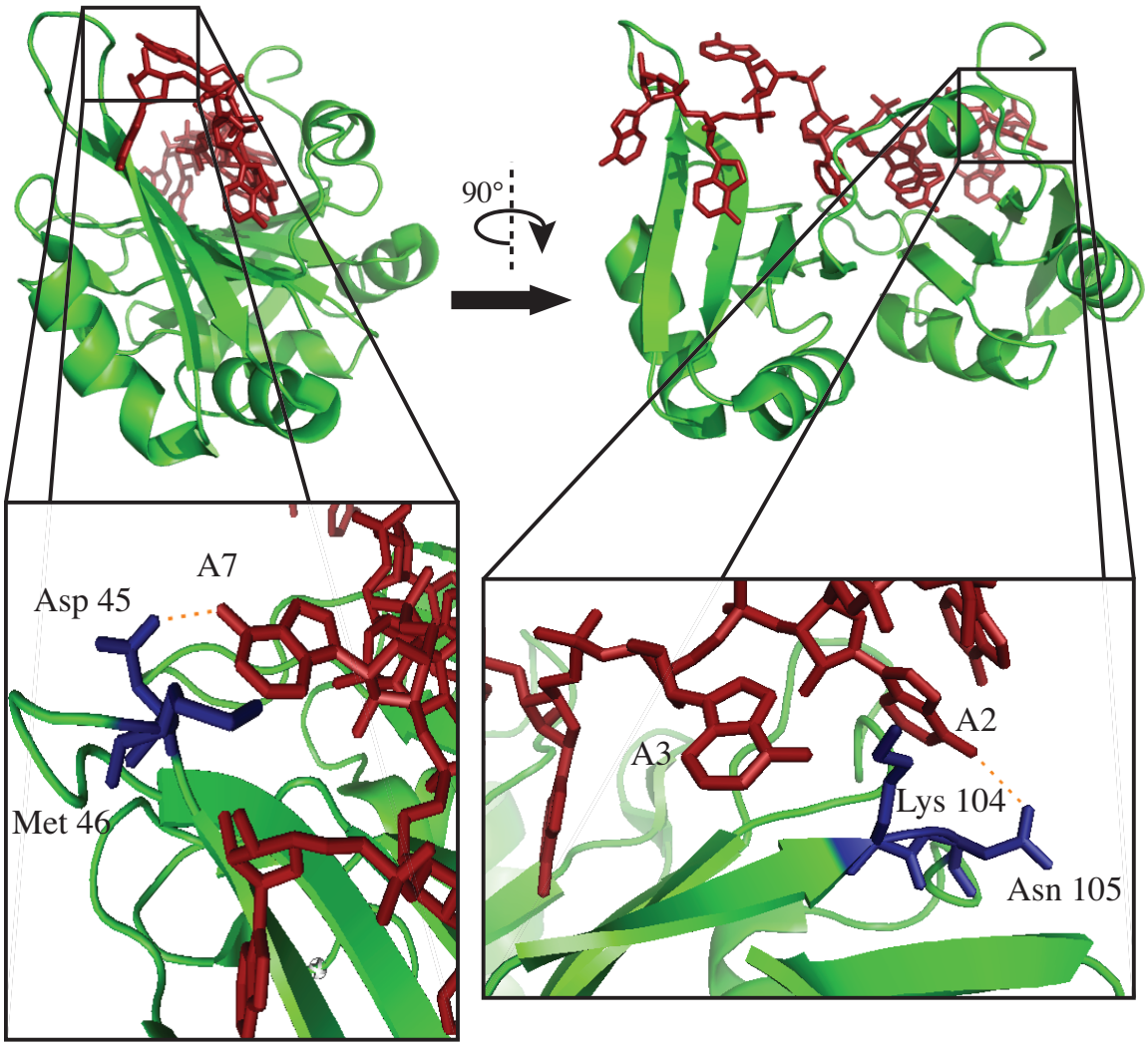

Figure 1.6: Recognition of polyadenosine RNA by the conserved RRMs of human PABPC1. This figure is adapted from the protein data bank file 1CVJ, published by Deo, *et al*. (104). RNA Recognition Motifs (RRMs) 1 and 2 from the human poly(A) binding protein, PABPC1, were co-crystallized with a poly(A) RNA oligonucleotide. The RRMs form the characteristic fold of 4 β-sheets packed tightly against two α-helices. Several hydrogen bonding interactions are highlighted. Specifically, The N6 nitrogen of adenine 7 (A7) forms a hydrogen bond (shown by an orange dotted line) with the carboxyl group of aspartic acid 45 (Asp 45 – shown in blue). Additionally, the double-bonded oxygen of asparagine 105 (Asn 105 – shown in blue) forms a hydrogen bond with the N6 nitrogen of adenine 2 (A2). Lysine 104 (Lys 104 – shown in blue) also interacts electrostatically with the phosphate backbone of the adenine 2.

recognition is mediated by base-stacking between conserved aromatic amino acids and specific bases within the oligonucleotide (Figure 1.6) as well as hydrogen bonding between the Watson-Crick edge of the bases and the protein backbone (108). Several other basic residues may ionically interact with RNA, but were unordered in the NMR solution structure (108).

In sum, a more thorough understanding of the mechanisms by which zinc fingers and other RNA binding domains specifically recognize their target RNAs is critical for our appreciation of the post-transcriptional regulation of gene expression.

The relevance of RNA binding proteins to disease

To underscore the importance of these post-transcriptional mechanisms of gene expression and the RNA binding proteins which facilitate them, several disease states have been identified in which mutations have been identified in RNA binding proteins critical for RNA biogenesis [Reviewed by (110)]. For instance, defects in the RNA binding protein, Quaking, have been implicated in neurological disorders including neuronal demyelination, ataxia, and schizophrenia [Reviewed by (111)]. All isoforms of Quaking encode a single KH (hnRNP K Homology) domain that has been shown to bind a defined recognition sequence found in over 1,400 different mRNA transcripts (112). These potential RNA targets are involved in a wide range of functions, from embryogenesis and cell adhesion to cell differentiation and cell growth (111, 112). Presumably, Quaking functions by controlling the stability or localization of its target RNA transcripts (111, 113).

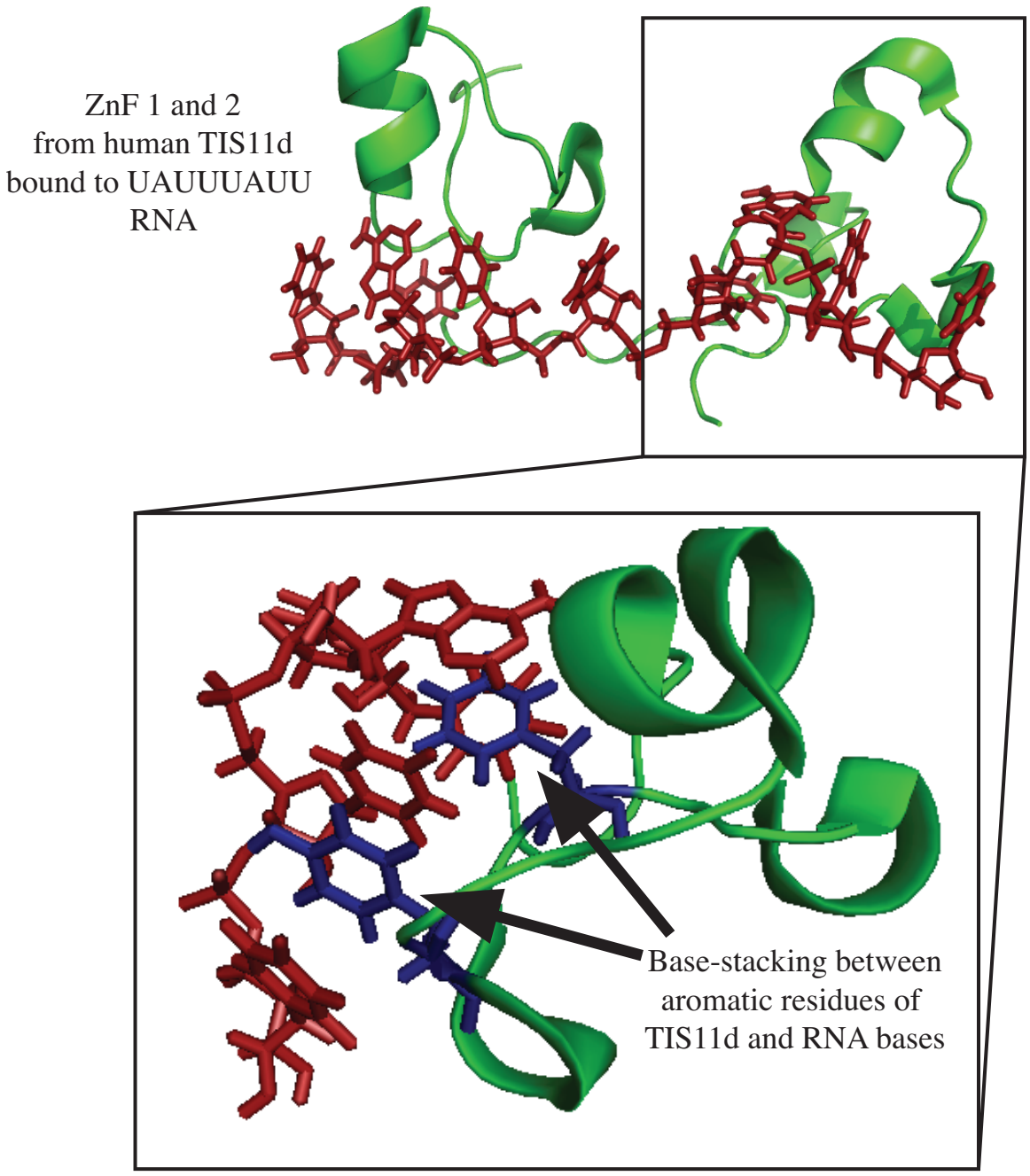

Figure 1.7: Recognition of AU-rich RNA sequences by the tandem CCCH zinc fingers of human TIS11d. This figure is adapted from the protein data bank file 1RGO, published by Hudson, *et al*. (108). The atomic resolution structure of the tandem zinc finger domain of TIS11d bound to an RNA oligonucleotide with the sequence UAUUUAUU was solved by NMR. Shown is the average of 20 solved structures generated using the average3d.py (114, 115) PyMOL script. A base stacking interaction between the UAUUUAUU RNA oligonucleotide and tyrosine 170 and phenylalanine 176 is highlighted.

In addition, Fragile X Syndrome, a common form of hereditary mental retardation is caused by defects in an RNA binding protein. A tri-nucleotide (CGG) expansion in the 5'-untranslated region (5'-UTR) of the *FMR1* gene (116) causes hypermethylation and gene silencing of *FMR1*. The gene product of *FMR1*, FMRP, seems to be responsible for the repression of translation in neuronal dendrites (116-119). Genome-wide identification of FMRP target RNAs will hopefully allow for a more thorough understanding of this disease.

Interestingly, several diseases have also been tied to poly(A) RNA binding proteins. An expansion within the nuclear poly(A) RNA binding protein, PABPN1, has been linked to oculopharyngeal muscular dystrophy (120, 121). PABPN1, which contains an RRM, specifically recognizes poly(A) RNA and stimulates poly(A) polymerase *in vitro*. Defects in another poly(A) binding protein (Pab), ZC3H14, have very recently been linked to a hereditary form of mental retardation (Andreas Kuss and colleagues, personal communication). In addition, preliminary evidence suggests that ZC3H14 may be over-expressed in certain breast cancer cell lines (Sara Leung and Callie Preast, unpublished data). Taken together, these links between defects in RNA binding proteins and disease demonstrate the importance of understanding the mechanisms by which RNA binding proteins post-transcriptionally regulate gene expression.

A brief summary of information known about Nab2 prior to this dissertation

The majority of the work presented in this dissertation precisely defines the polyadenosine RNA binding activity of the CCCH zinc fingers encoded by the yeast RNA binding protein, Nab2, and its human orthologue, ZC3H14. Prior to my work,

previous studies had defined domains of the Nab2 protein required for nuclear export of poly(A) RNA and for import of the Nab2 protein into the nucleus (71, 73, 122). These studies utilized Nab2 variants containing gross deletions of entire domains of Nab2 (Figure 1.8) and therefore were quite likely to disrupt the overall three-dimensional structure of the protein. Previous work also demonstrated that Nab2 could be UV-crosslinked to bulk polyadenylated RNA transcripts (70, 73) and that Nab2 bound poly(A) RNA oligonucleotides *in vitro* (70, 71). Little work, however, had been done to investigate the specificity of interaction between Nab2 and polyadenosine RNA. Previous studies had also demonstrated that Nab2 was required for poly(A) RNA export from the nucleus and that RNA isolated from cells expressing Nab2 mutants had longer poly(A) tails than RNA isolated from wild-type cells (71-73). Importantly, no orthologues of Nab2 had been identified in higher eukaryotes, such as humans, so Nab2 was often considered a budding yeast specific protein.

Scope and significance of the dissertation

Given all this information about Nab2, I hypothesized that Nab2 *specifically* recognized poly(A) RNA and that specific amino acid changes in Nab2 that disrupted RNA binding would result in defects in poly(A) tail length and/or poly(A) RNA export from the nucleus. Our hope was that these amino acid changes in the RNA binding domain of Nab2 would surgically impair one function without totally disrupting protein structure. In chapter 2, I present data demonstrating that the CCCH zinc finger domain of the *S. cerevisiae* protein, Nab2, is a novel polyadenosine RNA recognition

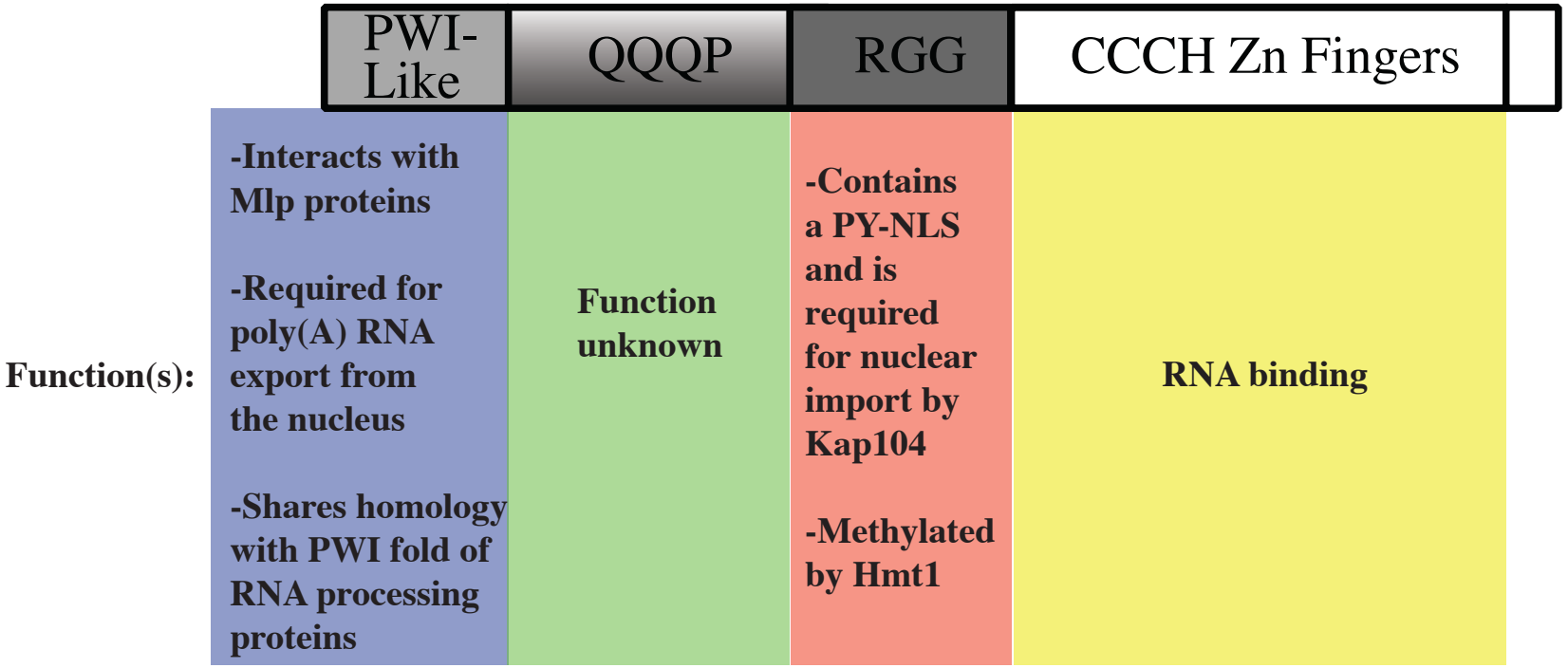

Figure 1.8: Domain diagram of the essential *S. cerevisiae* Nab2 protein. The essential Nab2 protein contains four domains. The N-terminal domain (residues 1-97) of Nab2 is required for poly(A) RNA export from the nucleus and also interacts with the myosin-like proteins, Mlp1 and Mlp2 (73, 123). Additionally, this domain folds into a PWI (phenylalanine-tryptophan-isoleucine)-like three dimensional structure (123). The function of the central glutamine-proline (QQQP) rich domain (residues 104-169) is unknown. The arginine-glycine-glycine (RGG) repeat domain (residues 201-265) is essential for the nuclear import of Nab2 by the importin, Kap104 (122). The C-terminal tandem CCCH zinc finger domain (residues 262-524) consists of seven tandem CCCH zinc fingers and is essential for polyadenosine RNA binding (67, 70, 71).

motif. All previously characterized poly(A) binding proteins (Pabs) recognize poly(A) RNA via RRMs, hence recognition of poly(A) RNA via CCCH zinc fingers constitutes a completely novel mechanism or molecular recognition. Importantly, the zinc finger domain of Nab2 is conserved between *S. cerevisiae* and humans. We also show that the human protein, ZC3H14, specifically interacts with polyadenosine RNA *in vitro*. While the Nab2 zinc finger domain contains seven tandem CCCH zinc fingers, ZC3H14 only contains five, suggesting that only a subset of Nab2 zinc fingers are required for polyadenosine RNA binding. Accordingly, in chapter 3, data is presented demonstrating that Nab2 zinc fingers 5-7 are sufficient for polyadenosine RNA recognition. Furthermore, I also demonstrate cysteine to alanine amino acid substitutions in the first cysteine of zinc fingers 5, 6, and 7 cause severe cold-sensitivity as well as extended poly(A) tails and accumulation of poly(A) RNA in the nucleus. Therefore, I conclude that the RNA binding activity of zinc fingers 5-7 is crucial for Nab2 function.

In order to gain an understanding of the molecular mechanism used by zinc fingers to specifically recognize polyadenosine RNA, we decided to use a structure/function approach to investigate the interaction between Nab2 and poly(A) RNA. In collaboration with Christoph Brockmann, a post-doctoral researcher in Murray Stewart's laboratory at MRC Cambridge, an atomic resolution structure of Nab2 zinc fingers 5-7 was solved using NMR. Critical analysis of the structure of Nab2 zinc fingers 5-7 revealed several conserved aromatic and basic residues that were solvent exposed and that might play roles in polyadenosine RNA binding. These predictions were tested by changing these residues to alanine either individually or in combination and testing the effects of these substitutions on RNA binding *in vitro*. From these studies, we conclude

that several specific aromatic and basic residues mediate the interaction between Nab2 and polyadenosine RNA.

Given the intimate relationships between Nab2 and the poly(A) tail of RNA transcripts, we hypothesized that Nab2 may genetically interact with components of the RNA processing machinery. In Chapter 4, data is presented demonstrating that Nab2 mutants containing amino acid substitutions that disrupt polyadenosine RNA binding genetically interact with these components, including proteins involved in the 3'-end formation of mRNA transcripts and components of the nuclear exosome. In particular, deletion of the nuclear exosome component Rrp6 suppresses deletion of the essential *NAB2* gene.

In sum, the data presented here demonstrate several important findings: 1) A subset of the CCCH zinc fingers specifically recognizes polyadenosine RNA; recognition of poly(A) RNA by anything other than an RNA Recognition Motif is completely novel 2) Specific aromatic and basic residues are involved in this interaction 3) Higher eukaryotic orthologues of Nab2 also contain this conserved CCCH zinc finger motif and specifically interact with polyadenosine RNA 4) Specific amino acid changes within the CCCH zinc fingers cause disruption of binding, growth defects *in vivo*, long poly(A) tails, and accumulation of poly(A) RNA in the nucleus of yeast cells and 5) Nab2 mutants genetically interact with components of the RNA processing machinery. Overall, these studies have provided us with a critical insight into the mechanism by which CCCH zinc fingers specifically recognize polyadenosine RNA. Furthermore, these studies reveal a new protein family that specifically recognizes polyadenosine RNA and plays a critical role in the post-transcriptional regulation of gene expression.

Chapter 2: Recognition of Polyadenosine RNA by a Conserved Family of Zinc Finger Proteins

This chapter is adapted from the following published paper:

Kelly, S.M.*, Pabit, S.*, Kitchen, C.M., Guo, P., Marfatia, K.A., Murphy, T., Corbett, A.H., and Berland, K.M. *Recognition of Polyadenosine RNA by Zinc finger proteins*. (2007) PNAS 104(30): 12306-11

*These authors contributed equally to this work

Fluorescence Correlation Spectroscopy (FCS) measurements were performed by Suzette Pabit, Ph.D. and Keith Berland, Ph.D. in the Department of Physics at Emory University

Introduction

The fate of an mRNA transcript is largely determined by its associated RNA binding proteins. A wide variety of RNA binding proteins associate with the nascent transcript co-transcriptionally and act as processing factors involved in capping, splicing, cleavage, and polyadenylation of the transcript (124). Additional RNA binding proteins package the mRNA into complexes that regulate transcript stability (1), promote export from the nucleus (2), and modulate translation (3). Accordingly, the protein constituents of these mRNA ribonucleoprotein (mRNP) complexes have been accurately equated to post-transcriptional activators and repressors (4, 5) of gene expression.

One family of proteins that are key post-transcriptional regulators of gene expression is composed of the poly(A) binding proteins (Pabs). Functional studies in a wide variety of organisms ranging from yeast to humans have demonstrated that members of this evolutionarily conserved protein family (reviewed by (62, 103, 125)) directly contact the poly(A) tail of mRNA transcripts to regulate transcript polyadenylation (65, 69), translation (126-130), stability (131, 132), and possibly nuclear export (63, 64). All known Pab family members specifically bind to poly(A) RNA via at least one RNA Recognition Motif (RRM) (133, 134). For example, the primary cytoplasmic *Saccharomyces cerevisiae* Pab, Poly(A) Binding protein 1 (Pab1), contains four RRMs that can each bind RNA with varying specificity and affinity (133-135).

Although all conventional Pabs interact with RNA through RRM domains (136), there is evidence to suggest that at least one other type of RNA binding motif may confer specific binding to polyadenosine RNA (49, 70, 71, 103). The yeast protein, Nuclear poly(A) Binding Protein 2 (Nab2), lacks RRM domains and instead contains two other

potential RNA binding motifs, an Arginine-Glycine-Glycine (RGG) repeat domain and seven tandem CCCH zinc fingers (70, 73). Nab2 was originally identified as an essential heterogeneous nuclear ribonucleoprotein (hnRNP) that co-purified with polyadenylated RNA transcripts (70). Subsequent studies revealed that Nab2 shuttles between the nucleus and the cytoplasm and is required for both nuclear export and proper polyadenylation of mRNA transcripts (71, 73).

The original co-purification of Nab2 with polyadenylated RNA transcripts merely indicated that Nab2 associates with RNA transcripts that contain poly(A) sequences and did not provide any information about the sequence specificity of this class of zinc finger proteins. Further characterization revealed that Nab2 bound to homopolymeric RNA and single-stranded DNA (70, 71). Domain analyses also suggested that the zinc finger domain could confer binding to poly(A) sepharose (70). A later study that purified several yeast hnRNPs and analyzed the co-purified RNA for consensus binding motifs revealed a Nab2 consensus of $(A)_{11}G$ (49), but did not directly examine binding specificity using purified Nab2 protein. Thus although Nab2 association with poly(A) sequences has been observed, the specificity of this interaction has not been thoroughly examined. Given that Nab2 modulates poly(A) tail length *in vivo*, specific recognition of polyadenosine could be a key aspect of Nab2 function. Taken together, these results suggest that Nab2 could be a member of a new class of poly(A) specific RNA binding proteins that recognizes poly(A) sequences in an RRM-independent manner.

To directly test whether a protein that lacked an RRM domain could bind specifically to polyadenosine RNA, we exploited a combination of conventional gel shift assays and fluorescence correlation spectroscopy (FCS) (137-140) to measure the

interaction between Nab2 and a variety of oligonucleotides *in vitro*. FCS measures the translational diffusion of fluorescently-labeled oligonucleotides in solution and can distinguish between rapidly diffusing free oligonucleotides and the more slowly diffusing fluorescent oligonucleotides bound to protein (139, 140). The relative concentration of bound and free oligonucleotides can be recovered from these measurements, allowing for the direct determination of binding constants. Results of these studies indicate that Nab2 binds with nanomolar affinity to fluorescently-labeled poly(A) RNA oligonucleotides. We have also investigated the specificity of this interaction through a series of competition experiments and find that Nab2 specifically binds to polyadenosine RNA as compared to other RNA or DNA sequences. Importantly, domain analyses reveal that the zinc finger domain of Nab2 mediates this sequence-specific RNA binding. To extend this study, we provide the first characterization of a human protein, ZC3H14 (zinc finger protein with CCCH motif #14), which contains CCCH zinc fingers similar to those found in Nab2. This zinc finger protein also specifically binds to polyadenosine RNA. Thus, our studies provide the first evidence to support the existence of evolutionarily conserved zinc finger polyadenosine RNA binding proteins.

Results

In order to test the hypothesis that a protein lacking an RRM domain could specifically bind polyadenosine RNA, we performed *in vitro* FCS-based binding experiments with purified, recombinant Nab2 and a Cy3-labeled 25-nucleotide (nt) poly(A) RNA oligonucleotide [Cy3-poly$(rA)_{25}$]. As described in Materials and Methods, a sample of concentrated Nab2 and Cy3-poly$(rA)_{25}$ was prepared (Figure 2.1). Nab2 was then

serially diluted while the concentration of Cy3-poly$(A)_{25}$ RNA remained constant. FCS measurements were taken for each concentration of Nab2, resulting in a series of autocorrelation curves (Figure 2.2A). As Nab2 is serially diluted, a smaller fraction of the Cy3-poly$(rA)_{25}$ is bound to the protein, resulting in shorter average diffusion times and the corresponding leftward shift of the correlation curves. FCS analysis of free oligonucleotide in solution and oligonucleotide bound to Nab2 yields diffusion coefficients of 1.25×10^{-6} and 4.5×10^{-7} cm^2/s for free RNA and Nab2-bound RNA, respectively. Global fitting of correlation curves to a multi-component diffusion model (Equation [1]) using these recovered diffusion coefficients returns the bound and free concentrations of Cy3-poly$(rA)_{25}$ at each protein concentration. The recovered concentration dependence of the protein-bound fraction of Cy3-poly$(rA)_{25}$ is shown in Figure 2.2B. Using a least-squares fitting routine, these data were fit to Equation [2] to recover an average value for the dissociation constant, K_d, of 29 ± 10 nM. This value is consistent with a previous study that examined GST-Nab2-His6 binding to poly$(A)_{25}$ RNA using a filter binding assay and measured a K_d of ~10 nM (71). As a control, titration experiments with both Pab1, a known yeast poly(A) binding protein (103), and ovalbumin, which is not expected to interact with nucleic acids, were performed (Figure 2.3). The Pab1 data yields a slower diffusion time ($D = 2.7 \times 10^{-7}$ cm^2/s) due to the interaction of Pab1 with the fluorescent oligonucleotides. In contrast, FCS experiments with ovalbumin are indistinguishable from pure oligonucleotide in buffer even at high protein concentrations (2.5 μM) indicating no interaction.

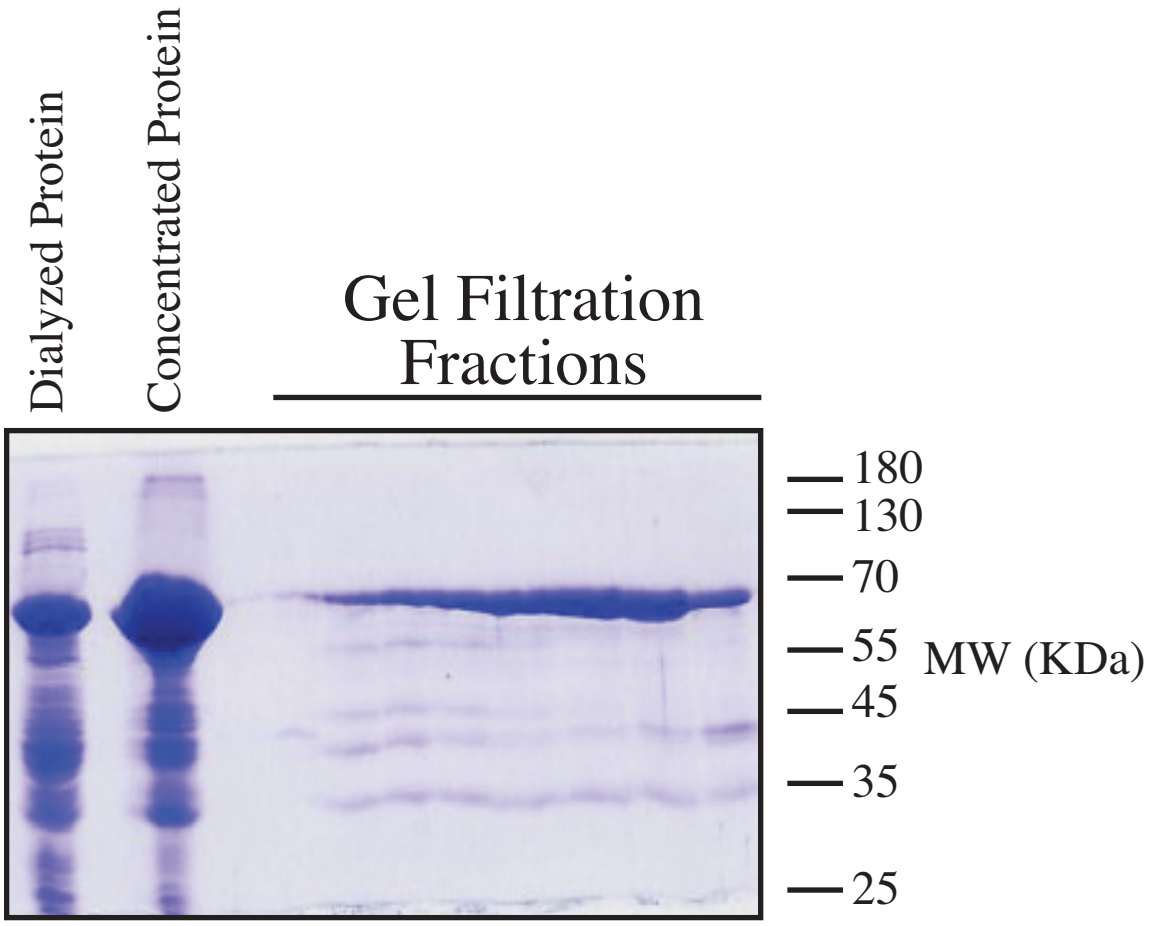

Figure 2.1: Purification fractions of untagged recombinant Nab2. Untagged recombinant Nab2 was purified from *E. coli* via ion exchange and gel filtration chromatography. Pooled Nab2 protein eluted from a HiTrap Q ion exchange column that has been dialyzed overnight into gel filtration buffer (Buffer C – See Experimental Procedures) is shown (Dialyzed Protein). Following dialysis, protein is concentrated using ammonium sulfate precipitation (Concentrated Protein) and fractionated using a Superdex 200 gel filtration column (Gel Filtration Fractions). Protein eluted from the gel filtration column is pooled and used for the subsequent binding experiments.

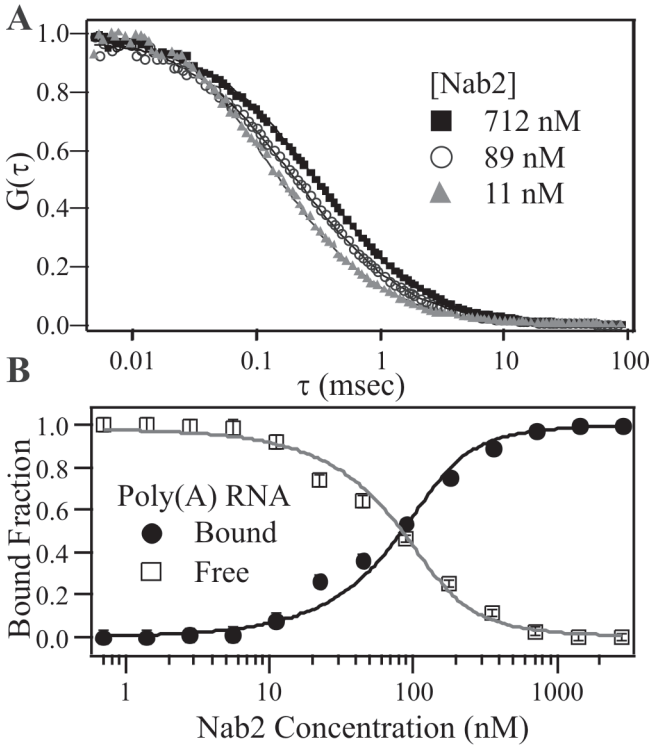

Figure 2.2: Nab2 binds polyadenosine RNA with high affinity. ***(A)*** Representative normalized FCS curves from a binding titration experiment where a concentrated sample of Nab2 (2.5 μM) was serially diluted while the Cy3-poly(A)$_{25}$ RNA concentration (~140 nM) remained constant. As Nab2 is diluted the FCS decay curves shift to the left, indicating a decrease in the fraction of bound oligonucleotide. ***(B)*** The FCS decay curves from *(A)* were used to determine the fraction of Nab2-bound Cy3-poly(A)$_{25}$ RNA. Data was fit with Equation (1) using global fit analysis to yield a K_d of 29 ± 10 nM.

To analyze the sequence specificity of the Nab2 interaction with polyadenosine RNA, a series of RNA competition experiments was performed (Table 2.1). For these experiments, a sample containing Nab2 and Cy3-poly$(A)_{25}$ RNA oligonucleotide was incubated with increasing amounts of a non-fluorescent competitor oligonucleotide. Oligonucleotides that compete with the Cy3-RNA for binding to Nab2 will displace the fluorescent RNA from the protein, resulting in faster average diffusion times and smaller bound fractions of Cy3-poly$(rA)_{25}$ in FCS measurements. Oligonucleotides that do not compete for binding do not produce any change in measured diffusion rates or bound fraction.

We first tested whether an unlabeled 25-nt poly(A) RNA oligonucleotide could efficiently compete with Cy3-poly$(A)_{25}$ RNA for binding to Nab2. As shown in Figure 2.4A, unlabeled poly$(rA)_{25}$ competes efficiently for binding to Nab2. The amount of competitor needed to displace 50% of the bound fluorescent oligo (IC_{50}) was determined by fitting the competitor concentration dependence of the bound fraction of Cy3-RNA to Equation [3]. Once the IC_{50} has been determined, the K_i of the competitor oligonucleotide can then be computed using Equation [4]. In several independent experiments, the K_i calculated for the unlabeled poly(A) RNA oligonucleotide (K_i=33 ± 12 nM) was virtually identical to the K_d calculated for the labeled poly$(A)_{25}$ RNA oligonucleotide (K_d=29 ± 10 nM). This analysis confirms that FCS-based competition assays can be used to assess binding to unlabeled oligonucleotides. Furthermore, these results demonstrates that the Cy3 label appears to have only minimal impact on the binding of the poly(A) oligonucleotide to Nab2.

Table 2.1. Summary of K_d and K_i values

Substrate/Competitor	K_d **(nM)**	K_i **(nM)**
Poly(A) RNA	29±10	33±12
Random ssDNA	ND*	770±320
Poly(A) ssDNA	400±170	690±270
Poly(N) RNA	ND	No comp[†]
UAUU RNA	ND	No comp

*ND – Not Determined
[†] - No Competition Observed

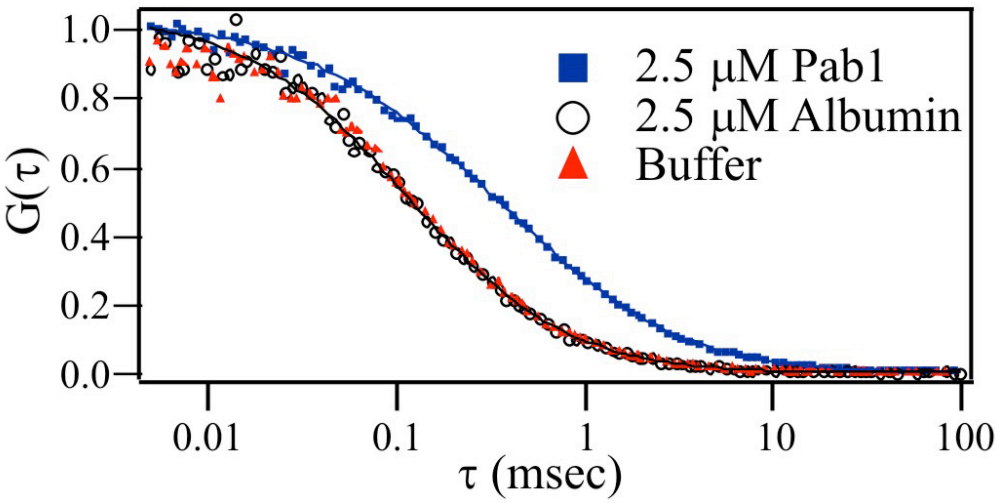

Figure 2.3: Pab1 binds poly(A) RNA. FCS control experiment showing a slower-decaying correlation curve for Cy3-labeled poly(A)-RNA oligonucleotide bound to poly(A) binding protein, Pab1, compared to FCS decay curves for Cy3-labeled poly(A) oligonucleotide in solution with a non-nucleic acid binding protein, ovalbumin. The Cy3-labeled poly(A) oligonucleotide in buffer alone is shown as a control.

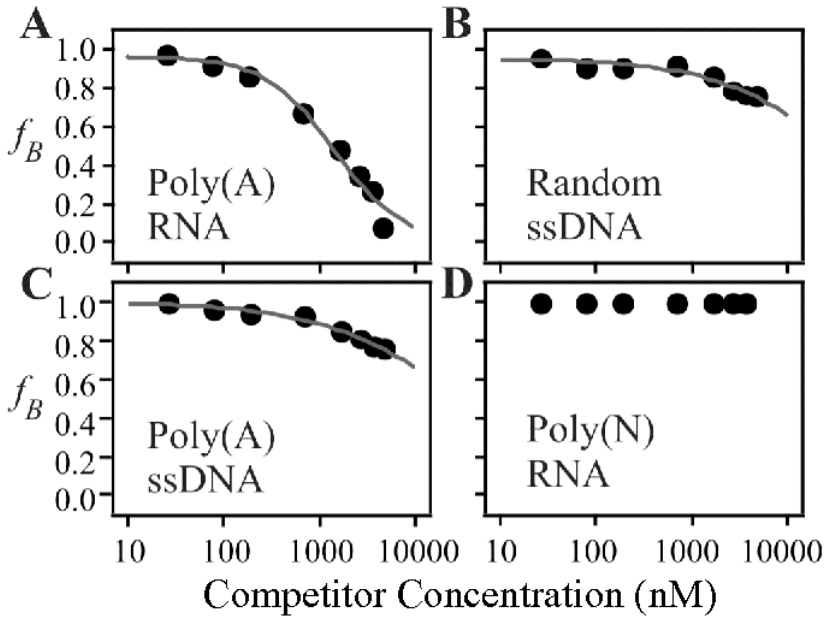

Figure 2.4: Nab2 binds preferentially to polyadenosine RNA. Nucleic acid binding properties of Nab2 investigated by competition experiments. Nab2 was incubated in binding buffer with Cy3-poly(A)$_{25}$ RNA and increasing amounts (up to 5 μM) of an unlabeled 25-nt competitor oligonucleotide: ***(A)*** poly(A) RNA; ***(B)*** random sequence ssDNA (CTTCTCTAGTTCAATCTTAGCATCG); **(*C*)** poly(A) DNA; or **(*D*)** poly(N) RNA (a pool of random 25-nt RNA oligonucleotides). Unlabeled poly(A)$_{25}$ RNA competes for binding to Nab2. Both DNA oligonucleotides showed very limited competition. No competition was observed using poly(N)$_{25}$ RNA oligonucleotide.

To assess the binding specificity of the zinc finger-containing protein, Nab2, we tested the ability of a 25-nt single stranded DNA oligonucleotide (Random ssDNA) to compete with Cy3-labeled poly$(A)_{25}$ RNA oligonucleotide for binding to Nab2. As illustrated by Figure 2.4B, the DNA oligonucleotide showed only minimal competition at very high concentrations of unlabeled competitor. This result indicates that Nab2 does not bind indiscriminately to nucleic acids. To determine whether Nab2 preferentially binds to RNA or single-stranded DNA, we analyzed binding of Nab2 to Cy3-labeled poly$(A)_{25}$ RNA when unlabeled poly$(A)_{25}$ DNA was added as competitor (Figure 2.4C). Again, only minimal competition was observed suggesting that Nab2 preferentially binds RNA rather than DNA. Since some weak non-sequence specific binding of DNA oligonucleotides to Nab2 was observed (Figure 2.4B,C), we used FCS to directly examine Nab2 binding to single-stranded DNA using a 25-nt Cy3-labeled poly(A) DNA oligonucleotide (Figure 2.5). Results of this analysis reveal that the interaction between Nab2 and DNA is highly dependent on salt concentration. In agreement with the competition experiments, Nab2 binding to Cy3-poly$(A)_{25}$ DNA was observed in buffer containing 50 mM NaCl (K_d = 400 ± 170 nM). However, in buffer containing 100 mM NaCl only very weak binding to DNA could be detected and the K_d was too weak to be determined. In contrast, binding to Cy3-poly$(A)_{25}$ RNA was virtually identical in buffer containing 50 mM (K_d = 29 ± 10 nM) or 100 mM (K_d = 39 ± 16 nM) NaCl. Thus, Nab2 does display some weak binding to single-stranded DNA, but it displays a much stronger affinity for poly(A) RNA.

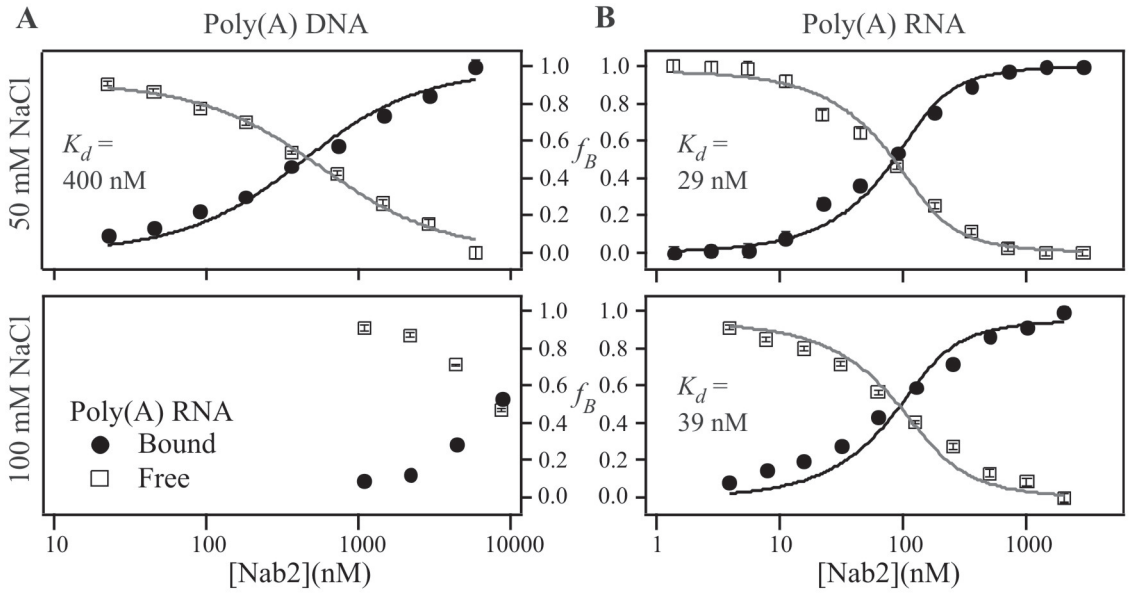

Figure 2.5: Nab2 preferentially binds RNA compared to DNA. FCS titration experiments were utilized to probe the ionic concentration dependence of Nab2 binding to either Cy3-poly(A) DNA or Cy3-poly(A) RNA. ***(A)*** Purified recombinant Nab2 was incubated with a 25-nt Cy3-poly(A) DNA oligonucleotide in binding buffer containing 50mM NaCl. Nab2 binds Cy3-poly(A) DNA with a K_d of 400 ± 170 nM in binding buffer containing 50 mM NaCl. ***(B)*** Nab2 binds Cy3-poly(A) RNA with a K_d of 29 ± 10 nM in binding buffer containing 50 mM NaCl. ***(C)*** Nab2 binding to DNA is highly dependent upon salt concentration. Nab2 binding to a Cy3-poly(A) DNA oligonucleotide is to weak to determine a K_d when incubated in binding buffer containing 100 mM NaCl. ***(D)*** Nab2 binding to poly(A) RNA is not dependent on salt concentration. Nab2 binds a Cy3-poly(A) RNA oligonucleotide with a K_d of 39 ± 16 nM in binding buffer containing 100 mM NaCl.

