THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

A BIOCHEMICAL ANALYSIS OF THE RNA BINDING ACTIVITY OF Not4 OF THE CCR4-NOT COMPLEX IN YEAST

BONNIE BOSHEN LU
Spring 2011

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degrees
in Biochemistry and Molecular Biology, Immunology and Infectious Disease, and Toxicology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Joseph C. Reese
Professor of Biochemistry and Molecular Biology
Thesis Supervisor

David S. Gilmour
Professor of Molecular and Cell Biology
Honors Adviser

Wendy Hanna-Rose
Associate Department Head for Undergraduate Studies
Department of Biochemistry and Molecular Biology

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Ccr4-Not is a multi-functional, multi-protein complex consisting of nine subunits (Ccr4p, Caf1p, Caf40p, Caf130p, and Not1-5p) and plays a role in nuclear RNA polymerase II transcription and cytoplasmic mRNA degradation. Not4, one of the subunits of the complex, is an E3 ubiquitin ligase with an RNA recognition motif (RRM). The RRM is a common protein domain in eukaryotes and is used for the binding of a variety of RNA sequences and proteins. However, it has not yet been shown whether Not4 binds RNA or whether the RRM of Not4 binds RNA.

In this research project, we cloned the full coding region of *NOT4* and the *NOT4* RRM coding region into expression vectors to over-express the proteins in *E. coli*. The expressed proteins were purified and used in *in vitro* RNA binding assays to determine RNA binding activity. The *in vitro* RNA binding assays used were the fluorescence polarization (FP) assay and electrophoretic mobility shift assay (EMSA). The Not4 full-length protein binding assays gave variable dissociation constants. The Not4 RRM binding assays did not show RNA binding under the conditions examined. These results show that Not4 binds RNA but does not eliminate the possibility that the RRM contributes to this function.

We propose that the variable RNA binding activity seen with the Not4 full-length protein may be due to either experimental conditions that interfere with protein folding or

protein–RNA binding, and that the lack of RNA binding activity seen with the Not4 RRM may be due to additional protein domain requirements for RNA binding.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

To Dr. Reese and Dr. Dutta, to all members of the Reese Lab and the B M B administrative staff, and to all who have aided and approved of me in the completion this project, Thank you.

CHAPTER 1

INTRODUCTION

1.1 Transcription cycle

Gene expression in the cell is achieved through intermediary messenger RNAs (mRNAs) that are transcribed from information contained in DNA by DNA-dependent RNA polymerases (RNAP). RNAP acts in a 5' – 3' direction in a transcription cycle divided into the initiation, elongation and termination phase (Figure 1.1).

1.1.1 Transcription pre-initiation and initiation. The initiation of transcription is marked by the recruitment of general transcription factors to the promoter to form the preinitiation complex (PIC). These transcription factors include IIA, IIB, IID, IIE, IIF, and IIH (Figure 1.2). The TATA binding protein (TBP) subunit of TFIID binds the promoter, causing the DNA to associate with proper positioning to RNAPII and TFIID. TFIID functions as the scaffold for the assembly of the remainder of the transcription complex. TFIIE then associates with the growing complex and recruits TFIIH. The ATPase and helicase activity of TFIIH creates the transcription bubble by generating negative superhelical tension in the DNA. Once the transcription bubble is formed, TFIIF binds the coding strand to maintain the open bubble conformation. The non-coding strand enters the RNAPII active site and transcription proceeds in the 5' – 3' direction.

1.1.2 Promoter clearance and transcription elongation. Once transcription progresses past six bases, TFIIB leaves the complex. After about 23 nucleotides, the transcription complex is stabilized, RNAPII clears the promoter, and transcription elongation begins. RNAPII then proceeds along the noncoding DNA strand and uses base pairing complementarity with the DNA template for mRNA synthesis.

Chromatin is in constant flux with nucleosomes and DNA-binding proteins being remodeled, removed or reassembled. DNA nucleosome structures are affected at a distance during transcription elongation as RNAPII creates a positive superhelical stress while translocating along the DNA template. This positive torque generates a wave downstream of the RNAPII, destabilizing nucleosome structures with a domino-like effect and allowing RNAPII to overcome the nucleosomal barrier (Zlatanova and Victor, 2009). When RNAPII encounters barriers such as DNA lesions, DNA-bound proteins, and mismatched nucleotides during transcription elongation, the polymerase will pause, arrest or terminate the transcription process. This disruption allows RNAPII to clear obstacles such as DNA lesions and DNA-bound proteins and for the inherent proofreading mechanisms of RNAPII to safeguard against mismatching nucleotides. RNAPII continues to transcribe until a termination signal is encountered on the noncoding DNA strand.

1.1.3 Transcription termination and pre-mRNA processing. Once transcription ends, RNAPII releases the pre-mRNA transcript and disassociates from the DNA template. The

eukaryotic pre-mRNA undergoes extensive processing (Figure 1.3). A 5' cap is added co-transcriptionally by the Capping Enzyme Complex (CEC), which associates with RNAPII prior to transcription initiation. A poly(A) tail is added to the 3' end immediately following transcription termination by polyadenylate polymerase and runs approximately 250 adenosine residues in length. Pre-mRNA splicing removes introns, and mRNA editing may further change the nucleotide composition of the final mature mRNA product. Mature mRNAs are recognized by their post-transcriptional modifications and transported from the nucleus to the cytoplasm through nuclear pores for translation.

1.2 mRNA decay

The processes of translation and mRNA decay are carefully balanced in the cell to regulate protein expression. Cytoplasmic P-bodies play a role in this balance by localizing mRNA decay proteins involved in mRNA turnover and by safekeeping mRNAs for later translation (Eulalio et al., 2007; Parker and Sheth, 2007). Prior to mRNA decay, eukaryotic initiation factors elF-4E and elF-4G stabilizes the mRNA by blocking the DCP2 decapping enzyme, and the poly(A) tail is stabilized by the poly(A)-binding protein (PABP), which blocks the exosome complex (Gingras et al., 1999; Sachs et al., 1997). The most prevalent pathway for mRNA decay in eukaryotic cells is the deadenylation-dependent pathway, which begins with the deadenylation of the poly(A) tail by specialized 3' – 5' exonucleases (Figure 1.4). In *Saccharomyces cerevisiae*, the poly(A) tail is deadenylated 60 – 80 nucleotides by Pan2-Pan3, a PABP-dependent poly(A) nuclease. The Ccr4-Not complex then completes the deadenylation of the poly(A) tail through the PABP-independent Ccr4 and Pop2 polyA nucleases (Goldstrohm

and Wickens, 2008). Removal of the poly(A) tail destabilizes the elF-4E and elF-4G cap-binding complex, leading to the degradation of the 5' cap by the decapping complex. The remaining transcript is then fully degraded by either the PM/Scl exosome complex or the decapping complex.

1.3 CCR4-Not complex

Ccr4-Not was first identified in *Saccharomyces cerevisiae* as a multi-functional, multi-protein global gene regulation complex consisting of nine subunits (Ccr4p, Caf1p, Caf40p, Caf130p, and Not1p to 5p) and plays a role in nuclear RNA polymerase II transcription and cytoplasmic mRNA degradation (Figure 1.5) (Chen et al., 2001; Collart, 2003; Collart and Timmers, 2004; Denis and Chen, 2003). Not1 is the scaffold protein for the complex and is the only subunit essential for complex formation (Bai et al., 1999; Maillet et al., 2000). The Ccr4 (carbon catabolite repressor) group of proteins interacts with the N terminus of Not1 while the Not group of proteins interacts with C terminus. The Ccr4, Caf1 (Ccr4-associated factor 1, also known as Pop2) and the Not proteins 1, 2, 4 and 5 all have identified roles in regulating transcription (Liu et al., 1998). Ccr4 and Pop2 also have identified roles in the regulation of mRNA decay (Collart and Timmers, 2004; Maillet and Collart, 2002; Tucker et al., 2001).

1.4 Not4 protein subunit

The Not4 protein (also called Mot2) is a 587 amino acid E3 ubiquitin ligase containing one C3H1-type zinc finger, one RING-type zinc finger and one RNA recognition motif (RRM) domain (Hanzawa et al., 2001; Laribee et al., 2007). Not4 is located on the

Chromosome V: 293,048-294,811 forward strand (Gene, 2011). A graphic of Not4 is shown in Figure 1.6 and a summary of the Not4 full-length protein is given in Table 1.1. The complete amino acid sequence of Not4 is given in Figure 1.7. The structure of the Not4 RING finger is shown in Figure 1.8. A schematic of a typical RRM is depicted in Figure 1.9.