To determine whether Nab2 binds in a sequence-nonspecific manner to RNA, we investigated the ability of an unlabeled poly$(N)_{25}$ RNA competitor oligonucleotide to compete for Nab2 binding. The poly(N) RNA sample consists of a pool of randomized 25-nt RNA oligonucleotides. Upon addition of increasing amounts of poly$(N)_{25}$ RNA , no competition was detected (Figure 2.4D). To determine whether Nab2 binds polyadenosine RNA specifically or merely stretches of polypurine, we used an RNA gel shift assay to determine whether a 25-nt unlabeled poly(G) or poly(AG) competitor oligonucleotide could compete with poly$(A)_{25}$ RNA for binding to Nab2. As indicated by the shift from free probe to bound complex, Nab2 binds to a radioactively-labeled 25-nt poly(A) RNA oligonucleotide (Figure 2.6). This binding is specific for poly(A) since unlabeled poly$(A)_{25}$ RNA oligonucleotide, but not poly(G) or poly(AG) RNA oligonucleotides, compete for binding to Nab2. These results further strengthen the argument that Nab2 is a specific polyadenosine RNA binding protein.

Nab2, unlike other poly(A) binding proteins, lacks an RRM RNA binding domain but instead contains two other domains, an RGG domain (141) and seven tandem CCCH zinc fingers (107), previously implicated in RNA binding. Hence, specific binding of either domain to poly(A) RNA constitutes a fundamentally different mechanism for polyadenosine RNA recognition than has been previously characterized. To determine which domain of Nab2 confers specific poly(A) RNA binding, we employed an RNA gel shift assay. Both full length Nab2 (Figure 2.7A) and the zinc finger domain alone (Figure 2.7B) bind to a radioactively-labeled 25-nt poly(A) RNA oligonucleotide specifically since unlabeled poly$(A)_{25}$ RNA competitor but not unlabeled poly$(N)_{25}$ RNA

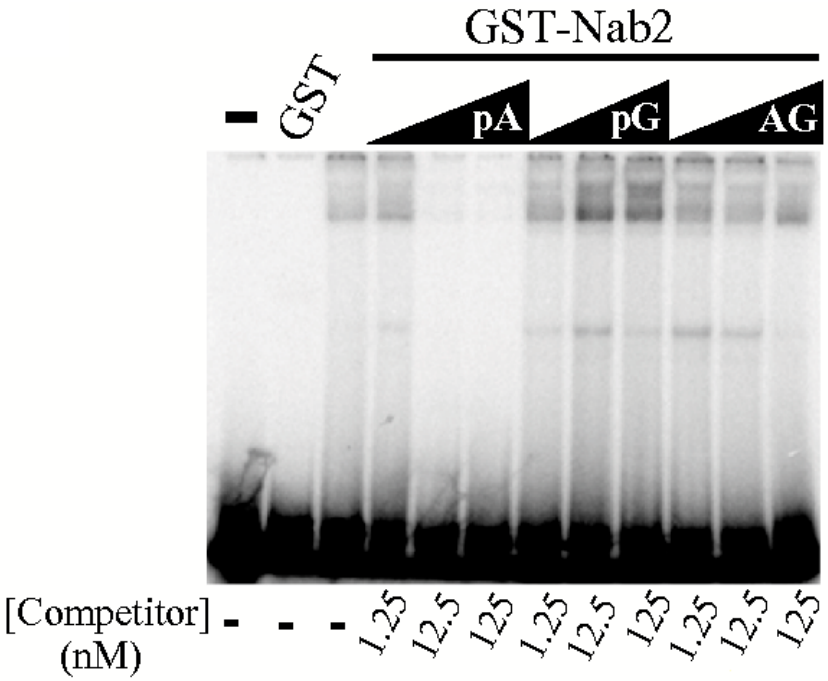

Figure 2.6: Nab2 binds preferentially to polyadenosine RNA vs. other polypurine rich RNA sequences. GST-Nab2 (50 nM) was incubated with a radioactively-labeled poly$(A)_{25}$ RNA oligonucleotide probe (~30 pM) and increasing amounts of unlabeled poly$(A)_{25}$, poly$(G)_{25}$, or poly$(AG)_{12}$ RNA competitor were added as indicated. RNA-protein complexes were then resolved from free probe by electrophoresis on a 5% non-denaturing polyacrylamide gel. No significant competition for Nab2 binding was observed upon addition of either poly$(G)_{25}$ or poly$(AG)_{12}$ RNA competitor oligonucleotide.

competitor can compete for binding. In contrast, a C-terminal truncation of Nab2 lacking the zinc fingers but still containing the RGG domain shows no binding to poly$(A)_{25}$ RNA (Figure 2.7A – GST-ΔCT).

To further probe the role of zinc fingers in mediating specific binding to polyadenosine RNA, we exploited a Nab2 variant, C437S, which contains a single conservative cysteine to serine amino acid change in the first cysteine of the sixth zinc finger. Since the last three zinc fingers of Nab2 have been implicated in Nab2 cross-linking to polyadenylated RNA transcripts *in vivo* (73), we predict that this substitution should disrupt the sixth zinc finger and alter the RNA binding properties of Nab2. To test this prediction, we used FCS to compare the binding affinity of wildtype Nab2 and C437S Nab2 for Cy3-poly$(rA)_{25}$ (Figure 2.6C,D). This analysis yielded a binding affinity of C437S Nab2 for Cy3-poly$(A)_{25}$ RNA (K_d =150 ± 40 nM) that is almost four-fold weaker than the affinity of wildtype Nab2 for Cy3-poly$(A)_{25}$ RNA (K_d = 39 ± 3 nM). Together, these gel shift and FCS experiments establish that a functional Nab2 zinc finger domain is both necessary and sufficient to confer preferential binding of Nab2 to polyadenosine RNA compared to random RNA.

Zinc fingers are common RNA binding motifs found in many proteins (107). One family of proteins that contain two CCCH zinc fingers, similar to those found in Nab2, consists of the human Tristetraproline (TTP)/Tis11 proteins, which specifically bind to the sequence UAUU to regulate the stability of specific mRNA transcripts (106, 142). In order to test whether the CCCH zinc finger motifs in Nab2 might also display affinity for the sequence recognized by this family of proteins, we used FCS-based competition

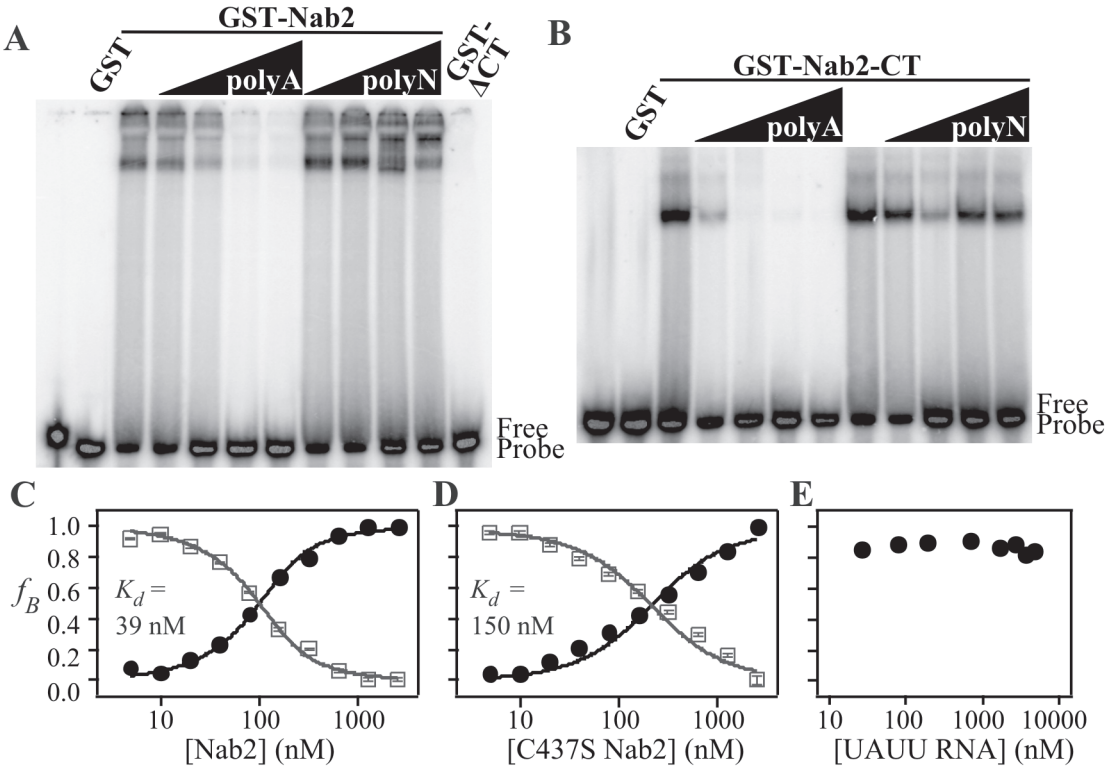

Figure 2.7: The zinc finger domain of Nab2 mediates polyadenosine RNA binding. Gel shift assays were used to determine which domain of Nab2 confers RNA binding. ***(A)*** 50 nM full-length GST-Nab2, GST-Nab2-ΔCT (amino acids 1-261), or ***(B)*** GST-Nab2-CT (amino acids 262-473) was incubated with a radioactively-labeled poly(A)$_{25}$ RNA oligonucleotide probe (~30 pM). RNA-protein complexes were then resolved from free probe by electrophoresis on a 5% non-denaturing polyacrylamide gel. To investigate binding specificity, either unlabeled poly(A)$_{25}$ RNA or poly(N)$_{25}$ RNA competitor oligonucleotides was added as indicated. No binding was observed to GST alone or Nab2 lacking the C-terminal zinc finger domain (GST-ΔCT). ***(C)*** and ***(D)*** Binding curves for wildtype ***(C)*** and C437S ***(D)*** Nab2 binding to Cy3-poly(A)$_{25}$ RNA generated from FCS analysis. ***(E)*** An oligonucleotide containing repeats of the TTP/TIS11 target sequence, UAUU, cannot compete with Cy3-labeled poly(A)$_{25}$ RNA for binding to Nab2.

assays to examine binding of Nab2 to a 25-nt RNA oligonucleotide (UAUU RNA) containing tandem repeats of the TTP/Tis11 target sequence. As shown in Figure 2.7E, no competition was observed, demonstrating that the zinc fingers of Nab2 show preferential binding to poly(A) RNA as compared to the target sequence of other proteins containing CCCH zinc fingers.

To begin to assess whether Nab2 is part of an evolutionarily conserved family of CCCH zinc finger proteins that preferentially binds polyadenosine RNA, we performed a BLAST search to identify other eukaryotic orthologues containing CCCH zinc fingers with similar spacing to those found in Nab2 ($CX_5CX_{4-6}CX_3H$). A survey of the database reveals a single human protein with zinc finger motifs that are closely related to Nab2, ZC3H14 [also known as NY-REN-37 (143) or UKp68]. While the yeast protein contains seven zinc finger domains, the human protein contains only five (Figure 2.8 A,B). However, only a subset of the yeast zinc fingers has been directly implicated in binding to RNA *in vivo* (73), suggesting that all seven zinc fingers in the yeast protein may not be critical for RNA binding. Nab2 and ZC3H14 also share homology within their N-terminal regions (Figure 2.8A), a domain required for Nab2 export from the nucleus (73) and association with other mRNA export factors (144). While Nab2 contains an arginine-glycine-glycine (RGG) repeat domain that mediates nuclear import (122), ZC3H14 lacks an RGG repeat and instead contains a predicted classical bipartite NLS. To test whether ZC3H14 is localized to the nucleus, we created a plasmid encoding ZC3H14-GFP and transiently transfected both HeLa and HEK cells. Similar to Nab2, the steady-state localization of ZC3H14 is nuclear (Figure 2.9A). Finally, to test whether the

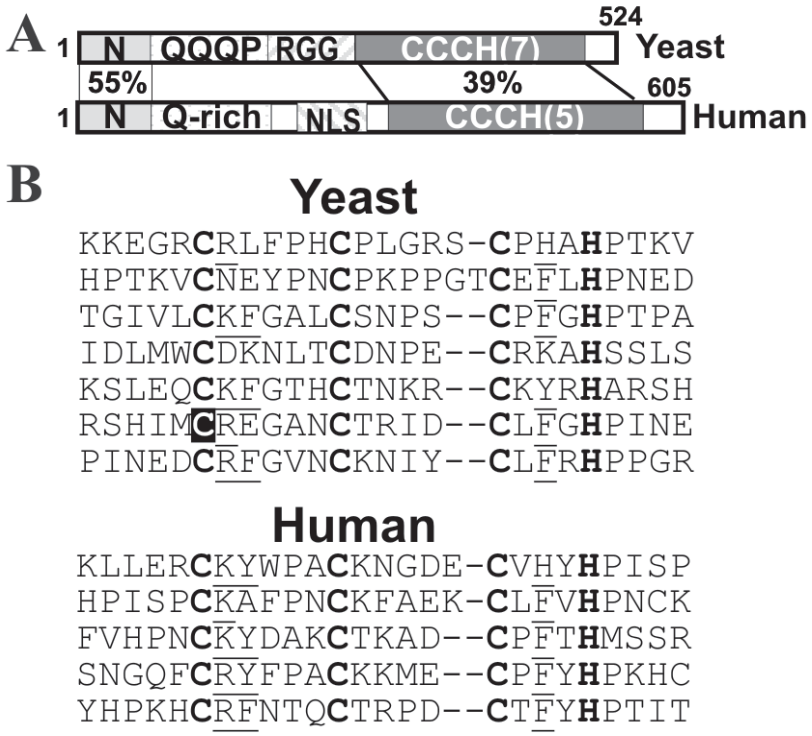

Figure 2.8: Nab2 is conserved in higher eukaryotes ***(A)*** Domain alignment of Nab2 and a putative human orthologue, ZC3H14. The percentage of similar amino acid residues between the N-terminal and the C-terminal zinc finger domains is indicated. ***(B)*** The cysteine and histidine residues in the zinc fingers of Nab2 (top-*S. cerevisiae*) and ZC3H14 (bottom-human) have a similar spacing pattern ($CX_5CX_{4-6}CX_3H$). Cysteine and histidine residues are shown in bold. The first cysteine of the sixth zinc finger, which is changed to serine in C437S Nab2, is boxed and conserved residues are underlined.

human CCCH zinc finger protein, ZC3H14, displays nucleic acid binding properties similar to Nab2, we examined the interaction of GST-ZC3H14 with a radioactively-labeled poly(A) RNA using a gel shift assay. As shown in Figure 2.9B, GST-ZC3H14 but not GST alone binds a radioactively-labeled 25-nt poly(A) RNA oligonucleotide. Furthermore, unlabeled 25-nt poly(A) RNA competitor competes for binding while unlabeled 25-nt poly(N) RNA competitor does not, indicating that, like the yeast Nab2, ZC3H14 preferentially binds poly(A) RNA.

To test whether human ZC3H14 is a functional orthologue of Nab2, we tested whether ZC3H14 could complement ΔNAB2 cells. First, we constructed a plasmid encoding ZC3H14 under the control of the *S. cerevisiae GAL1* promoter. This promoter allows us to induce expression of ZC3H14 by growing cells in the presence of galactose. Yeast cells deleted for the essential *NAB2* gene and complemented by a wild-type Nab2 maintenance plasmid were transformed with plasmids encoding wild type Nab2, ΔN-Nab2 (lacking amino acids 3-97), an unrelated control protein, Srp1 (Importin α), or ZC3H14 all under the control of a yeast *GAL1* promoter. Transformants were grown to saturation, serially diluted, and spotted onto control plates and plates containing 5-floroorotic acid (5-FOA) to select against the wild-type maintenance plasmid. As shown in Figure 2.10, Δ*NAB2* cells overexpressing Nab2 under the control of the *GAL1* promoter (pGAL-Nab2) grows normally at 30°C. However, cells overexpressing ZC3H14 show no growth, suggesting that either ZC3H14 is not expressed or cannot perform the essential function of Nab2. As expected, cells expressing ΔN-Nab2 or Srp1

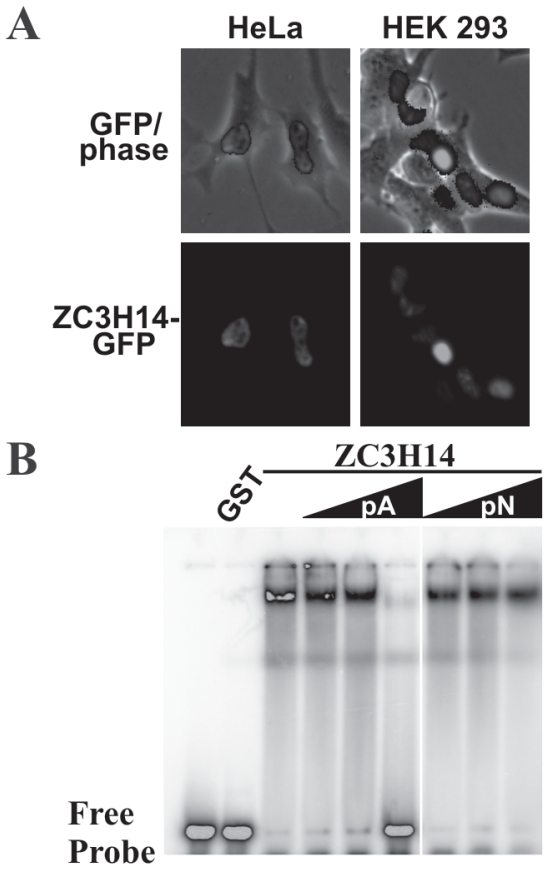

Figure 2.9: A human CCCH zinc finger protein, ZC3H14, binds specifically to polyadenosine RNA. ***(A)*** Localization of ZC3H14-GFP is shown in both HeLa (left) and HEK (right) cells. ***(B)*** RNA binding properties of ZC3H14 analyzed by gel shift assay. GST-ZC3H14, but not GST, binds poly$(A)_{25}$ RNA oligonucleotides in a gel shift assay. GST-ZC3H14 (1.2 μM) was incubated with a radioactively-labeled poly$(A)_{25}$ RNA oligonucleotide probe (~30 pM) and increasing amounts of unlabeled poly$(A)_{25}$ or poly$(N)_{25}$ RNA competitor oligonucleotides was added as indicated. RNA-protein complexes were then resolved from free probe by electrophoresis on a 5% non-denaturing polyacrylamide gel. Unlabeled poly$(A)_{25}$ RNA competitor efficiently competes for binding to ZC3H14 while unlabeled poly$(N)_{25}$ competitor oligonucleotide does not.

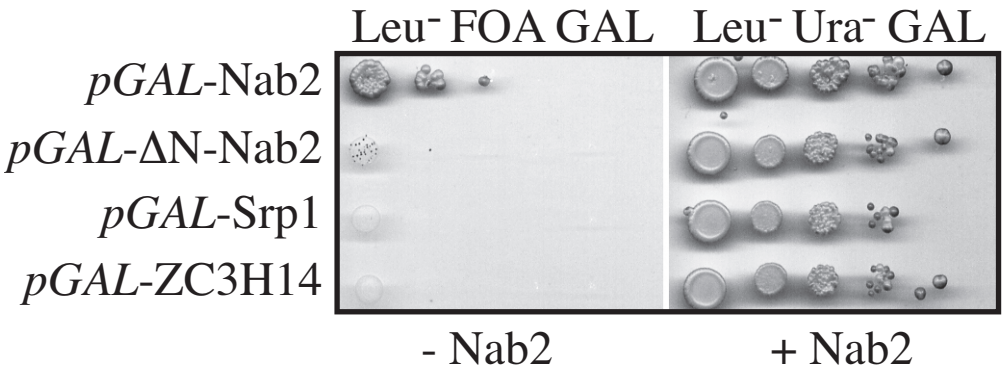

Figure 2.10: Overexpression of human ZC3H14 does not rescue Δ*NAB2 S. cerevisiae* cells. A plasmid shuffle assay was performed in Δ*NAB2* cells. Wild-type Nab2, ΔN-Nab2 (lacking amino acids 3-97), an unrelated control protein, Srp1 (Importin α), or human ZC3H14 were cloned into yeast expression vectors under the control of the yeast *GAL1* promoter. These plasmids were then transformed into Δ*NAB2* cells complemented by a plasmid-borne wild-type copy of Nab2. Cells were grown to saturation in liquid culture, serial diluted, and spotted onto control media (Leu$^-$ Ura$^-$ GAL) or onto media containing 5'-FOA (Leu$^-$ FOA GAL) to select against the wild-type maintenance plasmid. The media also contained galactose to induce expression of Nab2, ΔN-Nab2, Srp1, or human ZC3H14.

also show no growth. Overexpression of ΔN-Nab2 has been shown previously to be deleterious to yeast cell growth (73).

In addition to understanding the precise mechanism by which CCCH zinc fingers recognize poly(A) RNA, we also sought to establish a higher eukaryotic model organism that would allow us to investigate the broader implications of defects in poly(A) RNA recognition by CCCH zinc finger containing proteins. Towards these ends we began to investigate the *Drosophila melanogaster* homologue of Nab2 (henceforth termed dNab2). dNab2 is an uncharacterized open reading frame (CG5720) of 1004 amino acids that contains a zinc finger domain similar to that found in Nab2 and human ZC3H14. Similar to human ZC3H14, dNab2 contains only five zinc fingers compared to the seven encoded by *NAB2*. As shown in Figure 2.11, an alignment of each of the seven zinc fingers from *S. cerevisiae* Nab2 with each of the five zinc fingers in dNab2 demonstrates that not only is the spacing of the cysteines and histidines relatively conserved from yeast to flies, but several of the intervening residues between the cysteines and histidines are also conserved (Figure 2.11A – shown in blue). Since no previous studies have investigated dNab2, we wanted to ensure that it was expressed. In order to test expression of dNab2, we obtained total cellular RNA from cultured *D. melanogaster* S2 cells and performed RT-PCR using two sets of gene specific primers for dNab2. As a control we also performed RT-PCR using primers specific for β-tubulin and slmb, a regulator of fruit fly egg chamber development (145). As shown in Figure 2.11B, the dNab2 transcript can be detected in cultured S2 cells. As expected, β-tubulin and slmb are also expressed. As a control, a plasmid obtained from the Bloomington Drosophila Stock Center at Indiana

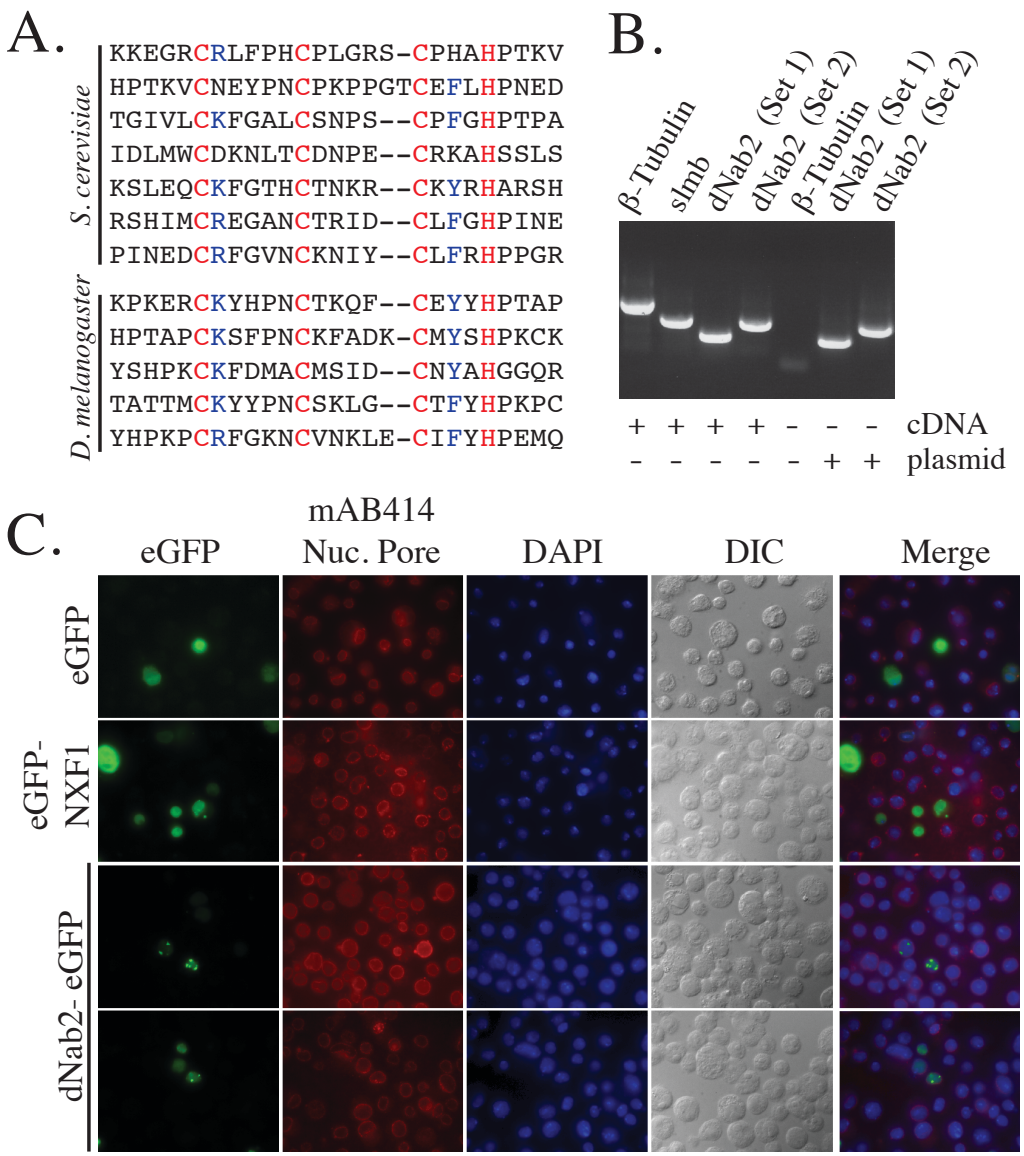

Figure 2.11: The *Drosophila melanogaster* gene, CG5720, encodes a putative orthologue of Nab2/ZC3H14. (A) Alignment of the seven *S. cerevisiae* Nab2 zinc fingers and the five *D. melanogaster* dNab2 zinc fingers. Cysteines and histidines are shown in red, while conserved basic and aromatic residues are shown in blue. (B) dNab2 is expressed in cultured *D. melanogaster* S2 cells. Total cellular RNA was isolated from S2 cells and subjected to RT-PCR using specific primers for β-tubulin, slmb, or dNab2. PCR amplification reactions were also carried out using a plasmid containing a predicted full-length dNab2 cDNA insert as a template. (C) dNab2-eGFP localizes to the nucleus of cultured S2 cells. S2 cells were transfected with plasmids encoding eGFP, eGFP-NXF1, or dNab2-eGFP. eGFP and dNab2-eGFP were expressed with $CuSO_4$ (as described in Experimental procedures) and cells were fixed, permeablized, and stained with the monoclonal antibody, mAB414, which recognizes several nuclear pore components. Nuclei were stained with DAPI.

University containing a full-length dNab2 cDNA was also used as a template for PCR amplification. PCR products obtained when using S2 cell cDNA are the same length as those obtained with using the control plasmid. Since both dNab2 primer sets 1 and 2 cross exon-exon boundaries, the fact that products generated from the RT-PCR reaction and the control cDNA are the same length demonstrates that cDNA was amplified and not contaminating fly genomic DNA.

Both Nab2 and human ZC3H14 localize to the nucleus at steady state (67, 70, 73, 146). We hypothesized that, similar to its orthologues, dNab2 should also localize to the nucleus. In order to investigate the localization of dNab2, we transfected a plasmid encoding a copper inducible copy of dNab2-eGFP into *Drosophila melanogaster* S2 cells and induced expression of dNab2-eGFP with 0.5 mM $CuSO_4$. As controls, we also transfected S2 cells with a plasmid encoding a copper inducible copy of eGFP and a plasmid encoding a constitutively expressed copy of eGFP-NXF1, the principle cellular mRNA export receptor (17). Following induction of eGFP and dNab2-eGFP, S2 cells were removed from culture dishes, fixed, permeablized, and incubated with mAB414 monoclonal antibodies specific to components of the nuclear pore complex. Nuclei were also stained with Hoechst dye to visualize chromatin. As demonstrated in Figure 2.11C, dZC3H14-eGFP localizes to discrete sub-domains within the nucleus. Interestingly, human ZC3H14 shows a similar localization and colocalizes with nuclear speckle markers (146). As expected, eGFP-NXF1 localizes to the nucleus of cultured *D. melanogaster* S2 cells (147) while eGFP alone localizes to the cytoplasm. Together, this expression and localization data suggests that *Drosophila melanogaster* may be a suitable

model organism for the investigation of the role of polyadenosine RNA recognition by CCCH zinc fingers in development and multi-cellular processes.

Discussion

In this study we report three important findings. First, the yeast zinc finger protein, Nab2, binds with high affinity (~30 nM) to polyadenosine RNA oligonucleotides *in vitro*. This interaction is specific for polyadenosine RNA, as poly(N), poly(G), poly(AG) RNA oligonucleotides or RNA oligonucleotides containing the TTP/TIS11 binding sequence, UAUU, could not compete for binding to Nab2. Second, we show that the zinc fingers of Nab2 are necessary and sufficient to mediate this specific interaction. The sixth zinc finger at least partially contributes to this interaction since a conservative amino acid change in the first cysteine of this zinc finger causes a four-fold decrease in Nab2 for poly$(A)_{25}$ RNA *in vitro*. Finally, we also present evidence to support the existence of a Nab2 orthologue in higher eukaryotes that contains highly homologous CCCH zinc finger motifs to those found in Nab2 and also preferentially binds polyadenosine RNA. Together, these results provide evidence for a previously uncharacterized family of CCCH zinc finger-containing poly(A) specific RNA binding proteins. To the best of our knowledge, these are the only poly(A) RNA binding proteins that lack RRM domains and bind specifically to poly(A) RNA, suggesting that they may be founding members of a new class of zinc finger-containing polyadenosine RNA binding proteins.

Numerous studies have demonstrated that the poly(A) tail and its associated proteins greatly influence the fate of an mRNA transcript and hence gene expression (4,

62, 103, 125). Transcripts lacking poly(A) tails are not properly translated and are quickly degraded (148). Additionally, *S. cerevisiae* transcripts that are terminated by a self-cleaving ribozyme element and thereby lack poly(A) tails are at least partially retained in the nucleus (58), suggesting that the poly(A) tail may also play a role in mRNA export from the nucleus. By regulating the stability, translatability, and even sub-cellular localization of mRNA transcripts, the poly(A) tail and its associated proteins can act as potent post-transcriptional regulators of gene expression. We speculate that Nab2 and possibly other zinc finger containing poly(A) binding proteins, such as ZC3H14, could function in some of the roles previously attributed to the canonical family of RRM containing Pabs. In fact, several pieces of evidence argue that Nab2 might be the more predominant yeast nuclear poly(A) tail binding protein. First, Nab2 shuttles between the nucleus and the cytoplasm, but unlike Pab1, is localized in the nucleus at steady state (70 2002, Green, 2003). Nab2 is also required for efficient poly(A) RNA export from the nucleus (73). Additionally, several studies have demonstrated that recombinant Nab2 can replace Pab1 in *in vitro* polyadenylation assays and that nab2 mutant cells accumulate transcripts with extended poly(A) tails (71), suggesting that Nab2 plays a crucial role in polyadenylation.

Poly(A) specific RNA binding proteins were once considered histone-like proteins that merely packaged the mRNA and prevented it from being prematurely degraded. More recent work has provided a better understanding of post-transcriptional control of gene expression and has convincingly established a direct link between the poly(A) tail, poly(A) binding proteins, and the stability and translatability of numerous mRNA transcripts. Interestingly, all recent studies investigating the role of the family of

poly(A) binding proteins in translation and transcript stability have solely focused on RRM-containing poly(A) binding proteins. With the identification of a zinc finger protein that displays poly(A) binding, the repertoire of this family of proteins has expanded. Given that zinc finger domains are one of the most abundant domains found in the human genome (149), this finding raises the possibility that there are additional proteins that could interact with polyadenylated mRNA transcripts and modulate gene expression.

Experimental Procedures

Chemicals, plasmids, and yeast manipulations –Chemicals were obtained from Fisher Scientific (Pittsburgh, PA), Sigma-Aldrich (St. Louis, MO), or US Biological (Swampscott, MA) unless otherwise noted. DNA manipulations were performed according to standard methods (150) and all media were prepared by standard procedures (151).

To create an untagged Nab2 bacterial expression clone, the *NAB2* open reading frame was amplified from yeast genomic DNA using PCR and cloned into the bacterial expression vector, pMW172 (152), creating pAC2133. The *NAB2* open reading frame was also subcloned from pAC2133 into pGEX-4T-1 (GE Healthcare) to create a GST-Nab2 expression clone (pAC2303). To create plasmids encoding GST-Nab2-ΔCT (amino acids 1-261) (pAC2315) and GST-Nab2-CT (amino acids 262-473) (pAC2304) the corresponding regions of the *NAB2* open reading frame were amplified by PCR and cloned into pGEX-4T-1. In order to create a plasmid encoding ZC3H14-GFP, the ZC3H14 open reading frame was amplified from a cDNA clone obtained from ATCC

and cloned into pEGFP-N1 (Clonetech/Invitrogen) to create pAC2111. A bacterial expression vector encoding the ZC3H14 open reading frame was created by subcloning ZC3H14 from pEGFP-N1 into pGEX-4T-1, creating pAC 2200. A plasmid (pAC 2317) encoding the putative *Drosophila melanogaster* Nab2 orthologue (dNab2) cDNA was obtained from the Bloomington Drosophila Stock Center at Indiana University. A plasmid encoding a copper sulfate inducible dNab2-eGFP was constructed by PCR amplification of dNab2 from pAC 2317. PCR products were digested with SpeI and EcoRI and ligated into a pMT/V5-His expression plasmid (Invitrogen). The eGFP open-reading frame was amplified from pEGFP-N1 (Clonetech/Invitrogen), digested with XhoI and XbaI, and ligated in-frame with dNab2 in pMT/V5-His. All clones were sequenced to ensure that no mutations were introduced during cloning.

In order to express human ZC3H14 in *S. cerevisiae* cells, ZC3H14 was amplified by PCR and cloned into a yeast expression vector under the control of the yeast *GAL1* promoter. Expression of human ZC3H14 could then be induced by the presence of galactose in the media. Plasmid shuffle assays were performed as described (153)

Oligonucleotides – RNA oligonucleotides were obtained from Dharmacon (Lafayette, CO). Fluorescent RNA oligonucleotides were labeled with Cy3 on the 5'-end and PAGE purified. All RNA oligonucleotides were deprotected using the manufacturer supplied buffer and protocol. Deprotected oligonucleotides were then evaporated to dryness using a speed-vac centrifuge and frozen at –20°C. Fluorescent and non-fluorescent competitor DNA oligonucleotides were obtained from IDT (Coralville, IA). All oligonucleotides used were 25 nucleotides long. Oligonucleotides were resuspended in binding buffer (20

mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAC, 2 μM $ZnCl_2$, 2% glycerol) prior to each experiment.

Protein Expression and Purification – To express recombinant Nab2 in *E. coli*, the expression plasmid, pAC2133, was transformed into BL21(DE3)pLYS cells (Novagen). Single colonies were inoculated into 2 ml of media and grown to saturation overnight. This culture was then used to inoculate 1 L of LB media. Cells were grown at 37°C until they reached an OD600 nm of 0.4 – 0.6. Cultures were then shifted to 30°C and induced with 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hrs. Cells were then centrifuged at 3800rpm for 20 minutes at 4°C and pelleted cells were frozen at -80°C. For purification of untagged Nab2 proteins from frozen cell pellets, cell pellets were thawed on ice and resuspended in an equal volume of lysis buffer (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 4 mM β-mercaptoethanol, 1mM phenylmethylsulfonylfluoride) and lysed using a French Press. The lysates were cleared by centrifugation, loaded onto a HiTRAP Q column (GE Healthcare) pre-equilibrated with Buffer A (20mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol), and bound protein was eluted using a linear gradient of Buffer B (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 2 mM β-mercaptoethanol, and 1 M NaCl). Fractions containing untagged recombinant Nab2 were dialyzed overnight into Buffer C (20 mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAC, 2 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol). The dialyzed protein was then concentrated by adding saturated ammonium sulfate to 25% saturation. Precipitated proteins were pelleted and resuspended in Buffer C. The concentrated protein was then loaded onto a Superdex S-200 gel filtration column pre-equilibrated with Buffer C and fractions

containing untagged recombinant Nab2 were pooled and concentrated using Centricon spin concentrators (Millipore). Since we found that freezing the purified protein led to loss of nucleic acid binding, freshly purified aliquots of protein were used for each binding experiment.

For purification of GST fusion proteins, plasmids encoding GST, GST-Nab2, GST-Nab2-CT, or GST-Nab2-ΔCT were transformed into BL21 (DE3) cells (Novagen). Additionally, plasmids encoding GST-ZC3H14 or GST-ZC3H14-CT were transformed into BL21 (DE3)RIL cells (Novagen). Expression was induced as described for untagged Nab2. Following induction, cells were pelleted and frozen at –80°C. Cell pellets were thawed on ice and resuspended in an equal volume of Lysis Buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM β-mercaptoethanol, 2% glycerol, 0.1% Triton X-100, 1 mM phenylmethylsulfonylfluoride) and lysed by sonication. Lysates were cleared by centrifugation and incubated with Glutathione Sepharose 4B (GE Healthcare Life Sciences) for 2 hours at 4°C. Bound proteins were washed three times with 5 ml of Lysis Buffer and eluted with Elution Buffer (20 mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAC, 2 μM $ZnCl_2$, 2% glycerol, 10 mM Reduced Glutathione). To ascertain the purity of these samples, a small amount of eluted protein was resolved by SDS-PAGE and stained with Coomassie Blue. The remainder of the protein was stored on ice at 4°C until used for RNA electrophoretic mobility shift assays.

Recombinant His-Pab1 was expressed and purified essentially as described (64).

FCS Measurements – Two-photon FCS data were acquired to measure the interaction between fluorescent oligonucleotides and Nab2 using a previously described home-built instrument (154, 155). Multi-component diffusion FCS analysis was applied to

determine the fraction of fluorescent oligonucleotides bound to Nab2 for different experimental conditions (140, 155, 156). All FCS curves were fit using a standard 3D Gaussian multi-component diffusion model (137, 157) with the volume calibrated using Rhodamine 6G. Each series of FCS experiments used global analysis to simultaneously fit all FCS curves acquired at different protein or competitor concentrations for a given fluorescent oligonucleotide. The global fits return the concentrations of bound and free fluorescent oligonucleotides for each Nab2 protein or competitor oligonucleotide concentration (140, 158, 159).

For all FCS measurements, samples were excited at 980 nm using a mode-locked Tsunami Ti:Sapphire laser pumped with a 532-nm 5W Millenia solid-state Nd:YVO_4 laser (Spectra-Physics, Mountain View, CA). The 5×-expanded laser beam was directed into a IX-71 Olympus microscope (Olympus America, Center Valley, PA) and focused into the sample with a 60× UPlanSApo 1.2 NA water-immersion objective lens (Olympus America, Center Valley, PA). Fluorescence collected through the objective lens passed through a dichroic mirror (675 DCSX) and a shortpass filter (E680SP, Chroma Technology, Rockingham, VT), and was focused onto a photon-counting avalanche photo diode (EG&G, Vaudreuil, Canada). The detector output was sent to an ALV hardware correlator (Langen, Germany). Power at the sample was controlled by rotating a λ/2-wave plate in front of a linear polarizer, and set at 7 mW for all experiments. The FCS observation volume was calibrated by fitting measured autocorrelation curves for Rhodamine 6G (J.T. Baker, Phillipsburg, NJ) in water. All samples were mounted in 8-well chambered cover glass containers (Nalge Nunc, Rochester, NY). Prior to loading, sample chambers were pre-treated with blocker casein (Pierce, Rockford, IL) and then

lightly washed with nanopure distilled water to prevent non-specific binding of proteins and oligonucleotides to the cover slip and container walls. Several FCS measurements were performed at each protein or competitor concentration and the average correlation values and associated standard deviations were computed using standard procedures (158).

Characterization of the Nab2-oligonucleotide binding interactions requires measurement of the fraction of Cy3-labeled fluorescent oligonucleotides bound to the Nab2 protein at different protein concentrations (156, 158, 159). A series of FCS measurements were thus performed using Cy3-labeled oligonucleotides and a wide range of Nab2 protein concentrations. Initially, a solution of concentrated recombinant Nab2 (~2.5 μM) and Cy3-labeled oligonucleotides (~140 nM) was prepared (total volume of 300 μL). Following each FCS measurement, one-half of the sample was removed from the chamber and replaced with an equal volume of fluorescent oligonucleotides at the same concentration, thus keeping the fluorescent nucleic acid concentration fixed while serially diluting the protein concentration until [Nab2] ≤ 0.1 nM. For each Nab2 concentration, FCS data was collected in eight separate correlation measurements of 60 seconds each. We use the standard 3D Gaussian multi-component diffusion model to fit measured FCS curves and recover the protein-bound and free nucleic acid concentrations for each sample condition

$$G(\tau) = \frac{2\sqrt{2}}{\pi^{3/2}\omega_0{}^2 z_0} \frac{\sum \psi_i{}^2 C_i A_i}{\left(\sum \psi_i{}^2 C_i\right)^2} \text{ with } A_i(\tau) = \left(1 + 8D_i\tau/\omega_0{}^2\right)^{-1}\left(1 + 8D_i\tau/z_0{}^2\right)^{-1/2} \quad [1]$$

The parameters C_i and D_i represent the concentration and diffusion coefficient for each fluorescent species (i.e. bound and free oligonucleotides), and τ is the correlation time. The axial and radial beam waists, ω_0, and z_0, respectively, specify the size of the FCS observation volume and were determined from the calibration measurements with rhodamine dye. The molecular brightness parameter ψ_i takes into account the absorption cross-section of the fluorescent dye, fluorescence quantum yield changes, and external excitation conditions (laser power, laser pulse width, etc.). We used global analysis to simultaneously fit all of the FCS curves acquired at different protein concentrations for a given fluorescent nucleic acid, and these fits return the concentrations of bound and free fluorescent oligonucleotides for each Nab2 concentration (140, 158, 159). For global fitting, the diffusion coefficient of the free oligonucleotide was held fixed at the known value determined from independent measurements.

The dissociation constants (K_d) for a given interaction were recovered by fitting the concentration dependence of the bound fraction, f_b, to the equation:

$$f_b = \frac{[PO]}{[O]_{tot}} = \frac{\left(K_d + [P]_{tot} + [O]_{tot}\right) - \sqrt{\left(K_d + [P]_{tot} + [O]_{tot}\right)^2 - 4 \cdot [O]_{tot} \cdot [P]_{tot}}}{2 \cdot [O]_{tot}} \qquad [2]$$

where [P], [O], and [PO] represent the concentrations of unbound protein, unbound fluorescent nucleic acid, and bound protein-nucleic acid. The unbound fraction is $f_u = 1 - f_b$.

A similar procedure was followed for competition experiments. Samples were prepared containing 1 μM purified recombinant Nab2 and 130 nM Cy3-labeled oligonucleotide. Concentrated non-fluorescent competitor oligonucleotide was added

incrementally to the sample, with the total competitor concentration titrated into the μM range (up to ~ 5 μM). Sequential addition of competitor to the sample resulted in a slight increase to the total sample volume (6% volume change at maximum dilution which was corrected for in the analysis). FCS curves were collected for each competitor concentration and the concentration of bound labeled oligonucleotide was determined by global fitting of all FCS curves as described above. The concentration of competitor needed to displace 50% of the bound fluorescently-labeled oligonucleotide (IC_{50}) is determined by fitting this data to the equation:

$$f_b = \frac{f_{b0}}{1 + \left([competitor] / IC_{50} \right)^m} \quad [3]$$

where f_b and f_{b0} are the fraction of bound labeled oligonucleotide in the presence or absence of a competitor, respectively (140), and m is a curvature parameter (160). The inhibition constant K_i of the competitor oligonucleotide can be estimated from:

$$K_i = \frac{IC_{50} \cdot K_d}{[P]_{tot} - 0.5 \cdot [O]_{tot} - K_d} \quad [4]$$

where $[P]_{tot}$ and $[O]_{tot}$ are total Nab2 and fluorescently-labeled oligonucleotide concentrations (161), respectively.