1.4.1 Role of Not4 in ubiquitination. Ubiquitination is a post-translational process that targets proteins for degradation in proteasomes through the attachment of ubiquitins to a lysine residue on the target protein via an isopeptide bond (Figure 1.10). The process is mediated by E1 ubiquitin-activating enzyme, E2 ubiquitin-conjugating enzyme and E3 ubiquitin ligase. E1 ubiquitin-activating enzyme activates ubiquition through an ATP-dependent mechanism and transfers the activated ubiquitin to an E2 ubiquitin-conjugating enzyme. An E3 ubiquitin ligase then interacts with the E2 ubiquitin-conjugating enzyme to label the targeted protein with ubiquitin. Not4 E3 ubiquitin ligase binds to E2 through its RING-type zinc finger domain (Laribee et al., 2007). Recent studies show that the turn-over of cytoplasmic nascent polypeptide-associated complex (NAC) is regulated through Not4-dependent ubiquitination (Panasenko et al., 2006). Not4 E3 ubiquitin ligase activity also regulates the stability of Jhd2, the H3K4me3 demethylase (Laribee et al., 2007; Mersman et al., 2009; Mulder et al., 2007). Additionally, Not4 E3 ligase is important for proteasomes integrity through its association with the proteasome chaperone Ecm29, a protein that stabilizes proteasomes. Not4 knock-out mutants showed reduced levels of free ubiquitin and an accumulation of polyubiquitinated proteins (Panasenko and Collart, 2011).

1.4.2 Not4 RNA recognition motif. The RNA Recognition Motif (RRM) is a common protein domain in eukaryotes, present in about 0.5-1.0% of human genes and is used for the binding of a variety of RNA sequences and proteins. The RRM is approximately 90 amino acids in length (Birney et al., 1993) and its domain structure is typically characterized by two α-helices packed on a four-stranded β-sheet with the β-sheets acting as the primary RNA binding surface (Figure 1.9). Binding affinity and specificity of the RRM to RNA is dependent on the β-strands, loops, and α-helices (Stefl et al., 2005). Binding and affinity may also be influenced by the neighboring RRM domains, which increase RNA-binding affinity and specificity (Maris et al., 2005).

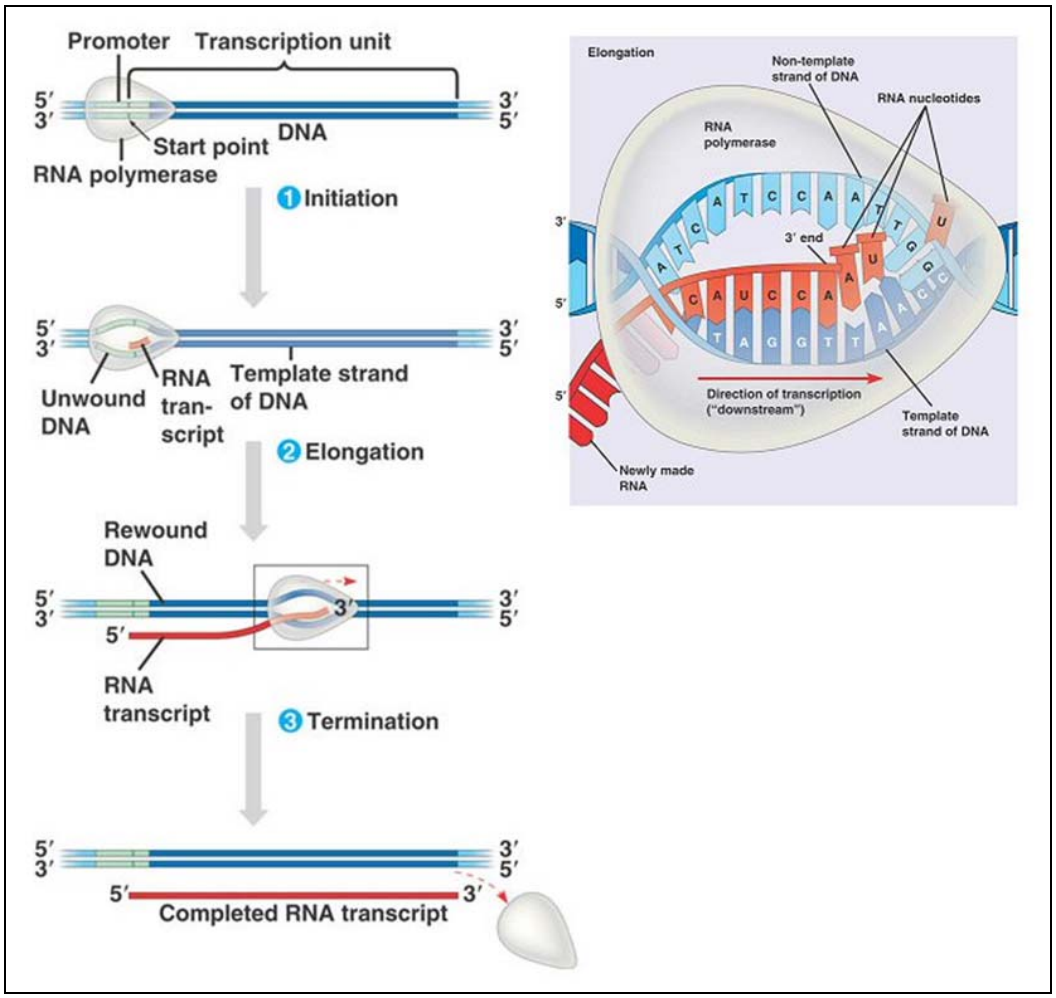
The human counterpart of the yeast Not4, hNot4, was recently identified and characterized (Albert et al., 2000). Like Not4, hNot4 interacts with yeast Not1 and is able to compensate for a *NOT4*-null mutation in yeast. Additionally, like Not4, hNot4 contains two protein motifs in the N-terminal domain: a RING finger and an RRM. The RRM domain of hNot4 has been shown to be involved in the binding of single-stranded nucleic acids (reviewed in Nagai et al., 1995).

1.5 Scope and significance of this project

This project studied the RNA binding activity of the Not4 protein subunit of the Ccr4-Not complex. We expressed and purified the Not4 full-length and the Not4 RRM and used *in vitro* RNA binding assays to study the RNA binding activity. Our results reveal that Not4 has RNA binding activity. Interestingly, the Not4 RRM did not show RNA binding

activity. Together, these results suggest that Not4 binding to RNA may be dependent on protein regions outside of the RRM domain.

Figure 1.1

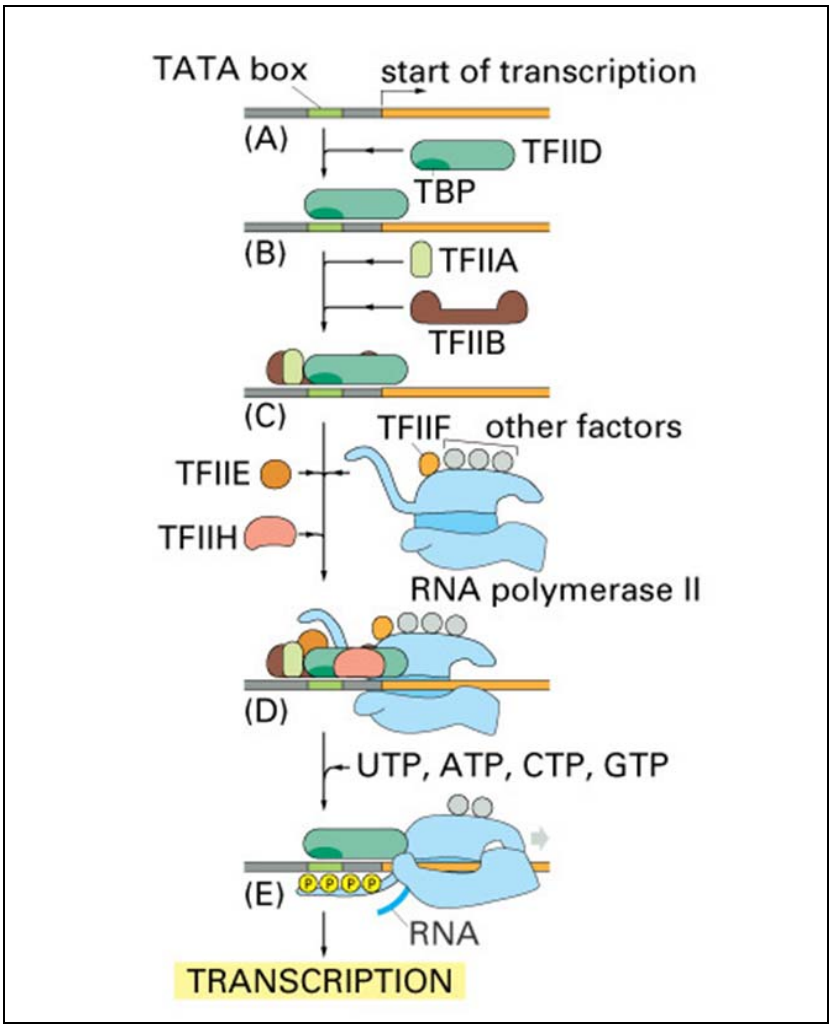


(Cambell et al., 2005)

Figure 1.1 Cell transcription cycle

The cell transcription cycle is characterized by three stages: initiation, elongation and termination.

Figure 1.2

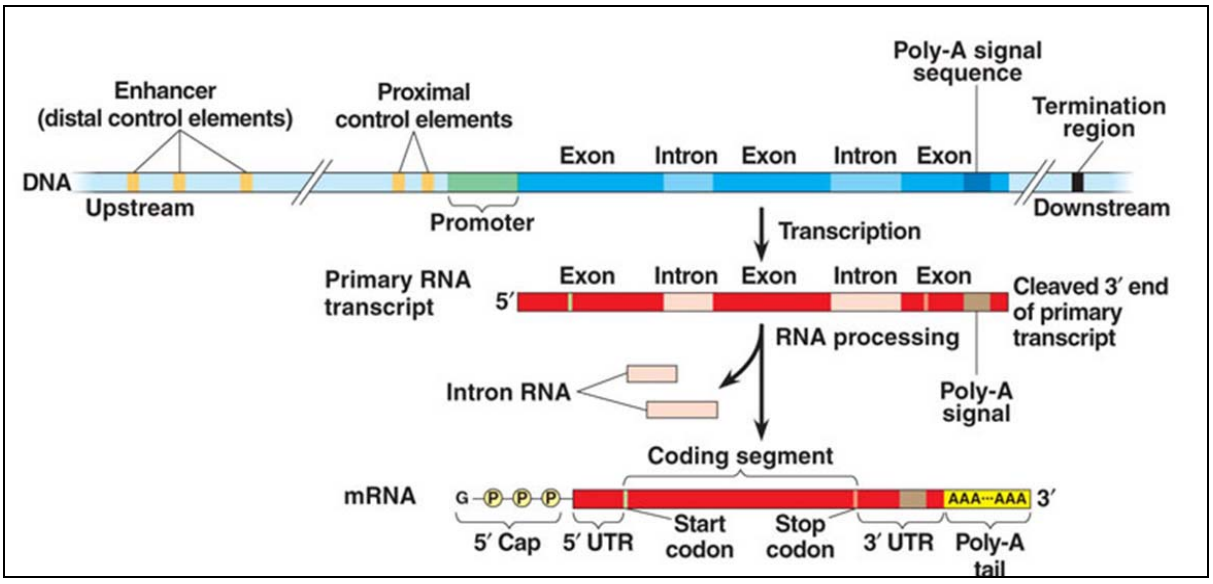


(Alberts et al., 2003)

Figure 1.2 Transcription initiation by RNAPII

The initiation of transcription is marked by the recruitment of general transcription factors to the promoter to form the preinitiation complex (PIC). These transcription factors include IIA, IIB, IID, IIE, IIF, and IIH. RNAPII transcription proceeds in the 5' – 3' direction.

Figure 1.3

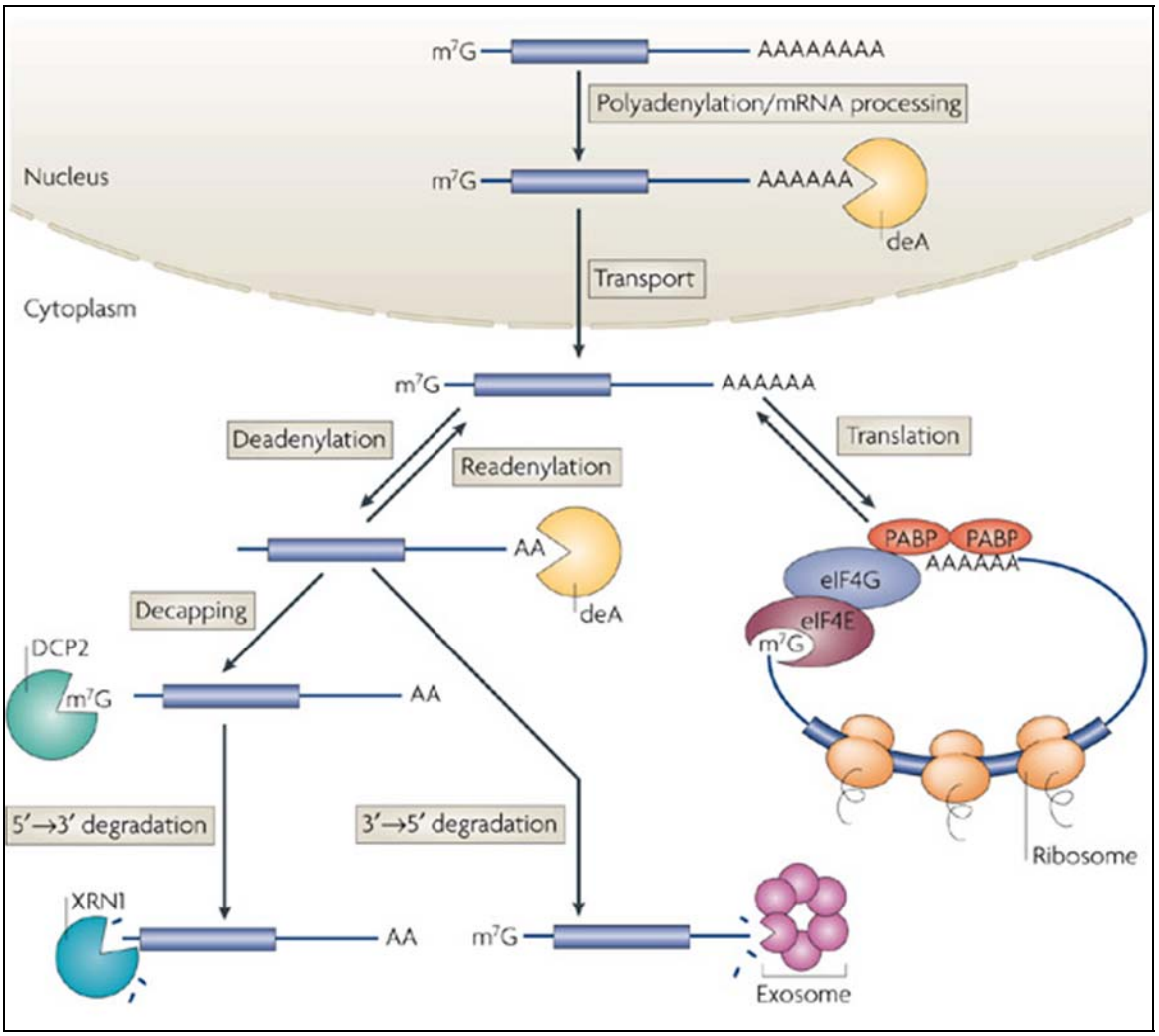


(Cambell et al., 2005)

Figure 1.3 mRNA processing

A 5’ cap is added co-transcriptionally by the Capping Enzyme Complex (CEC), which associates with RNAPII prior to transcription initiation. A poly(A) tail is added to the 3’ end immediately following transcription termination by polyadenylate polymerate and runs approximately 250 adenosine residues in length. Pre-mRNA splicing removes introns, and mRNA editing may further change the nucleotide composition of the final mature mRNA product.

Figure 1.4



(Goldstrohm and Wickens, 2008)

Figure 1.4 mRNA deadenylation decay

The most prevalent pathway for mRNA decay in eukaryotic cells is the deadenylation-dependent pathway, which begins with the deadenylation of the poly(A) tail by specialized 3' – 5' exonucleases. Removal of the poly(A) tail destabilizes the elF-4E and elF-4G cap-binding complex, leading to the degradation of the 5' cap by the decapping complex. The remaining transcript is then fully degraded by either the PM/Scl exosome complex or the decapping complex.