Gel Shift RNA Binding Assays – Synthetic 25-nt poly(A) RNA oligonucleotides (Dharmacon) were 5'-end labeled with [γ-^{32}P]ATP (GE Healthcare Life Sciences) using T4 polynucleotide kinase (Promega). RNA electrophoretic mobility shift assays were performed by incubating 50 nM recombinant GST, GST-Nab2, GST-Nab2-CT, or GST-Nab2ΔCT with approximately 30 pM radioactively-labeled poly(A) RNA oligonucleotide

and an increasing amount of unlabeled 25-nt competitor RNA oligonucleotide in binding buffer for 30 min at 20°C. For binding reactions containing GST-ZC3H14, a solution of 1.2 μM GST or GST-ZC3H14 was incubated with approximately 30 pM radioactively-labeled poly(A) RNA oligonucleotide and an increasing amount of unlabeled competitor RNA oligonucleotide. Binding reactions were loaded onto a 5% native polyacrylamide gel and electrophoresed at 30 mA in 0.3x T.B.E for 30 min to separate free oligonucleotide from protein-RNA complexes. Gels were dried and exposed overnight using a phosphorimager (Amersham).

ZC3H14 localization – A plasmid encoding ZC3H14-GFP expressed from a constitutive CMV promoter was transiently transfected into either HEK or HeLa cells. Two days following transfection, ZCH14-GFP localization was assessed by direct fluorescence microscopy.

Drosophila melanogaster S2 cell culture and immunofluorescence – D. melanogaster S2 cells were cultured in Schneider's Drosophila medium (Gibco) in the presence of fetal calf serum (Invitrogen) and Penicillin/Streptomycin/Glutamine (Invitrogen). S2 cells were passaged every 3-5 days and were typically grown in 75 cm^2 culture flasks (Fisher). In order to localize dZ3H14, a plasmid encoding either a copper sulfate-inducible copy of eGFP or dZC3H14-eGFP or a constitutively expressed copy of eGFP-NXF1 was transfected into *D. melanogaster* S2 cells. Two days following transfection, expression of eGFP and dZC3H14-eGFP was induced by 0.5 mM copper sulfate for 12 hours. Since eGFP-NXF1 was induced from a constitutively active promoter, copper sulfate was not added to cells transfected with plasmids encoding eGFP-NXF1. Following the 6 hour induction, cells were gently resuspended in each well by pipetting, transferred to

eppendorf tube and centrifuged at 1500 rpm for 3 minutes at 4°C. Cells were washed once with ice-cold 1X PBS and centrifuged as before. Cell pellets were resuspended in ice-cold 1X PBS and allowed to adhere to 0.3% poly-D-lysine coated cover slips for 10 minutes. Excess cells were aspirated off and cells were fixed with 4% paraformaldehyde at room temperature for 10 minutes. Following fixation, cells were washed twice with ice cold 1X PBS and then permeablized in 1X PBS 0.5% Triton X-100 for 10 minutes. Cells were again washed twice with ice cold 1X PBS and incubated in blocking solution (3% bovine serum albumin (BSA) in 1X PBS 0.02% Triton X-100) at 4°C overnight. In the morning, cells were incubated with the mAB414 nuclear pore specific antibody diluted 1:1000 in blocking solution for 1 hour in a humidified chamber. Cover slips were then washed four times with antibody wash buffer (1.5% BSA in 1X PBS 0.02% Triton X-100) and then incubated for 1 hour at room temperature with an anti-mouse secondary antibody conjugated with Texas-Red diluted 1:1000 in blocking solution. Cover slips were washed as before in antibody wash buffer and nuclei were stained with Hoechst dye (1 mg/ml) diluted in antibody wash buffer. Cover slips were then mounted using Vectashield and sealed with nail polish.

RNA isolation from D. melanogaster S2 cells – Drosophila melanogaster S2 cells were harvested by centrifugation and cell pellets were resuspended in 1-2 ml of Trizol (Invitrogen) and 100µl of bromo-3-chloropropane. Resuspended cell pellets were vortexed for 15 seconds and incubated at room temperature for 15 minutes to solubilize total cellular RNA. Samples were then centrifuged at 21,000g for 15 minutes at 4°C. The upper layer was transferred to a fresh eppendorf and RNA was precipitated with 500µl of isopropanol. Samples were incubated at room temperature for 10 minutes and

then centrifuged at 21,000g for 8 minutes at 4°C. The supernatant was removed and the pellet washed with 1 ml of 70% ethanol (in DEPC-treated dH_2O). Samples were centrifuged again at 21,000g for 5 minutes at 4°C. The supernatant was removed and RNA pellets were air-dried. RNA pellets were resuspended in 20 μl of DEPC-treated dH_2O.

Semi-quantitative reverse transcriptase PCR (RT-PCR) – RT-PCR was performed using a One-Step RT-PCR kit (Qiagen) and gene specific primers for β-tubulin, slmb, and ZC3H14 according to the manufacturer's directions.

Chapter 3: Poly(A) RNA binding by Nab2 is required for correct mRNA 3'-end formation

This chapter is adapted from the following manuscripts:

Kelly, S.M., Leung, S.W., Bramley, A.M., Tran, E.J., Chekanova, J.A., Wente, S.R., and Corbett, A.H., *Poly(A) RNA binding by Nab2 is required for correct mRNA 3'-end formation*. (In Revision)

Brockmann, C., Kelly, S.M., Soucek, S., Noto, J.J., Corbett, A.H., and Stewart, M. *The NMR structure of a CCCH zinc finger-containing poly(A) RNA binding protein* (In Preparation)

Structural studies were performed by Christoph Brockmann, Ph.D. and Murray Stewart, Ph.D. at the MRC Laboratory of Molecular Biology in Cambridge, UK.

Introduction

Although all cells within a eukaryotic organism contain the same genetic material, each cell must produce a unique combination of proteins that are necessary for that cell to perform its specific function. Each of these proteins is translated from a messenger RNA (mRNA) transcript that is transcribed within the nucleus. Following transcription by RNA polymerase II, these mRNA transcripts must be spliced and polyadenylated, exported from the nucleus, and perhaps even transported to a distant location within the cytoplasm (4). During each of these steps mRNA transcripts are coated from cap to tail by numerous RNA binding proteins. Importantly, each of these RNA binding proteins has the potential to post-transcriptionally regulate gene expression by dictating changes in RNA processing, RNA export from the nucleus, or RNA stability. Hence knowledge of the mechanisms by which proteins interact with RNA and dictate changes in the fate of mRNA transcripts is crucial for our overall understanding of the control of gene expression.

Although the processing steps that occur during mRNA biogenesis are often presented as disconnected parts of a tangled web of processing, they are actually tightly coupled to one another and normally function as one integrated machine for producing transcripts ready for translation in the cytoplasm (162). Like any good assembly line, early steps in the pathway ready the transcript for subsequent processes. For example, cleavage and polyadenylation factors, such as Rna15 [CstF64 in humans (163)] and Rna14 [CstF77 in humans (15, 164)] are recruited to the nascent mRNA transcripts during transcription via interactions with the C-terminal domain of RNA polymerase II

(165). Following transcription, these factors help to correctly position the 3'-end processing machinery on the transcript and add the poly(A) tail (50).

Once a transcript has been processed in the nucleus, it is then exported to the cytoplasm. A complex series of protein rearrangements occur on the mRNA transcript during and immediately following nuclear export to replace export proteins with proteins that are crucial for translation (166). This "molecular wardrobe change" (17) is orchestrated by an RNA helicase named Dbp5 which is localized on the cytoplasmic face of the nuclear pore complex (NPC) (167). Dbp5 is thought to remodel the mRNA ribonucleoprotein (mRNP) complexes as they exit from the nucleus (91). It has become increasingly evident that the intricate timing of these protein rearrangements is critical for efficient mRNA export. For example, if Dbp5 is activated too early and export proteins are displaced prior to translocation through the NPC, the transcripts will not be exported efficiently and will accumulate in the nucleus.

While the exact mechanism by which Dbp5 displaces proteins from mRNA transcripts has not been clearly defined, several Dbp5 substrates have been identified in *Saccharomyces cerevisiae* (91, 92). One of these substrates is the essential yeast protein, Nuclear Poly(A) Binding protein 2 (Nab2). Both Nab2 and the higher eukaryotic family member, ZC3H14, localize to the nucleus at steady state (67, 73, 146). ZC3H14, however, has not yet been investigated as a substrate of human Dbp5. These proteins are members of a novel class of poly(A) binding (Pab) proteins that recognize polyadenosine RNA via tandem Cys-Cys-Cys-His (CCCH) zinc fingers (67). Interestingly, all other characterized Poly(A) RNA binding proteins recognize polyadenosine RNA via at least one well-conserved RNA Recognition Motif (RRM) (62). Thus, Nab2 and other

members of this evolutionarily conserved CCCH zinc finger-containing family must recognize polyadenosine RNA through a fundamentally different mechanism than the RRM-containing family of poly(A) RNA binding proteins.

Nab2 has been implicated in two separate, but coupled, steps in mRNA biogenesis. First, Nab2 modulates poly(A) tail length as defects in Nab2 cause extended poly(A) tails *in vivo* (71, 72) and the addition of recombinant Nab2 to *in vitro* polyadenylation assays limits poly(A) tail length (71, 72, 168). Second, Nab2 is implicated in the export of poly(A) RNA from the nucleus (70, 71, 73, 74). Nab2 mutant cells show poly(A) RNA accumulation within the nucleus and *nab2* alleles show genetic interactions with mRNA export factors including *YRA1* and *MEX67* (74, 75). Furthermore, Nab2 is exported from the nucleus via a mechanism that is dependent upon ongoing RNA polymerase II transcription (73). A logical model based on this information is one in which Nab2 plays a role in polyadenylation, associates with the poly(A) tail of mRNA transcripts following polyadenylation, and then exits the nucleus in complex with the poly(A) tail of mRNA transcripts destined for the translation machinery in the cytoplasm.

Due to the fact that Nab2 presumably plays a role in both polyadenylation and mRNA export from the nucleus, as well as the fact that it recognizes polyadenosine RNA (67), a ubiquitous element of all mRNA transcripts, it has the potential to post-transcriptionally regulate the expression of a wide variety of mRNA transcripts. To gain insight into the mechanisms by which Nab2 post-transcriptionally regulates gene expression as well as the novel molecular mechanism by which this family of CCCH zinc finger containing proteins specifically recognizes polyadenosine RNA, we have precisely

defined Nab2 zinc fingers 5-7 as the high affinity poly(A) RNA binding domain within Nab2. Furthermore, we have determined the atomic resolution structure of this domain by NMR and shown that specific amino acid substitutions within Nab2 zinc fingers 5-7 abolish poly(A) RNA binding *in vitro*. Cells that express such Nab2 RNA binding mutants display cold sensitive growth defects as well as an increase in poly(A) tail length and nuclear accumulation of polyadenosine RNA. These results provide critical insight into the molecular mechanism underlying polyadenosine RNA recognition by CCCH zinc fingers as well as the essential function of Nab2.

Results

In order to define the primary function of the evolutionarily conserved Nab2 protein in mRNA biogenesis, we sought to generate Nab2 variants with defects in polyadenosine RNA binding. To address this question, we needed to first more precisely define the mode of polyadenosine RNA recognition by tandem CCCH zinc fingers. Our initial approach exploited a previously described genetic interaction between Nab2 and the nuclear pore associated RNA helicase, Dbp5 (91). Importantly, an amino acid substitution in the RNA binding domain of Nab2, nab2-C437S, which decreases the affinity of Nab2 for polyadenosine RNA 4-fold (67), suppresses the growth defect and poly(A) RNA accumulation of the helicase mutant *dbp5-2* (91). This finding suggests that if the RNA binding affinity of a Dbp5 substrate, such as Nab2, is decreased, as in nab2-C437S, the partially functional mutant dbp5-2 helicase may retain sufficient activity to remove the weakly bound Nab2 protein from the mRNA transcript. This model implies that other Nab2 mutants that weaken the interaction of Nab2 with polyadenosine

RNA may also suppress the temperature sensitive phenotype of the *dbp5-2* mutant. Hence, this genetic assay provides a simple approach to identify putative residues that are critical for Nab2 binding to RNA.

As a first step to determine which of the seven zinc fingers in Nab2 are important for RNA binding, we generated point mutants that encode cysteine to alanine amino acid changes in the first cysteine of each of the seven Nab2 zinc fingers (Figure 3.1A) and tested whether these mutants could suppress the temperature sensitive growth phenotype of the *dbp5-2* mutant at 33°C. As shown in Figure 3.1B, individual cysteine to alanine substitutions within each of the first cysteines of Nab2 zinc fingers 1-4 did not suppress the temperature sensitive phenotype of the *dbp5-2* mutant. In contrast, individual cysteine to alanine substitutions in each of the first cysteines of zinc fingers 5-7 did suppress the temperature sensitive growth phenotype of the *dbp5-2* mutant. This result suggests that only the last three zinc fingers of Nab2 may be necessary for polyadenosine RNA binding.

As a complement to the genetic suppression assay, we used a biochemical approach to identify a domain of Nab2 that binds to polyadenosine RNA. According to our results from the *dbp5-2* suppression assay (Figure 3.1B), zinc fingers 5-7 are implicated in binding to polyadenosine RNA. To test this prediction, we expressed and purified recombinant Nab2 protein fragments containing zinc fingers 1-7 (amino acids 262-473), zinc fingers 1-4 (amino acids 262-387), or zinc fingers 5-7 (amino acids 262-477) as GST fusion proteins. In order to assess RNA binding for each of these domains, RNA gel shifts were used to examine binding to a Cy3-labeled poly$(A)_{25}$ RNA

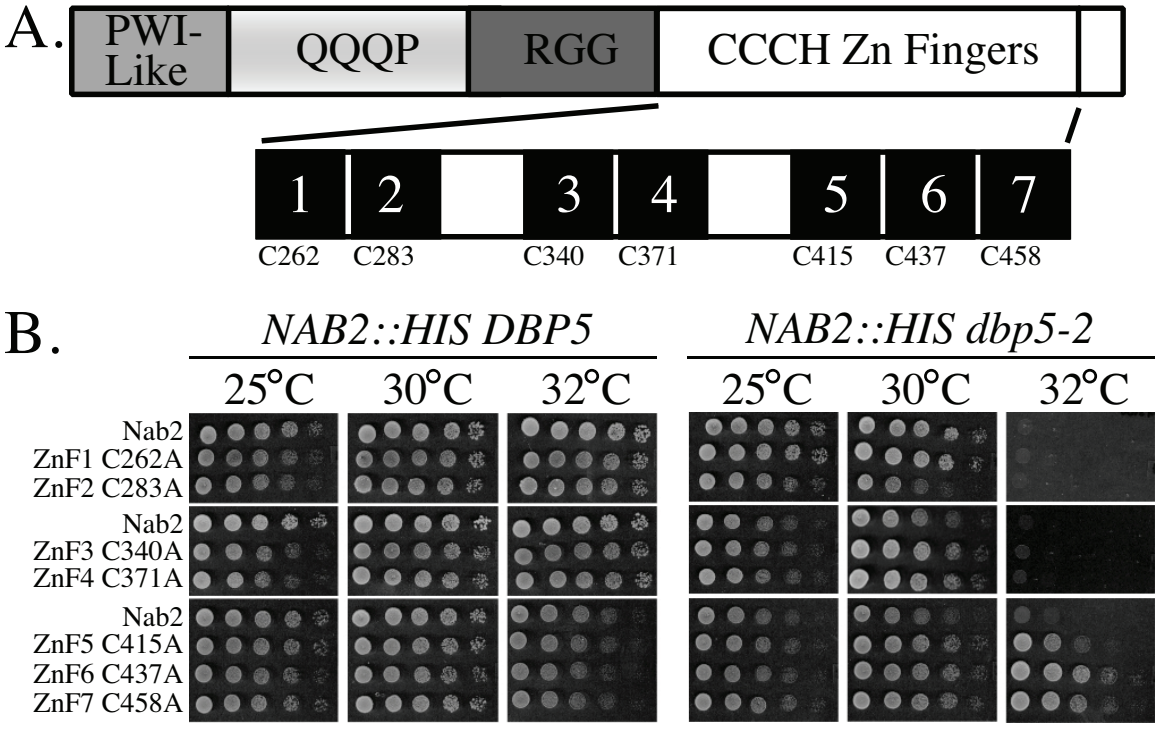

Figure 3.1: Cysteine to alanine substitutions in Nab2 zinc fingers 5-7 suppress the temperature sensitive phenotype of *dbp5-2* cells. (A) Schematic depicting the domains of *S. cerevisiae* Nab2. The C-terminal CCCH zinc finger domain contains seven tandem zinc fingers (black boxes). The position of the first cysteine of each zinc finger is indicated below the corresponding zinc finger. (B) *S. cerevisiae* plasmid shuffle assay analyzing the suppression of the temperature sensitive growth phenotype of *dbp5-2* cells by the indicated *nab2* mutants. Δ*NAB2* cells harboring a *URA3* plasmid encoding wild-type Nab2 and expressing either wild-type Dbp5 (left panel) or the mutant dbp5-2 (right panel) were transformed with *LEU2* plasmids encoding either wild-type Nab2 or Nab2 proteins that contain the denoted individual cysteine to alanine substitution in zinc fingers 1-7 (ZnF 1-7).

Table 3.1: Strains and plasmids used in chapter 3

Strain/Plasmid	Description	Source
ACY427	*ΔNAB2::HIS* (pAC636) *MAT**a** leu2 ura3*	(73)
ACY1669	*dbp5-2 ΔNAB2::HIS MAT**a** leu2 ura3 trp1*	(91)
pAC636	*NAB2, CEN, URA3*	(73)
pAC717	*NAB2, CEN, LEU2*	(73)
pAC1945	*pGEX-4T-1* (containing TEV cleavage site instead of Thrombin)	GE Healthcare
pAC2027	*nab2-C262A, CEN, LEU2*	This study
pAC2028	*nab2-C283A, CEN, LEU2*	This study
pAC2029	*nab2-C340A, CEN, LEU2*	This study
pAC2030	*nab2-C371A, CEN, LEU2*	This study
pAC2031	*nab2-C415A, CEN, LEU2*	This study
pAC2033	*nab2-C437A, CEN, LEU2*	This study
pAC2035	*nab2-C458A, CEN, LEU2*	This study
pAC2203	*nab2-C415A,C437A,C458A, CEN, LEU2*	This study
pAC2222	*nab2-C415R,C437R,C458R, CEN, LEU2*	This study
pAC2304	*GST-NAB2* Zinc Fingers 1-7	(67)
pAC2305	*GST-NAB2* Zinc Fingers 1-4	This study
pAC2307	*nab2-C437S, CEN, LEU2*	(91)
pAC2502	*nab2-Y428A, CEN, LEU2*	This study
pAC2503	*nab2-F450A, CEN, LEU2*	This study
pAC2504	*nab2-F460A, CEN, LEU2*	This study
pAC2505	*nab2-F471A, CEN, LEU2*	This study
pAC2522	*GST-NAB2* Zinc Fingers 5-7	This study
pAC2597	*GST-nab2*-Zinc Fingers 5-7 *F450A, F471A*	This study
pAC2618	*nab2-R438A, CEN, LEU2*	This study
pAC2620	*nab2-K416A, CEN, LEU2*	This study
pAC2624	*GST-nab2*-Zinc Fingers 5-7 *Y428A, F450A, F471A*	This study
pAC2647	*nab2-R459A, CEN, LEU2*	This study
pAC2648	*nab2-K416A, R438A, R459A, CEN, LEU2*	This study
pAC2672	*nab2-Y428A, F450A, F471A, CEN, LEU2*	This study
pAC2742	*GST-nab2-Zinc fingers 5-7 K416A*	This study
pAC2743	*GST-nab2-Zinc fingers 5-7 R438A*	This study
pAC2744	*GST-nab2-Zinc fingers 5-7 R459A*	This study

oligonucleotide [Cy3-r$(A)_{25}$]. As shown in Figure 3.2A, neither GST alone nor GST-Zinc Fingers (ZnF) 1-4 bound the Cy3-labeled poly(r$A)_{25}$ oligonucleotide, but GST-ZnF 1-7 and GST-ZnF 5-7 both bound to Cy3-r$(A)_{25}$. Binding of the Nab2 domain containing zinc fingers 5-7 is specific for stretches of polyadenosine RNA because only unlabeled poly(A) RNA oligonucleotide [r$(A)_{25}$] and not a randomized 25-nt RNA oligonucleotide [r$(N)_{25}$] could compete for binding (Figure 3.2B).

Our data argues that Nab2 zinc fingers 5-7 are sufficient for binding polyadenosine RNA *in vitro*. Therefore, we hypothesized that amino acid substitutions within this domain that are predicted to disrupt the structural integrity of the CCCH zinc fingers would impair the essential function of Nab2 *in vivo*. In order to test this hypothesis and examine the impact of these amino acid substitutions on Nab2 function *in vivo*, a plasmid shuffle assay was performed (153). Cells deleted for the essential *NAB2* gene and complemented by a wild-type *NAB2* maintenance plasmid were transformed with plasmids encoding either a control wild-type Nab2 or various nab2 mutants containing individual cysteine to alanine (C→A) as well as more drastic cysteine to arginine (C→R) amino acid changes in the first cysteine of zinc fingers 5-7. These cells were then serially diluted and spotted onto control plates and plates containing 5-fluoroorotic acid (5-FOA) to select against the wild-type maintenance plasmid and reveal any phenotype of the nab2 mutant proteins. Single C→A or C→R substitutions at the first cysteines of individual CCCH zinc fingers 1-7 did not impair Nab2 function as the cells expressing these mutants as the sole copy of Nab2 grew in a manner

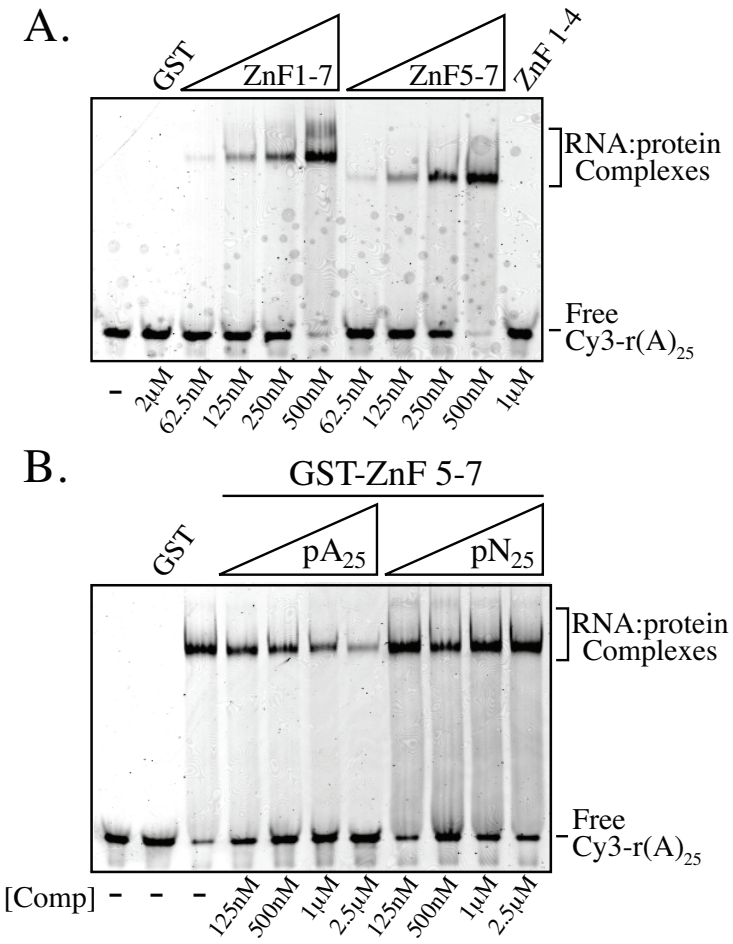

Figure 3.2: Nab2 zinc fingers 5-7 are necessary and sufficient to mediate high affinity specific binding to polyadenosine RNA. (A) GST, GST ZnF 1-7, GST ZnF 5-7, or GST ZnF 1-4 was incubated with ~125 nM Cy3-labeled poly(A) RNA oligonucleotide and RNA:protein complexes were resolved from free probe in a 5% non-denaturing polyacrylamide gel. The position of the free probe is evident (Free Cy3-$r(A)_{25}$) in the control lane with no protein added (-). (B) The binding specificity of GST ZnF 5-7 was investigated by incubating 250 nM GST ZnF 5-7 with 115 nM Cy3-$r(A)_{25}$ in the presence of unlabeled 25-nt poly(A) RNA oligonucleotide (pA_{25}) competitor or unlabeled 25-nt randomized RNA oligonucleotide (pN_{25}) competitor. RNA:protein complexes were resolved from free probe in a 5% non-denaturing polyacrylamide gel and the position of free Cy3-pA is indicated. Competitor concentration ([Comp]) and those samples containing no competitor (-) are indicated. As a control, no binding to Cy3-$r(A)_{25}$ is observed with 2 μM GST.

indistinguishable from wild-type cells (Figure 3.1B and Figure 3.3). As a control, *nab2-1* cells, which express a nab2 variant lacking the N-terminal 97 amino acids of Nab2 and display a severe growth defect (73), were also serially diluted and spotted on the same plates. To assess the consequence of impairing the function of multiple zinc fingers, we constructed plasmids encoding combinations of C→A or C→R amino acid substitutions at the first cysteine of zinc fingers 5-7 (nab2-$C_{5\text{-}7}$→A/R). Cells expressing *nab2* alleles containing these triple C→A or C→R amino acid substitutions showed greatly diminished viability at 18°C, suggesting that a combination of these amino acid changes impacts the essential function of the Nab2 protein by disrupting RNA binding. This loss of Nab2 function is not due to changes in the level of the Nab2 protein as immunoblotting reveals that all Nab2 proteins are expressed at approximately equal levels in exponentially growing yeast cells at either 30°C or 18°C (Data not shown).

In order to more completely assess the *in vivo* impact of disrupting the interaction of Nab2 with polyadenosine RNA, the *nab2-$C_{5\text{-}7}$→A* allele was integrated into the *NAB2* genomic locus using the method described in Experimental Procedures. Briefly, the zinc finger domain of nab2-$C_{5\text{-}7}$→A and the natamycin resistance (NAT^R) cassette were first amplified by PCR. These products were incubated together and overlap PCR was performed to amplify the entire nab2-$C_{5\text{-}7}$→A-NAT^R combinatorial product. The nab2-$C_{5\text{-}7}$→A-NAT^R combinatorial PCR product was then integrated into the *NAB2* genetic locus using conventional methods (169, 170). This integration yielded a mutant yeast strain that expresses nab2-$C_{5\text{-}7}$→A at the endogenous Nab2 locus.

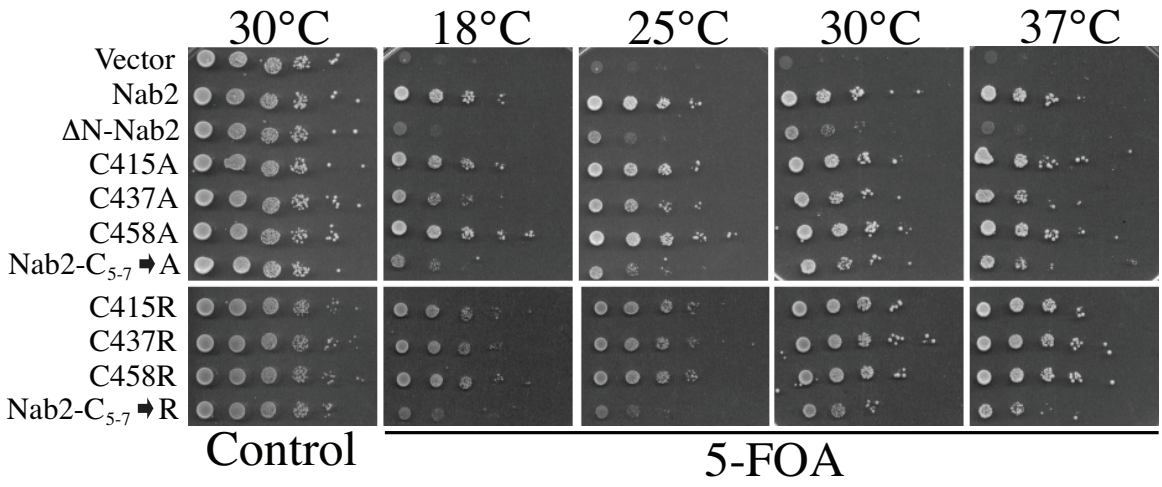

Figure 3.3: Combinatorial cysteine substitutions in Nab2 zinc fingers 5, 6, and 7 impair Nab2 function. Δ*NAB2* cells harboring a *URA3* plasmid encoding wild-type Nab2 were transformed with *LEU2* plasmids encoding either wild-type Nab2, ΔN-Nab2 (missing amino acids 3-97) or Nab2 proteins that contain the denoted individual or combinatorial cysteine to alanine or cysteine to arginine substitution(s) in zinc fingers 5-7. Nab2-C_{5-7}→A/R encodes full-length Nab2 proteins with cysteine to alanine or cysteine to arginine amino acid substitutions in the first cysteine of zinc fingers 5, 6, and 7. Transformants were inoculated into liquid media, grown to saturation at 30°C, and serially diluted. Dilutions were then spotted onto either control media lacking both uracil and leucine or media containing 5-fluoroorotic acid (5-FOA) to eliminate the wild-type Nab2 maintenance plasmid. The control plate was incubated at 30°C for 2 days while plates containing 5-FOA were incubated at the indicated temperatures for 3-5 days.

To assess the ability of integrated nab2-C_{5-7}→A function under the endogenous *NAB2* promoter, cells expressing an integrated copy of nab2-C_{5-7}→A were then transformed with either an empty vector or a plasmid encoding wild-type Nab2. Transformants were grown to saturation at room temperature, serially diluted, and spotted onto selective media. As shown in Figure 3.4, wild-type yeast cells transformed with either vector alone or a wild-type Nab2 plasmid grew at all temperatures tested. However, cells expressing nab2-C_{5-7}→A and transformed with an empty vector showed severe cold- and temperature-sensitive growth defects. Importantly, nab2-C_{5-7}→A cells transformed with a wild-type Nab2 plasmid grew similar to wild-type yeast cells, confirming that the growth defects of nab2-C_{5-7}→A cells are due to defects in Nab2 and not an extragenic mutation elsewhere in the genome.

In order to test whether a combination of cysteine to alanine substitutions within zinc fingers 5-7 disrupts RNA binding *in vitro*, we expressed and purified recombinant wild-type GST-ZnF 1-7; GST-ZnF 1-7 C415A, C437A, C458A (GST-ZnF 1-7 C_{5-7}→A); wild-type GST-ZnF 5-7; and GST-ZnF 5-7 C415A, C437A, C458A (GST-ZnF 5-7 C_{5-7}→A) and analyzed binding of each of these proteins to a Cy3-labeled polyadenosine RNA oligonucleotide in RNA gel shifts assays. As diagrammed in Figures 3.5A and 3.6A, respectively, both GST-ZnF 5-7 C_{5-7}→A and GST-ZnF 1-7 C_{5-7}→A contain cysteine to alanine amino acid changes in the first cysteines of zinc fingers 5 (C415A), 6 (C437A) and 7 (C458A).

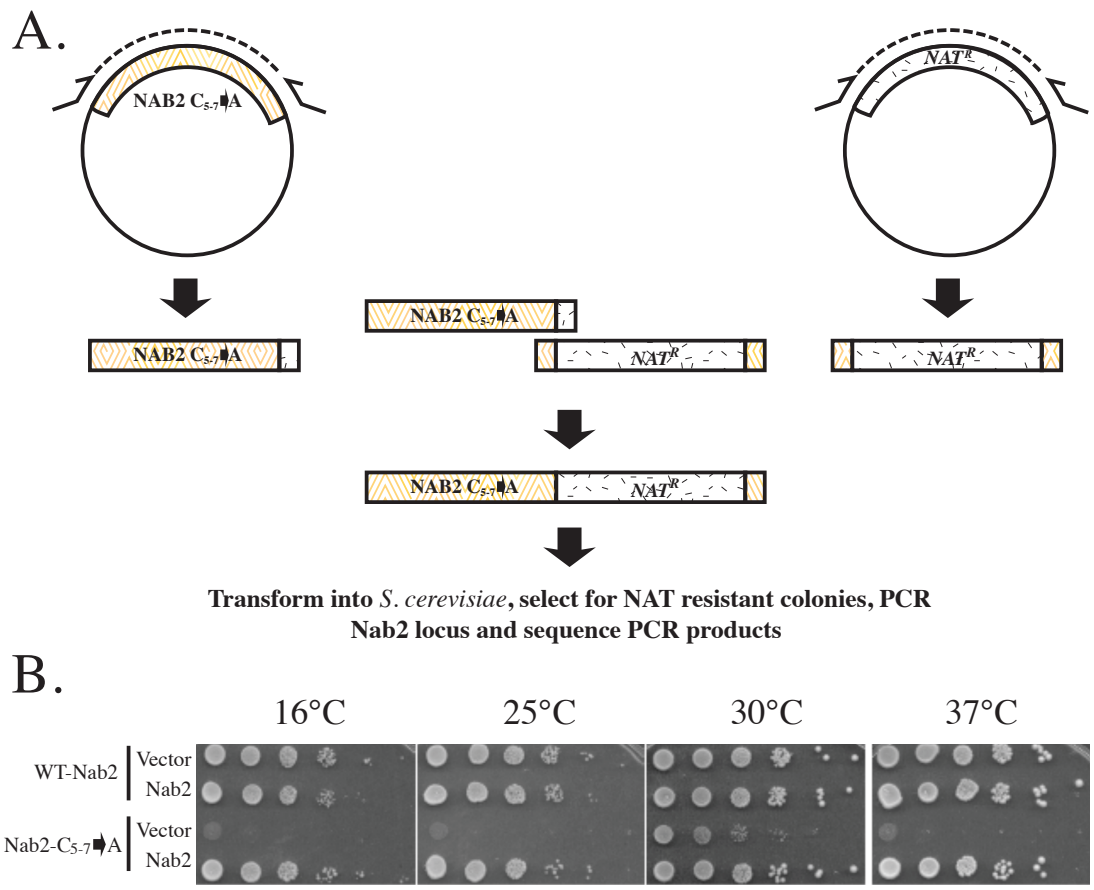

Figure 3.4: Integration of nab2-C_{5-7}→A into *S. cerevisiae* cells. (A) Diagram of integration strategy. First, the C-terminus of nab2-C_{5-7}→A and the natamycin resistance (NAT^R) cassette were both amplified by PCR. The 3'-end of the 3' primer used for amplification of nab2-C_{5-7}→A is complementary to the 5'-end of the NAT^R cassette. Similarly, the 5'-end of the 5' primer used for amplification of the NAT^R cassette is complementary to 3'-end of the nab2-C_{5-7}→A PCR product. Equal amounts of PCR products were mixed and overlap PCR was performed. The resulting combinatorial PCR products containing nab2-C_{5-7}→A::NAT^R are transformed into wild-type *S. cerevisiae* cells and integration occurs by homologous recombination of the PCR product into the *NAB2* genomic locus. Integrants are selected by growth on media containing the drug natamycin. Genomic DNA is isolated from natamycin resistant colonies and the Nab2 locus is amplified by PCR to check for correct integration. PCR products are sequenced to ensure correct recombination of *nab2-C_{5-7}→A* into the *NAB2* locus. (B) Integrated nab2-C_{5-7}→A confers a severe cold-sensitive growth defect. Cells containing a wild-type copy of Nab2 or integrated nab2-C_{5-7}→A were transformed with either an empty vector or a plasmid encoding a wild-type copy of Nab2. Cells were then grown to saturation in liquid culture, normalized according to cell density at OD_{600}, serially diluted, and spotted onto selective media. Plates were grown at 16°C, 25°C, 30°C, and 37°C as indicated.

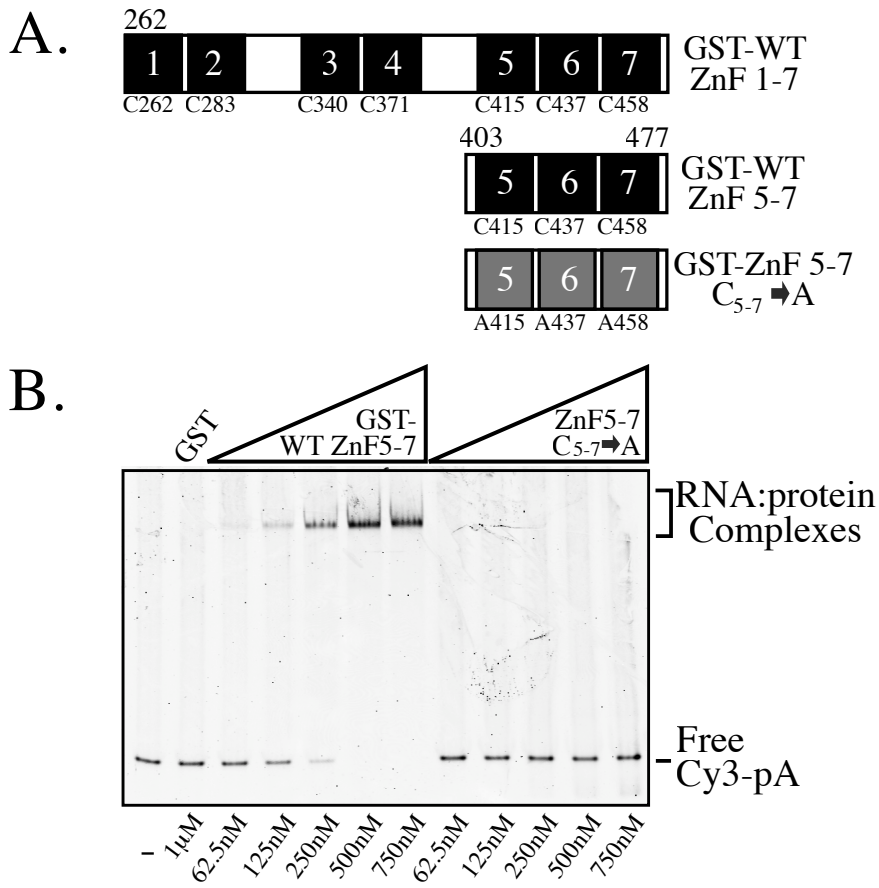

Figure 3.5: Combinatorial cysteine to alanine substitutions in Nab2 zinc fingers 5, 6, and 7 cause defects in polyadenosine RNA binding. (A) Schematic of proteins used in RNA gel shift assays. GST ZnF 5-7 proteins contain Nab2 residues 403-477 while GST ZnF 1-7 proteins contain Nab2 residues 262-477. GST ZnF 5-7 C_{5-7}→A and GST ZnF 1-7 C_{5-7}→A contain cysteine to alanine substitutions in the first cysteines of zinc fingers 5 (C415A), 6 (C437A), and 7 (C458A). (B) Combinatorial cysteine to alanine substitutions in ZnF 5-7 abolish RNA binding. GST (1 μM), GST ZnF 5-7 (62.5 nM – 500 nM), or GST ZnF 5-7 C_{5-7}→A (62.5 nM – 500 nM) was incubated with ~125 nM Cy3-labeled poly(A) RNA oligonucleotide. RNA:protein complexes were resolved from free probe by electrophoresis in a 5% non-denaturing polyacrylamide gel. The position of the free probe (Cy3-r$(A)_{25}$) is evident in the control lane with no protein added (-).

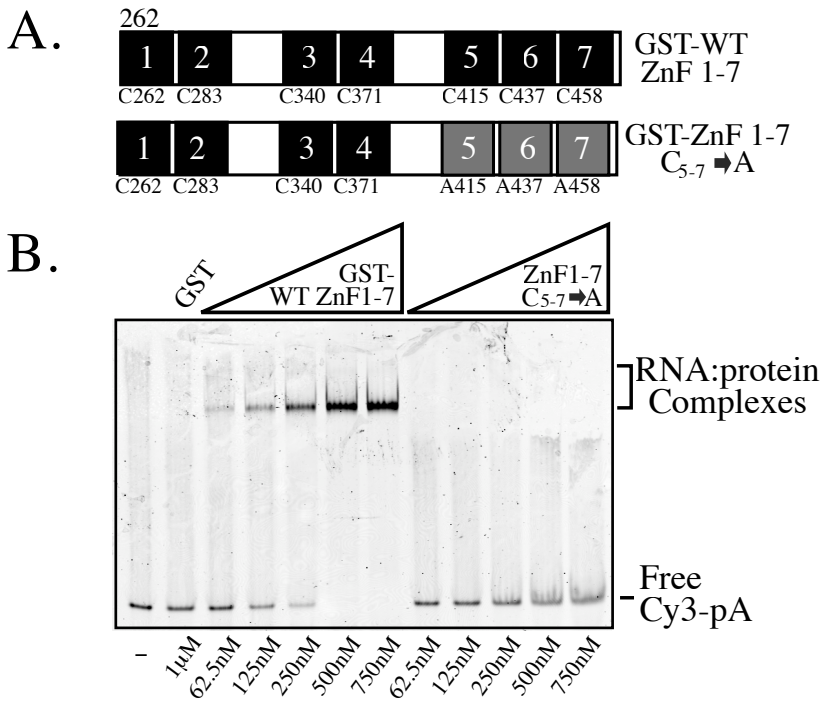

Figure 3.6: Combinatorial cysteine to alanine substitutions in Nab2 zinc fingers 5, 6, and 7 cause defects in polyadenosine RNA binding. (A) Schematic of proteins used in RNA gel shift assays. GST ZnF 5-7 proteins contain Nab2 residues 403-477 while GST ZnF 1-7 proteins contain Nab2 residues 262-477. GST ZnF 5-7 C_{5-7}→A and GST ZnF 1-7 C_{5-7}→A contain cysteine to alanine substitutions in the first cysteines of zinc fingers 5 (C415A), 6 (C437A), and 7 (C458A). (B) Combinatorial cysteine to alanine substitutions in ZnF 5-7 abolish RNA binding in the context of the entire Nab2 zinc finger domain. GST (1 μM), GST ZnF 1-7 (62.5 nM – 500 nM), or GST ZnF 1-7 C_{5-7}→A (62.5 nM – 500 nM) were incubated with ~125 nM Cy3-labeled poly(A) RNA oligonucleotide. RNA:protein complexes was resolved from free probe (Cy3-pA) by electrophoresis in a 5% non-denaturing polyacrylamide gel. Competitor concentration ([Comp]) and those samples containing no competitor (-) are indicated. As a control, no binding to Cy3-$r(A)_{25}$ is observed with 1 μM GST.

As shown in Figure 3.5B, wild-type GST ZnF 5-7 bound Cy3-labeled poly(A) RNA oligonucleotides; however, GST-ZnF 5-7 C_{5-7}→A showed no detectable binding to Cy3-r$(A)_{25}$ in this assay. To determine whether Nab2 zinc fingers 5-7 were necessary for polyadenosine RNA binding, we changed the first cysteine of zinc fingers 5-7 to alanine in the context of the full zinc finger domain (ZnF 1-7). As expected [Figure 3.2 and (67)], wild-type GST-ZnF 1-7 bound the Cy3-labeled polyadenosine oligonucleotide, however GST-ZnF 1-7 C_{5-7}→A did not bind Cy3-r$(A)_{25}$ (Figure 3.6B). This *in vitro* binding data strongly supports our conclusion that Nab2 zinc fingers 5, 6, and 7 are required for high affinity Nab2 binding to polyadenosine RNA.

The RNA gel shift experiments utilize a concentration of fluorescently labeled RNA oligonucleotides (~125 nM) that is approximately 4 times the published K_d (~30 nM) of Nab2 for polyadenosine oligos (67, 71) and therefore could not be used to accurately calculate dissociation constants. Furthermore, due to the high concentration of labeled RNA, weaker binding events would be less readily detected in this assay due to the high protein concentrations needed to cause an observable mobility shift in the Cy3-labeled poly$(A)_{25}$ RNA oligonucleotides. Therefore, to more quantitatively analyze the interaction between Nab2 zinc fingers 5-7 and polyadenosine RNA, we used fluorescence anisotropy to determine dissociation constants for both full-length Nab2 and a Nab2 protein that lacked zinc fingers 5-7 (Nab2-ΔZnF 5-7). Both proteins were expressed recombinantly in *E. coli* as untagged proteins and purified using the same purification strategy (See Experimental Procedures). As demonstrated in Figure 3.7, full-length Nab2 bound fluorescein-labeled poly$(A)_{25}$ RNA oligonucleotides with an apparent affinity of ~70 nM. Interestingly, Nab2-ΔZnF 5-7 bound fluorescein-labeled

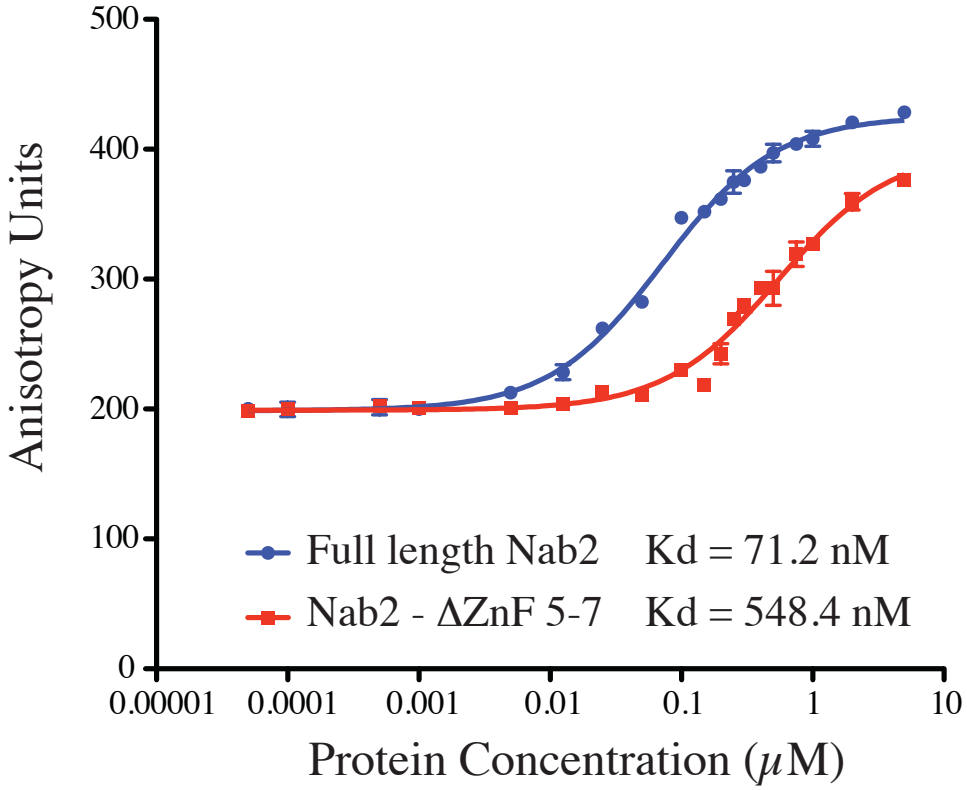

Figure 3.7: Nab2 ΔZnF 5-7 retains partial binding activity to polyadenosine RNA. Full-length Nab2 and Nab2 ΔZnF 5-7 were expressed and purified from BL21 (DE3) pLYS *E. coli* cells. The affinity of each protein for a fluorescein-labeled poly(A)$_{25}$ RNA oligonucleotide was determined using fluorescence anisotropy. Fluorescein-poly(A)$_{25}$ RNA (2 nM) was incubated with increasing protein concentrations (50 pM – 5 μM). Each sample was incubated for ~1 hour to reach equilibrium and then placed into one well of a 384 well plate. This analysis was done in triplicate for each protein concentration as described in Experimental Procedures. Anisotropy values were then determined using a fluorescence plate reader.

poly(A)$_{25}$ RNA oligonucleotides with an apparent K_d of ~540nM. While our previous data suggested that Nab2 zinc fingers 5-7 alone were sufficient for binding to polyadenosine RNA *in vitro*, weak binding of Nab2-ΔZnF 5-7 to fluorescein-labeled poly(A)$_{25}$ RNA oligonucleotides suggests that other domains within the protein may also harbor some affinity for polyadenosine RNA. Binding of Nab2-ΔZnF 5-7 to polyadenosine RNA is also consistent with a previous study in which Nab2 deletion mutants were cross-linked to poly(A) RNA *in vitro* (73). Although greatly diminished compared to wild-type Nab2, Nab2-ΔZnF 5-7 was still cross-linked to poly(A) RNA, suggesting that other domains within Nab2 can mediate this interaction (73).

As an unbiased biochemical approach to define an independently folding domain within Nab2 that binds polyadenosine RNA and would be amenable to structural studies, GST-tagged recombinant full-length Nab2 was expressed and purified from *E. coli* and subjected to partial trypsin proteolysis. Trypsinized Nab2 fragments were then incubated with poly(A) sepharose and bound fragments were eluted, separated by SDS-PAGE, and transferred to a PVDF membrane. A prominent band of approximately 10 kDa was then N-terminally sequenced to define the high affinity Nab2 poly(A) binding domain (Figure 3.8). This analysis led to the identification of three nested protein fragments with N-termini immediately upstream of zinc finger 5. Based upon the mass of these protein fragments, each was predicted to contain zinc fingers 5-7. This data supports the conclusion that the high affinity Nab2 polyadenosine RNA recognition motif is contained within zinc fingers 5-7 and that these zinc fingers alone are sufficient for specific recognition of polyadenosine RNA.

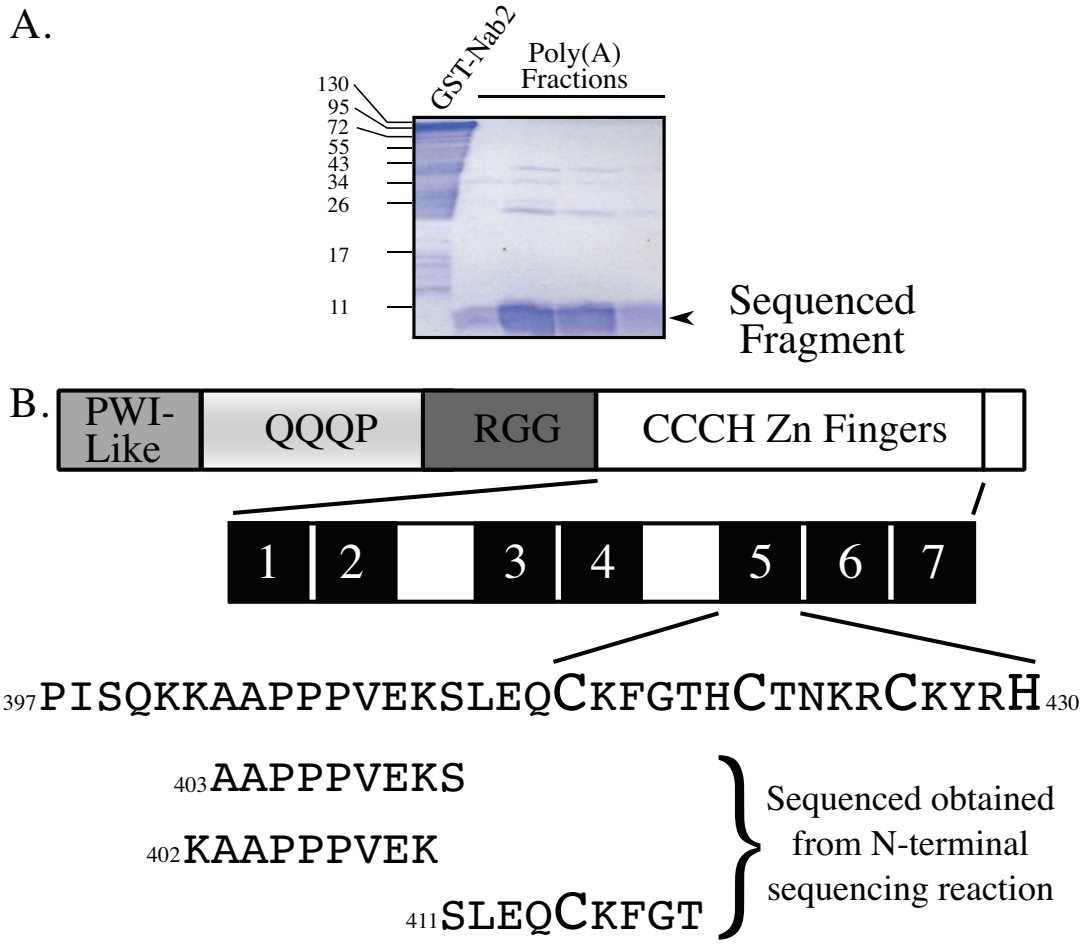

Figure 3.8: Nab2 proteolytic fragments recognize poly(A) sepharose. Partial trypsin proteolysis of Nab2 was used to identify an independently folding polyadenosine RNA recognition domain. (A) Purified recombinant GST-Nab2 was incubated with 20 μg of trypsin at room temperature for 30 min. The entire reaction was then incubated with poly(A) sepharose 4B, the sepharose was washed, and the bound RNA binding proteins were eluted by four sequential incubations with binding buffer containing increasing amounts of NaCl (Poly(A) Fractions). Eluted protein fragments were resolved by SDS-PAGE and stained with Coomassie Blue or transferred to PVDF membranes for N-terminal sequencing. As indicated, the ~10 kDa band was sequenced by Edmann degradation N-terminal sequencing. (B) The domains contained within Nab2. The C-terminal zinc finger domain consists of seven tandem CCCH zinc fingers. The amino acid sequence of zinc finger five as well as the upstream 18 amino acids are shown with numbers corresponding to the amino acid number of the full length Nab2 protein. The Cys and His residues of the CCCH zinc finger five are shown larger than the surrounding residues. All N-termini sequences obtained from the N-terminal protein sequencing reaction aligned with zinc finger five or the amino acid sequence immediately N-terminal to zinc finger five as shown. The most prevalent Nab2 protein fragment that was eluted from the poly(A) sepharose began at Nab2 residue 403. Less abundant protein fragments beginning at Nab2 residues 402 and 411 were also eluted from the poly(A) sepharose.

In order to understand the precise mechanism by which Nab2 zinc fingers 5-7 specifically recognize polyadenosine RNA, we collaborated with Christoph Brockmann, a post-doctoral researcher in Murray Stewart's lab at MRC Cambridge, to solve the solution structure of Nab2 zinc fingers 5-7 (Nab2 ZnF 5-7) using NMR. Large amounts of Nab2 Zn 5-7 (corresponding to residues 404 – 483) were expressed in *E. coli* and purified by ion exchange and gel filtration chromatography. The NMR structure was determined using standard 3D-NMR spectroscopy in conjunction with CYANA and X-plor NIH. A preliminary ensemble of 20 calculated structures is shown in Figure 3.9. The peptide backbone of residues Ser 411 – Pro 480 are well ordered in all 20 solved structures (Figure 3.9A). However, residues 404-411 and, to a lesser extent, 480-483 were relatively unordered in the ensemble of structures. The unordered ends of the Nab2 zinc finger domain are consistent with our previous limited trypsin proteolysis experiment demonstrating that the N-terminus of this domain could begin at several residues in close proximity with one another (Figure 3.7). Interestingly, the linker amino acids between zinc fingers 5, 6, and 7 (black ribbons in Figure 3.10) are also well structured. Hetero-NOE enhancement experiments indicate no major inter-domain flexibility. The side chains of the amino acids in zinc fingers 5-7 show some flexibility, but overall are in relatively similar conformations (Figure 3.10B). Unfortunately, the addition of poly(A) RNA oligonucleotides caused complete loss of discernable NMR spectra, presumably due to protein aggregation (data not shown). Therefore, no meaningful data could be obtained about the interaction of the entire zinc finger domain and poly(A) RNA oligonucleotides.

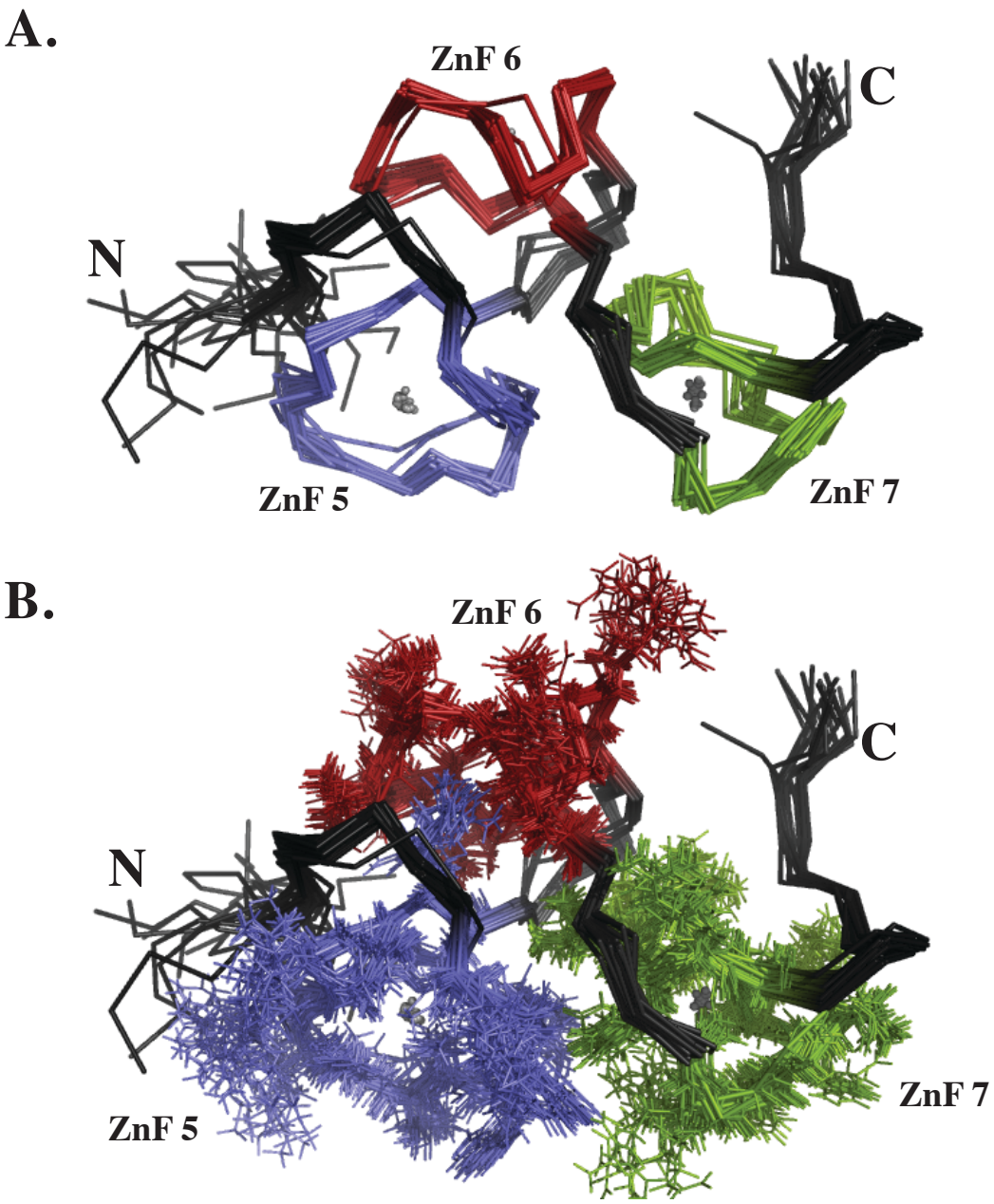

Figure 3.9: Ensemble of NMR solution structures for Nab2 ZnF 5-7. (A) Ribbon diagram of the protein backbone only from the ensemble of 20 solved NMR structures of Nab2 ZnF 5-7 (residues 404 – 483). The NMR structures were determined using standard 3D-NMR spectroscopy in conjunction with CYANA and X-plor NIH. Zinc fingers 5 (blue-purple), 6 (red), and 7 (green) as well as the N- and C-termini are labeled. Shown in black are the residues between each CCCH zinc finger. (B) Same as (A) except that stick representations are shown for the residues within each zinc finger.

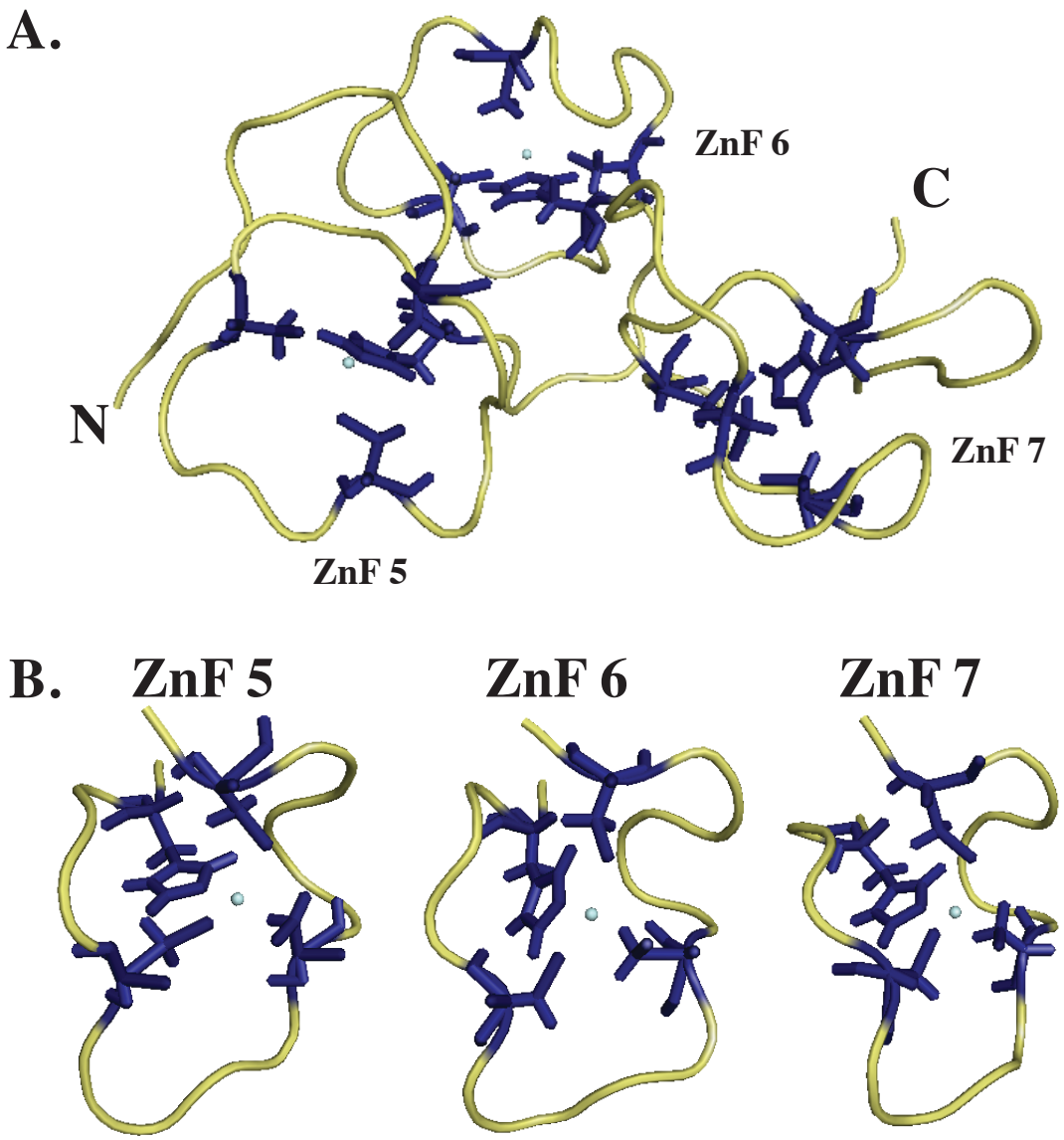

Figure 3.10: Ribbon diagram of the atomic resolution structure of Nab2 zinc fingers 5-7. (A) A preliminary average NMR structure was calculated from the ensemble of 20 solved lowest energy structures using the PyMOL script, average3d.py (115). While the N- and C-termini of this mean structure remain somewhat unreliable due to the flexibility of this domain, the residues within each zinc finger are well ordered. The backbone RMSD to the mean for all three domains is 1.46 Å. (B) The residues of zinc fingers 5, 6, and 7 are well ordered. The three-dimensional fold of each zinc finger is also very similar to one another. The RMSD for the individual fingers to the mean is 1.01 Å, 0.32 Å and 0.54 Å for fingers 5, 6, and 7, respectively.

From the ensemble of 20 structures, an average structure was generated using the average3d.py PyMOL script (115). As seen in Figure 3.10A, Nab2 zinc fingers 5-7 are composed entirely of loops; no α-helical or β-sheet secondary structure exists in this domain as determined by these structural analyses. Interestingly, the cysteine and histidine residues of these domains are also well ordered. As is apparent in Figure 3.10B, where the structural cysteines and histidines are displayed in navy blue, the three-dimensional structure of each zinc finger is very similar. Each zinc finger consists of the amino acid side chains of three cysteines and one histidine chelating one zinc ion. The backbone RMSD to the mean for all three domains is 1.46 Å, while the RMSD for the individual fingers to the mean is 1.01 Å, 0.32 Å and 0.54 Å for fingers 5, 6, and 7, respectively.

In order to identify specific amino acids that have the potential to directly contact polyadenosine RNA, we examined a model of the three dimensional structure of Nab2 zinc fingers 5-7 to identify conserved aromatic and charged residues that were solvent exposed (Figure 3.11). We identified several aromatic residues (Tyr 428 and Phe 450 – shown in purple) and several positively charged residues (Lys 417, Arg 438, and Arg 459 – shown in red) within zinc fingers 5-7 of Nab2 that are both solvent exposed and conserved within each of the zinc fingers in Nab2 and other members of this zinc finger protein family (Figure 3.11A). To analyze the contribution of these aromatic and charged residues to the RNA binding function of Nab2, each residue was changed to alanine, either individually or in combination. Mutant nab2 proteins containing these amino acid substitutions were then tested for suppression of the temperature-sensitive growth defect

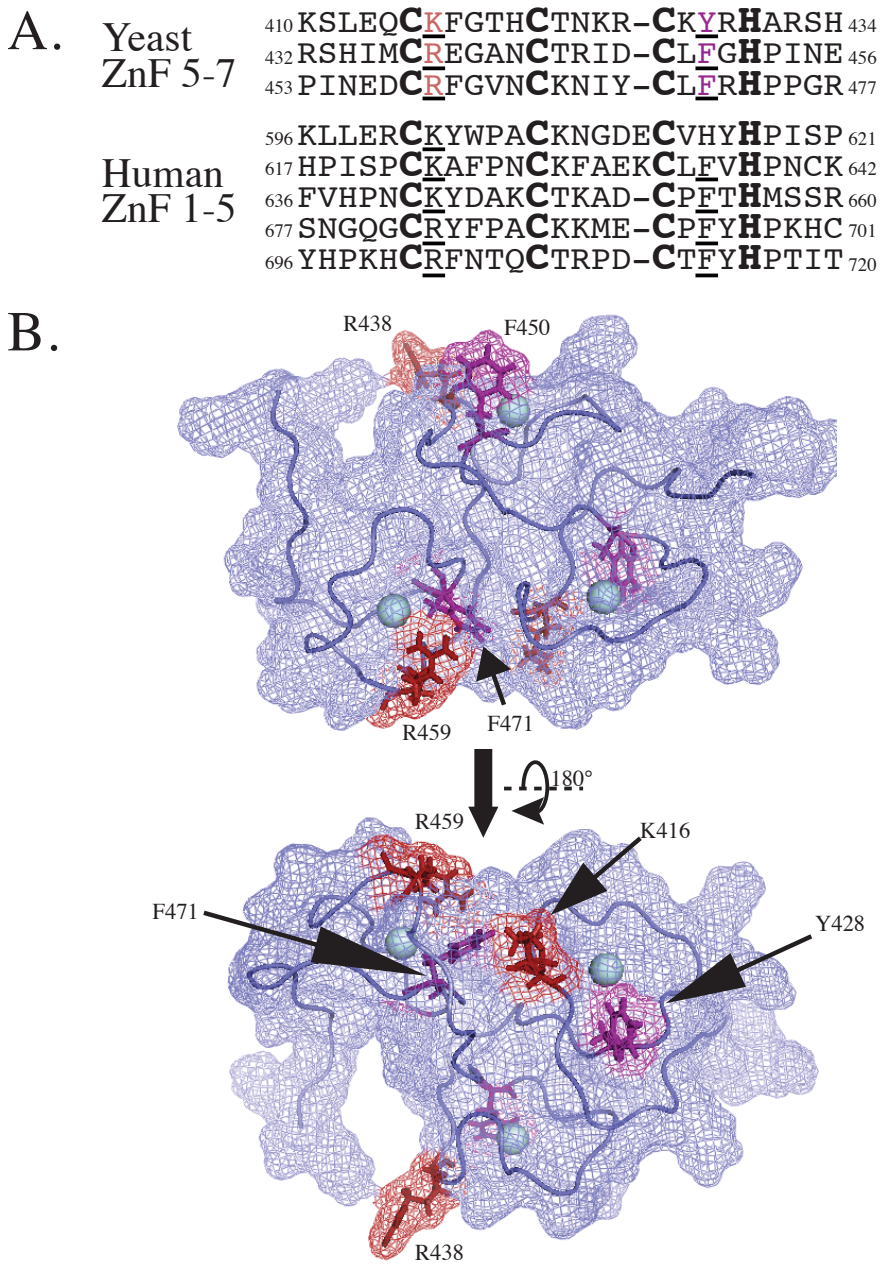

Figure 3.11: Analysis of Nab2 ZnF 5-7 structure reveals the presence of several conserved charged and aromatic residues that may interact with poly(A) RNA. (A) Amino acid sequence alignment of *S. cerevisiae* Nab2 zinc fingers 5-7 and the human ZC3H14 zinc fingers 1-5. Conserved positively charged residues and aromatic residues are underlined. Numbers of the corresponding residues within yeast Nab2 and human ZC3H14 are shown. (B) Mesh surface representation of the average structure of Nab2 zinc fingers 5-7. Stick representations of conserved positively charged residues (K416, R438, and R459) and conserved aromatic residues (Y428, F450, and F471) are shown in red and purple, respectively. Phe 471 (F471) is buried in within the globular fold of the proteins, while the remainder of the charged and aromatic residues are solvent exposed.

of *dbp5-2* cells, as well as growth defects when mutant alleles of *nab2* were the only copy of Nab2 present. Finally we tested whether these changes in the zinc finger domain impacted Nab2 binding to poly(A) RNA. As a control, another conserved phenylalanine in zinc finger 5, Phe 471, that is not solvent exposed but instead buried within the globular fold of the domain (Figure 3.11B) was also targeted. We also tested the contributions of one additional solvent-exposed phenylalanine in zinc finger 7, Phe 460, to polyadenosine RNA binding. With the exception of F471A, these alanine substitutions are not predicted to change the overall structure of the Nab2 zinc finger domain.

To begin to assess the contributions of each of these residues to polyadenosine RNA recognition, we first analyzed the ability of each of these residues to suppress the temperature-sensitive growth defect of *dbp5-2*. As described previously, this genetic assay allows us to analyze the putative contributions of individual residues to polyadenosine RNA binding. As expected (91), nab2-C437S suppresses *dbp5-2*. Expression of nab2-F450A as the only copy of Nab2 suppressed the temperature sensitive growth phenotype of *dbp5-2* cells (Figure 3.12B), while expression of nab2-Y428A or nab2-F460A did not suppress the temperature sensitive growth phenotype of *dbp5-2* cells, suggesting that specific aromatic residues may make direct contacts with polyadenosine RNA while other aromatic residues are not required for polyadenosine RNA recognition. Interestingly, expression of the buried nab2-F471A did suppress the temperature sensitive growth defect of *dbp5-2*, suggesting that changing this buried hydrophobic residue to alanine may be affecting the overall three-dimensional structure of the tandem zinc finger domain and cause a defect in poly(A) RNA recognition. As

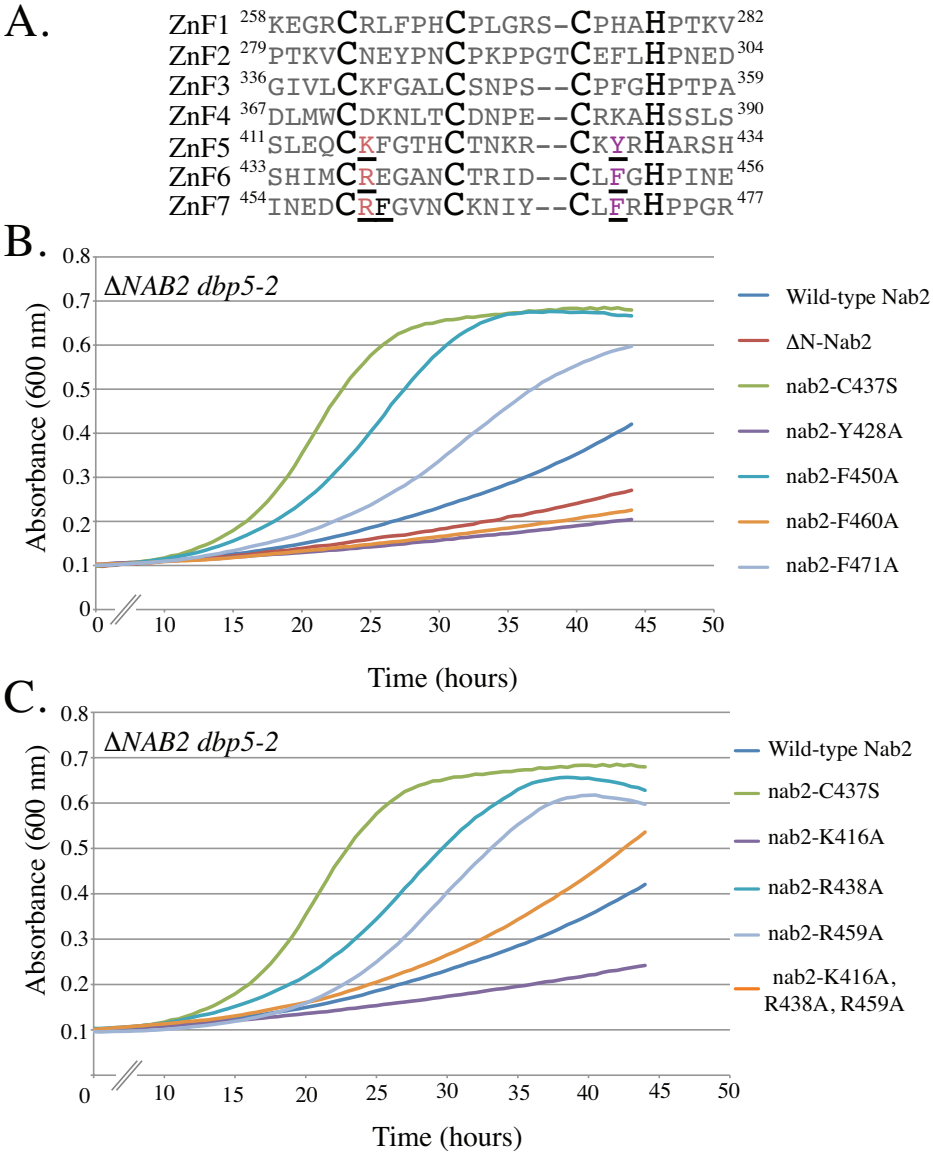

Figure 3.12: Nab2 mutants containing aromatic and positively charged amino acid substitutions in Nab2 zinc fingers 5-7 suppress *dbp5-2*. (A) Amino acid alignment of Nab2 zinc fingers 1-7. Conserved positively charged and aromatic residues are underlined and highlighted in red and purple, respectively. (B) and (C) Plasmid shuffle assay analyzing the suppression of the temperature sensitive growth phenotype of *dbp5-2* cells by the indicated *nab2* mutants. Δ*NAB2* cells harboring a *URA3* plasmid encoding wild-type Nab2 and expressing mutant dbp5-2 were transformed with *LEU2* plasmids encoding either wild-type Nab2 or Nab2 proteins that contain the denoted aromatic or positively charged amino acid substitution in zinc fingers 5-7. Cells were grown on media containing 5-FOA to select against the wild-type maintenance plasmid. Cells were then grown to saturation in liquid culture at 30°C and cell density was normalized by absorbance at 600 nm. Cells were diluted 1:50 into selective media and cell growth at 32°C was monitored in a plate reader every 30 minutes for ~48 hours by absorbance at 600 nm.

shown in Figure 3.12C, expression of nab2-R438A, nab2-R459A, and nab2-K416A, R438A, R459A also suppressed *dbp5-2*, while nab2-K416A did not, suggesting that only the positively charged residues (Arg 438 and Arg 459) within zinc fingers 6 and 7 interact with polyadenosine RNA. Interestingly, Phe 471 is adjacent to Arg 459 in the three-dimensional structure of the zinc fingers 5-7 (Figure 3.11), suggesting that F471A-nab2 may suppress *dbp5-2* by changing the positioning of Arg 459.

We also tested the function of each of these mutant proteins by determining whether they could function of the only copy of Nab2 *in vivo*. Cell growth at 22°C was monitoring by absorbance at 600 nm over ~48 hours (Figures 3.13B and 3.13C). As shown in Figure 3.13B, yeast cells expressing the individual aromatic mutants Y428A, F450A, F460A, or F471A as the only copy of the essential *NAB2* gene grows in a manner indistinguishable from cells expressing a wild-type *NAB2* allele. However, cells expressing an allele of *nab2* that encodes a triple (Y428A, F450A, F471A) mutant show a significant cold-sensitive growth defect. Similar to the individual aromatic amino acid substitutions, nab2-K416A, nab2-R438A, and nab2-R459A also each grew indistinguishably from wild-type Nab2. ΔN-Nab2 showed significant growth defects at 22°C and nab2-C437S was only slightly cold-sensitive, as expected (73, 91). Interestingly, the combinatorial nab2-K416A, R438A, R459A mutant also shows cold-sensitive growth defects, but not to the same extent as nab2-Y428A, F450A, F471A (Figure 3.13C). In sum, none of the individual alanine substitutions significantly impair Nab2 function and cell growth. Only nab2 mutants containing combinatorial aromatic or basic amino acid substitutions in zinc fingers 5, 6, and 7 confer cold-sensitive growth defects.

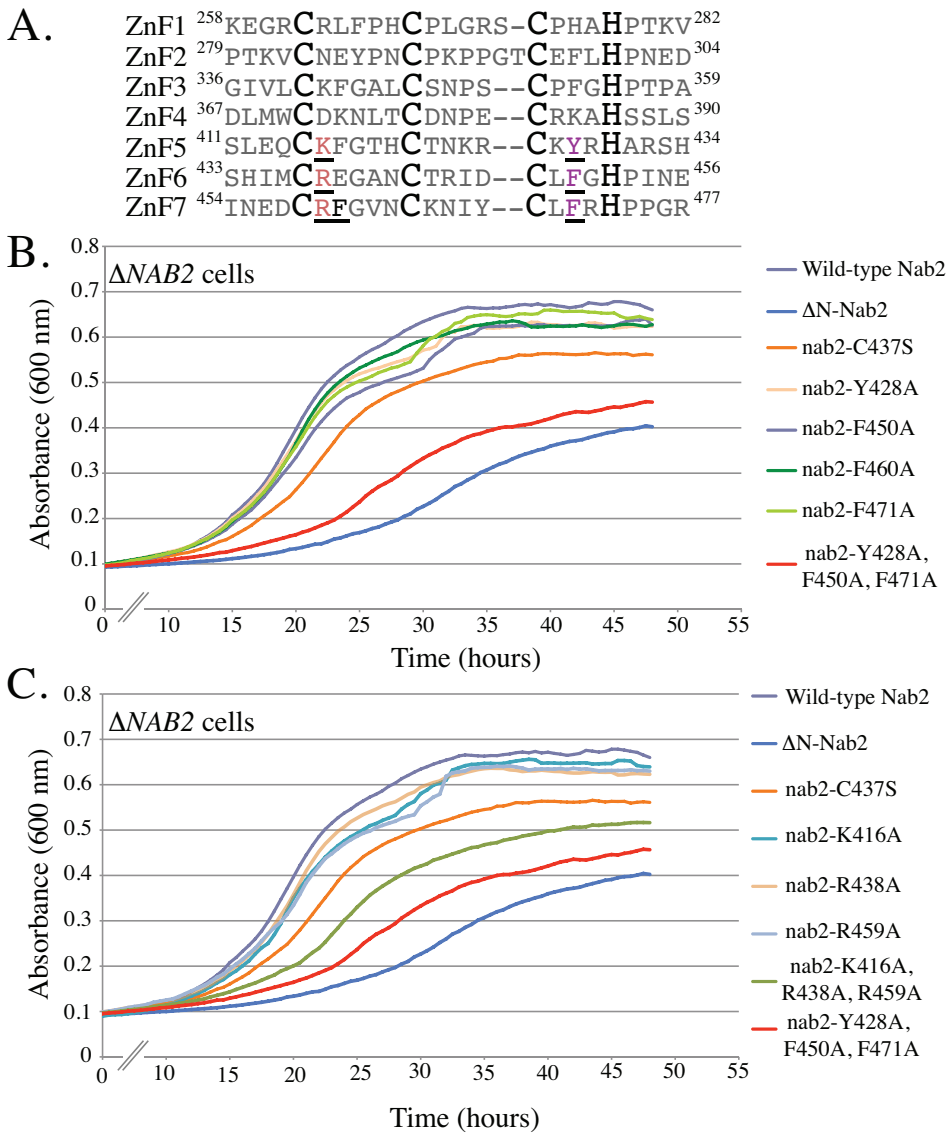

Figure 3.13: Nab2 mutants containing combinations of aromatic and charged amino acid substitutions in Nab2 zinc fingers 5-7 confer cold sensitive growth defects. (A) Amino acid alignment of Nab2 zinc fingers 1-7. Conserved positively charged and aromatic residues are underlined and highlighted in red and purple, respectively. (B) and (C) *S. cerevisiae* plasmid shuffle assay analyzing the cold-sensitive growth phenotype of *nab2* mutant cells. Δ*NAB2* cells harboring a *URA3* plasmid encoding wild-type Nab2 were transformed with *LEU2* plasmids encoding either wild-type Nab2 or Nab2 proteins that contain the denoted aromatic or positively charged amino acid substitution in zinc fingers 5-7. Cells were grown on media containing 5-FOA to select against the wild-type maintenance plasmid. Cells were then grown to saturation in liquid culture at 25°C and cell density was normalized by absorbance at 600 nm. Cells were diluted 1:100 into selective media and cell growth at 22°C was monitored in a plate reader every 30 minutes for ~48 hours by absorbance at 600 nm.

Finally, in order to more precisely assess the contributions of these aromatic residues to polyadenosine RNA recognition, wild-type GST-ZnF 5-7, GST-ZnF 5-7 F450A/F471A, GST-ZnF 5-7-K416A, GST-ZnF 5-7-R438A, and GST-ZnF 5-7-R459A were expressed and purified from *E. coli* and binding of each of these proteins to Cy3-$r(A)_{25}$ was analyzed using an RNA gel shift assay (Figures 3.14 – 3.17). As expected, wild-type GST-ZnF 5-7 binds Cy3-$r(A)_{25}$. However, introduction of the two amino acid substitutions that suppressed *dbp5-2*, F450A and F471A, completely abolishes binding between the Nab2 zinc finger domain and Cy3-$r(A)_{25}$. In addition, nab2-K416A (Figure 3.15), nab2-R438A (Figure 3.16), or nab2-R459A (Figure 3.17) did not bind Cy3-labeled poly(A) RNA oligonucleotides, suggesting that each of these residues is critical for polyadenosine RNA recognition as well. Taken together, this data suggests that poly(A) RNA recognition by Nab2 requires specific aromatic and basic amino acids within zinc fingers five and seven.

Our data thus far suggests that Nab2 zinc fingers 5-7 are sufficient to specifically recognize polyadenosine RNA. Furthermore, we have solved the solution structure of this domain using NMR. Amino acid substitutions predicted by this structure to contribute to recognition of poly(A) RNA impair binding between Nab2 ZnF 5-7 and Cy3-labeled poly(A) RNA oligonucleotides and confer cold-sensitive growth defects. However, the basis for this growth defect and the exact functional significance of disrupting Nab2 binding to polyadenosine RNA *in vivo* are not known. As mentioned previously, Nab2 has been implicated in the control of poly(A) tail length as well as poly(A) RNA export from the nucleus (71, 73). Presumably, the growth defects seen in

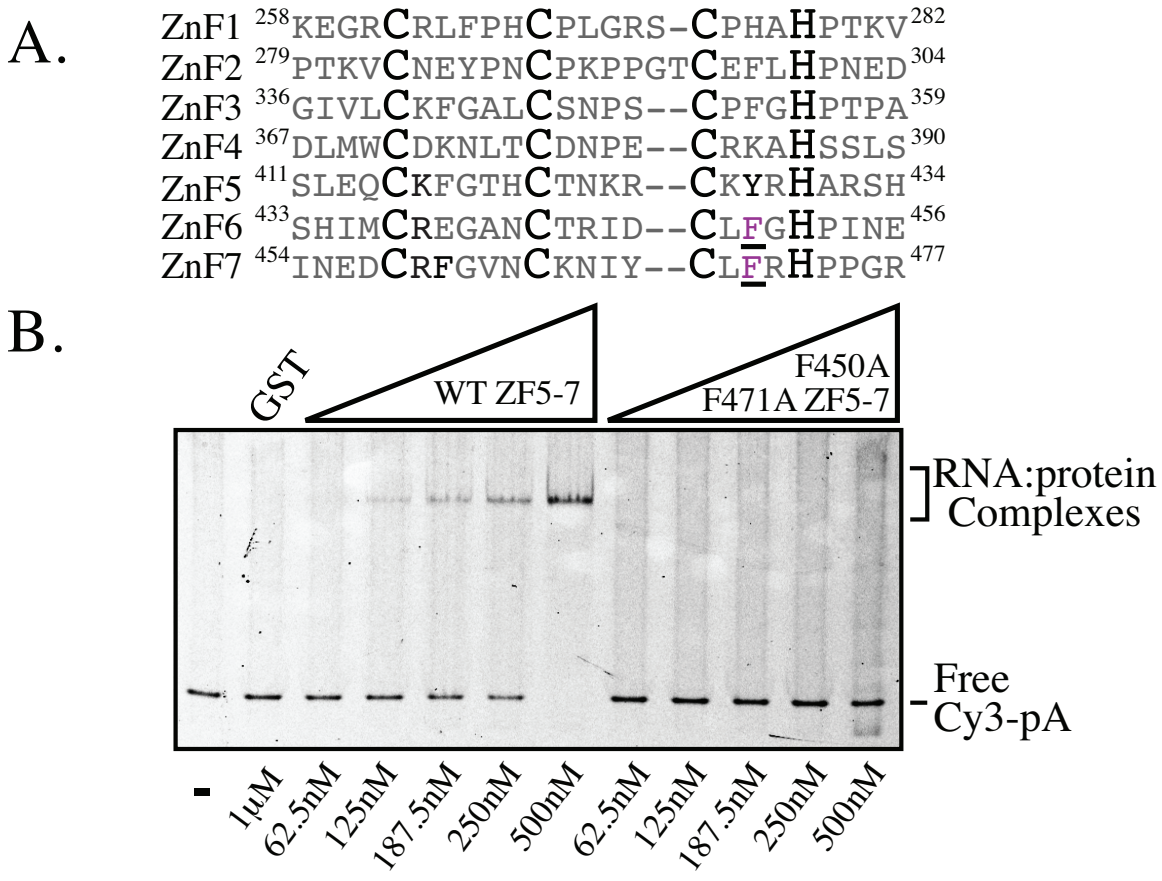

Figure 3.14: Evolutionarily conserved aromatic residues in Nab2 zinc fingers 5-7 are critical for Nab2 binding to polyadenosine RNA. ***(A)*** Alignment of the seven Nab2 CCCH zinc fingers. Numbers correspond to the first and last amino acid are shown surrounding each zinc finger. Cysteines and histidines are larger and shown in bold while those conserved aromatic residues (F450 and F471) that may impact RNA binding, are shown in purple and underlined. ***(B)*** Aromatic amino acid substitutions in Nab2 ZnF 5-7 abolish RNA binding. GST (1 μM), GST-ZnF 5-7 (62.5 nM – 500 nM), or GST-ZnF 5-7 F450A, F471A (62.5 nM – 500 nM) was incubated with ~125 nM Cy3-r$(A)_{25}$. Free Cy3-r$(A)_{25}$ was then separated from RNA:protein complexes by electrophoresis on a non-denaturing 5% polyacrylamide gel.