Figure 1.5

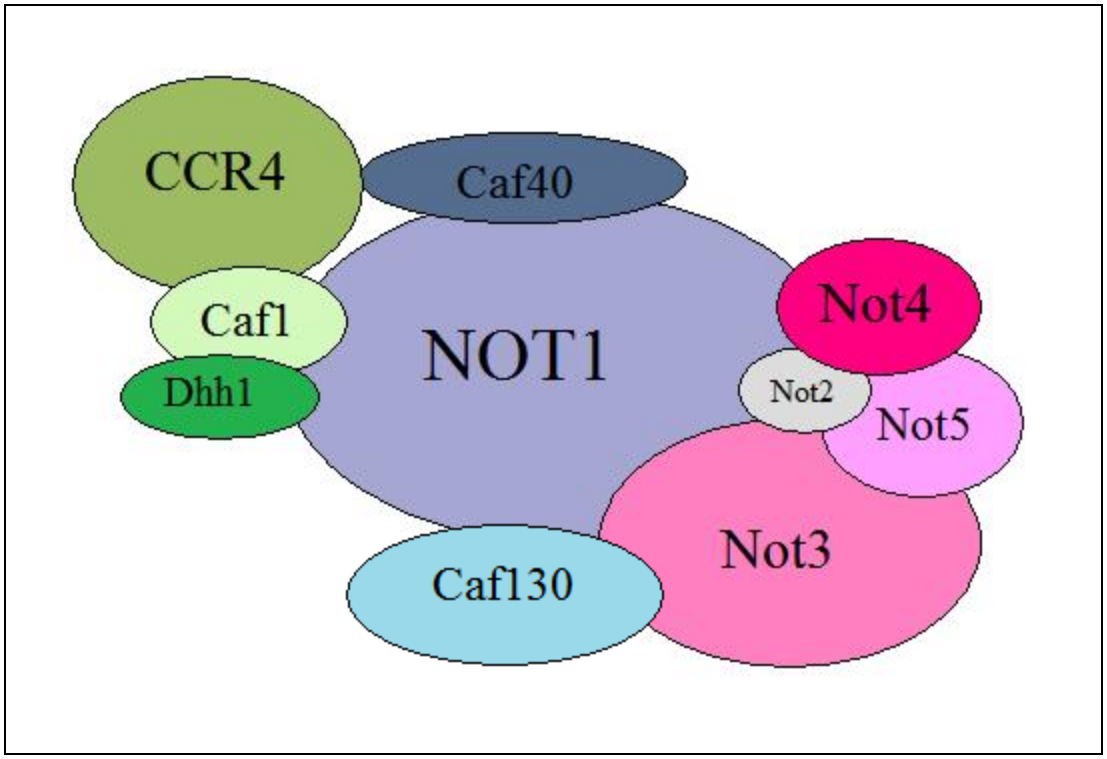
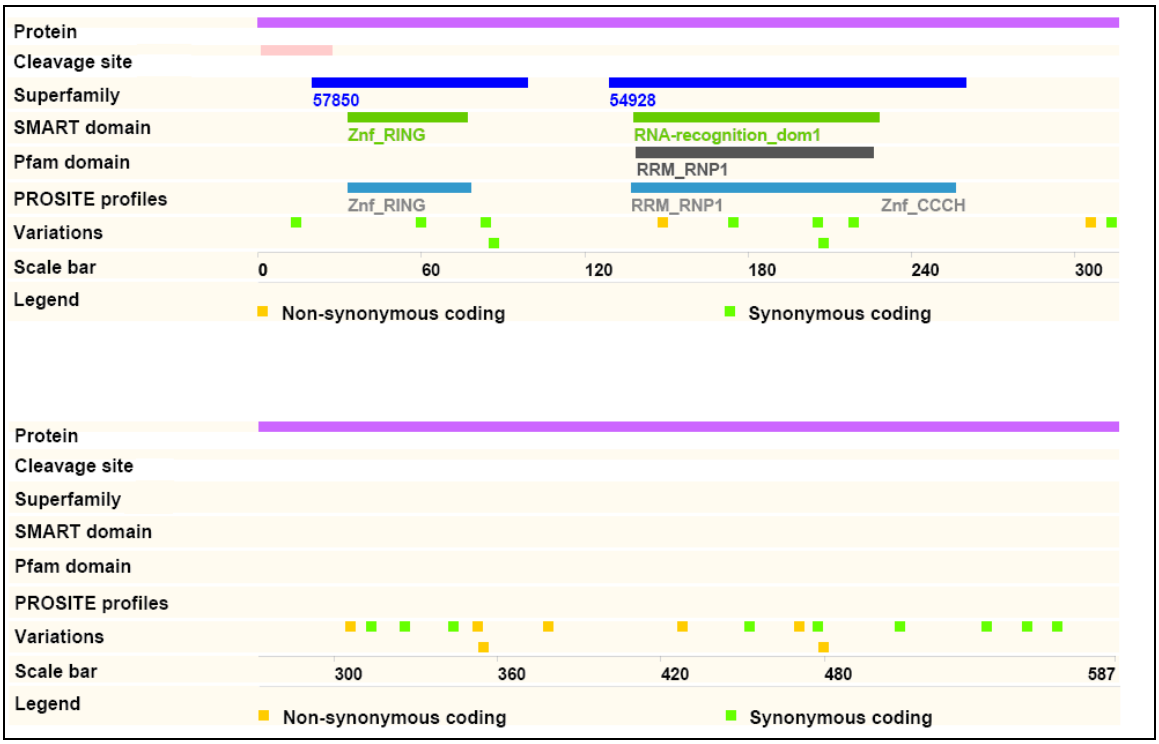


Figure 1.5 Ccr4-Not complex with protein subunits

Ccr4-Not is a multi-functional, multi-protein global gene regulation complex consisting of nine subunits (Ccr4p, Caf1p, Caf40p, Caf130p, and Not1-5p). The Ccr4 group of proteins interacts with the N terminus of Not1 while the Not group of proteins interacts with C terminus. The Ccr4, Caf1 (also known as Pop2) and the Not proteins 1, 2, 4 and 5 all have identified roles in regulating transcription. Ccr4 and Pop2 also have identified roles in the regulation of mRNA decay.

Figure 1.6



(Gene, 2011)

Figure 1.6 Not4 full-length summary

NOT4 located on the Chromosome V: 293,048-294,811 forward strand. Not4 contains two protein motifs in its N-terminal region: a RING finger and an RRM. The N-terminal of the protein is evolutionarily conserved, in contrast to the C-terminal, which is not (Albert et al., 2000).

Figure 1.7

MMNPH VQENL QAIHN ALSNF DTSFL SEDEE DYCPL CIEPM - 40

DITDK NFFPC PCGYQ ICQFC YNNIR QNPEL NGRCP ACRRK - 80

YDDEN VRYVT LSPEE LKMER AKLAR KEKER KHREK ERKEN - 120

EYTNR KHLSG TRVIQ KNLVY VVGIN PPVPY EEVAP TLKSE - 160

KYFGQ YGKIN KIVVN RKTPH SNNTT SEHYH HHSPG YGVYI - 200

TFGSK DDAAR CIAQV DGTYM DGRLI KAAYG TTKYC SSYLR - 240

GLPCP NPNCM FLHEP GEEAD SFNKR ELHNK QQAQQ QSGGT - 280

AFTRS GIHNN ISTST AGSNT NLLSE NFTGT PSPAA MRAQL - 320

HHDSH TNAGT PVLTP APVPA GSNPW GVTQS ATPVT SINLS - 360

KNSSS INLPT LNDSL GHHTT PTTEN TITST TTTTN TNATS - 400

HSHGS KKKQS LAAEE YKDPY DALGN AVDFL DARLH SLSNY - 440

QKRPI SIKSN IIDEE TYKKY PSLFS WDKIE ASKKS DNTLA - 480

NKLVE ILAIK PIDYT ASVVQ FLQSV NVGVN DNITI TDNTK - 520

TPTQP IRLQT VSQQI QPPLN VSTPP PGIFG PQHKV PIQQQ - 560

QMGDT SSRNS SDLLN QLING RKIIA GN - 587

Figure 1.7 Not4 full-length protein sequence

Not4 is 587 amino acids in length. The Not4 RING finger extends from a.a. 33 to a.a. 77 and is shown with a single underline. The Not4 RRM extends from a.a. 138 to a.a. 228 and is shown with a double underline (Gene, 2011).

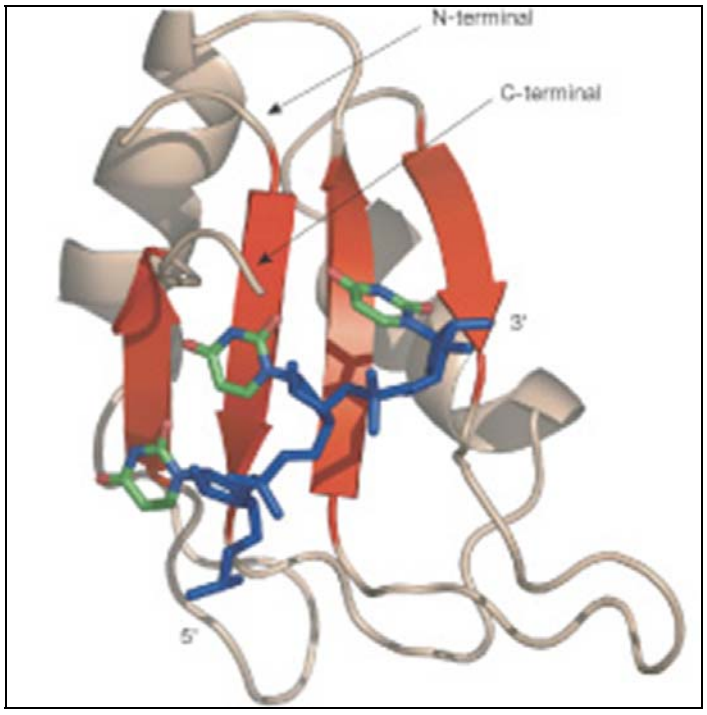
Figure 1.8

(Hanzawa et al., 2001)

Figure 1.8 Schematic view of the RING finger structure of Not4

The conserved α-helix is seen in the center of the structure. The conserved zinc ions are indicated as balls. Cysteine residues that coordinate the zinc ions are seen in close proximity to the ions.