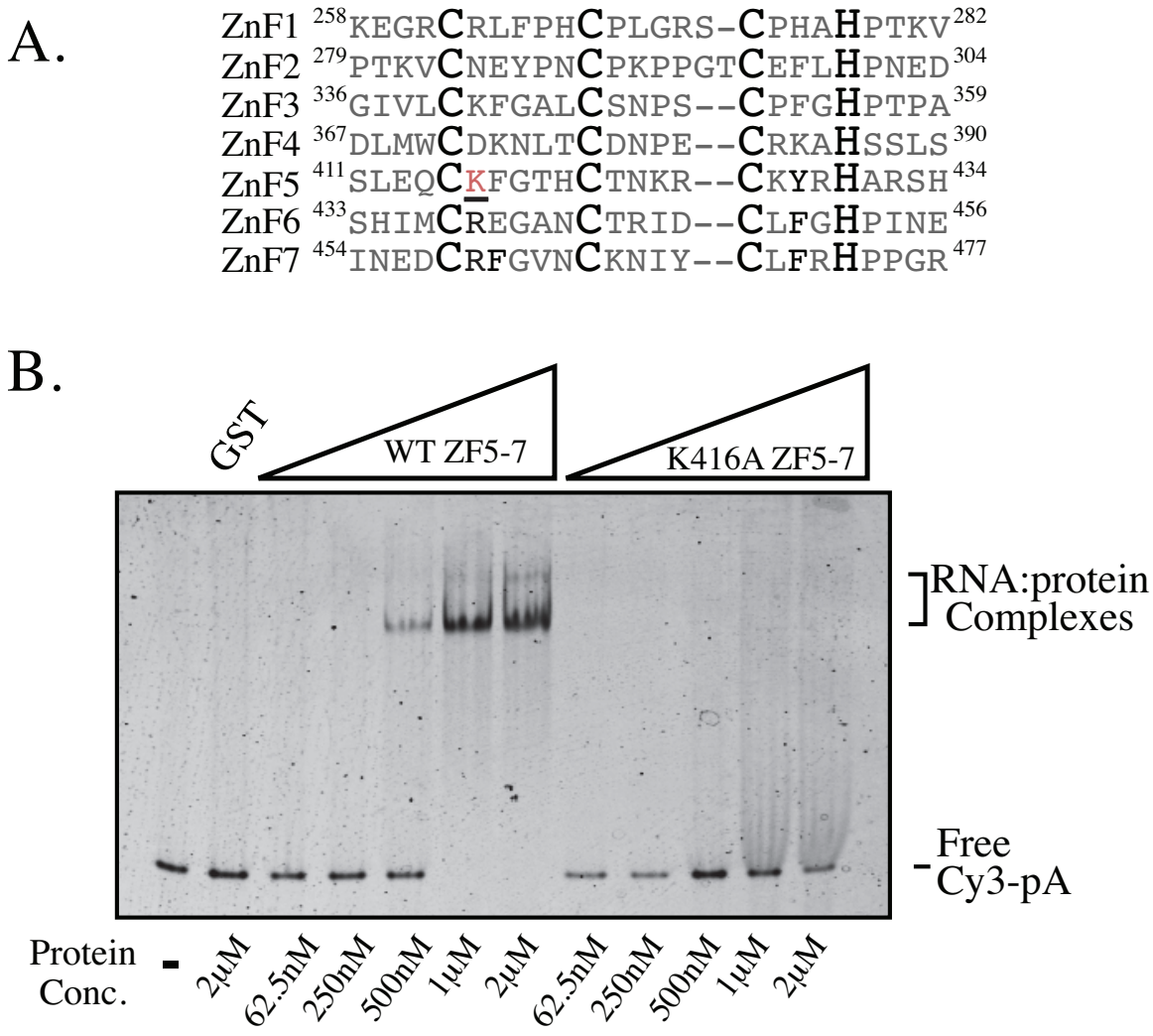

Figure 3.15: The evolutionarily conserved positively charged residue, Lys 416, in Nab2 zinc fingers 5 is critical for Nab2 binding to polyadenosine RNA. ***(A)*** Alignment of the seven Nab2 CCCH zinc fingers. Numbers corresponding to the first and last amino acid are shown surrounding each zinc finger. Cysteines and histidines are larger and shown in bold while the conserved basic residue (K416) that may impact RNA binding is shown in red and underlined. Other conserved positively charged and aromatic residues are also shown in bold. ***(B)*** The K416A amino acid substitution in Nab2 ZnF 5-7 abolishes RNA binding. GST (2 μM), GST-ZnF 5-7 (62.5 nM – 2 μM), or GST-ZnF 5-7 K416A (62.5 nM – 2 μM) was incubated with ~125 nM Cy3-r$(A)_{25}$. Free Cy3-r$(A)_{25}$ was then separated from RNA:protein complexes by electrophoresis on a non-denaturing 5% polyacrylamide gel.

A.

ZnF1 ^{258}KEGR**C**RLFPH**C**PLGRS-**C**PHA**H**PTKV^{282}
ZnF2 ^{279}PTKV**C**NEYPN**C**PKPPGT**C**EFL**H**PNED^{304}
ZnF3 ^{336}GIVL**C**KFGAL**C**SNPS--**C**PFG**H**PTPA^{359}
ZnF4 ^{367}DLMW**C**DKNLT**C**DNPE--**C**RKA**H**SSLS^{390}
ZnF5 ^{411}SLEQ**C**KFGTH**C**TNKR--**C**KYR**H**ARSH^{434}
ZnF6 ^{433}SHIM**C**REGAN**C**TRID--**C**LFG**H**PINE^{456}
ZnF7 ^{454}INED**C**RFGVN**C**KNIY--**C**LFR**H**PPGR^{477}

B.

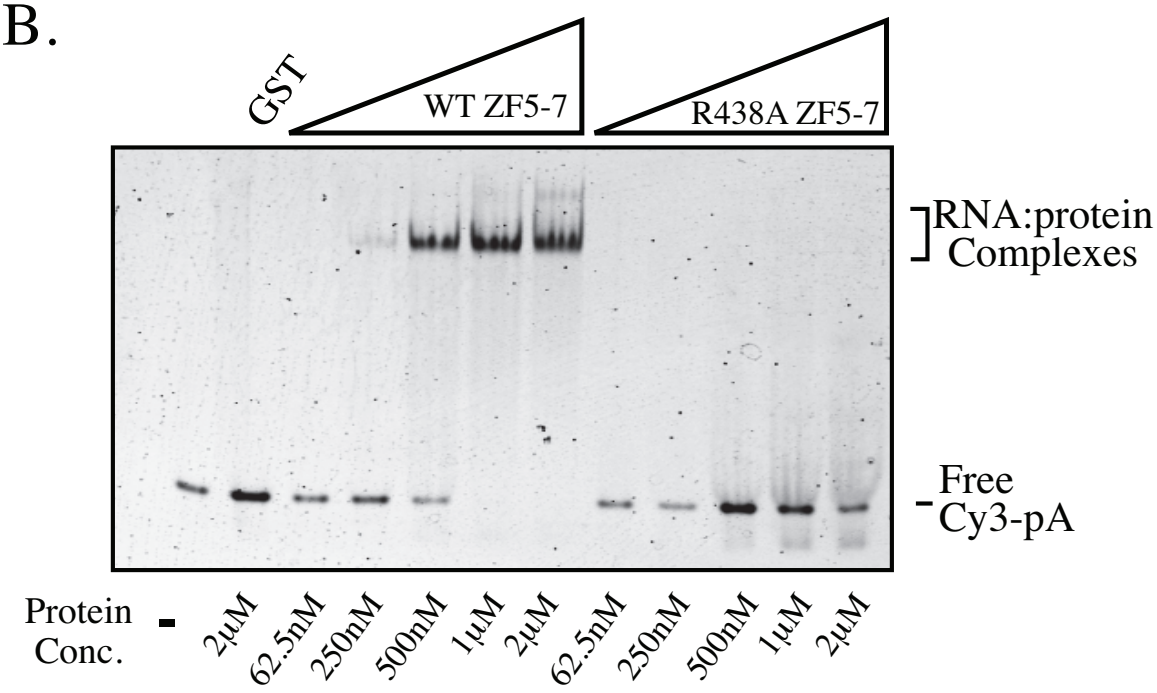

Figure 3.16: The evolutionarily conserved positively charged residue, Arg 438, in Nab2 zinc fingers 6 is critical for Nab2 binding to polyadenosine RNA. ***(A)*** Alignment of the seven Nab2 CCCH zinc fingers. Numbers corresponding to the first and last amino acid are shown surrounding each zinc finger. Cysteines and histidines are larger and shown in bold while the conserved basic residue (R438) that may impact RNA binding is shown in red and underlined. Other conserved positively charged and aromatic residues are also shown in bold. ***(B)*** The R438A amino acid substitution in Nab2 ZnF 5-7 abolishes RNA binding. GST (2 μM), GST-ZnF 5-7 (62.5 nM – 2 μM), or GST-ZnF 5-7 R438A (62.5 nM – 2 μM) was incubated with ~125 nM Cy3-r$(A)_{25}$. Free Cy3-r$(A)_{25}$ was then separated from RNA:protein complexes by electrophoresis on a non-denaturing 5% polyacrylamide gel.

A.

ZnF1 ^{258}KEGR**C**RLFPH**C**PLGRS-**C**PHA**H**PTKV^{282}
ZnF2 ^{279}PTKV**C**NEYPN**C**PKPPGT**C**EFL**H**PNED^{304}
ZnF3 ^{336}GIVL**C**KFGAL**C**SNPS--**C**PFG**H**PTPA^{359}
ZnF4 ^{367}DLMW**C**DKNLT**C**DNPE--**C**RKA**H**SSLS^{390}
ZnF5 ^{411}SLEQ**C**KFGTH**C**TNKR--**C**KYR**H**ARSH^{434}
ZnF6 ^{433}SHIM**C**REGAN**C**TRID--**C**LFG**H**PINE^{456}
ZnF7 ^{454}INED**C**RFGVN**C**KNIY--**C**LFR**H**PPGR^{477}

B.

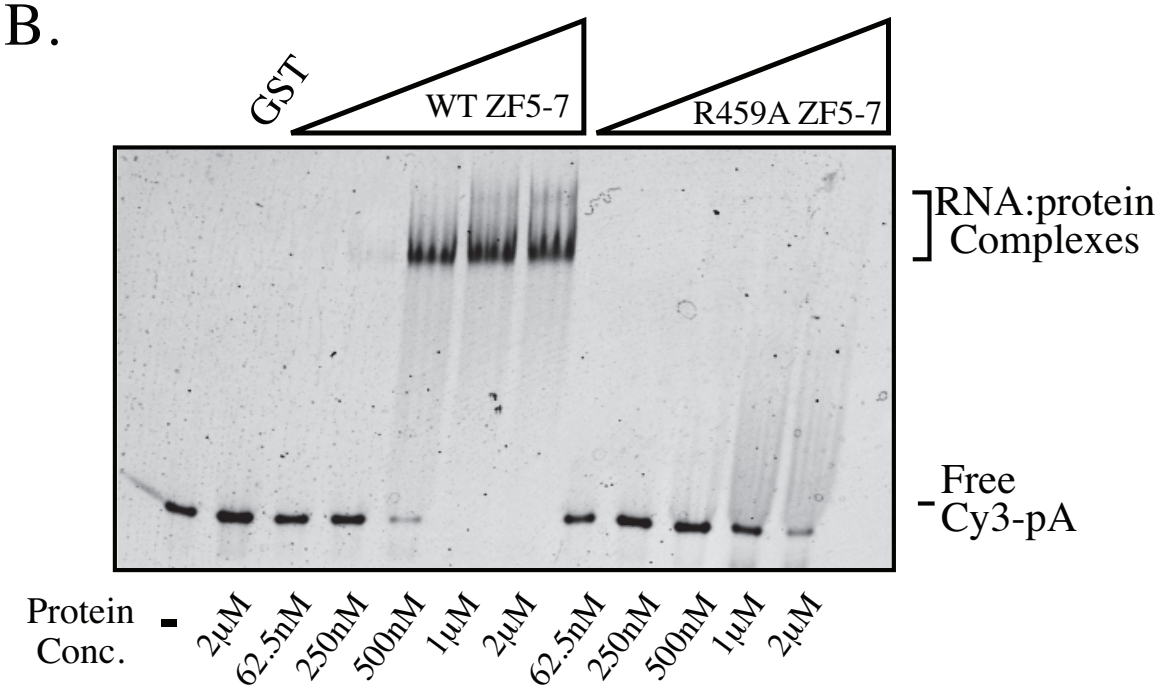

Figure 3.17: The evolutionarily conserved positively charged residue, Arg 459, in Nab2 zinc fingers 7 is critical for Nab2 binding to polyadenosine RNA. ***(A)*** Alignment of the seven Nab2 CCCH zinc fingers. Numbers corresponding to the first and last amino acid are shown surrounding each zinc finger. Cysteines and histidines are larger and shown in bold while the conserved basic residue (R459) that may impact RNA binding is shown in red and underlined. Other conserved positively charged and aromatic residues are also shown in bold. ***(B)*** The R459A amino acid substitution in Nab2 ZnF 5-7 abolishes RNA binding. GST (2 μM), GST-ZnF 5-7 (62.5 nM – 2 μM), or GST-ZnF 5-7 R459A (62.5 nM – 2 μM) was incubated with ~125 nM Cy3-r$(A)_{25}$. Free Cy3-r$(A)_{25}$ was then separated from RNA:protein complexes by electrophoresis on a non-denaturing 5% polyacrylamide gel.

Nab2 RNA binding mutants can be attributed to defects in one or both of these essential functions. To examine the direct requirement for Nab2 binding to RNA in each of these processes, we tested whether cells expressing Nab2 variants that are defective for RNA binding show changes in poly(A) tail length, poly(A) RNA export from the nucleus, or both processes.

RNA transcripts isolated from cells depleted for Nab2 or cells containing mutant alleles of *NAB2* display longer tracts of poly(A) than wild-type cells (71, 171). We therefore postulated that the growth defect seen in cells expressing nab-C415A,C437A,C458A (nab2-C_{5-7}→A) and nab2-C415R,C437R,C458R (nab2-C_{5-7}→R) could be due to deregulation of polyadenylation. To test this hypothesis, we analyzed poly(A) tail length in yeast cells expressing Nab2, nab2-1, nab2-C_{5-7}→A, nab2-C_{5-7}→R, or nab2-C437S as the only copy of *NAB2* (Figure 3.18). As expected, cells expressing wild-type Nab2 showed normal poly(A) tail length, while those expressing nab2-1 showed extended poly(A) tails at both 30°C and 18°C (171). Yeast cells expressing nab2-C_{5-7}→A, nab2-C_{5-7}→R, or nab2-C437S all showed longer poly(A) tails than cells expressing wild-type Nab2 at both temperatures tested. These results suggest that Nab2 RNA binding is critical for proper control of poly(A) tail length.

In order to examine whether changes within the CCCH zinc finger binding domain of Nab2 cause defects in poly(A) RNA export, we used fluorescence *in-situ* hybridization (FISH) to localize bulk polyadenylated RNA transcripts in cells expressing Nab2, nab2-1 (which displays a poly(A) RNA export defect) (73), and the triple zinc

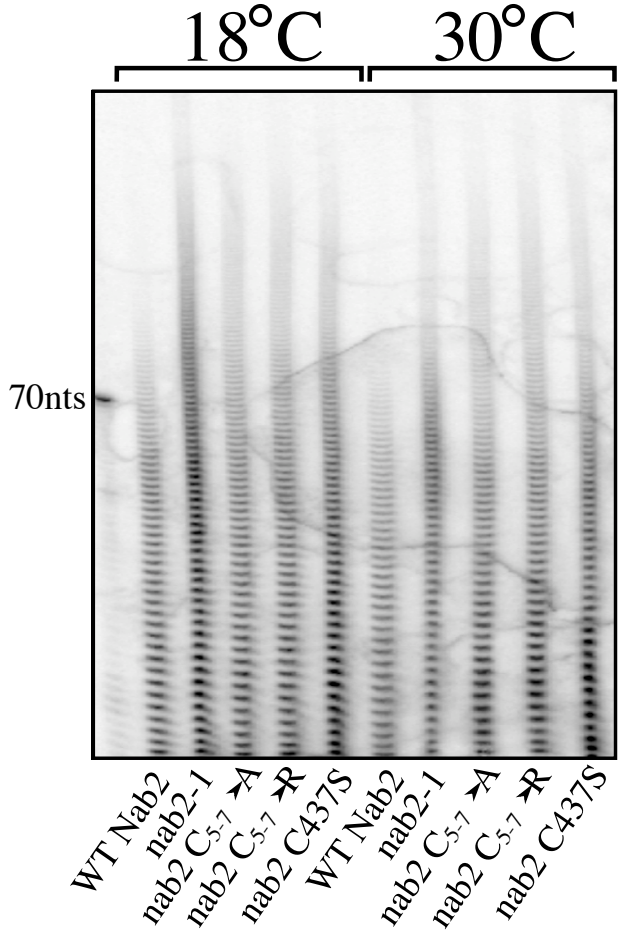

Figure 3.18: Cells that express Nab2 mutants defective for RNA binding show extended poly(A) tails. Total RNA was isolated from yeast cells grown at 30°C or 18°C expressing wild-type Nab2, nab2-1, nab2-C_{5-7}→A, nab2-C_{5-7}→R, or nab2-C437S and end-labeled with ^{32}pCp and T4 RNA ligase. RNA was then digested with RNase A and RNase T1 to digest non-poly(A) RNA tracts. Resulting stretches of poly(A) RNA were then resolved by denaturing urea-polyacrylamide gel electrophoresis and visualized by autoradiography. The position of the 70 nucleotide (nt) typical poly(A) tail length marker is indicated to the left.

finger mutants, nab2-C_{5-7}→A and nab2-C_{5-7}→R. Cells were grown at 30°C for 4 hours, cultures were divided and shifted to either 18°C or 12°C for 6 hours before being analyzed for poly(A) RNA localization. Cells shifted to 18°C and expressing wild-type Nab2 showed no accumulation of poly(A) RNA in the nucleus, while *nab2-1* cells displayed a significant accumulation of poly(A) RNA in the nucleus (Figure 3.19). Cells expressing nab2-C_{5-7}→A or nab2-C_{5-7}→R also showed no detectable nuclear accumulation of poly(A) RNA when shifted to 18°C. However, when cells expressing nab2-C_{5-7}→A are shifted to 12°C, poly(A) RNA accumulates within the nucleus (Figure 3.20). As a control, a mutant of Nab2, nab2-21, which lacks half of zinc finger 6 and all of zinc finger 7 and was previously shown to accumulate poly(A) RNA in the nucleus at lower temperatures (71), also shows nuclear accumulation of poly(A) RNA at 12°C. Cells expressing a wild-type copy of Nab2, however, showed no nuclear accumulation of poly(A) RNA at 12°C.

Discussion

In the current study, we define the high affinity RNA binding domain within the tandem CCCH zinc fingers of the polyadenosine RNA binding protein Nab2. We show through a combination of genetic and biochemical methods that Nab2 zinc fingers 5-7 are sufficient to mediate specific and high affinity binding to polyadenosine RNA. In collaboration with Murray Stewart's laboratory, the atomic resolution structure of the Nab2 zinc finger domain was solved by NMR. Using this structural data, specific solvent exposed aromatic and basic amino acids were identified that were predicted to interact

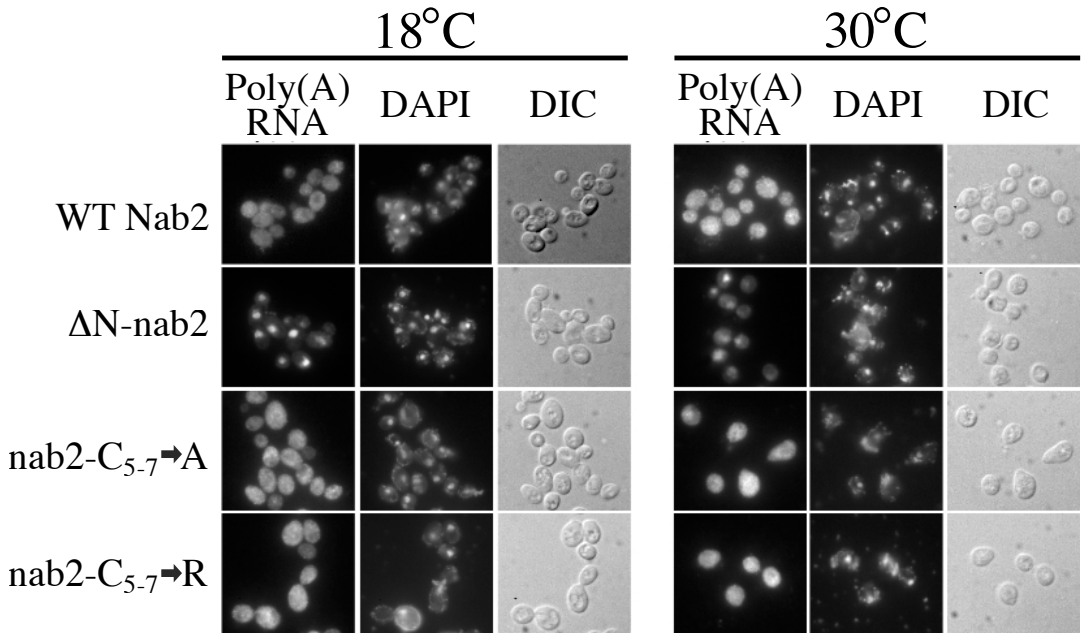

Figure 3.19: Cysteine to alanine amino acid substitutions within the Nab2 polyadenosine RNA binding domain do not cause accumulation of poly(A) RNA in the nucleus at 18°C. Fluorescence *in-situ* hybridization (FISH) using an oligo-d(T) probe to detect poly(A) RNA was performed on yeast cells expressing wild-type Nab2, ΔN-Nab2, nab2-C_{5-7}→A, or nab2-C_{5-7}→R as described in Materials and Methods. Corresponding DAPI and differential interfering contrast (DIC) images are shown.

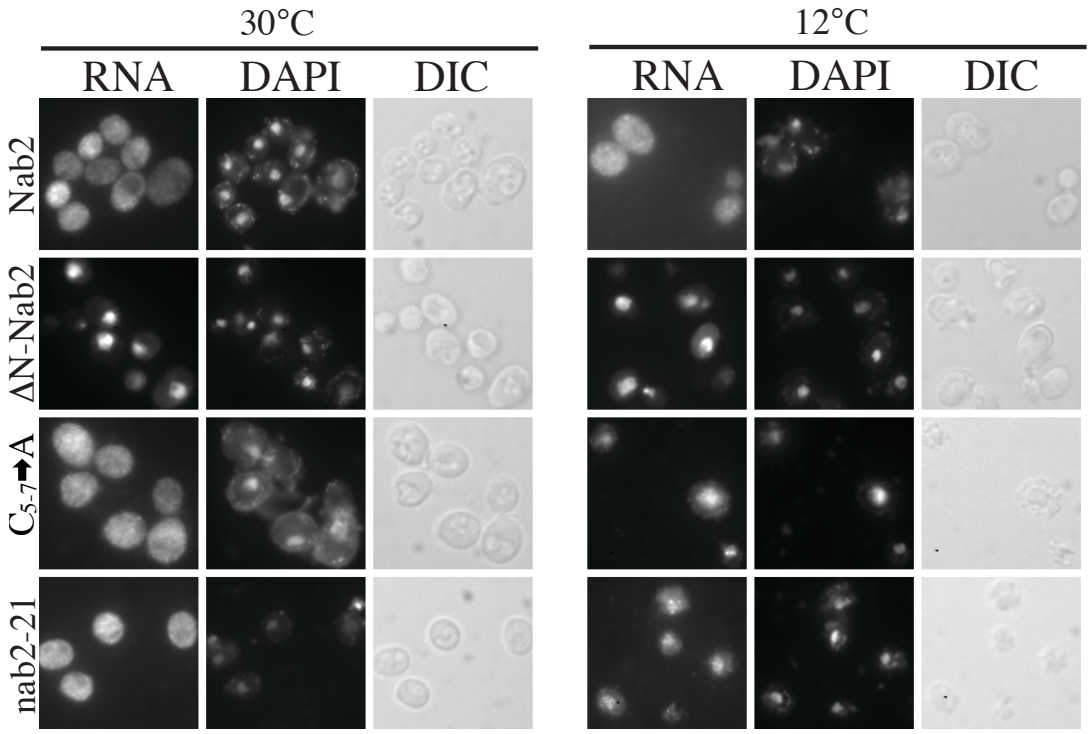

Figure 3.20: Cysteine to alanine amino acid substitutions within the Nab2 polyadenosine RNA binding domain cause accumulation of poly(A) RNA in the nucleus at 12°C. Fluorescence *in-situ* hybridization (FISH) using an oligo-d(T) probe to detect poly(A) RNA was performed on yeast cells expressing wild-type Nab2, ΔN-Nab2, nab2-C_{5-7}→A, or nab2-21 as described in Materials and Methods. Corresponding DAPI and differential interfering contrast (DIC) images are shown.

with polyadenosine RNA. Substitution of several of these residues by alanine disrupts Nab2 binding to poly(A) RNA *in vitro* and confers growth defects *in vivo*. Another Nab2 mutant, nab2-$C_{5\text{-}7}$→A, which contains cysteine to alanine substitutions at the first cysteine of zinc fingers 5-7, impairs Nab2 function *in vivo* and also causes aberrant control of poly(A) tail length and accumulation of poly(A) RNA in the nucleus at lower temperatures.

While *S. cerevisiae* Nab2 contains seven zinc fingers, putative Nab2 orthologues in higher eukaryotes contain only five tandem CCCH zinc fingers (67), which suggested to us that only a subset of Nab2 zinc fingers may be necessary for specific recognition of polyadenosine RNA. Accordingly, we show here that Nab2 CCCH zinc fingers 5-7 encode the Nab2 high affinity poly(A) RNA binding domain. Although Nab2 mutants that lack all seven zinc fingers do not bind polyadenosine RNA [Figure 2.7 and (67)], Nab2 mutants that lack only zinc fingers 5-7 (Nab2-ΔZnF 5-7) retain some affinity for polyadenosine RNA (Figure 3.6). However, Nab2-ΔZnF 5-7 binds poly(A) RNA with an apparent K_d of 548 nM, approximately 8-10 fold higher than the established binding affinity of full-length Nab2. This difference suggests that while the high affinity poly(A) RNA binding site within Nab2 is located within zinc fingers 5-7, an additional, lower affinity poly(A) RNA binding site, may be located within zinc fingers 1-4. Nab2 also encodes an arginine-glycine-glycine (RGG) repeat domain. Although the RGG domain of other proteins has been shown to bind nucleic acids (141), a Nab2 construct containing the RGG domain but none of the zinc fingers (Nab2 ΔZnF 1-7) does not bind polyadenosine RNA *in vitro* [Figure 2.7 and (67)]. Weak binding of Nab2-ΔZnF 5-7 to poly(A) RNA *in vitro* is also consistent with a previous study which demonstrated that

Nab2 mutants lacking zinc fingers 5-7 could be UV-crosslinked to poly(A) RNA *in vivo*, although at a greatly reduced efficiency compared to full-length Nab2 (73). Future experiments will endeavor to measure the binding affinities of zinc fingers 1-4 versus 5-7 *in vitro* using fluorescence anisotropy.

A close inspection of the Nab2 amino acid sequence reveals that the seven zinc fingers are arranged in three zinc finger clusters (diagrammed in Figure 3.1). The zinc fingers within each cluster are arranged in extremely close proximity to one another, consistent with the idea that they might function as independently folding domains. The spacing between the cysteine and histidine residues in zinc fingers 1-2 is $CX_5CX_{5-6}CX_3H$ while the spacing between zinc fingers 3-7 is $CX_5CX_4CX_3H$. These slight differences in spacing and, presumably, the resulting three-dimensional zinc finger structure could explain the differences seen in binding affinity to polyadenosine RNA between zinc fingers 1-4 and zinc fingers 5-7 (172). Zinc fingers 1 and 4 also lack an aromatic residue (such as Phe or Tyr) between the third cysteine and the histidine, while zinc fingers 2 and 4 lack a positively charged residue (such as Arg or Lys) after the first cysteine. We show here that aromatic and positively charged residues at these positions are critical for Nab2 binding to polyadenosine RNA. The absence of aromatic or positively charged residues at these particular positions may be another contributing factor leading to the lack of high affinity poly(A) RNA binding detected for Nab2-ΔZnF 5-7. Although our data suggests that zinc fingers 1-4 could bind poly(A) RNA with relatively low affinity, in fact the specificity of zinc fingers 1-4 alone has not been tested. Nab2 zinc fingers 1-4 could interact preferentially with another RNA sequence or, alternatively, could interact with

RNA in a non-sequence specific manner. Further experiments will be necessary to examine the potential contributions of these zinc fingers to RNA binding.

Importantly, we present evidence here that cysteine to alanine amino acid substitutions in zinc fingers 5-7 disrupt polyadenosine RNA binding *in vitro* and also cause defects in poly(A) tail length control and nuclear export of poly(A) RNA. This finding suggests that Nab2 plays a critical role in the control of both poly(A) tail length and export of poly(A) RNA from the nucleus. The question remains however, as to whether Nab2 plays a direct role in one or both of these processes. Several studies have demonstrated a reciprocal nature to the defects seen in both mRNA export proteins and 3'-end processing factors. Specifically, defects in many nuclear export factors cause not only nuclear accumulation of poly(A) RNA, but also extended poly(A) tails, while mutants in the mRNA 3'-end processing machinery show defects in cleavage and polyadenylation as well as nuclear accumulation of poly(A) RNA (52, 55).

These surprising results demonstrate the intricacy of the coupling between mRNA processing and export in eukaryotes but also begs the question as to whether extended poly(A) tails cause RNA export defects or whether defects in the transport machinery result in extended poly(A) tails. The Nab2 mutants analyzed in this study provide compelling evidence to support either model to the exclusion of the other, and thus a combination of both models is likely the most accurate. If the cause of extended poly(A) tails is the nuclear accumulation of poly(A) RNA, then mRNA transcripts from cells expressing nab2-$C_{5\text{-}7}$→A, where poly(A) RNA does not accumulate in the nucleus at the permissive temperature of 30°C (Figures 3.19 and 3.20), should not have long poly(A) tails at 30°C. However, RNA isolated from cells expressing nab2-$C_{5\text{-}7}$→A has longer

poly(A) tails compared to wild-type cells even at the permissive temperature of 30°C (Figure 3.18). In other words, Nab2 mutants still show extended poly(A) tails even in the absence of poly(A) RNA accumulation in the nucleus. Therefore, mRNA transcripts are not hyper-polyadenylated simply because they are "stuck" in the nucleus. Conversely, extended poly(A) tails, alone, are not sufficient to cause the nuclear accumulation of poly(A) RNA under normal growth conditions. We propose that only under stressed growth conditions, such as growth of nab2-$C_{5\text{-}7}$→A cells at the suboptimal temperature of 18°C, are longer poly(A) tails detrimental to cell viability. Accordingly, we also suggest that the poly(A) RNA export defect of nab2-$C_{5\text{-}7}$→A cells is a secondary effect of a primary defect in RNA processing.

While previous results have suggested that the essential function of Nab2 is in mRNA export from the nucleus, a primary role for Nab2 in mRNA 3'-end processing is consistent with evidence obtained in other model organisms. A recent genome-wide study investigating the components required for proper export of poly(A) RNA in *Drosophila melanogaster* S2 cells demonstrated that siRNA knock-down of the putative *D. melanogaster* Nab2 orthologue, CG5720, did not cause nuclear accumulation of poly(A) RNA (173). In addition, recent work characterizing the putative human Nab2 orthologue, ZC3H14, demonstrates that ZC3H14 co-localizes with the nuclear speckle marker, SC35 (146). Nuclear speckles are thought to house the workers for the mRNA assembly line as they contain many splicing factors as well as transcription-related proteins that play critical roles in RNA metabolism [reviewed by (174)]. Hence, localization of ZC3H14 in nuclear speckles strongly suggests a role in mRNA processing. The evidence presented here strongly supports the idea that carefully defining a specific

protein function, such as RNA binding, and making mutants to disrupt that function, allows for precise definition of the cellular role of that protein.

Experimental Procedures

Chemicals, Plasmids, and Saccharomyces cerevisiae Manipulations – Chemicals were obtained from Fisher Scientific (Pittsburgh, PA), Sigma-Aldrich (St. Louis, MO), or US Biological (Swampscott, MA) unless otherwise noted. DNA manipulations were performed according to standard methods (150) and all media were prepared by standard procedures (151). All *S. cerevisiae* strains and plasmids used in this study are described in Table 1. Plasmids encoding mutant nab2 proteins were created by site-directed mutagenesis of a wildtype Nab2 plasmid (pAC717) using the QuikChange Site-directed mutagenesis kit (Stratagene). All plasmids were fully sequenced to ensure that no mutations other than those targeted were introduced during site-directed mutagenesis. The *NAB2* open reading frame was deleted from *S. cerevisiae* strains encoding the 3'-end processing mutants *pap1-1* or *rna15-2* using PCR-based gene disruption essentially as described (175)

In vivo Functional Analysis – The *in vivo* function of each nab2 mutant was tested using a plasmid shuffle assay (153). *S. cerevisiae* cells deleted for *NAB2* (ACY427) and containing a wild-type *NAB2 URA3* plasmid (pAC636) were transformed with *LEU2* plasmids encoding various nab2 mutants. Transformants were grown to saturation and cells were then serially diluted and spotted onto control ura^- leu^- glucose plates or plates containing 5-fluoroorotic acid (5-FOA). Plates were then incubated at 18°C, 25°C, 30°C, or 37°C for 3-5 days. The toxic uracil analog, 5-FOA, kills cells that contain a functional

uracil biosynthetic pathway (153). Hence, only those cells that lose the *URA3* plasmid encoding wild-type Nab2 and retain a functional copy of Nab2 on the *LEU2* plasmid will grow on media containing this drug.

For analysis of *dbp5-2* suppression, a plasmid shuffle assay was performed using a Δ*NAB2* strain that also harbored the *dbp5-2* allele (ACY1669) as described above with the following modifications. Instead of directly spotting Δ*NAB2 dbp5-2* transformants onto selective media containing 5-FOA, cells were first struck to selective media containing 5-FOA, then struck to rich media (YPD), and finally grown to saturation, serially diluted, and spotted onto YPD. Plates were then incubated at 16°C, 18°C, 25°C, 30°C, or 32°C for 3-5 days.

Genomic Integration of Cysteine to Alanine Mutations – Genomic integration of the nab2-$C_{5\text{-}7}$→A allele into wild-type yeast cells were performed essentially as described (169, 170) (For a pictorial representation of this description see Figure 3.4). Briefly, the C-terminal zinc finger domain of nab2-$C_{5\text{-}7}$→A along with ~100 base pairs of the Nab2 3'-untranslated region (3'-UTR) was initially amplified by PCR from pAC 2203 (See Table 3.1) yielding PCR product A. The natamycin resistance cassette (NAT^R) was also amplified (See Table 3.1) yielding PCR Product B. As shown in Figure 3.4, the 3'-end of PCR product A was homologous to the 5'-end of PCR product B, which allowed us to perform overlap PCR. In order to perform overlap PCR, products from the previous two reactions were diluted 1:1000 and 1 μl of each diluted PCR product was added as a template to a third PCR reaction. The first two rounds of the final PCR reaction were performed without oligonucleotide primers and under higher annealing temperatures (65°C for rounds 1-2, vs. 56°C for rounds 3-30) that favored specific annealing of the

complementary ends of PCR products A and B. After the initial two rounds of PCR reaction #3, oligonucleotide primers were added to the reaction and the annealing temperature was lowered to 56°C to specifically amplify the entire nab2-C_{5-7}→A-NAT^R combinatorial product. *S. cerevisiae* cells were transformed with the resulting PCR products from the overlap PCR reaction, plated onto rich media (YEPD), and grown at 30°C for 1 day. Cells were then replica plated to media containing natamycin (50 μg/ml concentration) and grown at 30°C. NAT^R colonies usually appeared within 3-4 days of replica plating. Genomic DNA was isolated from NAT^R colonies and the *NAB2* genomic locus was amplified by PCR to check for the insertion of the NAT^R cassette. Insertion of the NAT^R cassette and the nab2-C_{5-7}→A mutations were confirmed by sequencing PCR products.

Protein Expression and Purification – GST-tagged proteins were expressed in *E. coli* and purified essentially as described (67). Briefly, plasmids encoding GST, GST-Nab2, GST-Nab2 Zinc Finger (ZnF) 1-4, or GST-Nab2-ZnF 5-7 were transformed into BL21(DE3) *E. coli* cells. Single colonies were inoculated into 2 ml cultures and grown overnight to saturation. These starter cultures were used to inoculate 50 ml of media. Cultures were grown at 37°C until they reached an OD 600nm of 0.4-0.6. Protein expression was induced with 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hrs at 30°C. Following protein induction, cells were pelleted and frozen at –80°C. Frozen cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 100 mM NaCl, 5% glycerol, 4 μM 2-mercaptoethanol, 2 μM $ZnCl_2$, 1 mM phenylmethylsulfonylfluoride) and lysed by sonication. GST-tagged proteins were purified on GST-sepharose 4B (GE Healthcare Life Sciences) according to the

manufacturer's directions. A small amount of purified protein was resolved by SDS-PAGE to determine the purity of the sample.

To express recombinant full-length Nab2 in *E. coli*, the expression plasmid, pAC2133, was transformed into BL21(DE3)pLYS cells (Novagen). Single colonies were inoculated into 2 ml of media and grown to saturation overnight. This culture was then used to inoculate 1 L of LB media. Cells were grown at 37°C until they reached an OD600 nm of 0.4 – 0.6. Cultures were then shifted to 30°C and induced with 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hrs. Cells were then centrifuged at 3800rpm for 20 minutes at 4°C and pelleted cells were frozen at -80°C. For purification of untagged Nab2 proteins from frozen cell pellets, cell pellets were thawed on ice and resuspended in an equal volume of lysis buffer (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 4 mM β-mercaptoethanol, 1mM phenylmethylsulfonylfluoride) and lysed using a French Press. The lysates were cleared by centrifugation, loaded onto a HiTRAP Q column (GE Healthcare) pre-equilibrated with Buffer A (20mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol), and bound protein was eluted using a linear gradient of Buffer B (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 2 mM β-mercaptoethanol, and 1 M NaCl). Fractions containing untagged recombinant Nab2 were dialyzed overnight into Buffer C (20 mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAC, 2 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol). The dialyzed protein was then concentrated by adding saturated ammonium sulfate to 25% saturation. Precipitated proteins were pelleted and resuspended in Buffer C. The concentrated protein was then loaded onto a Superdex S-200 gel filtration column pre-equilibrated with Buffer C and fractions containing untagged recombinant Nab2 were pooled and

concentrated using Centricon spin concentrators (Millipore). Since we found that freezing the purified protein led to loss of nucleic acid binding, freshly purified aliquots of protein were used for each binding experiment.

RNA Oligonucleotides and RNA Electrophoretic Mobility Shift Assays – All RNA oligonucleotides were obtained from Dharmacon (Lafayette, CO) and were deprotected in the supplied deprotection buffer according to the manufacturer's instructions. Deprotected oligos were evaporated to dryness and stored at –20°C. For RNA gel shift assays, a Cy3-labeled 25-nt poly(A) RNA oligonucleotide was incubated in binding buffer (20 mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAc, 2 μM $ZnCl_2$, 2% glycerol) with GST or GST-tagged proteins as indicated for 20 min at room temperature. Binding reactions were loaded onto a native 5% polyacrylamide gel and electrophoresed in 0.3x TBE for 30-60 min to resolve free Cy3-labeled oligonucleotide from RNA-protein complexes. The position of the Cy3-labeled oligonucleotide within the native gel was determined using a Typhoon phosphorimager (Amersham) equipped with a laser capable of exciting Cy3. RNA gel shift competition experiments were performed as described previously (67).

RNA labeling protocol – RNA 3'-end labeling was performed essentially as described (176). Briefly, RNA oligonucleotides were incubated with 100 mM NaOAc pH 5.1, 100 μM $NaIO_4$ at room temperature for 90 minutes and then precipitated with 2.5 μL of 5 M NaCl, 1 μl of 20μg/μl glycogen (Invitrogen), and 100 μl of ice-cold 100% ethanol for 20 minutes at -20°C. Precipitated RNA was pelleted by centrifugation and resuspended in 50 μl of labeling solution (1.5 mM fluorescein 5-thiosemicabazide, 100 mM NaOAc pH 5.1. Samples were covered and RNA was labeled overnight at 4°C. To remove

unreacted fluorescein, labeling reactions were added to a G-25 sephadex column. Labeled RNA was eluted by centrifugation. Labeling efficiency was determined using the following equation:

$$\text{Efficiency} = \frac{OD_{490}/\varepsilon_{fluorescein}}{OD_{260}/\varepsilon_{oligonucleotide}} \qquad [1]$$

where OD_{490} and OD_{260} is the absorbance at 490 nm and 260 nm, respectively and ε is the extinction coefficient for either fluorescein or the oligonucleotide as noted. Typical labeling efficiencies were 60-85%. The purity of labeled RNA oligonucleotides was analyzed by agarose gel electrophoresis.

Limited Trypsin Proteolysis and Poly(A) Sepharose Binding Assays – For partial proteolysis, 3.2 mg of recombinant purified GST-Nab2 was then incubated with 20 μg of trypsin for 30 min at room temperature. The entire reaction was then incubated with poly(A) sepharose 4B (Sigma), the sepharose was washed and bound complexes were eluted. Bound fractions were resolved by SDS-PAGE and either stained with Coomassie Blue or transferred to PVDF membranes for N-terminal sequencing. Edmann degradation N-terminal sequencing was performed by the Emory Microchemical Facility to identify the N-terminal eight amino acids of the eluted Nab2 RNA binding fragments. Please refer to the supplemental Materials and Methods for a more complete description of this procedure.

Fluorescence Anisotropy Assay and Dissociation Constant Determination – 2 nM fluorescein-labeled RNA oligonucleotides were incubated with increasing amounts (protein concentrations range from 50 pM to 5 μM) of recombinant purified proteins in 384-well black plates as described (176). Polarization data were fit to Equation 2:

$$\theta = base + \frac{(\max - base)}{1 + \left(\frac{K_d}{P_t}\right)^N} \qquad [2]$$

where θ is the experimentally observed polarization value, base is the minimum polarization value when no protein is added, max is the maximum polarization value at saturating concentrations of protein, K_d is the apparent dissociation constant, and P_t is the protein concentration. Reported K_d values are the average of at least three experiments.

Poly(A) Tail Length Determination – NAB2, nab2-1, nab2-C437S, nab2-C_{5-7}→A, or *nab2-C_{5-7}→R* cells were inoculated into YPD media and grown to saturation at 30°C. Cells were then diluted into 50 ml of YPD and grown at either 30°C or 18°C until they reached OD 600nm of 0.6 – 0.8. A total of 20 OD units of cells were harvested from each culture and poly(A) tail length was determined as described previously (171, 177). Briefly, total RNA was end-labeled with ^{32}P-pCp and T4 RNA ligase. In order to digest non-poly(A) RNA, the labeled RNA was simultaneously digested with RNases A/T1 and then ethanol precipitated. Resuspended RNA was then resolved by denaturing urea-acrylamide gel electrophoresis and imaged using a phosphorimager.

Fluorescence in-situ Hybridization (FISH) – *NAB2, nab2-1*, *nab2-C_{5-7}→A*, or *nab2-C_{5-7}→R* cells were initially grown in 2 ml YPD cultures to saturation at 30°C. These starter cultures were then used to inoculate 5 ml YPD cultures that were grown overnight (approx. 16 hours) at either 30°C or 18°C. Cells were then fixed by the addition of 700 μl of 37% formaldehyde and incubated at 30°C or 18°C for 90 min and FISH using an oligo d(T) probe to detect poly(A) RNA was performed as described (178). Cells were

also stained with 4', 6-diamidino-2-phenylindole (DAPI) to visualize DNA within the nucleus. Please refer to the supplemental Materials and Methods for a more complete description of this procedure.

Cloning and Purification of proteins used for NMR analyses – Yeast Nab2 residues 409 – 483 were cloned into pGEX6P-1 (GE-Healthcare) using BamHI, XhoI, resulting in the additional N-terminal sequence GPLGS left on the final protein construct. The plasmid was transformed into E. coli strain BL21 DE3 and cells were grown in M9 minimal medium at 37C to an optical density of 0.6. Protein expression was induced by addition of 200 μM IPTG and 250 μM $ZnCl_2$. Expression was carried out at 20 C over night. The protein was then purified using standard GST-purification methods in 50 mM TRIS pH 8.0, 200 mM NaCl, 10 μM $ZnCl_2$ and 5 mM β-mercaptoethanol. After a first wash bound nucleotides were digested with a treatment of DNAse / RNase and a subsequent wash with the above buffer containing 1M NaCl. The GST-fusion protein was eluted using reduced glutathione and cleaved overnight at 4C using PreScission-protease. After addition of glutamic acid and arginine to 50mM the cleaved protein was concentrated and subjected to gel-filtration chromatography on a S75 column (GE-Healthcare) in 50 mM TRIS-HCl pH 6.75, 50 mM NaCl, 50 mM Glu/Arg, 10 μM $ZnCl_2$ and 5mM β-ME. Fractions containing the desired Nab2 fragment were pooled and concentrated up to 1.2 mM protein concentration.

NMR spectroscopy – All data was acquired on Bruker DMX600 and DRX500 spectrometers, each equipped with a triple resonance ($^{1}H/^{15}N/^{13}C$) cryoprobe. ^{1}H, ^{15}N and ^{13}C chemical shifts were calibrated using sodium 3,3,3-trimethylsilylpropionate (TSP) as an external ^{1}H reference. Unless otherwise stated, all NMR experiments for the free

protein were performed at 17°C using ^{15}N- or ^{15}N, ^{13}C-labelled protein samples in the buffer used for gel-filtration supplemented with 5% D_2O.

Resonance assignments were made using a standard suite of triple resonance NMR experiments. For experiments used to derive structural constraints the samples comprised 1.2 mM ^{15}N, ^{13}C-labelled solutions of Nab2 409-483. The following spectra were acquired: 2D: [^{15}N-^{1}H] HSQC, long-range-optimized [^{15}N-^{1}H] HMQC to correlate histidine ring ^{1}H and ^{15}N signals,(179) [^{13}C-^{1}H] HSQC covering the full ^{13}C spectral width, constant-time [^{13}C-^{1}H] HSQC covering only the aliphatic ^{13}C region, constant-time [^{13}C-^{1}H] HSQC covering only the aromatic ^{13}C region; 3D data sets: CBCANH, CBCACONH, HBHACONH, [^{1}H-^{13}C-^{1}H] HCCH-TOCSY, [^{13}C-^{13}C-^{1}H] HCCH-TOCSY, ^{15}N NOESY-HSQC (t_m = 120 ms and t_m = 50 ms), ^{13}C NOESY-HSQC (t_m = 150 ms), separate datasets acquired for ^{13}C aliphatic and aromatic spectral regions. Residual dipolar couplings were measured using a 0.6 mM ^{15}N, ^{13}C-labelled solution of Nab2 409-483, to which Tobacco Mosaic Virus was added to a final concentration of 25 mg/ml; splittings were measured in F_1 cross-sections of [^{15}N-^{1}H] HSQC IPAP and [^{13}C-^{15}N-^{1}H] HNCO IPAP spectra.

Structure Calculations – Initial structures for the free protein fragments were calculated using the semi-automatic program CYANA, for which the input comprises the protein sequence (residues 409-483), the full resonance assignment and the following 3D NOESY datasets: ^{15}N NOESY-HSQC (t_m = 120 ms), ^{13}C aliphatic region NOESY-HSQC (t_m = 150 ms) and ^{13}C aromatic region NOESY-HSQC (t_m = 150 ms). During the CYANA calculations no metal was represented explicitly, but the effect of metal binding

was approximated by including inter-ligand distance constraints as follows: Sγ to Sγ, 3.7-4.0Å; Sγ to histidyl-N, 3.4-3.8Å; histidyl-N to histidyl-N, 3.1-3.5Å.

In order to be able to employ explicit zinc bonding and geometry terms in the force-field for the calculations (including bond-angle and, for the histidines, in-plane constraints), as well as constraints based on RDC measurements, we next calculated structures using XPLOR-NIH. As input, these calculations used the set of NOE restraints generated by the final (seventh) cycle of CYANA, re-imported into CCPNMR Analysis and groomed by manual inspection. Since the XPLOR-NIH calculations employed r^{-6} summation for all groups of equivalent protons and non-stereospecifically assigned prochiral groups, and since no stereoassignments were made (and the assignment-swapping protocol within XPLOR-NIH for deriving stereoassignments indirectly during the structure calculation itself was not applied), the constraints for all such groups were converted to group constraints (i.e. such groups were specified using wildcards such as HB*). All lower bounds were set to zero. The pattern of zinc connectivities to the histidine residues was established using long-range ^{15}N-HMQC experiments as described by Legge *et al*. which showed unambiguously that the $N^{\varepsilon 2}$ atom binds the zinc in all three cases. Structures were calculated from polypeptide chains with randomized Φ and Ψ torsion angles using a two-stage simulated annealing protocol within the program XPLOR-NIH, essentially as described elsewhere. A final stage of refinement against measured values of amide group ^{15}N-^{1}H residual dipolar couplings was employed using the ISAC protocol of Sass *et al*.

CCPNMR Analysis was used for resonance assignment and inspection of the CYANA NOE-assignments. RDCs were derived in Sparky. The program

CLUSTERPOSE was used to calculate the mean rmsd of ensembles to their mean structure. Structures were visualized using the program PyMOL (http://www.pymol.org)

Chapter 4: Nab2 genetically interacts with RNA processing components

A portion of this chapter is adapted from the following manuscript:

Kelly, S.M., Leung, S.W., Bramley, A.M., Tran, E.J., Chekanova, J.A., Wente, S.R., and Corbett, A.H., *Poly(A) RNA binding by Nab2 is required for correct mRNA 3'-end formation*. (In Revision)

Introduction

The lifecycles of distinct RNA transcripts are extremely complex. For instance, messenger RNA (mRNA) transcripts are transcribed by RNA polymerase II (Pol II) in the nucleus and are then extensively processed by enzymes that add the 5'-7-methylguanosine cap, splice out introns, and cleave and polyadenylate the 3'-end of the transcript. Once processed, mRNA transcripts are then exported from the nucleus and are translated into protein in the cytoplasm. The error-free completion of each of these steps is critical in producing mature mRNA transcripts that will eventually be competent for translation by ribosomes in the cytoplasm. Beyond mRNA transcripts, other small RNA transcripts are also produced by Pol II (180), including short nucleolar RNAs (snoRNAs) and cryptic unstable transcripts (CUTs) (181, 182). Although the function of snoRNA transcripts in ribosomal RNA (rRNA) biogenesis is well documented (180, 183), the cellular function of CUTs is less well understood. In the budding yeast, *Saccharomyces cerevisiae*, CUTs are transcribed from intergenic regions of the genome and are hypothesized to function in the control of gene expression (184, 185).

Although RNA Polymerase II initially synthesizes both mRNAs and short RNA transcripts, such as snoRNAs and CUTs, the processing and final destination of these two classes of RNA transcripts varies greatly. For example, in *S. cerevisiae*, mRNA 3'-end cleavage and polyadenylation occurs once components of the cleavage machinery, including Rna14 and Rna15, recognize specific sequences within the pre-mRNA 3'-untranslated region (3'-UTR). Rna14 and Rna15 are both conserved RNA binding proteins that function to correctly position the cleavage machinery (15). Following cleavage, poly(A) polymerase 1 (Pap1) processively synthesizes the poly(A) tail, which

contains 70-90 adenosines on most *S. cerevisiae* transcripts (15). snoRNA transcripts, however, are not normally polyadenylated. These short 60 – 300 nucleotide transcripts (183) are initially synthesized by RNA polymerase II, as mentioned previously, and are retained in the nucleolus to function in the processing of precursor rRNAs. Interestingly, several snoRNAs are transcribed as polycistronic messages and must be riboendonucleolytically cleaved following transcription. Following cleavage, the 5'- and 3'-ends of precursor snoRNA transcripts are trimmed by components of the nuclear exosome (100, 186, 187), a multi-protein complex containing 3'→5' riboexonuclease activity (reviewed by (187)). Although functionally mature snoRNA transcripts are not polyadenylated, recent evidence (100) suggests that prior to trimming by the exosome, short oligo(A) tails are added to snoRNA transcripts to facilitate recruitment of the nuclear exosome (See below).

The termination of RNA polymerase II-mediated transcription also varies between mRNA transcripts and short RNAs. Although the exact mechanism of mRNA or snoRNA transcription termination remains contentious, several key proteins play roles in these processes. In particular, snoRNA (and CUT) transcription termination is facilitated by two RNA binding proteins, Nab3 and Nrd1, as well as an RNA helicase, Sen1 (188-191). Nrd1 is recruited to nascent transcripts via its direct interaction with the C-terminal domain of RNA polymerase II (192, 193). Nrd1 and Nab3 bind the sequences GUAA/G and UCUU, respectively, both of which are commonly found in the 3'-end of many snoRNA transcripts (194, 195). Sen1 is then recruited and may modulate transcription termination (191, 196-198). Interestingly, the termination of mRNA transcription is facilitated by a completely separate set of proteins, including the 3'-end

processing machinery and the 5'→3' RNA exonuclease, Rat1. According to the best current model for mRNA transcription termination, once 3'-end processing has been completed, Rat1 binds the transcript still associated with RNA Pol II and essentially chases down the polymerase, eventually causing termination of transcription [reviewed by (15, 199-201)].

During the many processing steps from transcription by RNA polymerase II to their final destination in the cytoplasm or nucleolus, mRNAs and snoRNAs, respectively, are also subjected to various quality control checkpoints that monitor the correctness of processing events (18, 19). Generally, the RNA quality control machinery recognizes error-containing transcripts and targets them for degradation via the exosome (18, 19). The majority of exosome components are localized to both the nucleus and cytoplasm (202). However, one component, Rrp6, localizes primarily to the nucleus at steady state and is therefore considered a principle component of only the nuclear exosome (203). The conserved Rrp6 protein contains 3'→5' riboexonuclease activity (203) and plays a role in a wide variety of exosome functions, including the degradation of error-containing mRNA, the trimming of precursor rRNAs and snoRNAs, and the degradation of intergenic transcripts such as CUTs (98, 99, 204-206).

Specific types of aberrant transcripts are recognized by several quality control pathways to monitor the correctness of mRNA transcription and processing (18, 19). In particular, mRNA transcripts containing premature stop-codons are detected by the nonsense mediate decay (NMD) pathway (207, 208), while those transcripts lacking a stop-codon are detected by the nonstop decay pathway (208). Neither of these pathways contains intrinsic RNA exonuclease activity and instead transcripts containing errors are

"marked" as such and targeted for degradation via the cytoplasmic 5'→3' riboexonuclease, Xrn1 (209, 210).

Just as mRNA transcripts are "marked" for degradation by the NMD or nonstop decay quality control machinery, snoRNA transcripts that contain errors are also "marked" as error containing and are targeted for degradation by the nuclear exosome (100, 186). Error containing snoRNAs are targeted for nuclear exosome-mediated degradation by the addition of a short oligo(A) tail on the 3'-end of the transcript (98, 100, 186, 211). The oligo(A) tails found on short RNA transcripts destined for degradation are added by one of the distributive poly(A) polymerases, Trf4 or Trf5 (97, 212-215). Trf4 and Trf5 are members of the recently identified TRAMP complex. The TRAMP complex consists of one of the two poly(A) polymerases, Trf4 or Trf5, one of two RNA binding proteins, Air1 or Air2, and an RNA helicase, Mtr4 (186, 216). Once bound to an error-containing transcript, members of the TRAMP complex add oligo(A) tails to the 3'-end of the transcript and recruit the nuclear exosome (186). These short oligo(A) tails may help to facilitate RNA degradation by essentially provided an unstructured poly(A) "landing pad" for the nuclear exosome.

The question arises, however, as to the mechanism by which the cell differentiates between the poly(A) tail of mRNA transcripts, which is a major determinant of transcript stability, and the shorter oligo(A) tails found on short transcripts, such as snoRNAs or CUTs, destined for degradation by the nuclear exosome. One possibility is that the cell can distinguish short oligo(A) tails from long poly(A) tails based on length alone or perhaps due to differences in the processivity of the polymerases that add the poly(A) tails. The poly(A) tails of mRNA transcripts are added in a very processive manner by

the canonical poly(A) polymerase, Pap1 (217, 218). However, the oligo(A) tails of short RNA transcripts destined for degradation are added by the more distributive polymerases, Trf4 and Trf5 (97, 212-215). Additionally, the polymerase itself could play an active role in dictating whether the transcript will be targeted for degradation. Trf4 and Trf5 are part of the TRAMP complex, along with the putative helicase Mtr4 (97). Interestingly, Mtr4 interacts with the exosome component Rrp6 (219, 220), suggesting that perhaps it is not the oligo(A) tails, per se, that lead to degradation, but instead the physical presence of the TRAMP complex that recruits the exosome. In further support of this model, recent evidence suggests that Trf4 promotes degradation of excised introns by the exosome-independent of polyadenylation (212).

The cell may also distinguish between long and short stretches of adenosines based on proteins bound to the poly(A) tail and the RNA sequences immediately proximal to the poly(A)/oligo(A) tail (e.g. the 3'-UTR of mRNA transcripts). The question remains, however, as to how a poly(A) binding protein would determine poly(A) tail length or, more specifically, the type of RNA transcript to which the tail was attached. One possibility is that poly(A) binding proteins are recruited to poly(A) tails by other proteins already bound to the transcript, including enzymes responsible for 3'-end processing, quality control, or even transcriptional termination. Furthermore, the "marker" Pab could also be recruited by the C-terminal domain of RNA polymerase II. The recruited poly(A) binding (Pab) proteins could then either recruit the exosome, if it is bound to the oligo(A) tails of short RNA transcripts, or protect against exosome-mediated RNA degradation, if the marker Pab protein interacts with the poly(A) tail of mRNA transcripts.

In *S. cerevisiae*, one likely "marker" Pab candidate is the essential Nuclear Poly(A) Binding protein 2 (Nab2). Nab2, which binds specifically to poly(A) RNA via a conserved tandem CCCH zinc finger motif, is required for poly(A) RNA export from the nucleus (67, 70, 71, 73). Cells that express Nab2 mutants, such as ΔN-Nab2, a nab2 mutant lacking residues 3-97, or nab2-$C_{5\text{-}7}$→A, which contains cysteine to alanine substitutions in the first cysteines of zinc fingers 5-7, also exhibit extended poly(A) tails on bulk RNA (71, 72), suggesting that Nab2 is intricately involved with the 3'-end processing machinery. Interestingly, the other primary *S. cerevisiae* Pab, Pab1, localizes to the cytoplasm, while Nab2 localizes to the nucleus at steady state (73), suggesting that Nab2, and not Pab1, may be a critical marker of polyadenylated transcripts within the nucleus. A previous study investigating the *in vivo* consensus RNA binding sequence of Nab2, revealed that Nab2 binds the sequence $(A)_{11}G$ (49). Since this study only investigated mRNA transcripts and not short RNAs, it is unclear whether Nab2 binds solely mRNA transcripts or is also associated with snoRNAs and CUTs destined for degradation by the nuclear exosome. Given this information, we hypothesized that Nab2 may interact with components of the mRNA and snoRNA biogenesis pathways.

In the current study, we present data demonstrating that *NAB2* genetically interacts with RNA processing components. Specifically, a Nab2 mutant, nab2-$C_{5\text{-}7}$→A, which shows decreased binding to polyadenosine RNA *in vitro* (Figure 3.4) and also leads to extended poly(A) tails (Figure 3.18) and nuclear accumulation of poly(A) RNA (Figure 3.20) *in vivo*, is synthetically lethal with mRNA cleavage and polyadenylation mutants. We also demonstrate that deletion of the normally essential *NAB2* gene can be suppressed by deletion of *RRP6*. Furthermore, Rrp6 active-site mutants that destroy the

catalytic activity of Rrp6 also suppress deletion of *NAB2*. Deletion of both *NAB2* and *RRP6* causes increased accumulation of poly(A) RNA in nuclear foci, suggesting that these foci may be processing "factories" where normal RNA transcripts are processed or degraded. Overall, these results demonstrate the intricate balance between mRNA quality control and mRNA processing.

Results

In order to test our hypothesis that Nab2 is involved in mRNA processing, we first analyzed genetic interactions between mutant alleles of *nab2* and mutant alleles of several RNA processing enzymes including those involved in mRNA 3'-end formation, mRNA degradation, and RNA transcription termination. To test for an interaction with the mRNA 3'-end processing machinery, the *NAB2* gene was deleted in *S. cerevisiae* cells expressing either a mutant allele of the poly(A) polymerase, *pap1-1* (221), or a mutant allele of a component of the cleavage machinery, *rna15-2* (177). Cells deleted for the essential *NAB2* gene were complemented by a wild-type *NAB2* maintenance plasmid and transformed with plasmids encoding wild-type Nab2, nab2-1, nab2-$C_{5\text{-}7}$→A, or nab2-C437S. Cells were then serially diluted and spotted onto either control plates or plates containing 5-FOA to select against the wild-type *NAB2* maintenance plasmid. As shown in Figure 4.1, cells expressing nab2-$C_{5\text{-}7}$→A in combination with either pap1-1 or rna15-2 showed a severe growth defect at 25°C, supporting the hypothesis that Nab2 RNA binding plays a critical role in 3'-end formation. Interestingly, expression of the Nab2 mutant, ΔN-Nab2, which lacks Nab2 residues 3-97, in *pap1-1* cells only slightly exacerbated the growth defect of ΔN-Nab2 cells. In addition, ΔN-Nab2 did not show any

genetic interaction with rna15-2, suggesting that any genetic interaction between Nab2 and Rna15 may be allele specific and perhaps directly related to the decreased polyadenosine RNA binding affinity of nab2-$C_{5\text{-}7}$→A for poly(A) RNA.

The extended poly(A) tails observed in cells expressing Nab2 $C_{5\text{-}7}$→A (Figure 3.18) as well as genetic interactions between Nab2 alleles and components of the 3'-end processing machinery, strongly suggest that Nab2 is involved in mRNA processing. Interestingly, previous work has demonstrated that the *pap1-1* mutation can be suppressed by the deletion of the nuclear exosome component, Rrp6 (203). Deletion of *RRP6* most-likely suppresses the mutant pap1-1 by restoring steady-state levels of poly(A) RNA (203).

We hypothesized that if Nab2 and Pap1 function in the same genetic pathway, deletion of *RRP6* might also suppress Nab2 mutations. To test whether Δ*RRP6* can suppress mutations within Nab2, a plasmid shuffle assay was performed on Δ*NAB2* Δ*RRP6* cells. Δ*NAB2* cells or Δ*NAB2* Δ*RRP6* cells complemented by a maintenance plasmid encoding a wild-type copy of Nab2 were transformed with either empty vector control plasmids or plasmids encoding wild-type Nab2, ΔN-Nab2, or a mutant copy of Nab2 containing a leucine to proline amino acid substitution in the N-terminus of Nab2 (Nab2- L18P). The L18P mutation is thought to phenocopy deletion of the entire N-terminal domain by severely disrupting the folding of the PWI-like domain (our unpublished results and (123)). Transformants were grown to saturation in liquid culture, serially diluted, and spotted onto control media lacking uracil and leucine or media containing 5-FOA to select against the wild-type maintenance plasmid. As shown in the

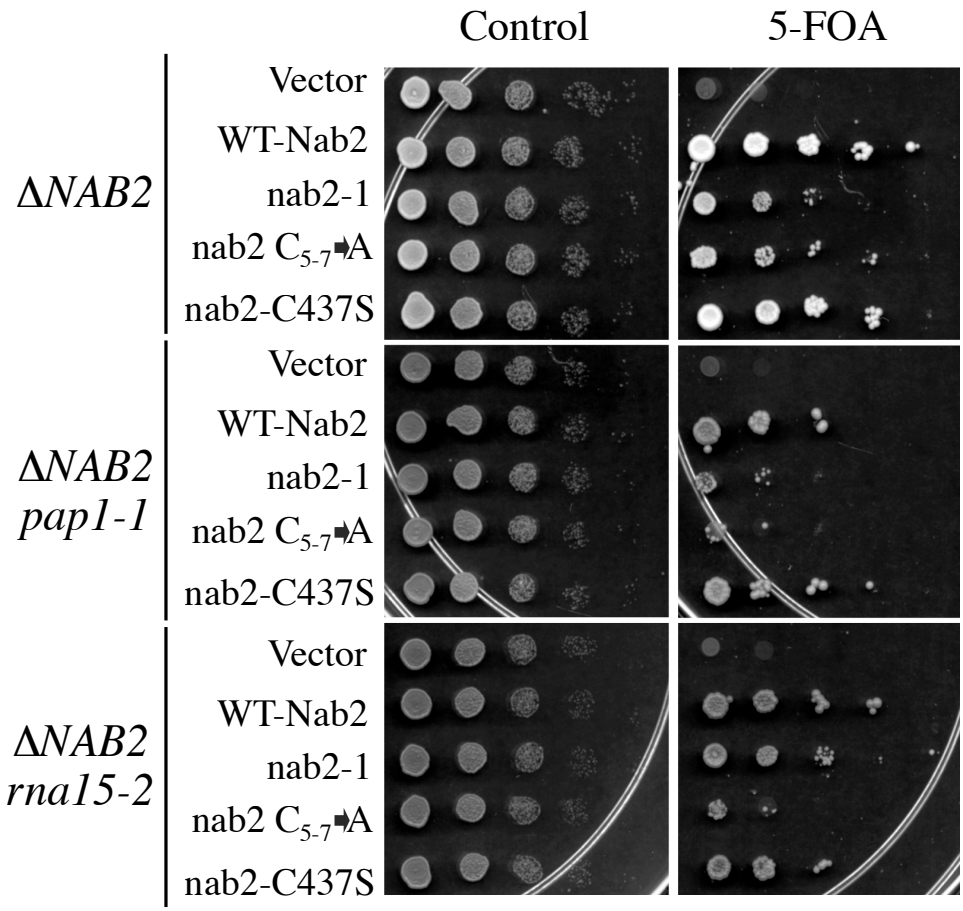

Figure 4.1: Nab2 genetically interacts with mRNA 3'-end processing components. Δ*NAB2* (top row), Δ*NAB2 pap1-1* (middle row), or Δ*NAB2 rna15-2* (bottom row), cells harboring a *URA3* plasmid encoding wild-type *NAB2* were transformed with *LEU2* plasmids encoding either wild-type Nab2, nab2-1, nab2-C_{5-7}→A, or nab2-C437S. Cells were spotted onto either control media lacking both uracil and leucine (Control) or media containing 5-fluoroorotic acid (5-FOA) to eliminate the wild-type Nab2 maintenance plasmid and incubated at 25°C for 3-5 days.

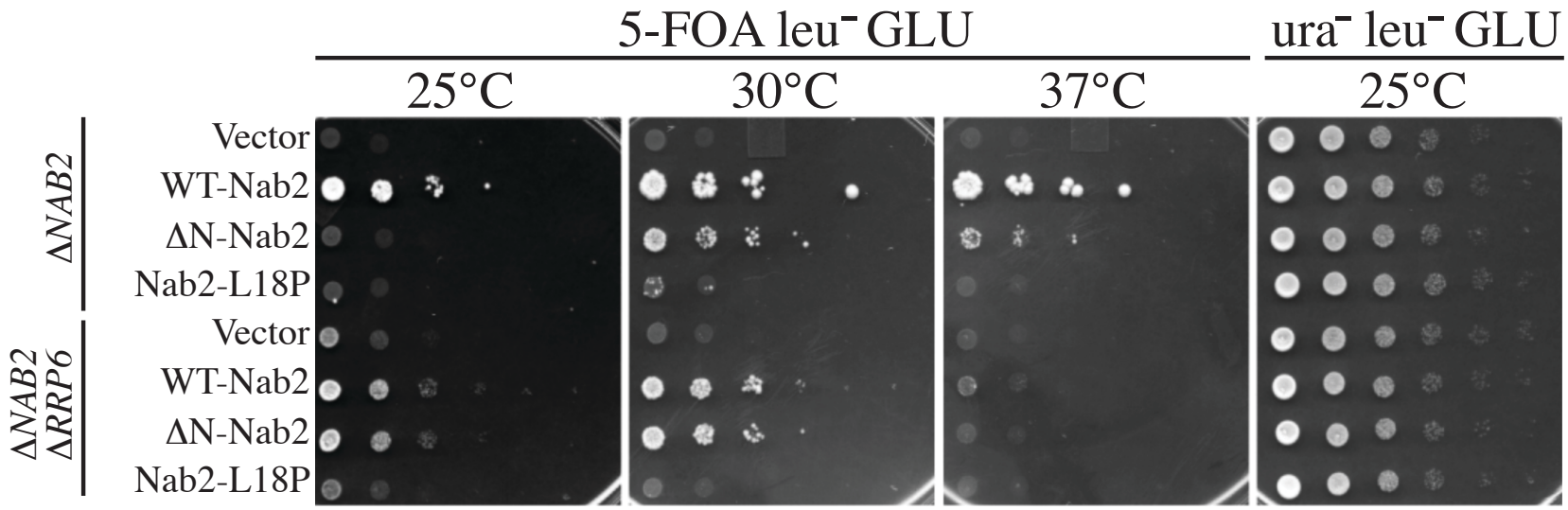

Figure 4.2: Deletion of the exosome component Rrp6 partially rescues the deletion of the essential *NAB2* gene. Δ*NAB2* or Δ*NAB2* Δ*RRP6* cells transformed with a *URA3* plasmid encoding wild-type Nab2 were transformed with *LEU2* plasmids encoding either wild-type Nab2, ΔN-Nab2 (missing amino acids 3-97) or Nab2-L18P. Transformants were inoculated into liquid media, grown to saturation at 25°C, and serially diluted. Dilutions were then spotted onto either control media lacking both uracil and leucine or media containing 5-fluoroorotic acid (5-FOA) to eliminate the wild-type Nab2 maintenance plasmid. The control plate was incubated at 25°C for 2 days while plates containing 5-FOA were incubated at the indicated temperatures for 3-5 days.

control for Figure 4.2, both Δ*NAB2* and Δ*NAB2* Δ*RRP6* cells were properly diluted and spotted across all plates as demonstrated by the uniform cell growth at 25°C on the control plate. As expected, when Δ*NAB2* cells were transformed with a control empty vector and grown on media containing 5-FOA, no cell growth was observed at any temperature tested. Interestingly, when Δ*NAB2* Δ*RRP6* cells containing only the empty vector were grown at 25°C on 5-FOA, a moderate amount of cell growth was observed (See also Figure 4.3), suggesting that deletion of *RRP6* could bypass the essential function of Nab2. Deletion of *RRP6* also suppressed ΔN-Nab2 and, to a lesser extent, Nab2-L18P. As previously reported (222), Δ*RRP6* cells were temperature sensitive and did not grow at or above 30°C.

In order to test whether an Rrp6 mutant that lacks enzymatic activity can also suppress deletion of *NAB2*, Δ*NAB2* or Δ*NAB2* Δ*RRP6* cells that were complemented with a wild-type Nab2 maintenance plasmid were transformed with plasmids encoding either wild-type Rrp6 or a mutant copy of Rrp6 containing an aspartic acid to alanine amino acid substitution (D238A Rrp6), which eliminates the catalytic activity of Rrp6 (222). Transformed cells were then also transformed with either empty vector or plasmids encoding wild-type Nab2, ΔN-Nab2, or Nab2-L18P. Single colonies were then restreaked to media containing 5-FOA to select against the wild-type *NAB2* maintenance plasmid. As shown in Figure 4.3, Δ*NAB2* cells (top left) transformed with a plasmid encoding wild-type Nab2 grew at 25°C. However, Δ*NAB2* cells transformed with an empty vector showed no growth. As demonstrated previously (Figure 4.2), Δ*NAB2* Δ*RRP6* cells transformed with an empty vector showed modest growth at 25°C (Figure

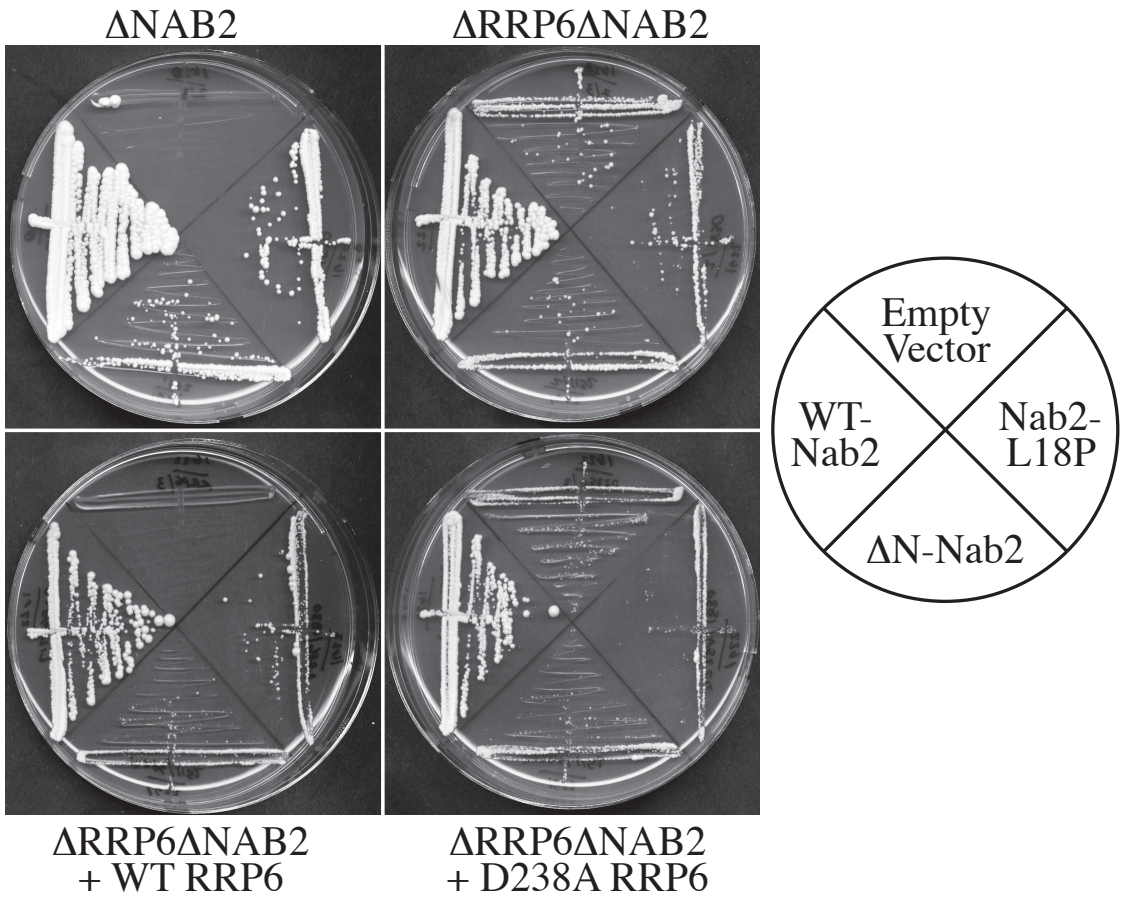

Figure 4.3: A catalytically inactive Rrp6 mutant also suppresses the deletion of the essential Nab2 gene. Δ*NAB2* and Δ*NAB2* Δ*RRP6* cells harboring a *URA3* plasmid encoding wild-type Nab2 were transformed with an empty *TRP3* plasmid (top row) or a *TRP3* plasmid encoding either wild-type Rrp6 (bottom left) or an active site mutant, D238A-Rrp6, which eliminates the catalytic activity of Rrp6 (bottom right). Cells were then also transformed with an empty *LEU2* plasmid or *LEU2* plasmids encoding wild-type Nab2, ΔN-Nab2 (missing amino acids 3-97) or Nab2-L18P. Transformants were streaked to media containing 5-fluoroorotic acid (5-FOA) to eliminate the wild-type Nab2 maintenance plasmid. Plates were incubated at 25°C for 7-10 days. The diagram on the right demonstrates the quadrant of each plate where cells expressing no Nab2, wild-type Nab2, ΔN-Nab2, or Nab2-L18P are streaked.

4.3, upper right). As a control, Δ*NAB2* Δ*RRP6* cells were also transformed with a plasmid encoding wild-type Rrp6 and an empty vector (and hence no Nab2 was expressed on media containing 5-FOA). In Δ*NAB2* Δ*RRP6* cells transformed with a plasmid encoding a wild-type copy of Rrp6 no growth was observed (bottom left), confirming that it is the deletion of Rrp6 that acts as a partial bypass suppressor of loss of Nab2. However, when Δ*NAB2* Δ*RRP6* cells were transformed with a plasmid encoding the Rrp6 active site mutant D238A and vector only (so that no Nab2 or functional Rrp6 was present in these cells), modest growth was observed, suggesting that in the absence of Rrp6 activity Nab2 is not essential.

Since Nab2 mutations cause accumulation of poly(A) RNA in the nucleus (71), we hypothesized that deletion of Rrp6 might rescue Δ*NAB2* cell growth by alleviating this phenotype. To test whether deletion of *RRP6* rescues the nuclear accumulation of poly(A) RNA, we used fluorescence *in situ* hybridization (FISH) to localize poly(A) RNA in Δ*NAB2* or Δ*NAB2* Δ*RRP6* cells transformed with an empty vector (and therefore expressing no Nab2), or plasmids encoding wild-type Nab2, ΔN-Nab2, or Nab2-L18P. As controls in Figure 4.4, Rrp6 wild-type cells expressing wild-type Nab2 show no accumulation of RNA in the nucleus, while cells expressing either ΔN-Nab2 or Nab2-L18P show profound nuclear accumulation of poly(A) RNA. This finding is consistent with previous studies demonstrating that cells expressing ΔN-Nab2 accumulate poly(A) RNA in the nucleus (73). Interestingly, Δ*NAB2* Δ*RRP6* cells transformed with an empty vector, and therefore not expressing any Nab2, show accumulation of poly(A) RNA in distinct nuclear foci (white arrows). This observation is also consistent with recent work

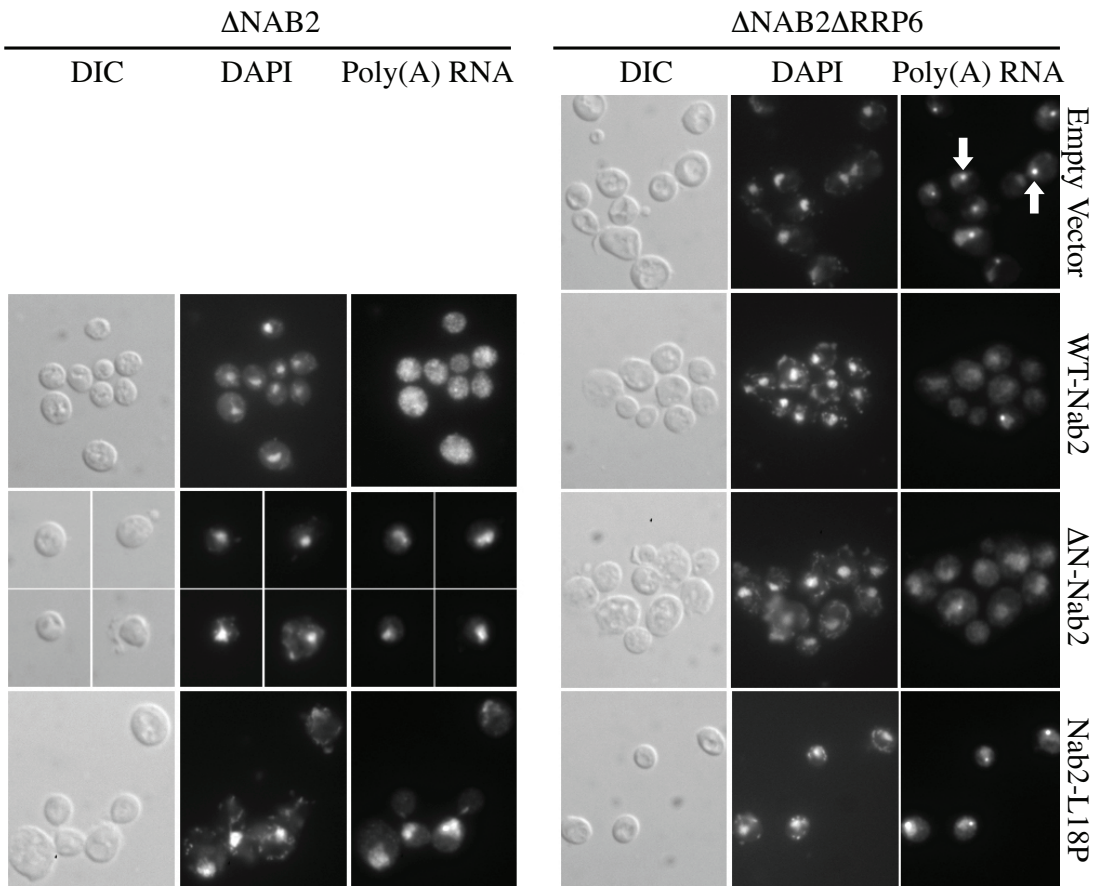

Figure 4.4: Cells deleted for both Nab2 and RRP6 show distinct nuclear foci of poly(A) RNA. Fluorescence *in-situ* hybridization (FISH) using an oligo-d(T) probe to detect poly(A) RNA was performed on Δ*NAB2* and Δ*NAB2* Δ*RRP6* yeast cells not expressing Nab2 (empty vector) or expressing wild-type Nab2, ΔN-Nab2, or Nab2-L18P. Nuclei were stained with DAPI. Corresponding differential interfering contrast (DIC) images are shown. White arrows demonstrate the presence of poly(A) foci in the nucleus of Δ*NAB2* Δ*RRP6* cells not expressing Nab2.

demonstrating that cells lacking Rrp6 show poly(A) RNA accumulation in the nucleolus (223). When Δ*NAB2* Δ*RRP6* cells express wild-type Nab2 or ΔN-Nab2 from a plasmid, these foci are still present but much less prevalent, suggesting that while deletion of Rrp6 may cause the formation of these poly(A) RNA foci, lack of functional Nab2 may make them more widespread. Although poly(A) RNA foci accumulate in cells lacking both Rrp6 and Nab2, the nuclear accumulation of poly(A) RNA in Δ*RRP6* Δ*NAB2* cells is not as robust as in cells lacking *NAB2* alone, supporting our hypothesis that deletion of Rrp6 suppresses mutations in Nab2 by partially alleviating the nuclear accumulation of poly(A) RNA.

Several recent studies have demonstrated that Rrp6 is involved, along with the mRNA 5'-end cap-binding protein, Cbc1, in the decay of normal mRNA when transcripts are abnormally retained in the nucleus (204-206). Accordingly, we hypothesized that Rrp6 and Cbc1 might act together to degrade mRNA transcripts that accumulate in the nucleus of Δ*NAB2* cells and therefore deletion of *CBC1*, like Δ*RRP6*, might suppress deletion of *NAB2*. To test this idea, a plasmid shuffle assay was performed with Δ*NAB2* or Δ*NAB2* Δ*CBC1* cells that were transformed with empty vector or plasmids encoding wild-type Nab2, ΔN-Nab2, or Nab2-L18P. As demonstrated in Figure 4.5, neither Δ*NAB2* cells nor Δ*NAB2* Δ*CBC1* cells transformed with empty vector grew on media containing 5-FOA. Both Δ*NAB2* and Δ*NAB2* Δ*CBC1* cells expressing wild-type Nab2 grew similarly. Interestingly, cells lacking Cbc1 and expressing either ΔN-Nab2 or Nab2-L18P showed slight growth defects compared to cells lacking Cbc1 and expressing wild-type Nab2. Further experimentation will be needed to verify this result. Overall,

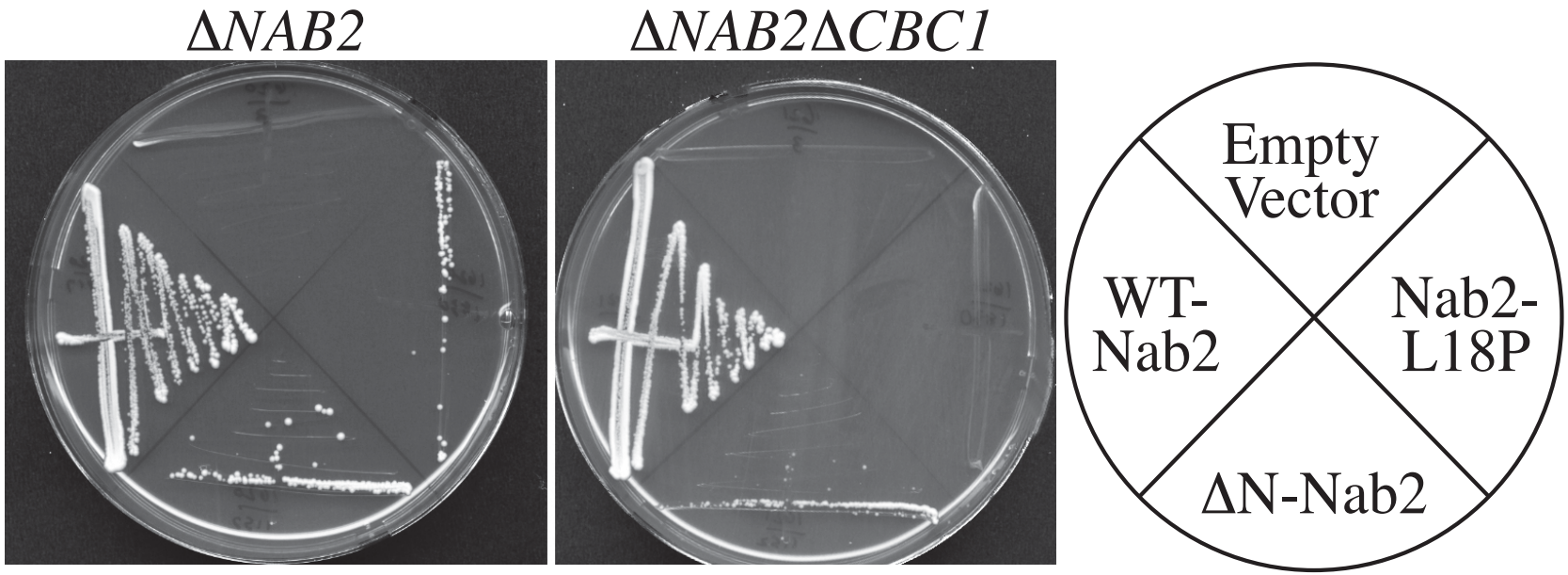

Figure 4.5: *NAB2* does not interact with another component of the Decay of RNA in the Nucleus (DRN) machinery, *CBC1*. Δ*NAB2* and Δ*NAB2* Δ*CBC1* cells harboring a *URA3* plasmid encoding wild-type Nab2 were transformed with an empty *LEU2* plasmid or *LEU2* plasmids encoding wild-type Nab2, ΔN-Nab2 (missing amino acids 3-97) or Nab2-L18P. Transformants were streaked to media containing 5-fluoroorotic acid (5-FOA) to eliminate the wild-type Nab2 maintenance plasmid. Plates were incubated at 30°C for 3-5 days. The diagram on the right demonstrates the quadrant of each plate where cells expressing no Nab2, wild-type Nab2, ΔN-Nab2, or Nab2-L18P are streaked.

these results suggest that suppression of Δ*NAB2* cells by deletion of *RRP6* is not simply due to a decrease in RNA degradation in the nucleus.