Figure 1.9

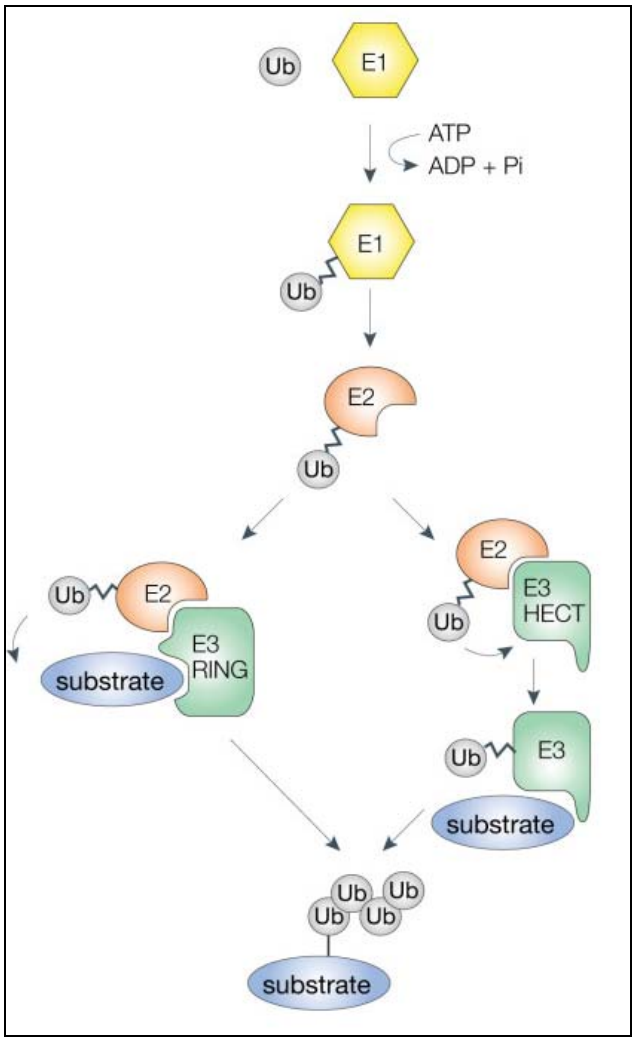


(Stefl et al., 2005)

Figure 1.9 Schematic view of the RRM

The schematic shows an RRM interacting with a triplet UUU sequence. The RRM domain structure is characterized by two α-helices packed on a four-stranded β-sheet with the β-sheets acting as the primary RNA binding surface.

Figure 1.10



(Woelk et al., 2007)

Figure 1.10 Protein ubiquitination

Protein ubiquitination is mediated by E1 ubiquitin-activating enzyme, E2 ubiquitin-conjugating enzyme and E3 ubiquitin ligase. E1 ubiquitin-activating enzyme activates ubiquition through an ATP-dependent mechanism and transfers the activated ubiquitin to an E2 ubiquitin-conjugating enzyme. An E3 ubiquitin ligase then interacts with the E2 ubiquitin-conjugating enzyme to label the targeted protein with ubiquitin.

Table 1.1 Not4 full-length statistics

Average residue weight	111.335
Charge	14.5
Isoelectric point	8.1690
Molecular weight	65,353.83
Number of residues	587

(Gene, 2011)

Table 1.1 Not4 full-length statistics. The Not4 protein (also called Mot2) is a 587 amino acid E3 ubiquitin ligase.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

The strain used in this work is BY4741 Not4-13myc::kanMX. All media were standard; all RNA oligonucleotides were from Dharmacon Research; PreScission protease was from GE Healthcare; all other reagents were from Sigma, Fisher Scientific or VWR.

2.2 Cloning of Not4 ORF and Not4 RRM into pGEX-6P-1 vector

The genomic DNA from a BY4741 Not4-13myc::kanMX strain was subjected to PCR to amplify the *NOT4* gene. The primers used for amplification of this sequence were 5'-GGCGAATTCATGATGAATCCACACGTTCA-3' (forward) and 5'-GGCCTCGAGTTAATTACCGGCGATAATTTTCC-3' (reverse). The amplified protein product was then purified through phenol-chloroform isoamyl (PCIAA) extraction and ethanol precipitation. The RNA Recognition Motif (RRM) of Not4 was purified through PCIAA and ethanol precipitation as well in preparation for sequential restriction digestion. Restriction enzyme Xho1 was used to digest *NOT4* ORF, the *NOT4* RRM, and the pGEX-6P-1 vector (Figure 2.1). PCIAA extraction and ethanol precipitation was performed, and then restriction digestion with enzyme EcoR1 was performed. Sequential restriction digestion was completed with the PCIAA extraction and ethanol precipitation of both the *NOT4* ORF and *NOT4* RRM inserts and the pGEX-6P-1 vector. Treatment

with calf intestine alkaline phosphatase (CIAP) was further used for the pGEX-6P-1 vector.

2.3 DNA gel purification

DNA fragment gel purification was performed in 1% Low Melt agarose in 1.5X TAE with STE buffer [10 mM Tris-HCl (pH 8), 100 mM NaCl, 1 mM EDTA] for *NOT4* ORF, *NOT4* RRM, and pGEX-6P-1. The purified plasmid product was concentrated through butanol extraction with an equal volume of anhydrous *n*-butanol to DNA samples, and DNA concentrations were measured with a Thermo Scientific Nanodrop^{TM} 2000 spectrophotometer. The ligation reaction was done with T4 DNA ligation buffer at a 3:1 ratio for the *NOT4* ORF and *NOT4* RRM insert to the pGEX-6P-1 vector.

2.4 Bacterial cell transformation for plasmid amplification

To amplify the recombinant plasmids in preparation for protein expression in competent BL21(DE3)pLysS *Escherichia coli* cells, the recombinant plasmids were transformed into DH5α competent *E. coli* cells in lysogeny broth (LB) media. A volume of 5 µl of plasmid DNA was added to 100 µl of DH5α competent cells in a 1.5 ml microfuge tube. The tube was immediately placed on ice for 30 minutes with intermittent mixing by flicking every 10 minutes. The cell suspension was then placed in a 42 °C water bath for 1.5 minutes. 1 ml of room temperature LB media was added to the mixture and the microfuge tube was placed in a 37 °C water bath for 30 minutes. The mixture was centrifuged at 5000 rpm for one minute, and 500 µl of the supernatant was pipetted out and discarded. The remaining contents of the microfuge tube were re-suspended gently

by pipetting 3-4 times and inoculated to three LB and ampicillin (LB/AMP) culture plates. The culture plates were incubated overnight in a 37 °C incubator and then stored at 4 °C.

2.5 Mini scale plasmid preparation

Single colonies from both the *NOT4* ORF / pGEX-6P-1 transformation plates and the *NOT4* RRM / pGEX-6P-1 transformation plates were each inoculated into separate 5 ml of LB containing 100 µg/ml of ampicillin for 20 cultures total. The cultures were grown overnight on a 37 °C shaker. The overnight cultures were then centrifuged for 10 minutes at 4 °C and 2200 rpm and the supernatant was discarded. Each cell pellet was re-suspended in 1 ml of 0.1X TE by vortexing and transferred to separate 1.7 ml microfuge tubes. The cell suspensions were centrifuged for 1 minute at 4 °C and 5000 rpm and the supernatants were aspirated. The pellets were re-suspended in 100 µl of cold Solution I [50 mM glucose, 25 mM Tris-Cl (pH 8.0), 10 mM EDTA (pH 8.0)] by vortexing. 200 µl of Solution II [0.2 M NaOH, 1% SDS] was then added and the contents were mixed by inversion and placed on ice for 5 minutes. 150 µl of cold Solution III [3 M potassium acetate, 5 M acetate] was added. The contents were then mixed by vortexing and placed on ice for 5 minutes. The microfuge tubes were then centrifuged for 10 minutes at 4 °C and 5000 rpm. Each supernatant was transferred into a separate fresh microfuge tube. 0.7 – 0.8 volume of isopropanol was added and mixed with the supernatant. The suspension was placed at room temperature for 15 minutes, after which the supernatants were discarded and the plasmid pellets were air dried and stored at 4 °C. Digestion with Xho1 and EcoR1 in 10X Buffer 2 was used to verify the plasmid prep product.

2.6 Bacterial cell transformation with BL21(DE3)pLysS *E. coli* cells

Bacterial cell transformations for both the Not4 ORF pGEX-6P-1 and Not4 RRM pGEX-6P-1 recombinant plasmid were performed beginning with 100 µl of competent BL21(DE3)pLysS *E. coli* cells and 5 µl of the desired recombinant plasmid in a 1.7 ml microfuge tube. The mixture was kept on ice for 30 minutes with intermittent mixing by flicking every 10 minutes. The mixture was placed in a 42 °C water bath for 1.5 minutes. 1 ml of room temperature LB media was added to the mixture and the microfuge tube was placed in a 37 °C waterbath for 30 minutes. The mixture was centrifuged at 5000 rpm for 1 minute, and 500 µl of the supernatant was pipetted out and discarded. The remaining contents of the microfuge tube were re-suspended gently by pipetting and plated over three LB/AMP plates. The culture plates were grown overnight at 37 °C and then stored at 4 °C.

2.7 Conditions for protein expression

The Not4 ORF pGEX-6P-1 recombinant plasmid was transformed into either E. coli BL21(DE3)pLysS or CODON+ competent cells. Protein expression was then performed under auto-induction (Fox, 2009), LB only, and isopropyl β-D-1-thiogalactopyranoside (IPTG) induction conditions to determine the most suitable competent cell strain and expression condition to use for Not4 protein expression. BL21(DE3)pLysS competent cells and IPTG induction was determined to be an appropriate choice.

2.7 *E. coli* protein expression, purification and extraction

To make the starter culture, a smear (20-30 count) of colonies from the overnight LB/AMP culture plates were used to inoculate 100 ml LB media containing 10 mg ampicillin and 2.5 mg chloramphenicol (LB/AMP/CAM media). The cell culture was grown for 2 hours in a 37° C shaker and the optical density (O.D.) was taken at the end of the 2 hours. The starter culture in its entirety was then introduced to 4 L of LB/AMP/CAM media and grown to an O.D. of 0.6 in a 37° C shaker (the doubling time for *E. coli* cells is approximately 20 minutes). The culture was then induced with IPTG at 0.5 mM and placed in a 22° C shaker for 2 hours. Culture cells were sampled just before and directly after IPTG induction and processed through Coomassie Brilliant Blue staining SDS-PAGE.

To harvest the cells, the *E. coli* cell culture was centrifuged for 12 minutes at 4° C and 3300 rpm in the Sorvall SLC-6000 centrifuge. The supernatant was discarded and the remaining cell sediment was washed 4 times with 10 ml of 1X PBS, centrifuging between each wash. For storage, cells transferred to 15 ml test tubes, flash-frozen under dry ice for 5 minutes, and stored at –80 °C.

GST protein purification and extraction was performed on ice with pre-chilled buffers to minimize protein degradation. Per 2 ml of harvested cells, 8 ml of 1 M NaCl GST buffer [20 mM Tris (pH 7.5), 0.5 mM EDTA, 10 mM β-mercaptoethanol, 1 M NaCl, 0.1% NP40] containing 1 mM benzamidine-HCl and 0.4 mM PMSF were used to re-suspend the cells. 40 µl of DDT was added to each 10 ml total volume and the cell suspension was

subjected to sonication using a Branson sonifier at power 6 and 60% duty for 10 pulses. Sonication was repeated 4 times with a 3 minute chill interval on ice between each session. After sonication, 0.8 ml of 10% Triton X-100 was added per 10 ml total volume cell lysate. The cells were then incubated in a 22 °C shaker for 30 minutes and centrifuged for 15 minutes at 4 °C and 15,000 rpm in the Sorvall SLC-6000 centrifuge. After centrifugation, the supernatant was sampled and processed through Coomassie Brilliant Blue staining SDS-PAGE.

The supernatant was introduced to 2 ml of Glutathione Sepharose4 Fast Flow beads pre-washed with 1M GST buffer. Binding was given 2 hours on a 4 °C test tube rocker. The beads were then centrifuged for 2 minutes at 4 °C and 1000 rpm. The flow-through was sampled. 8 ml of GST buffer was used to wash the beads 4 times with a 5 minute interval on the test tube rocker between each wash. The beads were then transferred to a fresh test tube, washed once with GST buffer, and re-suspended in 1 ml GST buffer. The re-suspended beads were sampled and processed through Coomassie Brilliant Blue staining SDS-PAGE.

2.10 PreScision protease protein digestion

The protein-bound beads were washed with 10 ml ice-cold PreScision protease buffer [50 mM Tris-HCl (pH 7.0), 150 mM NaCl, 1 mM EDTA, 1 mM DTT] for 5 minutes on a 4 °C test tube rotator and centrifuged for 2 minutes at 4 °C and 1000 rpm. 1 mL of PreScision protease buffer was used to re-suspend and transfer the protein-bound beads to 1.7 ml microcentrifuge tubes. The protein-bound beads were centrifuged for 2 minutes at

4 °C and 1000 rpm and the supernatant was discarded. 400 µl of PreScision protease buffer and 25 µl of PreScision protease were then added to each microcentrifuge tube and the digestion was set at 4 °C for 20 hours.

After 20 hours, the beads were centrifuged for 3 minutes at 4 °C and 3000 rpm. The supernatant was collected in a fresh 1.7 ml microcentrifuge tube. The remaining beads were first washed with 200 µl of PreScision protease buffer and then washed three more times with 500 µl of PreScision protease buffer. A sample of the supernatant and each subsequent wash was taken and run on an SDS gel. Since the likelihood of protein elution is highest in the first wash, the 200 µl wash would be combined with the supernatant if protein was detected on the SDS gel. Additional washes typically did not elute protein and were discarded.

The pooled protein preparation was concentrated using a Vivaspin 500 centrifugal concentrator as described by the Vivaspin 500 protocol (Sartorius Stedim Biotech, 2008) until total volume was reduced to 500 µl. The protein was then stored in 60 µl aliquots at –80 °C.

2.12 SP-sepharose column purification for Not4 full-length

The pooled and concentrated Not4 protein was passed over a 0.5 ml SP-sepharose column. The column was washed with 20 ml of buffer A [20 mM HEPES-NaOH (pH 7.4), 75 mM NaCl, 10% glycerol] followed by a wash with 10 ml of buffer B [20 mM HEPES-NaOH (pH 7.4), 150 mM NaCl, 10% glycerol]. The protein was eluted using

buffer C [20 mM HEPES-NaOH (pH 7.4), 500 mM NaCl, 10% glycerol]. The fractions free of impurities were pooled and concentrated as described by the Vivaspin 500 protocol. The protein was stored at -80 °C.

2.13 Bio-Rad protein assay

Protein concentration was determined by the Bio-Rad Protein Assay, which is based on the Bradford method. Acidic Coomassie Brilliant Blue G-250 dye is added to the protein solution and measured at 595 nm with a spectrophotometer. A relative measurement of protein concentration is then determined by comparison to a standard curve generated with bovine serum albumin (BSA).

2.14 Fluorescence polarization assay

The fluorescence polarization (FP) assay was performed on a Beacon^{TM} 2000 fluorescence polarization system by GE Healthcare with minimal light exposure by combining 0.1 nM 20mer 3'-fluorescein-labeled polyU RNA (Dharmacon) with increasing concentrations of either the Not4 full-length or the Not4 RRM derivative proteins in 1X RNA binding buffer [25 mM HEPES-NaOH (pH 7.4), 10 mM $MgCl_2$, 1 mM β-mercaptoethanol and 50 mM NaCl] and measuring the flurescence polarization mP value at each protein concentration. The volume of the protein mix was added at 10% of the total reaction volume, the RNA mix was added at 50% of the total reaction volume, and the RNA binding buffer was added at 5% of the total reaction volume. The mP value was plotted against the protein concentration and a hyperbolic curve was fitted to the data

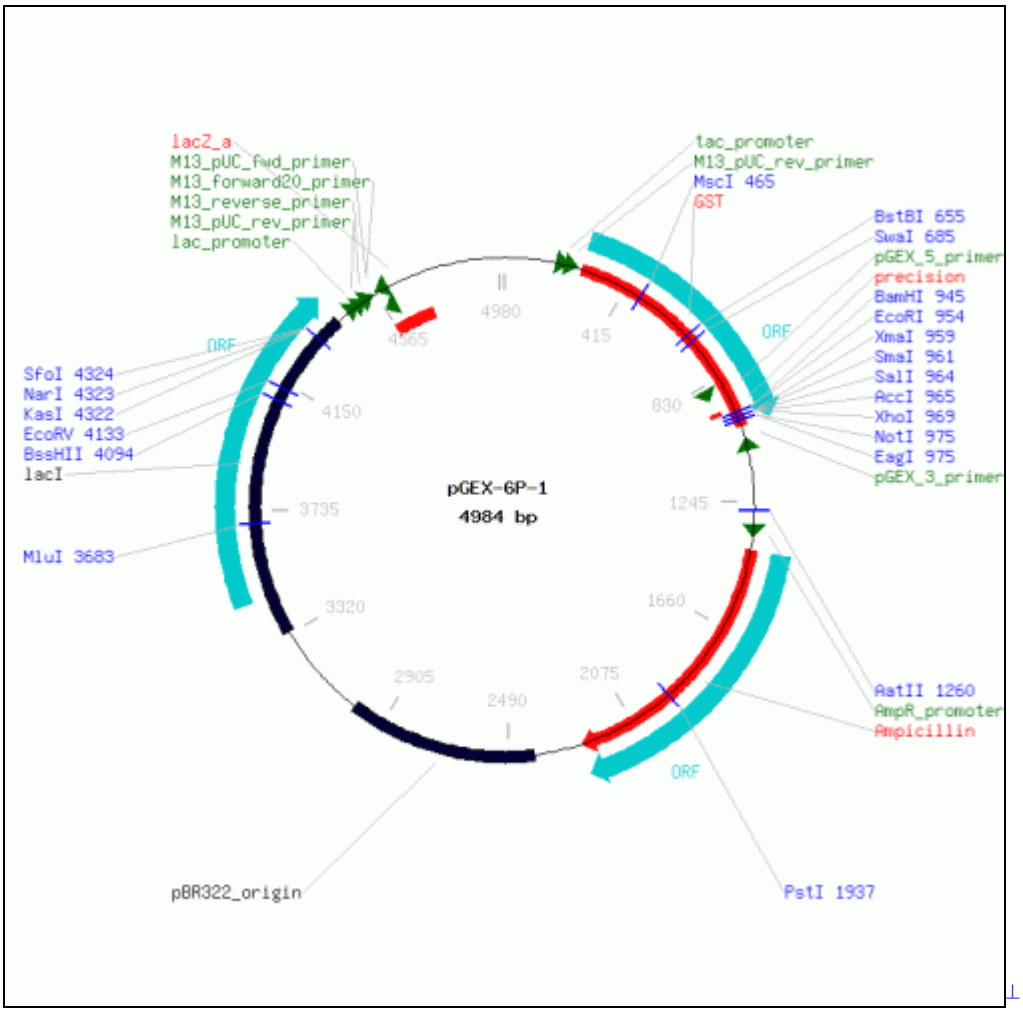
using the Synergy Software KaleidaGraph program to determine the K_d dissociation constant.