Since Nab2 genetically interacts with Rrp6 and Rrp6 is important for the processing of snoRNAs, we hypothesized that Nab2 might genetically interact with proteins involved in the biogenesis of snoRNA transcripts. To test whether Nab2 genetically interacts with a component of the snoRNA transcription termination machinery, the essential *NAB2* gene was deleted in cells expressing a temperature sensitive allele of *NAB3*, *nab3-11*, and complemented by a wild-type Nab2 maintenance plasmid. A plasmid shuffle assay (153) was performed with Δ*NAB2* cells or Δ*NAB2* cells that also expressed the temperature sensitive nab3-11 protein. Δ*NAB2* or Δ*NAB2 nab3-11* cells were transformed with plasmids encoding either wild-type Nab2, ΔN-Nab2, nab2-C_{5-7}→A, or mutant Nab2 containing a single cysteine to serine amino acid substitution in the first cysteine of zinc finger six (nab2-C437S). As shown in Figure 4.6, in the presence of a wild-type copy of Nab3 (top), wild-type Nab2 cells grow at all temperatures, nab2-C_{5-7}→A and nab-C437S show cold-sensitivity at 16°C and 25°C, and ΔN-Nab2 shows dramatic growth defects at all tested temperatures. These results are consistent with our previous data which demonstrates that nab2-C_{5-7}→A displays cold sensitive growth defects at 16°C (Figure 3.3). Interestingly, the combination of nab3-11 and nab2-C_{5-7}→A resulted in enhanced cold-sensitivity. However, neither *nab2-C437S* nor *ΔN-Nab2* demonstrated any genetic interaction with *nab3-11*, suggesting that this phenotype is specific to *nab2-C_{5-7}→A*. Although nab2-C437S binds slightly more weakly to polyadenosine RNA than wild-type Nab2 (67),

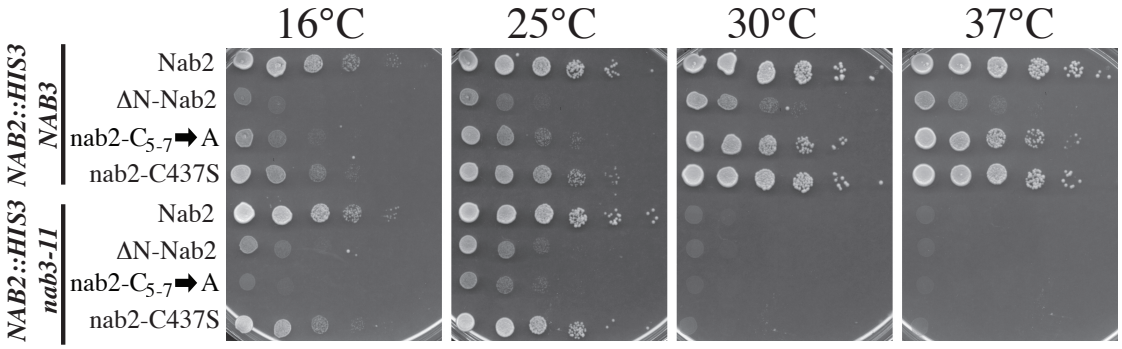

Figure 4.6: *NAB2* mutants that do not bind polyadenosine RNA genetically interact with *nab3-11*. Δ*NAB2* or Δ*NAB2 nab3-11* cells harboring a maintenance plasmid encoding wild-type Nab2 were transformed with *LEU2* plasmids encoding wild-type Nab2, ΔN-Nab2, nab2-C_{5-7}→A, and nab2-C437S. Cells were then struck to media containing the drug 5-FOA to select against the wild-type Nab2 maintenance plasmid. Single colonies were then inoculated into liquid culture, grown to saturation, serially diluted, and spotted onto selective media lacking leucine. Plates were grown at selected temperatures for 5-10 days.

binding of nab2-$C_{5\text{-}7}$→A to poly(A) RNA is even more greatly impaired (Figure 3.4). Therefore, we postulate that the cause behind this genetic interaction is the greatly reduced binding of nab2-$C_{5\text{-}7}$→A to polyadenosine RNA compared to wild-type Nab2 (or nab2-C437S).

Since Nab2 genetically interacts with components of the snoRNA transcriptional termination machinery as well as an exosome component, Rrp6, required for trimming of snoRNA 3'-ends, we wanted to investigate whether Nab2 could interact with these short oligo(A) tails added to the 3'-end of snoRNA transcripts. To begin to test the hypothesis that Nab2 binds short oligo(A) tails, we needed to accurately determine both the length of the oligo(A) tails added to snoRNA transcripts destined for degradation as well as the shortest stretch of adenosines to which Nab2 could bind. A previous study (223), demonstrated that Δ*RRP6* cells accumulate an oligoadenylated U14 (also known as SNR128) snoRNA transcript in the nucleolus. Normally, U14 is transcribed as part of a polycistronic message, cleaved by the endonuclease Rnt1, and the 3'-end of the snoRNA is trimmed by the nuclear exosome (224, 225). These longer, oligoadenylated U14 transcripts accumulate in cells lacking Rrp6 (223). Therefore, we used this as a candidate snoRNA to determine the length of the oligo(A) tails added to snoRNA transcripts. In order to determine the length of the oligo(A) tails added to U14 snoRNA transcripts in Δ*RRP6* cells, we utilized the method described by Couttet *et al*. (226). As shown in Figure 4.7, this method begins by removing the 5'-cap of all RNA transcripts using tobacco acid pyrophosphatase (TAP). Decapped RNA is then circularized using T4 RNA ligase. The circularized RNA is then used as a template for cDNA synthesis during an RT-PCR reaction. Gene-specific primers are then used to amplify the 5'-end/3'-end

boundary of the cDNA of interest. PCR products are then cloned and sequenced. Using this technique, we found that the majority of U14 snoRNA transcripts are properly processed in wild-type cells at both 25°C and 37°C, as expected (Figure 4.8). However, in cells lacking the exosome component Rrp6, U14 transcripts were extended and oligoadenylated at both 25°C and following a shift to 37°C. Although more transcripts will need to be sequenced to increase the accuracy of this assessment, the shortest oligo(A) tail we identified was 17 adenosines.

In order to investigate the shortest stretch of adenosines bound by Nab2, we analyzed binding of full-length recombinant Nab2 to fluorescein-labeled poly(A) RNA oligonucleotides *in vitro* using fluorescence anisotropy. As shown in Figure 4.9, the relative affinity of Nab2 for fluorescein-labeled $poly(A)_{25}$ and $poly(A)_{20}$ RNA oligonucleotides was approximately 15 nM for $poly(A)_{25}$ and 16.3 nM for $poly(A)_{20}$. Nab2 bound fluorescein-labeled $poly(A)_{15}$ RNA oligonucleotides with a relative affinity of 31.2 nM, about two-fold more weakly than either the 20 or 25 nucleotide oligonucleotides. Interestingly, Nab2 affinity for $poly(A)_{10}$ was significantly lower (K_d = 202.5 nM), suggesting that the smallest length of adenosines to which Nab2 can bind with high affinity is roughly between 10 and 15 nucleotides. Along with the fact that shortest oligo(A) tail added to the 3'-end of U14 snoRNA transcripts was at least 17 adenosines long, this binding data suggests that Nab2 could bind to the short oligo(A) tails of snoRNAs destined for exosome-mediated degradation.

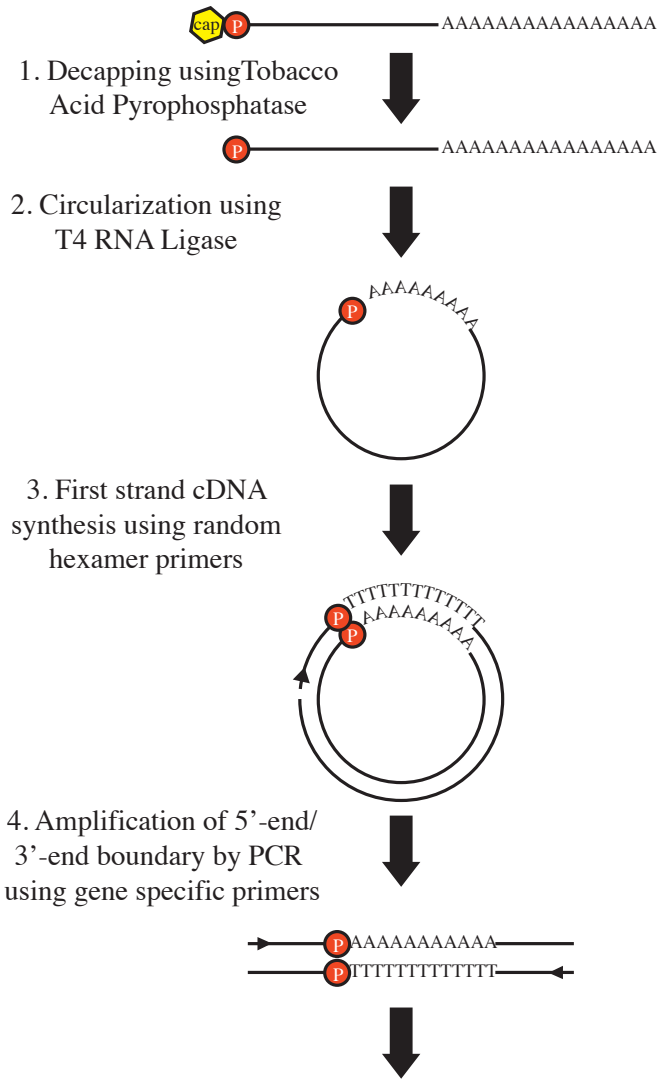

Figure 4.7: Schematic for circular RT-PCR. Figure adapted from (226). As described in Experimental Procedures, total RNA is isolated from *S. cerevisiae* cells and decapped using tobacco acid pyrophosphatase. Decapped RNA is then circularized using T4 RNA ligase. First strand cDNA is synthesized using random hexamer primers. cDNA is then used as a template for PCR amplification of the 5'/3'-end boundary of the transcript of interest, in this case the U14 (SNR128) snoRNA. PCR products are cloned into Qiagen's pDRIVE vector and sequenced to determine the exact processing state of both the snoRNA 5'- and 3'-ends.

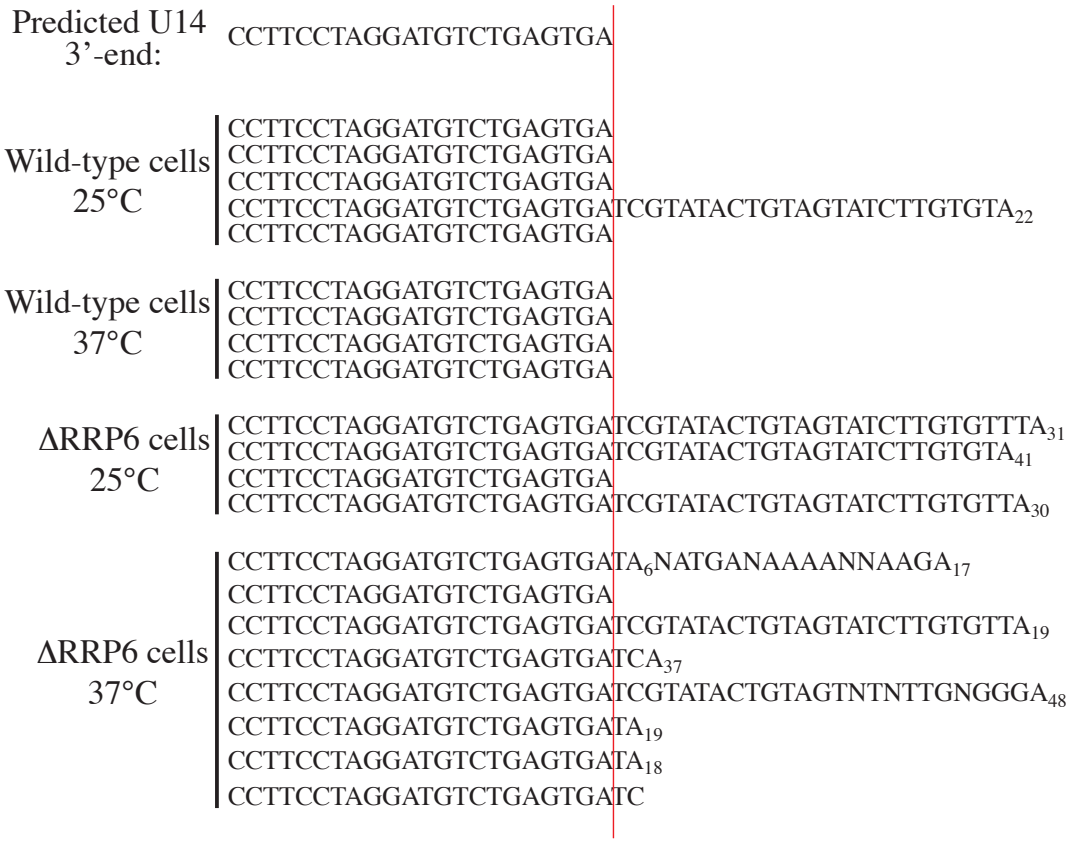

Figure 4.8: Cells lacking Rrp6 accumulate misprocessed U14 snoRNA transcripts. Δ*RRP6* cells have previously been shown to accumulate an elongated and oligoadenylated version of the U14 transcript (223), however the exact length of the oligo(A) tail was not known. Total RNA was isolated from isogenic wild-type (BY4741) and Δ*RRP6 S. cerevisiae* cells grown only at 25°C or cells grown at 25°C and shifted to 37°C for 1 hour. The results of a circular-RT-PCR reaction using gene specific primers for the 5'-end/3'-end overlap of the U14 snoRNA transcript are shown here. The published 3'-end of the U14 transcript is shown for comparison. At least 4 cloned cDNAs were sequenced for each condition. The red line denotes the normal 3' -end of the U14 snoRNA transcript. Nucleotides to the right of the red line would normally be trimmed by the exosome.

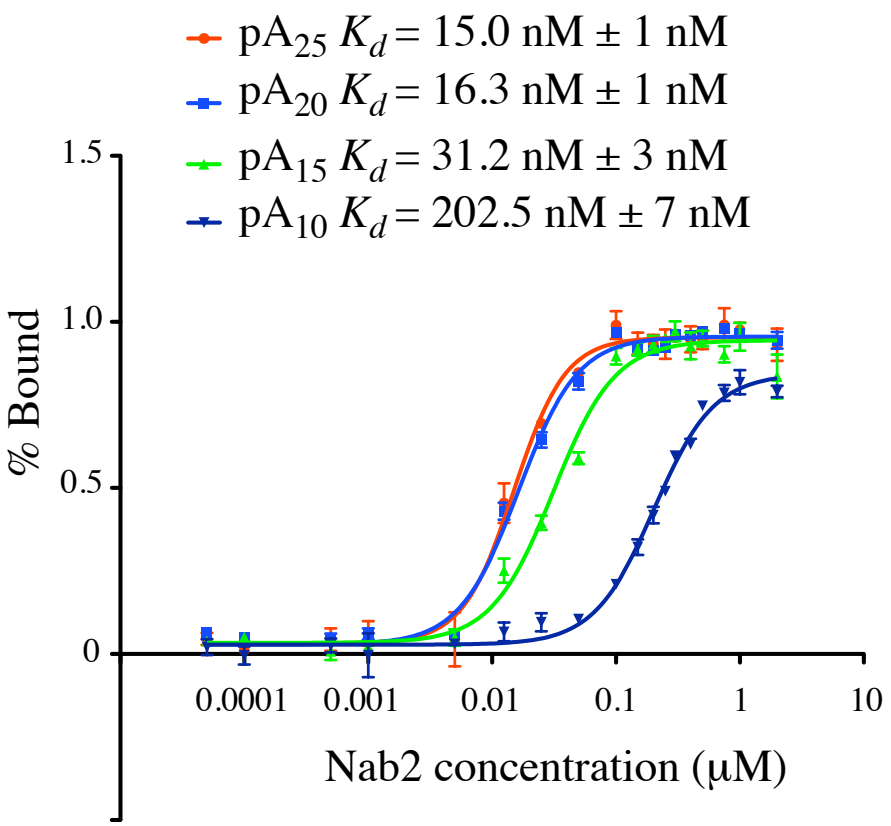

Figure 4.9: The Nab2 minimal binding element is between 10 – 15 adenosines. Recombinant full-length Nab2 was expressed and purified from BL21 (DE3) pLYS *E. Coli* cells. The affinity of Nab2 for fluorescein-labeled poly(A)$_{25}$, poly(A)$_{20}$, poly(A)$_{25}$, poly(A)$_{15}$, and poly(A)$_{10}$ RNA oligonucleotides was determined using fluorescence anisotropy. 2 nM fluorescein-labeled RNA was incubated with increasing protein concentrations (50pM – 5 μM). Each sample was incubated for ~1 hour to reach equilibrium and then placed into one well of a 384 well plate. Each protein concentration was done in triplicate. Anisotropy values were then determined using a fluorescence plate reader. Apparent K_d values of Nab2 for each of the fluorescein-labeled RNA oligonucleotides are shown.

Discussion

Here we present data demonstrating that mutants in the essential poly(A) RNA binding protein, Nab2, genetically interact with critical components of the mRNA 3'-end processing machinery, including both Pap1 and Rna15. Furthermore, we also demonstrate that deletion of *RRP6* suppresses deletion of the essential *NAB2* gene. A mutant of Rrp6, Rrp6-D238A, which lacks enzymatic riboexonuclease activity, also bypasses the lethality of deletion of *NAB2*. Δ*NAB2* Δ*RRP6* cells also accumulate poly(A) RNA in nuclear foci. We also investigate the length of the oligo(A) tails found on the 3'-end of a candidate snoRNA, U14, in cells lacking Rrp6. Our results demonstrate that at a minimum 17 adenosines are added to the 3'-end of the U14 transcript in Δ*RRP6* cells. Using fluorescence anisotropy, we also demonstrate that the minimal stretch of adenosines to which Nab2 can bind with high affinity is between 10-15 nucleotides.

Together these results, suggest several non-mutually exclusive scenarios. The first scenario is one in which Nab2 interacts with only the poly(A) tails of mRNA transcripts and prevents degradation by the nuclear exosome, specifically Rrp6. In support of this model, Nab2 mutants show extended poly(A) tails and those same mutants genetically interacts with two components of the 3'-end cleavage and polyadenylation machinery, Pap1 and Rna15. A previous study investigating the transcripts preferentially associated with Nab2 also revealed that Nab2 bound to several different functional classes of mRNA transcripts (49). Deletion or inactivation of Rrp6 would suppress Δ*NAB2* due to an increase in steady-state levels of mRNA transcripts. Without Rrp6, there would be no need for an mRNA protector, making the function of Nab2 non-essential. Interestingly, however, this model is not consistent with the fact that cells

expressing mutant version of Nab2 show extended poly(A) tails compared to the poly(A) tail length of RNA isolated from wild-type cells (Figure 3.18 and (71)). If Nab2 were protecting transcripts from degradation via the 3'→5' riboexonuclease activity of the nuclear exosome, then shorter poly(A) tails would be expected when Nab2 was not efficiently recruited.

A second model can also be formulated in which Nab2 binds to the oligo(A) tails of short RNA transcripts and actually recruits Rrp6 and the nuclear exosome. Although Nab2 has never been found associated with snoRNAs or CUTs, these RNAs are remarkable unstable and therefore extremely difficult to detect when they are oligoadenylated and destined for degradation via the exosome (181, 182, 223, 227). Indeed, the mere presence of these short oligoadenylated transcripts was only uncovered upon deletion of the exosome component Rrp6 (181, 182, 223, 227). Furthermore, the previous study investigating transcripts associated with Nab2 used microarrays containing only mRNA transcripts (49). According to this model, deletion or mutation of Nab2 would lead to the nuclear accumulation of error-containing transcripts, which could be deleterious to the cell. Although the mechanism by which deletion of Rrp6 partially suppresses Δ*NAB2* remains unclear, this model is also more consistent with the fact that Nab2 mutants show longer poly(A) tails, as would be expected if the exosome were not recruited. While Rna15 and Pap1 are typically thought of solely as components of the mRNA 3'-end processing machinery, several studies suggest that they may also play roles in the 3'-end formation of other RNA polymerase II transcripts (100, 211, 228, 229). Specifically, components of the yeast cleavage complexes CFIA and CFIB may be required for the correct 3'-end formation of specific snoRNA transcripts (228, 229).

Genetic interactions between *NAB2* and both *RNA15* and *NAB3* suggest that perhaps Nab2 could also play a role in snoRNA biogenesis.

While often overlooked, snoRNAs play a critical role in the processing of numerous ribosomal RNAs (180, 183). Although little is known about the factors involved in 3'-end processing of these small RNA transcripts, several proteins, which were previously thought to only function in mRNA 3'-end processing, have recently been implicated. The quality control of snoRNAs is also not well understood. Short oligo(A) tails are added to error-containing transcripts and these faulty transcripts are then targeted for degradation via the nuclear exosome. The mechanism by which the cell differentiates between the short oligo(A) tails of error containing snoRNAs and the longer poly(A) tails of mRNA transcripts also remains a mystery, but the data presented here provides a possible link between all of these unknowns. Overall, this data suggests that the yeast Nab2 protein may play a role in the complicated processing and quality control of these short RNA polymerase II transcripts. Interestingly, Nab2 is a member of a family of CCCH zinc finger containing poly(A) RNA binding proteins (67, 146). The putative human orthologue of Nab2, ZC3H14, localizes to nuclear speckles, which are thought to house RNA processing components (146). Future studies will be needed to cement the link between Nab2, and perhaps its higher eukaryotic orthologues, and small RNA biogenesis.

Experimental Procedures

Chemicals, Plasmids, and Saccharomyces cerevisiae Manipulations – Chemicals were obtained from Fisher Scientific (Pittsburgh, PA), Sigma-Aldrich (St. Louis, MO), or US

Biological (Swampscott, MA) unless otherwise noted. DNA manipulations were performed according to standard methods (150) and all media were prepared by standard procedures (151). All S. cerevisiae strains and plasmids used in this study are described in Table 4.1. Plasmids encoding mutant nab2 proteins were created by site-directed mutagenesis of a wildtype Nab2 plasmid (pAC717) using the QuikChange Site-directed mutagenesis kit (Stratagene). All plasmids were fully sequenced to ensure that no mutations other than those targeted were introduced during site-directed mutagenesis. Double deletions strains were constructed by mating *S. cerevisiae* cells deleted for NAB2 (ACY427) and deletions strains obtained from the genome-wide deletion collection (230). All deletions were checked by PCR amplification of the deleted locus from genomic DNA.

In vivo Functional Analysis – The in vivo function of each nab2 mutant was tested using a plasmid shuffle assay (175). *S. cerevisiae* cells deleted for NAB2 (ACY427) and containing a wild-type NAB2 URA3 plasmid (pAC636) were transformed with LEU2 plasmids encoding various nab2 mutants. Transformants were grown to saturation and cells were then serially diluted and spotted onto control ura- leu- glucose plates or plates containing 5-fluoroorotic acid (5-FOA). Plates were then incubated at 16°C, 18°C, 25°C, 30°C, or 37°C for 3-5 days. The toxic uracil analog, 5-FOA, kills cells that contain a functional uracil biosynthetic pathway (175). Hence, only those cells that lose the URA3 plasmid encoding wild-type Nab2 and retain a functional copy of Nab2 on the LEU2 plasmid will grow on media containing this drug.

Fluorescence in-situ Hybridization (FISH) – FISH was performed as described previously (231). Briefly, yeast cells expressing wild-type *NAB2, nab2-1*, *nab2-*$C_{5\text{-}7}$$\rightarrow$*A*,

or *nab2-C_{5-7}→R* were initially grown in 2 ml YPD cultures to saturation at 30°C. These starter cultures were then used to inoculate 5 ml YPD cultures that were grown overnight (approx. 16 hours) at either 30°C or 18°C. Cells were then fixed by the addition of 700 μl of 37% formaldehyde and incubated at 30°C or 18°C for 90 min. Cells were pelleted, washed three times with 0.1 M $K_4(PO_4)_3$ pH 6.5, and resuspended in P-solution (0.1 M $K_4(PO_4)_3$ pH 6.5, 1.2 M sorbitol). Fixed cells were then seated onto poly-L-lysine coated multi-well slides and permeabilized with 0.1% Triton-X-100. Permeabilized cells were then treated with pre-hybridization solution (50% deionized formamide, 4x SCC, 1x Denhardt's Solution, 125 μg/ml tRNA, 10% dextran sulfate) and hybridized with a digoxygenin end-labeled oligo(dT_{50}) probe overnight at 37°C in a humidified chamber. Unhybridized probe was washed off the cells with increasing amounts of salt. Digoxigenin-labled oligo (dT_{50}) was detected using a rhodamine labeled anti-digoxygenin antibody. Cells were also stained with 4', 6-diamidino-2-phenylindole (DAPI) to visualize DNA within the nucleus.

Total RNA isolation – Total cellular RNA was isolated from *S. cerevisiae* as previously described (232). Briefly, single colonies were inoculated into rich media (YEPD) and cells were grown to saturation at the permissive temperature of either 25°C (for ΔRRP6 cells) or 30°C (for nab2-C_{5-7}→A) for two days. Cultures were then diluted into 10 ml of rich medium and grown at the permissive temperature for 4 hours. Cells were then split into two 5 ml cultures. One 5 ml culture was replaced at the permissive temperature while the other 5 ml culture was shifted to the restrictive temperature for 6 hours (37°C for ΔRRP6 cells and 18°C for nab2-C_{5-7}→A). Following the 6 hr. temperature shift, cells were harvested by centrifugation and cell pellets were frozen at -80°C. In order to isolate

RNA, cell pellets were thawed on ice and resuspended in extraction buffer (200mM Tris-HCl pH 7.5, 0.5M LiCl, 5mM EDTA, 1% SDS) containing 400µl of phenol/chloroform/isoamyl alcohol (PCI) and acid-washed glass beads. Samples were then lysed by bead-beating for three 1 minute pulses with 1 minute on ice between each pulse. Cell debris was pelleted by centrifugation at 13,000 rpm for 10 minutes at 4°C. The top layer, containing the RNA, was transferred to a fresh eppendorf tube and RNA was extracted twice with an equal volume of PCI. RNA was then precipitated with three volumes of ice cold 100% ethanol at -20°C for 1 hour. RNA was pelleted by centrifugation at 13,000 rpm for 10 minutes at 4°C. RNA pellets were washed once with 70% ethanol (made with DEPC-treated dH_2O) and air-dried. Dried pellets were dissolved in 100µl of 1X TE. RNA quality was analyzed by formaldehyde-agarose gel electrophoresis and RNA yield was calculated by absorbance at 260 nm.

Circular RT-PCR (cRT-PCR) analysis of poly(A) tail length – In order to analyze the presence and/or length of the poly(A) tail of several small nucleolar RNA (snoRNA) transcripts we utilized circular RT-PCR (cRT-PCR) as described previously (226). This method ligates the 5' and 3'-ends of a transcript and therefore allows for close inspection of both 5'- and 3'-end processing. As diagrammed in Figure 4.7, this multi-step protocol began with the isolation of total RNA from yeast cells (as described above). 15μg of total RNA was incubated with 5 units of tobacco acid pyrophosphatase (Epicentre) at 37°C for 1 hour. RNA was then precipitated with 3 volumes of ice-cold 100% ethanol and 1/10th volume of 3M NaOAc for 1 hour at -20°C. Precipitated RNA was pelleted and then resuspended in 10µl of DEPC-treated dH_2O. Decapped RNA was then circularized by incubation with 20 units of T4 RNA ligase at 16°C overnight. RNA was then

extracted once with PCI and precipitated as before. RNA pellets were resuspended in 25μl of DEPC-treated dH_2O and RNA yield was analyzed by absorbance at 260 nm.

Once the RNA was circularized, first strand cDNA was synthesized using Qiagen Quantitect RT-PCR kit according to the manufacturer's directions. Briefly, samples were initially treated with gDNA wipe-out buffer to remove contaminating yeast genomic DNA. Then RT-PCR was performed using random hexamer oligonucleotides as primers. 1μl of cDNA was used as a template for PCR amplification of the U14 snoRNA. Oligonucleotide primers were designed so that they amplified the 5'-3' boundary. PCR products were resolved by agarose gel electrophoresis. To directly analyze 5'- and 3'-end processing PCR products were also cloned into Qiagen's pDrive TA cloning vector according to the manufacturer's directions. Plasmids containing inserts were then sequenced.

Protein expression and purification – To express recombinant full-length Nab2 in *E. coli*, the expression plasmid, pAC2133, was transformed into BL21(DE3)pLYS cells (Novagen). Single colonies were inoculated into 2 ml of media and grown to saturation overnight. This culture was then used to inoculate 1 L of LB media. Cells were grown at 37°C until they reached an OD600 nm of 0.4 – 0.6. Cultures were then shifted to 30°C and induced with 200 μM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 hrs. Cells were then centrifuged at 3800rpm for 20 minutes at 4°C and pelleted cells were frozen at -80°C. For purification of untagged Nab2 proteins from frozen cell pellets, cell pellets were thawed on ice and resuspended in an equal volume of lysis buffer (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 4 mM β-mercaptoethanol, 1mM phenylmethylsulfonylfluoride) and lysed using a French Press. The lysates were cleared

by centrifugation, loaded onto a HiTRAP Q column (GE Healthcare) pre-equilibrated with Buffer A (20mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol), and bound protein was eluted using a linear gradient of Buffer B (20 mM piperazine pH 9.5, 1 μM $ZnCl_2$, 2% glycerol, 2 mM β-mercaptoethanol, and 1 M NaCl). Fractions containing untagged recombinant Nab2 were dialyzed overnight into Buffer C (20 mM Tris-HCl pH 8.5, 50 mM NaCl, 2 mM MgOAC, 2 μM $ZnCl_2$, 2% glycerol, and 2 mM β-mercaptoethanol). The dialyzed protein was then concentrated by adding saturated ammonium sulfate to 25% saturation. Precipitated proteins were pelleted and resuspended in Buffer C. The concentrated protein was then loaded onto a Superdex S-200 gel filtration column pre-equilibrated with Buffer C and fractions containing untagged recombinant Nab2 were pooled and concentrated using Centricon spin concentrators (Millipore). Since we found that freezing the purified protein led to loss of nucleic acid binding, freshly purified aliquots of protein were used for each binding experiment.

RNA labeling protocol – RNA 3'-end labeling was performed essentially as described (176). Briefly, RNA oligonucleotides were incubated with 100 mM NaOAc pH 5.1, 100 μM $NaIO_4$ at room temperature for 90 minutes and then precipitated with 2.5 μL of 5 M NaCl, 1 μl of 20μg/μl glycogen (Invitrogen), and 100 μl of ice-cold 100% ethanol for 20 minutes at -20°C. Precipitated RNA was pelleted by centrifugation and resuspended in 50 μl of labeling solution (1.5 mM fluorescein 5-thiosemicabazide, 100 mM NaOAc pH 5.1. Samples were covered and RNA was labeled overnight at 4°C. To remove unreacted fluorescein, labeling reactions were added to a G-25 sephadex column.

Labeled RNA was eluted by centrifugation. Labeling efficiency was determined using the following equation:

$$Efficiency = \frac{OD_{490} / \varepsilon_{fluorescein}}{OD_{260} / \varepsilon_{oligonucleotide}} \quad [1]$$

where OD_{490} and OD_{260} is the absorbance at 490 nm and 260 nm, respectively and ε is the extinction coefficient for either fluorescein or the oligonucleotide as noted. Typical labeling efficiencies were 60-85%. The purity of labeled RNA oligonucleotides was analyzed by agarose gel electrophoresis.

Fluorescence Anisotropy Assay and Dissociation Constant Determination – 2 nM fluorescein-labeled RNA oligonucleotides were incubated with increasing amounts (protein concentrations range from 50 pM to 5 μM) of recombinant purified proteins in 384-well black plates as described (176). Polarization data were fit to Equation 2:

$$\theta = base + \frac{(\max - base)}{1 + \left(\frac{K_d}{P_t}\right)^N} \quad [2]$$

where θ is the experimentally observed polarization value, base is the minimum polarization value when no protein is added, max is the maximum polarization value at saturating concentrations of protein, K_d is the apparent dissociation constant, and P_t is the protein concentration. Reported K_d values are the average of at least three experiments.

Chapter 5: Conclusion and Discussion

A brief review

In their seminal work, Beadle and Tatum originally made the breakthrough finding that one gene encodes one enzyme (233). However, we now know that genetic networks and the regulation of gene expression are much more complex than the simplicity implied by this and other early studies. Even the definition of a gene itself has become more nebulous as we learn more about small gene products, such as short non-coding RNAs transcripts, that play essential roles in the regulation of gene expression. Typically, most of us still think of a gene as Beadle and Tatum did, as a stretch of DNA that encodes a protein. However, in order for the information encoded in that stretch of DNA to be made into protein, an essential intermediary must be made. In eukaryotes, this intermediary, a messenger RNA (mRNA) transcript, is transcribed from DNA in the nucleus by RNA polymerase II. The mRNA is essentially a blueprint from nuclear headquarters that dictates to the translation machinery in the cytoplasm how to make a certain protein. Like any important blueprint sent from a foreman to his or her workers, the original transcript is edited (i.e. processed) and monitored for correctness before leaving the nucleus and being translated into protein in the cytoplasm. It has become increasingly apparent that many of these processing and quality control checkpoints can play large roles in determining the amount of a protein produced as well as the spatial and temporal regulation of gene expression.

The mRNA transcript is never "naked" inside of a cell. From the instant the 5'-end of the nascent transcript emerges from RNA polymerase II, it is immediately coated with RNA binding proteins. These proteins complete the processing steps, signal to the quality control checkpoints that the transcript is properly (or improperly) processed, and

help to package the transcript for the next step in the assembly line. In sum, RNA binding proteins dictate the fate of the transcript to which they are bound. RNA binding proteins perform functions such as facilitating processing, determining transcript stability, functioning in the nuclear export or retention of a certain transcript, and controlling transcript localization in the cytoplasm. Each of these processes, in turn, affects the final protein product. Thus, a more holistic approach, beyond simply studying changes in transcription, needs to be taken in order to more fully understand the regulation of gene expression. Specifically, an accurate account of the mechanisms by which RNA binding proteins specifically recognize target RNAs and the cellular implications of disrupting those interactions is necessary to fully comprehend the post-transcriptional regulation of gene expression.

In the preceding dissertation I have investigated the molecular recognition of polyadenosine RNA by the essential yeast Nuclear poly(A) Binding protein 2 (Nab2). Nab2, like its putative human orthologue, ZC3H14, encodes tandem CCCH zinc fingers. Both proteins bind specifically to polyadenosine RNA via their zinc finger domains. While ZC3H14 contains five CCCH zinc fingers, Nab2 contains seven, suggesting that only a subset of the Nab2 zinc fingers is required for high affinity polyadenosine RNA binding. Accordingly we have demonstrated that only Nab2 zinc fingers 5-7 are required for high affinity binding to poly(A) RNA. In collaboration with Christoph Brockmann and Murray Stewart, the atomic resolution structure of Nab2 zinc fingers 5-7 was solved. However, this structure did not contain RNA and therefore did not allow us to investigate the precise mechanism of molecular recognition. Using this structural information, however, we were able to identify several conserved positively charged and aromatic

residues that were solvent exposed. These residues could mediate interactions with polyadenosine RNA. When several of these amino acids were changed to alanine, defects in polyadenosine RNA binding *in vitro* were seen. These amino acid substitutions also conferred cold-sensitive growth defects *in vivo*. Future structural studies will be required to more accurately assess the mechanism by which Nab2 zinc fingers 5-7 recognize polyadenosine RNA.

Interestingly, we also demonstrate here that Nab2 genetically interacts with components of the RNA processing machinery. Specifically Nab2 genetically interacts with components of the mRNA 3'-end processing machinery (Pap1 and Rna15), a nuclear exosome component (Rrp6), and a factor involved in transcriptional termination of RNA polymerase II (Nab3). These genetic interactions demonstrate the highly intricate and interconnectedness of RNA processing and biogenesis and suggest that Nab2 (and perhaps ZC3H14) may play a critical role in RNA processing.

For the remainder of this chapter I will discuss the implications of these results on the current models of mRNA processing and nuclear export. In addition, several questions that arise from this data about the function of Nab2 and its putative higher eukaryotic orthologue, ZC3H14, will also be addressed. Finally, I will also expound upon several overall conclusions about the impact of this research as a whole on the field of mRNA transport.

The specificity of CCCH zinc fingers

Until our recent studies, the only known RNA binding proteins to specifically recognize polyadenosine RNA did so via at least one RNA Recognition motif. However,

we present data here that a family of CCCH zinc finger containing proteins can also specifically recognize polyadenosine RNA. The atomic resolution structure of Nab2 zinc fingers 5-7, although lacking RNA, has allowed us to identify conserved solvent exposed residues that are important for the interaction between Nab2 and polyadenosine RNA.

One interesting question that arises from these studies is whether an RNA binding domain, such as the CCCH zinc finger domain of Nab2 can be engineered to specifically recognize a different sequence. While the answer to this question may sound straight forward, in most cases a multitude of interactions exist between the protein and the RNA target sequence that dictate specificity. For example, the tandem CCCH zinc fingers of TIS11d are remarkably similar to those found within Nab2 and ZC3H14 and therefore might theoretically recognize similar sequences. However, in Chapter 2 we demonstrated that the UAUUUAUU target sequence of TIS11d could not compete with a poly(A) RNA oligonucleotide for binding to Nab2. A brief comparison between the sequences of the two CCCH zinc fingers can explain these differences [Figure 5.1 – adapted from (108)]. First, the tandem CCCH zinc fingers of TIS11d have a spacing of $CX_8CX_5CX_3H$, while the zinc fingers of Nab2 have a slightly different spacing of $CX_5CX_{4-6}CX_3H$. Second, the (R/K)YTEL amino acid sequence upstream of each CCCH zinc finger within TIS11d forms hydrogen bonds with the AU-rich RNA sequence (108). Nab2 and ZC3H14 do not contain these upstream sequences. Finally, while Nab2 and ZC3H14 encode several aromatic residues shared by TIS11d and important for base stacking interactions (denoted by asterisks and highlighted in purple in Figure 5.1), these poly(A) specific proteins lack two other tyrosines within TIS11d (denoted by asterisks and highlighted in light green

Figure 5.1: An alignment of tandem zinc finger domains reveals conserved aromatic and positively charged residues. Alignment of CCCH zinc fingers from human ZC3H14, *S. cerevisiae* Nab2, and human TIS11d. Figure is adapted from (108). Structural cysteines and histidines are highlighted in yellow, while conserved positively charged residues and aromatic residues are highlighted in red and purple, respectively. Sequences in TIS11d important for AU-rich sequence recognition are highlighted in green, while aromatic residues required for base stacking interactions are denoted by asterisks.

in Figure 5.1) that form specific base stacking interactions with AU-rich RNA sequences (108). Hence, it is only through a collection of interactions that this molecular recognition is achieved. Future co-crystallographic studies between the CCCH zinc finger domain of Nab2 and poly(A) RNA will hopefully provide us with the same level of critical insight into poly(A) RNA recognition by CCCH zinc fingers.

The implications of coupling between mRNA 3'-end processing and nuclear export

Some RNA binding proteins, such as FMRP, the protein implicated in the Fragile X mental retardation syndrome (116), bind to a specific recognition sequence found in a handful of RNA transcripts, while others, like the poly(A) RNA binding proteins PABPN1, ZC3H14, and the putative yeast ZC3H14 orthologue, Nab2, bind an RNA sequence ubiquitous to all mRNA transcripts. Specifically, the poly(A) tail and the proteins bound to it play a significant role in the transport, stability, and translation of most mRNA transcripts (16). Given that the poly(A) tails of mRNA transcripts are involved in so many crucial processes, it is perhaps not too surprising that in yeast, proteins involved in 3'-end formation, such as the cleavage machinery component, Rna15, and the poly(A) polymerase, Pap1, have also been implicated in mRNA export from the nucleus (52-54). Defects in both Rna15 and Pap1 cause poly(A) RNA accumulation within the nucleus (52-54). These findings are confounded even further by the fact that defects in proteins not absolutely linked to either 3'-end formation or mRNA export, such as Nab2, cause both extended poly(A) tails and poly(A) RNA accumulation in the nucleus. While these results emphasize the interconnections between mRNA 3'-end formation and mRNA export in *S. cerevisiae*, uncovering the exact molecular

function of proteins, like Nab2, has been problematic because of the reciprocity of these phenotypes. In other words, due to the fact that defects in processing factors cause defects in export and defects in export factors cause extended poly(A) tails, distinguishing which of these is the primary defect or whether both of these mutant phenotypes are caused by another, as of yet, unidentified defect is extremely difficult.

Distinguishing poly(A) tails from one another

Beyond mRNA transcripts, short oligo(A) tails are also added to other RNA polymerase II transcripts, such as short nucleolar RNAs (snoRNAs) and cryptic unstable transcripts (CUTs), when those transcripts are being processed or targeted for degradation (100, 186, 187). One question that arises from these findings is how these stretches of polyadenosine RNA are distinguished from one another. One possibility is that the RNA binding proteins bound to the poly(A) tail and the immediately surrounding sequences (e.g. the 3'-UTR of mRNA transcripts) distinguish the poly(A) tail of mRNA transcripts from the oligo(A) tail of snoRNA transcripts or CUTs destined for degradation. In chapter 4 we demonstrate that the short oligo(A) tails found on U14 transcripts in exosome deficient cells are at least 17 adenosines long. In addition, we also demonstrate that recombinant Nab2 can bind an RNA oligonucleotide 15 adenosines long *in vitro*, suggesting that the nuclear poly(A) binding protein, Nab2, could bind to these short RNA transcripts. However, the question as to how Nab2 recognizes the poly(A) tail of mRNA transcripts versus the short oligo(A) tail added to the 3'-end of snoRNAs remains unanswered. One possibility is that Nab2 is recruited to transcripts via interactions with other RNA binding proteins or perhaps interactions with RNA polymerase II. Numerous

RNA processing factors are recruited to nascent mRNA and snoRNA transcripts via the C-terminal domain of RNA Polymerase II (234). Although previous experiments investigating the protein-protein interaction partners of Nab2 have not found RNA polymerase II components or snoRNA binding proteins associated with Nab2 (235), these would most likely be extremely transient interactions and most likely not visible using conventional immunoprecipitation techniques. In accordance with this idea, the RNA helicase Dbp5 removes Nab2 from poly(A) RNA *in vitro* and yet no physical interaction between Nab2 and Dbp5 has been demonstrated (236). In addition, Nab2 mutants show extended poly(A) tails *in vivo* and the addition of recombinant Nab2 to *in vitro* polyadenylation reactions limits poly(A) tail length, but no physical interaction has been identified with any of the numerous components of the mRNA 3'-end cleavage and polyadenylation machinery (71, 72). Future work will be needed to investigate these interactions and to investigate whether Nab2 associates with short RNA transcripts destined for degradation. Specifically, *in vivo* protein-protein interaction techniques such as fluorescence resonance energy transfer (FRET) could be used to analyze the spatial and temporal interaction between Nab2 and components of the mRNA transcription and processing machinery. Nab2 has also never been found associated with snoRNA transcripts or CUTs. However, no studies have precisely looked for these interactions. A previous study investigating the specific mRNA transcripts associated with Nab2 used a microarray containing only mRNA open-reading frames. In order to investigate Nab2 association with these short, non-coding RNAs, these experiments will most likely need to be repeated using tiling microarrays covering the entire genome or high throughput "deep-read" or 454 pyro-sequencing. In addition, the nuclear exosome is incredibly

efficient, so if Nab2 does associate with these transcripts destined for degradation these transcripts may not be very abundant. In order to observe these interactions between Nab2 and unstable transcripts, these experiments may need to be done in an exosome deficient background.

A putative cytoplasmic function of Nab2

Beyond the documented nuclear functions of Nab2 in mRNA export from the nucleus and control of poly(A) tail length, indirect evidence also suggests that Nab2 may play a role in the cytoplasm (237). Although Nab2 shuttles from the nucleus to the cytoplasm in a manner dependent upon ongoing RNA polymerase II transcription (73), it is not associated with polyribosomes (87), suggesting that Nab2 must be displaced from RNA transcripts prior to translation. However, a cytoplasmic function for Nab2 has never been identified. As demonstrated by van den Bogaart, *et al*., Kap104, the nuclear import receptor of Nab2, as well as the RNA helicase Dbp5, asymmetrically localizes to the tip and neck of budding daughter yeast cells (237). The authors of this study propose a simple model whereby the asymmetric distribution of both these proteins, but specifically Kap104, allows for enhanced translation in the daughter yeast cell (237). If indeed Nab2 does bind specifically to mRNA transcripts and not the shorter oligo(A) tails of snoRNAs or CUTS, presumably Nab2 would bind the poly(A) tail of mRNA transcripts and Kap104 in a mutually exclusive manner, preventing futile rounds of RNA export. Binding of Kap104 to Nab2, as well as the "remodelase" action of the RNA helicase, Dbp5, may help to facilitate the release of Nab2 from these transcripts in the cytoplasm (122, 236). In addition, an enhanced Kap104-dependent site of translation is

also present in the tip of the budding daughter cell. Together, the asymmetric localization of Kap104 and Dbp5 in the bud tip as well as the localized site of translation in the daughter cell bud tip suggest that Kap104 and Dbp5 may be removing a translational repressor protein from mRNA transcripts. Interestingly, overexpression of the Nab2-RGG-NLS inhibits this localized site of translation, presumably by acting as a dominant negative and sequestering Kap104 (237). These results suggest that Nab2, which is both a Dbp5 substrate and Kap104 cargo, may function as a translational repressor in the cytoplasm by binding to the poly(A) tail of mRNA transcripts. Interestingly, the principle cytoplasmic poly(A) RNA binding protein in *S. cerevisiae*, Pab1, stimulates translation (16). One logical hypothesis, therefore, is that following export from the nucleus, Nab2 remains bound to the poly(A) tail of select mRNA transcripts destined for the daughter cell until it reaches a high localized concentration of Kap104 and Dbp5. Kap104 and Dbp5, facilitate the removal of Nab2 from these transcripts and reimport Nab2 into the (daughter) nucleus. Once Nab2 has been removed, Pab1 can then bind the poly(A) tail and stimulate translation. While this model might seem somewhat far-fetched, Dbp5 has been implicated in the control of translation initiation and termination (236, 238). In addition, an isoform of ZC3H14, which contains the C-terminal zinc finger domain and a unique N-terminus localizes to the cytoplasm and is expressed in a brain and testes specific manner (146). Many neuronal specific RNA transcripts are transported along the length of the neuronal axon in a repressed state until they reach a localized site of translation in the cytoplasm. It is tempting to hypothesize that this particular ZC3H14 isoform, as well as Nab2 perhaps, could function in the translational repression of select RNA transcripts. The targeted knock-down of ZC3H14 isoforms and

analysis of RNA localization and/or translation status in neuronal cells lines may help to answer these questions. In addition, the isolation of Nab2 and its associated transcripts in a cell-cycle dependent manner may also give further insight into whether Nab2 is associated with certain transcripts destined for the growing yeast daughter cell.