2.15 Electrophoretic mobility shift assay

The electrophoretic mobility shift assay (EMSA) was performed by mixing 100 nM 20mer 3'-fluorescein-labeled polyU RNA (Dharmacon) with increasing concentrations of either the Not4 full-length or the Not4 RRM derivative proteins in 1X RNA binding buffer. 10X native agarose gel loading buffer [15% ficoll, 0.25% xylene cyanol, 0.25% bromophenol blue] was added to the RNA samples to a final concentration of 1X. The final binding volume was 10 µl. Samples were run on a 5% native gel in a running buffer [50 mM Tris (pH 8.5), 0.38 M Glycine, 2 mM EDTA (pH 8.0), 5 mM $MgCl_2$]. The 5% native polyacrylamide gel was composed of 5% polyacrylamide with 29:1 acrylamide/bis-acrylamide and 0.5X Tris-Borate-EDTA (TBE) running buffer [44.5 mM Tris (pH 8.5), 44.5 mM boric acid, 1 mM EDTA (pH 8.0)].

The gel was run at 100 V for 1 hour and analyzed using a PhosphorImager. The gel was scanned with the Molecular Dynamics Typhoon system and a Fluorescence 526 SP emission filter, and ImageQuant was used to view the scan.

Figure 2.1



(Lablife, 2011)

Figure 2.1 pGEX-6P-1 vector plasmid map

Restriction enzymes EcoR1 and Xho1 were used to digest the pGEX-6P-1 vector at sites 954 and 969 respectively.

CHAPTER 3

RESULTS

3.1 Not4 full-length shows RNA binding activity *in vitro*

The Not4 protein is an E3 ubiquitin ligase containing one C3H1-type zinc finger, one RING-type zinc finger and one RNA recognition motif (RRM) domain. While the three domains in Not4 have all been described, only the function of the RING domain has been identified (Panasenko and Collart, 2011). The RRM domain is conserved from humans to yeast and is implicated in binding a variety of RNA sequences and proteins (Maris et al., 2005). Interestingly, recent studies have linked the process of ubiquitination to mRNA metabolism, and there is evidence to support the idea that RNA-binding activities of E3 ubiquitin ligases help to regulate RNA stability (Cano et al., 2010). However, the Not4 protein and the Not4 RRM domain have not been shown to bind RNA. It is possible that as an E3 ubiquitin ligase with an RRM domain, the Not4 protein is able to bind RNA through its RRM domain.

To study the RNA binding activity of Not4, the Not4 ORF was first cloned into pGEX-6P-1 and the recombinant protein was expressed in either BL21-DE3 pLysS or CODON+ competent cells. Protein expression was then tried under auto-induction, LB only, and IPTG induction conditions (Figure 3.1 A) to determine a suitable expression condition to use for Not4 protein expression. BL21(DE3)pLysS competent cells and IPTG induction was chosen for Not4 full-length protein expression (Figure 3.1 B) because the cells showed strong Not4 expression relative to protein products not of interest to this study.

GST protein purification and extraction was used to process the expressed Not4 full-length protein and Glutathione Sepharose4 Fast Flow beads were used for protein binding. Protein digestion was performed with PreScision protease (Figure 3.1 C).

The Not4 protein was concentrated with a Vivaspin 500 centrifugal concentrator (Figure 3.1 D). However, after protein concentration, the degradation products of the Not4 full-length protein of interest became concentrated as well. A SDS-gel protein sample run shows protein purity at approximately 60% (Figure 3.1 D). To increase percent purity, the concentrated Not4 protein was passed through a 0.5 ml SP-sepharose column and eluted at 500 mM NaCl (Figure 3.1 E). Six fractions were collected. Fractions one and three (lane 2 and lane 4 of Figure 3.1 E respectively) were free of impurities and were pooled and concentrated again with a Vivaspin 500 centrifugal concentrator to be used in RNA binding assays.

Next, fluorescence polarization was used to determine the RNA binding activity of Not4. Assays were performed by combining 0.1 nM 20mer 3'-fluorescein-labeled polyU RNA with increasing concentrations of the Not4 full-length protein. Data was collected by monitoring the change in the fluorescence polarization mP value. The K_d dissociation constant for Not4 binding to RNA was determined by plotting fluorescence polarization as a function of protein concentration and fitting the experimental data to a rectangular hyperbola with the following equation: y=(m1*m0)/(m2+m0)+m3, where m0 is the protein concentration independent variable, m1 is the maximum fluorescence polarization

mP value, m2 is the K_d dissociation constant, m3 is the y-axis shift, and y is the fluorescence polarization mP dependent variable.

We observed variable RNA binding activity from assay to assay for the Not4 full-length protein. The results from six assays are shown in Figure 3.3 A – F and typify the range of binding activity seen in the fluorescence polarization assays. K_d values range from 385.62 ± 202.8 (Figure 3.3 D) to 1787.1 ± 1087.4 (Figure 3.3 B). Additionally, assays showing no RNA binding activity were obtained such as that seen in Figure 3.3 E, where the increase in mP value with increasing protein concentration appears with a linear correlation that is likely due to protein aggregation and non-specific binding rather than specific binding. The mean K_d value as determined from the fluorescence polarization binding assays showing RNA binding (those shown in Figure 3.3 A – F, excluding that shown in Figure 3.3 E) is 1174.38 with a standard deviation of 520.50. The variable binding assay results suggest possible discrepancies in protein quality between assays.

The binding assay graph shown in Figure 3.3 F with a K_d value of 1221.2 ± 412.35 was re-plotted with the two terminal data points removed because the curvature of the plotted data points suggests a smooth binding saturation and because the hyperbolic fit yields the highest R-value, 0.99182. The two terminal data points were removed as outliers possibly caused by protein aggregation at high protein concentrations and the plot was re-fitted to a rectangular hyperbola (Figure 3.4). The re-plotted data yielded a better curve fit as seen by an increase in R-value to 0.99707. The K_d value determined from this re-adjusted plot is 594.46 ± 114.28 (Table 3.1).

To see if a different type of binding assay would yield more consistent results and further confirm the binding activity of the Not4 protein, an electrophoretic mobility shift assay (EMSA) was performed with 0.1 nM 20mer 3'-fluorescein-labeled polyU RNA and increasing amounts of the purified Not4 protein. The EMSA did not show RNA binding activity (Figure 3.6), suggesting that experimental conditions under which the EMSA assay was performed was not as conducive to protein – RNA binding as that of the fluorescence polarization assay. Factors that may have prevented protein – RNA binding include the ionic strength and pH of the binding buffer and the temperature and time of the binding reaction among other factors. Because the RNA binding activity of Not4 had been demonstrated through fluorescence polarization, further Not4 full-length binding assays through EMSA with adjusted conditions were not pursued.

3.2 Not4 RRM does not show RNA binding activity *in vitro*

The RRM is a common protein domain in eukaryotes, present in about 0.5-1.0% of human genes and is involved in the binding of a variety of RNA sequences and proteins (Maris et al., 2005). Its domain structure is characterized by two α-helices packed on a four-stranded β-sheet with the β-sheets acting as the primary RNA binding surface (Figure 1.9). Binding affinity and specificity of the RRM to RNA is dependent on the β-strands, loops, and α-helices (Stefl et al., 2005). Neighboring RRM domains influence binding and affinity by increasing RNA-binding affinity and specificity (Maris et al., 2005). However, little information is known about the binding activity of the RNA recognition motif of Not4 (Panasenko and Collart, 2011). Since we have demonstrated

that Not4 has RNA binding activity through fluorescence polarization assays, we hypothesized that binding occurs via the evolutionarily conserved RRM region of Not4.