Final conclusions and future directions

In sum, these studies have greatly advanced our knowledge of poly(A) RNA recognition. We have shown that CCCH zinc fingers can specifically recognize polyadenosine RNA with high affinity. Furthermore, we have demonstrated that the recognition of poly(A) RNA by CCCH zinc fingers occurs in multiple species, including the budding yeast *S. cerevisiae*, the fruit fly, *D. melanogaster*, and humans. ESTs representing various ZC3H14 isoforms are also present in mice and rats, suggesting that ZC3H14 may play an evolutionarily conserved role in the molecular recognition of polyadenosine RNA. In collaboration with Ken Moberg at Emory University, we have also established the fruit fly, *D. melanogaster*, as a model system in which to study defects in polyadenosine RNA recognition on an organism-wide level. The use of *D. melanogaster* as a model system will also allow us to investigate the potential role of dNab2 (*D. melanogaster* ZC3H14) in development. Furthermore, recent evidence also suggests that mutations within human ZC3H14 lead to a familial form of mental retardation found within consanguineous families (Andreas Kuss and colleagues, personal communication). Interestingly, at least one mutation in these patients is an early stop codon resulting in a protein lacking the CCCH zinc finger domain, suggesting that ZC3H14 may post-transcriptionally regulate the expression of genes involved in proper

neuronal development. It will be interesting to use *D. melanogaster* as a model of this disease state in order to investigate the link between ZC3H14 and neuronal development and the post-transcriptional control of gene expression.

These studies also highlight the intricately connected web of RNA processing and export from the nucleus. Each step during RNA maturation, from initial processing in the nucleus to the final demise in the cytoplasm is controlled by a different set of RNA binding proteins (4). In fact, RNA binding proteins and their associated transcripts have been likened to "transcriptional operons" due to the fact that they can have such a large affect on the final protein product (4, 5). Nab2 and ZC3H14 recognize stretches of polyadenosine, such as the poly(A) tail, that are ubiquitous to all mRNA transcripts. Thus, these poly(A) binding proteins have the potential to post-transcriptionally regulate not just a handful of transcripts, but *every single mRNA transcript that is ever synthesized*. Overall, the results of this work demonstrate that the recognition of polyadenosine RNA is a critical component in the post-transcriptional control of gene expression.

References

1. Beelman CA, Parker R. Degradation of mRNA in eukaryotes. Cell 1995;81(2):179-183.
2. Dimaano C, Ullman KS. Nucleocytoplasmic transport: integrating mRNA production and turnover with export through the nuclear pore. Mol Cell Biol 2004;24(8):3069-3076.

3. Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature 2005;433(7025):477-480.

4. Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science 2005;309(5740):1514-1518.

5. Keene JD, Tenenbaum SA. Eukaryotic mRNPs may represent posttranscriptional operons. Mol Cell 2002;9(6):1161-1167.

6. Corbett A, Ferrigno P, Henry M, Kahana J, Koepp D, Lee M, Nguyen L, Schlenstedt G, Seedorf M, Shen E, Taura T, Wong D, Silver P. Genetic analysis of macromolecular transport across the nuclear envelope. Exp Cell Res 1996;229(2):212-216.

7. Kohler A, Hurt E. Exporting RNA from the nucleus to the cytoplasm. Nat Rev Mol Cell Biol 2007;8(10):761-773.

8. Parenteau J, Durand M, Veronneau S, Lacombe AA, Morin G, Guerin V, Cecez B, Gervais-Bird J, Koh CS, Brunelle D, Wellinger RJ, Chabot B, Abou Elela S. Deletion of many yeast introns reveals a minority of genes that require splicing for function. Mol Biol Cell 2008;19(5):1932-1941.

9. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, *et al*. Initial sequencing and analysis of the human genome. Nature 2001;409(6822):860-921.

10. Lopez PJ, Seraphin B. Genomic-scale quantitative analysis of yeast pre-mRNA splicing: implications for splice-site recognition. RNA 1999;5(9):1135-1137.

11. Juneau K, Palm C, Miranda M, Davis RW. High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. Proc Natl Acad Sci U S A 2007;104(5):1522-1527.

12. Newbury SF. Control of mRNA stability in eukaryotes. Biochem Soc Trans 2006;34(Pt 1):30-34.

13. Meyer S, Temme C, Wahle E. Messenger RNA turnover in eukaryotes: pathways and enzymes. Crit Rev Biochem Mol Biol 2004;39(4):197-216.

14. Pestova TV, Kolupaeva VG, Lomakin IB, Pilipenko EV, Shatsky IN, Agol VI, Hellen CU. Molecular mechanisms of translation initiation in eukaryotes. Proc Natl Acad Sci U S A 2001;98(13):7029-7036.

15. Proudfoot N. New perspectives on connecting messenger RNA 3' end formation to transcription. Curr Opin Cell Biol 2004;16(3):272-278.

16. Mangus DA, Smith MM, McSweeney JM, Jacobson A. Identification of factors regulating poly(A) tail synthesis and maturation. Mol Cell Biol 2004;24(10):4196-4206.

17. Kelly SM, Corbett AH. Messenger RNA Export from the Nucleus: A Series of Molecular Wardrobe Changes. Traffic 2009.

18. Fasken MB, Corbett AH. Mechanisms of nuclear mRNA quality control. RNA Biol 2009;6(3).

19. Fasken MB, Corbett AH. Process or perish: quality control in mRNA biogenesis. Nat Struct Mol Biol 2005;12(6):482-488.

20. Segref A, Sharma K, Doye V, Hellwig A, Huber J, Luhrmann R, Hurt E. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. Embo J 1997;16(11):3256-3271.

21. Katahira J, Strasser K, Podtelejnikov A, Mann M, Jung JU, Hurt E. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. EMBO J 1999;18(9):2593-2609.

22. Kang Y, Bogerd HP, Yang J, Cullen BR. Analysis of the RNA binding specificity of the human tap protein, a constitutive transport element-specific nuclear RNA export factor. Virology 1999;262(1):200-209.

23. Strasser K, Hurt E. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. Embo J 2000;19(3):410-420.

24. Stutz F, Bachi A, Doerks T, Braun IC, Seraphin B, Wilm M, Bork P, Izaurralde E. REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. RNA 2000;6(4):638-650.

25. Chavez S, Beilharz T, Rondon AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T, Aguilera A. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae. EMBO J 2000;19(21):5824-5834.

26. Abruzzi KC, Lacadie S, Rosbash M. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. EMBO J 2004;23(13):2620-2631.

27. Reed R, Cheng H. TREX, SR proteins and export of mRNA. Curr Opin Cell Biol 2005;17(3):269-273.

28. Strasser K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodriguez-Navarro S, Rondon AG, Aguilera A, Struhl K, Reed R, Hurt E. TREX is a conserved complex coupling transcription with messenger RNA export. Nature 2002;417(6886):304-308.

29. Jensen TH, Patricio K, McCarthy T, Rosbash M. A block to mRNA nuclear export in S. cerevisiae leads to hyperadenylation of transcripts that accumulate at the site of transcription. Mol Cell 2001;7(4):887-898.

30. Strasser K, Hurt E. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. Nature 2001;413(6856):648-652.

31. Libri D, Graziani N, Saguez C, Boulay J. Multiple roles for the yeast SUB2/yUAP56 gene in splicing. Genes Dev 2001;15(1):36-41.

32. Portman DS, O'Connor JP, Dreyfuss G. YRA1, an essential Saccharomyces cerevisiae gene, encodes a novel nuclear protein with RNA annealing activity. RNA 1997;3(5):527-537.

33. Zenklusen D, Vinciguerra P, Strahm Y, Stutz F. The yeast hnRNP-Like proteins Yra1p and Yra2p participate in mRNA export through interaction with Mex67p. Mol Cell Biol 2001;21(13):4219-4232.

34. Chavez S, Aguilera A. The yeast HPR1 gene has a functional role in transcriptional elongation that uncovers a novel source of genome instability. Genes Dev 1997;11(24):3459-3470.

35. Rondon AG, Jimeno S, Garcia-Rubio M, Aguilera A. Molecular evidence that the eukaryotic THO/TREX complex is required for efficient transcription elongation. J Biol Chem 2003;278(40):39037-39043.

36. Zenklusen D, Vinciguerra P, Wyss JC, Stutz F. Stable mRNP formation and export require cotranscriptional recruitment of the mRNA export factors Yra1p and Sub2p by Hpr1p. Mol Cell Biol 2002;22(23):8241-8253.

37. Johnson SA, Cubberley G, Bentley DL. Cotranscriptional recruitment of the mRNA export factor Yra1 by direct interaction with the 3' end processing factor Pcf11. Mol Cell 2009;33(2):215-226.

38. Gwizdek C, Iglesias N, Rodriguez MS, Ossareh-Nazari B, Hobeika M, Divita G, Stutz F, Dargemont C. Ubiquitin-associated domain of Mex67 synchronizes recruitment of the mRNA export machinery with transcription. Proc Natl Acad Sci U S A 2006;103(44):16376-16381.

39. Gilbert W, Guthrie C. The Glc7p nuclear phosphatase promotes mRNA export by facilitating association of Mex67p with mRNA. Mol Cell 2004;13(2):201-212.

40. Bucheli ME, Buratowski S. Npl3 is an antagonist of mRNA 3' end formation by RNA polymerase II. EMBO J 2005;24(12):2150-2160.

41. Kress TL, Krogan NJ, Guthrie C. A single SR-like protein, Npl3, promotes pre-mRNA splicing in budding yeast. Mol Cell 2008;32(5):727-734.

42. Lee MS, Henry M, Silver PA. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. Genes Dev 1996;10(10):1233-1246.

43. Bucheli ME, He X, Kaplan CD, Moore CL, Buratowski S. Polyadenylation site choice in yeast is affected by competition between Npl3 and polyadenylation factor CFI. RNA 2007;13(10):1756-1764.

44. Barilla D, Lee BA, Proudfoot NJ. Cleavage/polyadenylation factor IA associates with the carboxyl-terminal domain of RNA polymerase II in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2001;98(2):445-450.

45. Licatalosi DD, Geiger G, Minet M, Schroeder S, Cilli K, McNeil JB, Bentley DL. Functional interaction of yeast pre-mRNA 3' end processing factors with RNA polymerase II. Mol Cell 2002;9(5):1101-1111.

46. Sadowski M, Dichtl B, Hubner W, Keller W. Independent functions of yeast Pcf11p in pre-mRNA 3' end processing and in transcription termination. EMBO J 2003;22(9):2167-2177.

47. Strasser K, Bassler J, Hurt E. Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. J Cell Biol 2000;150(4):695-706.

48. Hieronymus H, Silver PA. Genome-wide analysis of RNA-protein interactions illustrates specificity of the mRNA export machinery. Nat Genet 2003;33(2):155-161.

49. Guisbert KK, Duncan K, Li H, Guthrie C. Functional specificity of shuttling hnRNPs revealed by genome-wide analysis of their RNA binding profiles. RNA-a Publication of the RNA Society 2005;11(4):383-393.

50. Proudfoot N, O'Sullivan J. Polyadenylation: a tail of two complexes. Curr Biol 2002;12(24):R855-857.

51. Coller JM, Gray NK, Wickens MP. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. Genes Dev 1998;12(20):3226-3235.

52. Hammell CM, Gross S, Zenklusen D, Heath CV, Stutz F, Moore C, Cole CN. Coupling of termination, 3' processing, and mRNA export. Mol Cell Biol 2002;22(18):6441-6457.

53. Brodsky AS, Silver PA. Pre-mRNA processing factors are required for nuclear export. Rna 2000;6(12):1737-1749.

54. Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH. Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. Mol Cell Biol 2002;22(23):8254-8266.

55. Hilleren P, Parker R. Defects in the mRNA export factors Rat7p, Gle1p, Mex67p, and Rat8p cause hyperadenylation during 3'-end formation of nascent transcripts. RNA 2001;7(5):753-764.

56. Saguez C, Olesen JR, Jensen TH. Formation of export-competent mRNP: escaping nuclear destruction. Curr Opin Cell Biol 2005;17(3):287-293.

57. Iglesias N, Stutz F. Regulation of mRNP dynamics along the export pathway. FEBS Lett 2008;582(14):1987-1996.

58. Dower K, Kuperwasser N, Merrikh H, Rosbash M. A synthetic A tail rescues yeast nuclear accumulation of a ribozyme-terminated transcript. Rna 2004;10(12):1888-1899.

59. Abruzzi KC, Belostotsky DA, Chekanova JA, Dower K, Rosbash M. 3'-end formation signals modulate the association of genes with the nuclear periphery as well as mRNP dot formation. EMBO J 2006;25(18):4253-4262.

60. Symons RH. Small catalytic RNAs. Annu Rev Biochem 1992;61:641-671.

61. Duvel K, Valerius O, Mangus DA, Jacobson A, Braus GH. Replacement of the yeast TRP4 3' untranslated region by a hammerhead ribozyme results in a stable and efficiently exported mRNA that lacks a poly(A) tail. Rna 2002;8(3):336-344.

62. Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. Genome Biol 2003;4(7):223.

63. Dunn EF, Hammell CM, Hodge CA, Cole CN. Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. Genes Dev 2005;19(1):90-103.

64. Brune C, Munchel SE, Fischer N, Podtelejnikov AV, Weis K. Yeast poly(A)-binding protein Pab1 shuttles between the nucleus and the cytoplasm and functions in mRNA export. Rna 2005;11(4):517-531.

65. Amrani N, Minet M, Le Gouar M, Lacroute F, Wyers F. Yeast Pab1 interacts with Rna15 and participates in the control of the poly(A) tail length in vitro. Mol Cell Biol 1997;17(7):3694-3701.

66. Minvielle-Sebastia L, Preker PJ, Wiederkehr T, Strahm Y, Keller W. The major yeast poly(A)-binding protein is associated with cleavage factor IA and functions in premessenger RNA 3'-end formation. Proc Natl Acad Sci U S A 1997;94(15):7897-7902.

67. Kelly SM, Pabit SA, Kitchen CM, Guo P, Marfatia KA, Murphy TJ, Corbett AH, Berland KM. Recognition of polyadenosine RNA by zinc finger proteins. Proc Natl Acad Sci U S A 2007;104(30):12306-12311.

68. Perreault A, Lemieux C, Bachand F. Regulation of the nuclear poly(A)-binding protein by arginine methylation in fission yeast. J Biol Chem 2007;282(10):7552-7562.

69. Kerwitz Y, Kuhn U, Lilie H, Knoth A, Scheuermann T, Friedrich H, Schwarz E, Wahle E. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. EMBO J 2003;22(14):3705-3714.

70. Anderson JT, Wilson SM, Datar KV, Swanson MS. NAB2: a yeast nuclear polyadenylated RNA-binding protein essential for cell viability. Mol Cell Biol 1993;13(5):2730-2741.

71. Hector RE, Nykamp KR, Dheur S, Anderson JT, Non PJ, Urbinati CR, Wilson SM, Minvielle-Sebastia L, Swanson MS. Dual requirement for yeast hnRNP Nab2p in mRNA poly(A) tail length control and nuclear export. Embo J 2002;21(7):1800-1810.

72. Viphakone N, Voisinet-Hakil F, Minvielle-Sebastia L. Molecular dissection of mRNA poly(A) tail length control in yeast. Nucleic Acids Res 2008;36(7):2418-2433.

73. Marfatia KA, Crafton EB, Green DM, Corbett AH. Domain analysis of the Saccharomyces cerevisiae heterogeneous nuclear ribonucleoprotein, Nab2p.

Dissecting the requirements for Nab2p-facilitated poly(A) RNA export. J Biol Chem 2003;278(9):6731-6740.

74. Fasken MB, Stewart M, Corbett AH. Functional significance of the interaction between the mRNA-binding protein, Nab2, and the nuclear pore-associated protein, Mlp1, in mRNA export. J Biol Chem 2008;283(40):27130-27143.

75. Vinciguerra P, Iglesias N, Camblong J, Zenklusen D, Stutz F. Perinuclear Mlp proteins downregulate gene expression in response to a defect in mRNA export. Embo J 2005;24(4):813-823.

76. Masuda S, Das R, Cheng H, Hurt E, Dorman N, Reed R. Recruitment of the human TREX complex to mRNA during splicing. Genes Dev 2005;19(13):1512-1517.

77. Cheng H, Dufu K, Lee CS, Hsu JL, Dias A, Reed R. Human mRNA export machinery recruited to the 5' end of mRNA. Cell 2006;127(7):1389-1400.

78. Zhou Z, Luo MJ, Straesser K, Katahira J, Hurt E, Reed R. The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature 2000;407(6802):401-405.

79. Jurica MS, Moore MJ. Pre-mRNA splicing: awash in a sea of proteins. Mol Cell 2003;12(1):5-14.

80. Fleckner J, Zhang M, Valcarcel J, Green MR. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. Genes Dev 1997;11(14):1864-1872.

81. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 2000;19(24):6860-6869.

82. Katahira J, Inoue H, Hurt E, Yoneda Y. Adaptor Aly and co-adaptor Thoc5 function in the Tap-p15-mediated nuclear export of HSP70 mRNA. EMBO J 2009.

83. Nojima T, Hirose T, Kimura H, Hagiwara M. The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J Biol Chem 2007;282(21):15645-15651.

84. Terry LJ, Wente SR. Nuclear mRNA export requires specific FG nucleoporins for translocation through the nuclear pore complex. J Cell Biol 2007;178(7):1121-1132.

85. Weis K. The nuclear pore complex: oily spaghetti or gummy bear? Cell 2007;130(3):405-407.

86. Strawn LA, Shen T, Wente SR. The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex. J Biol Chem 2001;276(9):6445-6452.

87. Windgassen M, Sturm D, Cajigas IJ, Gonzalez CI, Seedorf M, Bastians H, Krebber H. Yeast shuttling SR proteins Npl3p, Gbp2p, and Hrb1p are part of the translating mRNPs, and Npl3p can function as a translational repressor. Mol Cell Biol 2004;24(23):10479-10491.

88. Snay-Hodge CA, Colot HV, Goldstein AL, Cole CN. Dbp5p/Rat8p is a yeast nuclear pore-associated DEAD-box protein essential for RNA export. EMBO J 1998;17(9):2663-2676.

89. Schmitt C, von Kobbe C, Bachi A, Pante N, Rodrigues JP, Boscheron C, Rigaut G, Wilm M, Seraphin B, Carmo-Fonseca M, Izaurralde E. Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. EMBO J 1999;18(15):4332-4347.

90. Tanner NK, Linder P. DExD/H box RNA helicases: from generic motors to specific dissociation functions. Mol Cell 2001;8(2):251-262.

91. Tran EJ, Zhou Y, Corbett AH, Wente SR. The DEAD-box protein Dbp5 controls mRNA export by triggering specific RNA:protein remodeling events. Mol Cell 2007;28(5):850-859.

92. Lund MK, Guthrie C. The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. Mol Cell 2005;20(4):645-651.

93. Vanacova S, Stefl R. The exosome and RNA quality control in the nucleus. EMBO Rep 2007;8(7):651-657.

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

95. Schneider C, Anderson JT, Tollervey D. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. Mol Cell 2007;27(2):324-331.

96. Dziembowski A, Lorentzen E, Conti E, Seraphin B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat Struct Mol Biol 2007;14(1):15-22.

97. LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell 2005;121(5):713-724.

98. Allmang C, Kufel J, Chanfreau G, Mitchell P, Petfalski E, Tollervey D. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J 1999;18(19):5399-5410.

99. 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 2000;20(2):441-452.

100. Grzechnik P, Kufel J. Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. Mol Cell 2008;32(2):247-258.

101. Messias AC, Sattler M. Structural basis of single-stranded RNA recognition. Acc Chem Res 2004;37(5):279-287.

102. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol 2008;18(3):290-298.

103. Kuhn U, Wahle E. Structure and function of poly(A) binding proteins. Biochim Biophys Acta 2004;1678(2-3):67-84.

104. Deo RC, Bonanno JB, Sonenberg N, Burley SK. Recognition of polyadenylate RNA by the poly(A)-binding protein. Cell 1999;98(6):835-845.

105. Kanaar R, Lee AL, Rudner DZ, Wemmer DE, Rio DC. Interaction of the sex-lethal RNA binding domains with RNA. EMBO J 1995;14(18):4530-4539.

106. Hall TM. Multiple modes of RNA recognition by zinc finger proteins. Curr Opin Struct Biol 2005;15(3):367-373.

107. Brown RS. Zinc finger proteins: getting a grip on RNA. Curr Opin Struct Biol 2005;15(1):94-98.

108. Hudson BP, Martinez-Yamout MA, Dyson HJ, Wright PE. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. Nature Structural & Molecular Biology 2004;11(3):257-264.

109. Lu D, Searles MA, Klug A. Crystal structure of a zinc-finger-RNA complex reveals two modes of molecular recognition. Nature 2003;426(6962):96-100.

110. Lukong KE, Chang KW, Khandjian EW, Richard S. RNA-binding proteins in human genetic disease. Trends Genet 2008;24(8):416-425.

111. Chenard CA, Richard S. New implications for the QUAKING RNA binding protein in human disease. J Neurosci Res 2008;86(2):233-242.

112. Galarneau A, Richard S. Target RNA motif and target mRNAs of the Quaking STAR protein. Nat Struct Mol Biol 2005;12(8):691-698.

113. Larocque D, Richard S. QUAKING KH domain proteins as regulators of glial cell fate and myelination. RNA Biol 2005;2(2):37-40.

114. Bolstad ES, Anderson AC. In pursuit of virtual lead optimization: the role of the receptor structure and ensembles in accurate docking. Proteins 2008;73(3):566-580.

115. Mura C. The average3d PyMOL module. 2005.

116. Garber KB, Visootsak J, Warren ST. Fragile X syndrome. Eur J Hum Genet 2008;16(6):666-672.

117. Laggerbauer B, Ostareck D, Keidel EM, Ostareck-Lederer A, Fischer U. Evidence that fragile X mental retardation protein is a negative regulator of translation. Hum Mol Genet 2001;10(4):329-338.

118. Li Z, Zhang Y, Ku L, Wilkinson KD, Warren ST, Feng Y. The fragile X mental retardation protein inhibits translation via interacting with mRNA. Nucleic Acids Res 2001;29(11):2276-2283.

119. Mazroui R, Huot ME, Tremblay S, Filion C, Labelle Y, Khandjian EW. Trapping of messenger RNA by Fragile X Mental Retardation protein into cytoplasmic granules induces translation repression. Hum Mol Genet 2002;11(24):3007-3017.

120. Goh KJ, Wong KT, Nishino I, Minami N, Nonaka I. Oculopharyngeal muscular dystrophy with PABPN1 mutation in a Chinese Malaysian woman. Neuromuscul Disord 2005;15(3):262-264.

121. Brais B, Bouchard JP, Xie YG, Rochefort DL, Chretien N, Tome FM, Lafreniere RG, Rommens JM, Uyama E, Nohira O, Blumen S, Korczyn AD, Heutink P, Mathieu J, Duranceau A, *et al*. Short GCG expansions in the PABP2 gene cause oculopharyngeal muscular dystrophy. Nat Genet 1998;18(2):164-167.

122. Lee DC, Aitchison JD. Kap104p-mediated nuclear import. Nuclear localization signals in mRNA-binding proteins and the role of Ran and Rna. J Biol Chem 1999;274(41):29031-29037.

123. Grant RP, Marshall NJ, Yang JC, Fasken MB, Kelly SM, Harreman MT, Neuhaus D, Corbett AH, Stewart M. Structure of the N-terminal Mlp1-binding domain of

the Saccharomyces cerevisiae mRNA-binding protein, Nab2. J Mol Biol 2008;376(4):1048-1059.

124. Maniatis T, Reed R. An extensive network of coupling among gene expression machines. Nature 2002;416(6880):499-506.

125. Gorgoni B, Gray NK. The roles of cytoplasmic poly(A)-binding proteins in regulating gene expression: a developmental perspective. Brief Funct Genomic Proteomic 2004;3(2):125-141.

126. Sachs AB, Davis RW. The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation. Cell 1989;58(5):857-867.

127. Tarun SZ, Jr., Sachs AB. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. Embo J 1996;15(24):7168-7177.

128. Le H, Tanguay RL, Balasta ML, Wei CC, Browning KS, Metz AM, Goss DJ, Gallie DR. Translation initiation factors eIF-iso4G and eIF-4B interact with the poly(A)-binding protein and increase its RNA binding activity. J Biol Chem 1997;272(26):16247-16255.

129. Gallie DR. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. Gene 1998;216(1):1-11.

130. Kahvejian A, Svitkin YV, Sukarieh R, M'Boutchou MN, Sonenberg N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. Genes Dev 2005;19(1):104-113.

131. Caponigro G, Parker R. Multiple functions for the poly(A)-binding protein in mRNA decapping and deadenylation in yeast. Genes Dev 1995;9(19):2421-2432.

132. Caponigro G, Parker R. Mechanisms and control of mRNA turnover in Saccharomyces cerevisiae. Microbiol Rev 1996;60(1):233-249.

133. Adam SA, Nakagawa T, Swanson MS, Woodruff TK, Dreyfuss G. mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. Mol Cell Biol 1986;6(8):2932-2943.

134. Sachs AB, Davis RW, Kornberg RD. A single domain of yeast poly(A)-binding protein is necessary and sufficient for RNA binding and cell viability. Mol Cell Biol 1987;7(9):3268-3276.

135. Deardorff JA, Sachs AB. Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast poly(A)-binding protein. J Mol Biol 1997;269(1):67-81.

136. Maris C, Dominguez C, Allain FH. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. Febs J 2005;272(9):2118-2131.

137. Hess ST, Huang SH, Heikal AA, Webb WW. Biological and chemical applications of fluorescence correlation spectroscopy: A review. Biochemistry 2002;41(3):697-705.

138. Maiti S, Haupts U, Webb WW. Fluorescence correlation spectroscopy: Diagnostics for sparse molecules. Proceedings Of The National Academy Of Sciences Of The United States Of America 1997;94(22):11753-11757.

139. Eigen M, Rigler R. Sorting Single Molecules - Application to Diagnostics and Evolutionary Biotechnology. Proceedings of the National Academy of Sciences of the United States of America 1994;91(13):5740-5747.

140. Wohland T, Friedrich K, Hovius R, Vogel H. Study of ligand-receptor interactions by fluorescence correlation spectroscopy with different fluorophores: Evidence that the homopentameric 5-hydroxytryptamine type 3(As) receptor binds only one ligand. Biochemistry 1999;38(27):8671-8681.

141. Burd CG, Dreyfuss G. Conserved structures and diversity of functions of RNA-binding proteins. Science 1994;265(5172):615-621.

142. Carrick DM, Lai WS, Blackshear PJ. The tandem CCCH zinc finger protein tristetraprolin and its relevance to cytokine mRNA turnover and arthritis. Arthritis Res Ther 2004;6(6):248-264.

143. Scanlan MJ, Gordan JD, Williamson B, Stockert E, Bander NH, Jongeneel V, Gure AO, Jager D, Jager E, Knuth A, Chen YT, Old LJ. Antigens recognized by autologous antibody in patients with renal-cell carcinoma. Int J Cancer 1999;83(4):456-464.

144. Suntharalingam M, Alcazar-Roman AR, Wente SR. Nuclear export of the yeast mRNA-binding protein Nab2 is linked to a direct interaction with Gfd1 and to Gle1 function. J Biol Chem 2004;279(34):35384-35391.

145. Muzzopappa M, Wappner P. Multiple roles of the F-box protein Slimb in Drosophila egg chamber development. Development 2005;132(11):2561-2571.

146. Leung SW, Apponi LH, Cornejo OE, Kitchen CM, Valentini SR, Pavlath GK, Dunham CM, Corbett AH. Splice variants of the human ZC3H14 gene generate

multiple isoforms of a zinc finger polyadenosine RNA binding protein. Gene 2009.

147. Stutz F, Izaurralde E. The interplay of nuclear mRNP assembly, mRNA surveillance and export. Trends Cell Biol 2003;13(6):319-327.

148. Edmonds M. A history of poly A sequences: from formation to factors to function. Prog Nucleic Acid Res Mol Biol 2002;71:285-389.

149. Andreini C, Banci L, Bertini I, Rosato A. Counting the zinc-proteins encoded in the human genome. J Proteome Res 2006;5(1):196-201.

150. Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. Second ed. Plainview, NY: Cold Spring Harbor Laboratory Press; 1989.

151. Adams A, Gottschling DE, Kaiser CA, Stearns T. Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual. 1997 Edition ed. Plainview, NY: Cold Spring Harbor Laboratory Press; 1997.

152. Way M, Pope B, Gooch J, Hawkins M, Weeds AG. Identification of a region in segment 1 of gelsolin critical for actin binding. Embo J 1990;9(12):4103-4109.

153. Boeke JD, Trueheart J, Natsoulis G, Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol 1987;154:164-175.

154. Berland KM, So PTC, Gratton E. 2-Photon Fluorescence Correlation Spectroscopy - Method And Application To The Intracellular Environment. Biophysical Journal 1995;68(2):694-701.

155. Berland KM. Detection of specific DNA sequences using dual-color two-photon fluorescence correlation spectroscopy. Journal Of Biotechnology 2004;108(2):127-136.

156. Van Craenenbroeck E, Engelborghs Y. Fluorescence correlation spectroscopy: molecular recognition at the single molecule level. Journal of Molecular Recognition 2000;13(2):93-100.

157. Magde D, Elson EL, Webb WW. Fluorescence Correlation Spectroscopy .2. Experimental Realization. Biopolymers 1974;13(1):29-61.

158. Meseth U, Wohland T, Rigler R, Vogel H. Resolution of fluorescence correlation measurements. Biophysical Journal 1999;76(3):1619-1631.

159. Berland KM. Fluorescence Correlation Spectroscopy: A New Tool for Quantification of Molecular Interactions. In: Fu H, editor. Methods in Molecular Biology: Protein-Protein Interactions: Methods and Applications. Totowa, New Jersey: Humana Press; 2004. p. 383-397.

160. Tuk B, vanOostenbruggen MF. Solving inconsistencies in the analysis of receptor-ligand interactions. Trends in Pharmacological Sciences 1996;17(11):403-409.

161. Kenakin TP. Pharmacologic Analysis of Drug/Receptor Interaction. 3rd ed. Philadelphia: Lippincott-Raven; 1997.

162. Pandit S, Wang D, Fu XD. Functional integration of transcriptional and RNA processing machineries. Curr Opin Cell Biol 2008;20(3):260-265.

163. Aranda A, Proudfoot N. Transcriptional termination factors for RNA polymerase II in yeast. Mol Cell 2001;7(5):1003-1011.

164. Gross S, Moore C. Five subunits are required for reconstitution of the cleavage and polyadenylation activities of Saccharomyces cerevisiae cleavage factor I. Proc Natl Acad Sci U S A 2001;98(11):6080-6085.

165. Misteli T, Spector DL. RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol Cell 1999;3(6):697-705.

166. Stewart M. Ratcheting mRNA out of the nucleus. Mol Cell 2007;25(3):327-330.

167. Strahm Y, Fahrenkrog B, Zenklusen D, Rychner E, Kantor J, Rosbach M, Stutz F. The RNA export factor Gle1p is located on the cytoplasmic fibrils of the NPC and physically interacts with the FG-nucleoporin Rip1p, the DEAD-box protein Rat8p/Dbp5p and a new protein Ymr 255p. EMBO J 1999;18(20):5761-5777.

168. Dheur S, Nykamp KR, Viphakone N, Swanson MS, Minvielle-Sebastia L. Yeast mRNA Poly(A) tail length control can be reconstituted in vitro in the absence of Pab1p-dependent Poly(A) nuclease activity. J Biol Chem 2005;280(26):24532-24538.

169. Van Driessche B, Tafforeau L, Hentges P, Carr AM, Vandenhaute J. Additional vectors for PCR-based gene tagging in Saccharomyces cerevisiae and Schizosaccharomyces pombe using nourseothricin resistance. Yeast 2005;22(13):1061-1068.

170. Hentges P, Van Driessche B, Tafforeau L, Vandenhaute J, Carr AM. Three novel antibiotic marker cassettes for gene disruption and marker switching in Schizosaccharomyces pombe. Yeast 2005;22(13):1013-1019.

171. Chekanova JA, Belostotsky DA. Evidence that poly(A) binding protein has an evolutionarily conserved function in facilitating mRNA biogenesis and export. Rna 2003;9(12):1476-1490.

172. Lunde BM, Moore C, Varani G. RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol 2007;8(6):479-490.

173. Farny NG, Hurt JA, Silver PA. Definition of global and transcript-specific mRNA export pathways in metazoans. Genes Dev 2008;22(1):66-78.

174. Handwerger KE, Gall JG. Subnuclear organelles: new insights into form and function. Trends Cell Biol 2006;16(1):19-26.

175. Wach A, Brachat A, Pohlmann R, Philippsen P. New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 1994;10(13):1793-1808.

176. Pagano JM, Farley BM, McCoig LM, Ryder SP. Molecular basis of RNA recognition by the embryonic polarity determinant MEX-5. J Biol Chem 2007;282(12):8883-8894.

177. Minvielle-Sebastia L, Winsor B, Bonneaud N, Lacroute F. Mutations in the yeast RNA14 and RNA15 genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. Mol Cell Biol 1991;11(6):3075-3087.

178. Wong DH, Corbett AH, Kent HM, Stewart M, Silver PA. Interaction between the small GTPase Ran/Gsp1p and Ntf2p is required for nuclear transport. Mol Cell Biol 1997;17(7):3755-3767.

179. Legge GB, Martinez-Yamout MA, Hambly DM, Trinh T, Lee BM, Dyson HJ, Wright PE. ZZ domain of CBP: an unusual zinc finger fold in a protein interaction module. J Mol Biol 2004;343(4):1081-1093.

180. Filipowicz W, Pogacic V. Biogenesis of small nucleolar ribonucleoproteins. Curr Opin Cell Biol 2002;14(3):319-327.

181. Davis CA, Ares M, Jr. Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2006;103(9):3262-3267.

182. Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Regnault B, Devaux F, Namane A, Seraphin B, Libri D, Jacquier A. Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. Cell 2005;121(5):725-737.

183. Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell 2002;109(2):145-148.

184. Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 2009;457(7232):1038-1042.

185. Camblong J, Iglesias N, Fickentscher C, Dieppois G, Stutz F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in S. cerevisiae. Cell 2007;131(4):706-717.

186. Schmid M, Jensen TH. The exosome: a multipurpose RNA-decay machine. Trends Biochem Sci 2008;33(10):501-510.

187. Butler JS. The yin and yang of the exosome. Trends Cell Biol 2002;12(2):90-96.

188. Kim M, Vasiljeva L, Rando OJ, Zhelkovsky A, Moore C, Buratowski S. Distinct pathways for snoRNA and mRNA termination. Mol Cell 2006;24(5):723-734.

189. Arigo JT, Eyler DE, Carroll KL, Corden JL. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. Mol Cell 2006;23(6):841-851.

190. Vasiljeva L, Buratowski S. Nrd1 interacts with the nuclear exosome for 3' processing of RNA polymerase II transcripts. Mol Cell 2006;21(2):239-248.

191. Steinmetz EJ, Conrad NK, Brow DA, Corden JL. RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts. Nature 2001;413(6853):327-331.

192. Conrad NK, Wilson SM, Steinmetz EJ, Patturajan M, Brow DA, Swanson MS, Corden JL. A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II. Genetics 2000;154(2):557-571.

193. Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A. The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. Nat Struct Mol Biol 2008;15(8):795-804.

194. Carroll KL, Ghirlando R, Ames JM, Corden JL. Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements. RNA 2007;13(3):361-373.

195. Carroll KL, Pradhan DA, Granek JA, Clarke ND, Corden JL. Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts. Mol Cell Biol 2004;24(14):6241-6252.

196. Steinmetz EJ, Warren CL, Kuehner JN, Panbehi B, Ansari AZ, Brow DA. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. Mol Cell 2006;24(5):735-746.

197. Rasmussen TP, Culbertson MR. The putative nucleic acid helicase Sen1p is required for formation and stability of termini and for maximal rates of synthesis

and levels of accumulation of small nucleolar RNAs in Saccharomyces cerevisiae. Mol Cell Biol 1998;18(12):6885-6896.

198. Ursic D, Himmel KL, Gurley KA, Webb F, Culbertson MR. The yeast SEN1 gene is required for the processing of diverse RNA classes. Nucleic Acids Res 1997;25(23):4778-4785.

199. Buratowski S, Moazed D. Gene regulation: expression and silencing coupled. Nature 2005;435(7046):1174-1175.

200. Rondon AG, Mischo HE, Proudfoot NJ. Terminating transcription in yeast: whether to be a 'nerd' or a 'rat'. Nat Struct Mol Biol 2008;15(8):775-776.

201. Rondon AG, Proudfoot NJ. Nuclear roadblocks for mRNA export. Cell 2008;135(2):207-208.

202. Buttner K, Wenig K, Hopfner KP. The exosome: a macromolecular cage for controlled RNA degradation. Mol Microbiol 2006;61(6):1372-1379.

203. Burkard KT, Butler JS. A nuclear 3'-5' exonuclease involved in mRNA degradation interacts with Poly(A) polymerase and the hnRNA protein Npl3p. Mol Cell Biol 2000;20(2):604-616.

204. Das B, Das S, Sherman F. Mutant LYS2 mRNAs retained and degraded in the nucleus of Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2006;103(29):10871-10876.

205. Kuai L, Das B, Sherman F. A nuclear degradation pathway controls the abundance of normal mRNAs in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 2005;102(39):13962-13967.

206. Das B, Butler JS, Sherman F. Degradation of normal mRNA in the nucleus of Saccharomyces cerevisiae. Mol Cell Biol 2003;23(16):5502-5515.

207. Silva AL, Romao L. The mammalian nonsense-mediated mRNA decay pathway: to decay or not to decay! Which players make the decision? FEBS Lett 2009;583(3):499-505.

208. Isken O, Maquat LE. Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev 2007;21(15):1833-1856.

209. Muhlrad D, Parker R. Aberrant mRNAs with extended 3' UTRs are substrates for rapid degradation by mRNA surveillance. RNA 1999;5(10):1299-1307.

210. Muhlrad D, Decker CJ, Parker R. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5'-->3' digestion of the transcript. Genes Dev 1994;8(7):855-866.

211. Carneiro T, Carvalho C, Braga J, Rino J, Milligan L, Tollervey D, Carmo-Fonseca M. Inactivation of cleavage factor I components Rna14p and Rna15p induces sequestration of small nucleolar ribonucleoproteins at discrete sites in the nucleus. Mol Biol Cell 2008;19(4):1499-1508.

212. San Paolo S, Vanacova S, Schenk L, Scherrer T, Blank D, Keller W, Gerber AP. Distinct roles of non-canonical poly(A) polymerases in RNA metabolism. PLoS Genet 2009;5(7):e1000555.

213. Egecioglu DE, Henras AK, Chanfreau GF. Contributions of Trf4p- and Trf5p-dependent polyadenylation to the processing and degradative functions of the yeast nuclear exosome. RNA 2006;12(1):26-32.

214. Haracska L, Johnson RE, Prakash L, Prakash S. Trf4 and Trf5 proteins of Saccharomyces cerevisiae exhibit poly(A) RNA polymerase activity but no DNA polymerase activity. Mol Cell Biol 2005;25(22):10183-10189.

215. Vanacova S, Wolf J, Martin G, Blank D, Dettwiler S, Friedlein A, Langen H, Keith G, Keller W. A new yeast poly(A) polymerase complex involved in RNA quality control. PLoS Biol 2005;3(6):e189.

216. Bernstein J, Patterson DN, Wilson GM, Toth EA. Characterization of the essential activities of Saccharomyces cerevisiae Mtr4p, a 3'->5' helicase partner of the nuclear exosome. J Biol Chem 2008;283(8):4930-4942.

217. Wahle E. Poly(A) tail length control is caused by termination of processive synthesis. J Biol Chem 1995;270(6):2800-2808.

218. Bienroth S, Keller W, Wahle E. Assembly of a processive messenger RNA polyadenylation complex. EMBO J 1993;12(2):585-594.

219. Krogan NJ, Peng WT, Cagney G, Robinson MD, Haw R, Zhong G, Guo X, Zhang X, Canadien V, Richards DP, Beattie BK, Lalev A, Zhang W, Davierwala AP, Mnaimneh S, *et al*. High-definition macromolecular composition of yeast RNA-processing complexes. Mol Cell 2004;13(2):225-239.

220. Peng WT, Robinson MD, Mnaimneh S, Krogan NJ, Cagney G, Morris Q, Davierwala AP, Grigull J, Yang X, Zhang W, Mitsakakis N, Ryan OW, Datta N, Jojic V, Pal C, *et al*. A panoramic view of yeast noncoding RNA processing. Cell 2003;113(7):919-933.

221. Patel D, Butler JS. Conditional defect in mRNA 3' end processing caused by a mutation in the gene for poly(A) polymerase. Mol Cell Biol 1992;12(7):3297-3304.

222. Phillips S, Butler JS. Contribution of domain structure to the RNA 3' end processing and degradation functions of the nuclear exosome subunit Rrp6p. RNA 2003;9(9):1098-1107.

223. Carneiro T, Carvalho C, Braga J, Rino J, Milligan L, Tollervey D, Carmo-Fonseca M. Depletion of the yeast nuclear exosome subunit Rrp6 results in accumulation of polyadenylated RNAs in a discrete domain within the nucleolus. Mol Cell Biol 2007;27(11):4157-4165.

224. Li HD, Zagorski J, Fournier MJ. Depletion of U14 small nuclear RNA (snR128) disrupts production of 18S rRNA in Saccharomyces cerevisiae. Mol Cell Biol 1990;10(3):1145-1152.

225. Zagorski J, Tollervey D, Fournier MJ. Characterization of an SNR gene locus in Saccharomyces cerevisiae that specifies both dispensible and essential small nuclear RNAs. Mol Cell Biol 1988;8(8):3282-3290.

226. Couttet P, Fromont-Racine M, Steel D, Pictet R, Grange T. Messenger RNA deadenylylation precedes decapping in mammalian cells. Proc Natl Acad Sci U S A 1997;94(11):5628-5633.

227. Thiebaut M, Kisseleva-Romanova E, Rougemaille M, Boulay J, Libri D. Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance. Mol Cell 2006;23(6):853-864.

228. Fatica A, Morlando M, Bozzoni I. Yeast snoRNA accumulation relies on a cleavage-dependent/polyadenylation-independent 3'-processing apparatus. EMBO J 2000;19(22):6218-6229.

229. Morlando M, Greco P, Dichtl B, Fatica A, Keller W, Bozzoni I. Functional analysis of yeast snoRNA and snRNA 3'-end formation mediated by uncoupling of cleavage and polyadenylation. Mol Cell Biol 2002;22(5):1379-1389.

230. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, *et al*. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science 1999;285(5429):901-906.

231. Taura T, Krebber H, Silver PA. A member of the Ran-binding protein family, Yrb2p, is involved in nuclear protein export. Proc Natl Acad Sci U S A 1998;95(13):7427-7432.

232. Kitchen CM, Leung SW, Corbett AH, Murphy TJ. The mating response cascade does not modulate changes in the steady-state level of target mRNAs through control of mRNA stability. Yeast 2009;26(5):261-272.

233. Beadle GW, Tatum EL. Genetic Control of Biochemical Reactions in Neurospora. Proc Natl Acad Sci U S A 1941;27(11):499-506.

234. Egloff S, Murphy S. Cracking the RNA polymerase II CTD code. Trends Genet 2008;24(6):280-288.

235. Nykamp KR. Coordination of mRNA 3'-End Formation and Nuclear Export by a Nuclear Poly(A) Binding Protein. 2003.

236. Bolger TA, Folkmann AW, Tran EJ, Wente SR. The mRNA export factor Gle1 and inositol hexakisphosphate regulate distinct stages of translation. Cell 2008;134(4):624-633.

237. van den Bogaart G, Meinema AC, Krasnikov V, Veenhoff LM, Poolman B. Nuclear transport factor directs localization of protein synthesis during mitosis. Nat Cell Biol 2009;11(3):350-356.

238. Gross T, Siepmann A, Sturm D, Windgassen M, Scarcelli JJ, Seedorf M, Cole CN, Krebber H. The DEAD-box RNA helicase Dbp5 functions in translation termination. Science 2007;315(5812):646-649.