To study whether Not4 RNA binding occurs through the Not4 RRM domain, a derivative of the *NOT4* gene containing only the RRM was cloned into pGEX-6P-1. The recombinant protein was expressed in *E. coli* BL21-DE3 pLyS competent cells (Figure 3.2 A). The protein was purified through GST protein purification using Glutathione Sepharose4 Fast Flow beads for protein binding. Protein digestion was performed with PreScision protease and the protein was concentrated with a Vivaspin 500 centrifugal concentrator. RRM protein purity is approximately 95% (Figure 3.2 B). SP-sepharose purification was not performed as was done for the Not4 full-length protein because protein degradation was minimal.

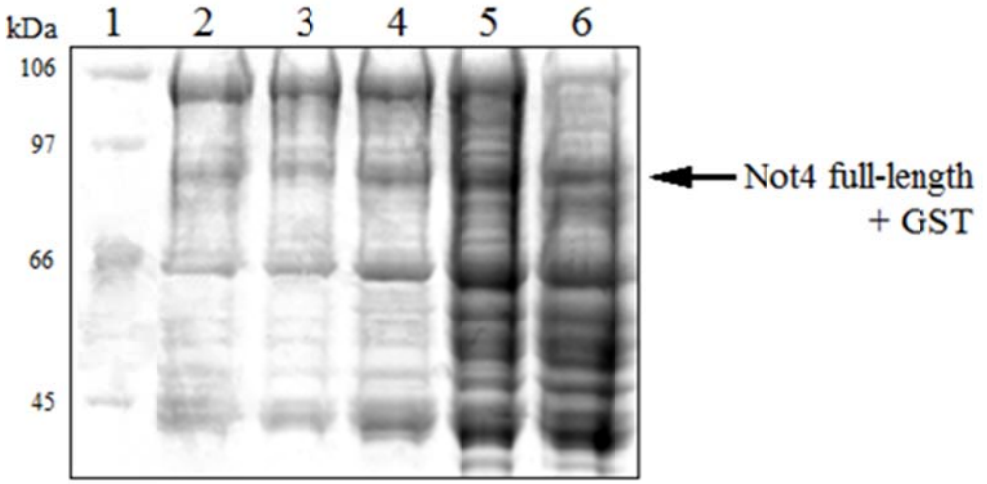
Fluorescence polarization was used to determine the RNA binding activity of the Not4 RRM. Assays were performed by combining 0.1 nM 20mer 3'-fluorescein-labeled polyU RNA with increasing concentrations of the Not4 RRM protein. Data was collected by monitoring the change in the fluorescence polarization mP value. A total of four fluorescence polarization assays were performed for the Not4 RRM derivative (Figure 3.5 A – D). In the 0.0 – 6.0 μM range of RRM protein concentrations tested, no binding was observed. The increase in mP value with increasing protein concentration appears with a linear correlation that is likely due to protein aggregation and non-specific binding rather than specific binding. An EMSA binding assay was performed with 0.1 nM 20mer 3'-fluorescein-labeled polyU RNA and increasing amounts of the purified Not4 RRM

protein in the concentration range of 0.0 – 10.0 μM. As with fluorescence polarization, no RNA binding activity was observed for the Not4 RRM with EMSA (Figure 3.6).

Having shown that the Not4 full-length protein, but not the Not4 RRM, has RNA binding activity, our results suggest that the Not4 RRM alone is not sufficient for RNA binding. However, the RRM may still contribute to RNA binding within the context of the full-length Not4 protein. Also, it is possible that the RNA binding activity of Not4 is dependent on domain(s) other than the RRM.

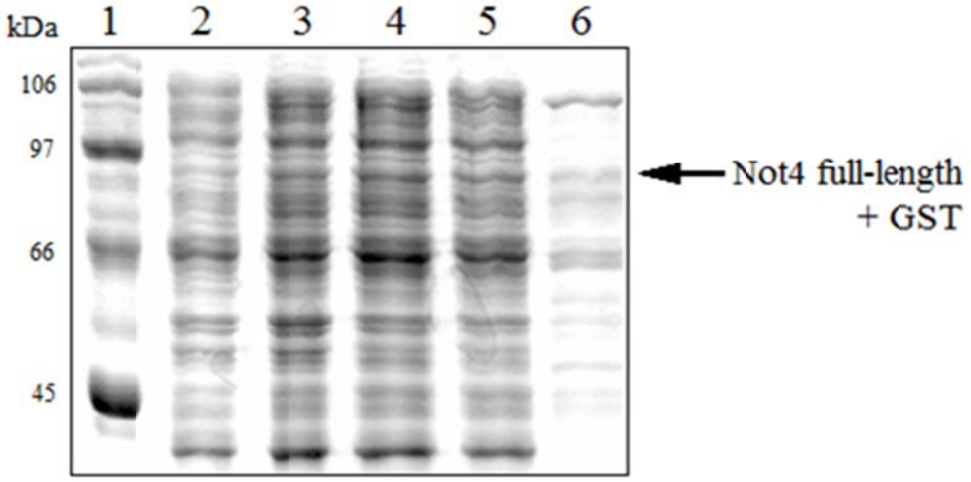
Figure 3.1 Not4 full-length protein expression and purification

A.



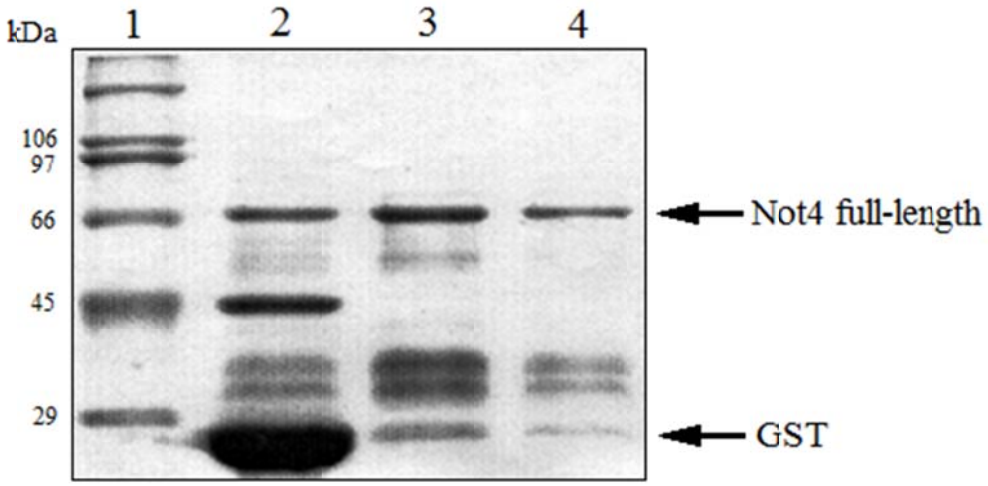
A. Not4 full-length protein expression under different conditions. Not4 / pGEX recombinant protein was expressed in BL21-DE3 pLysS cells under auto-induction (lane 2) and IPTG induction (lane 3) conditions and in CODON+ competent cells under LB (lane 4), auto-induction (lane 5), and IPTG induction (lane 6) conditions. A high molecular weight marker is shown in lane 1.

B.



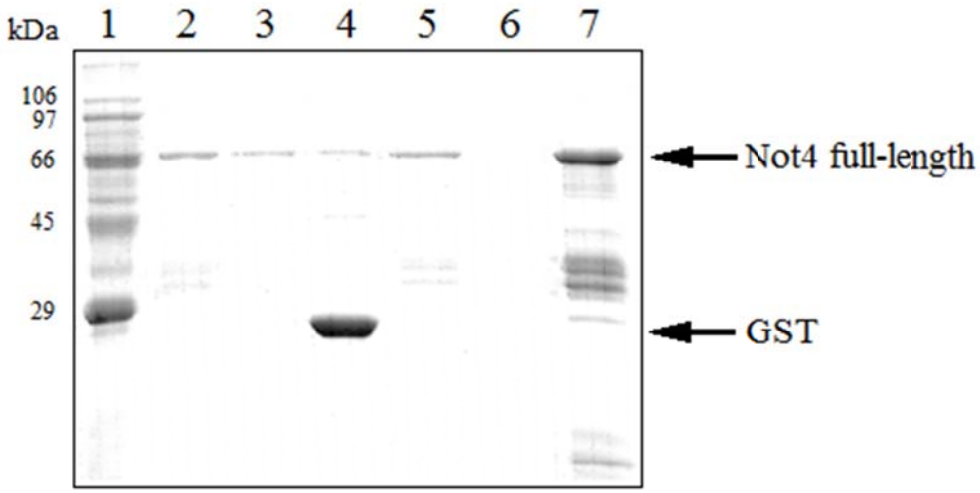
B. Not4 full-length protein expression in *E. coli* BL21-DE3 pLysS cells. Culture cells were sampled before IPTG induction (lane 2) and after IPTG induction (lane 3). GST protein purification and extraction was performed on the IPTG induced cell culture and the supernatant was samples (lane 4). The protein was bound to Glutathione Sepharose4 Fast Flow beads (lane 6). The flow-through from the beads shows loss of protein (lane 5). A high molecular weight marker is shown in lane 1.

C.



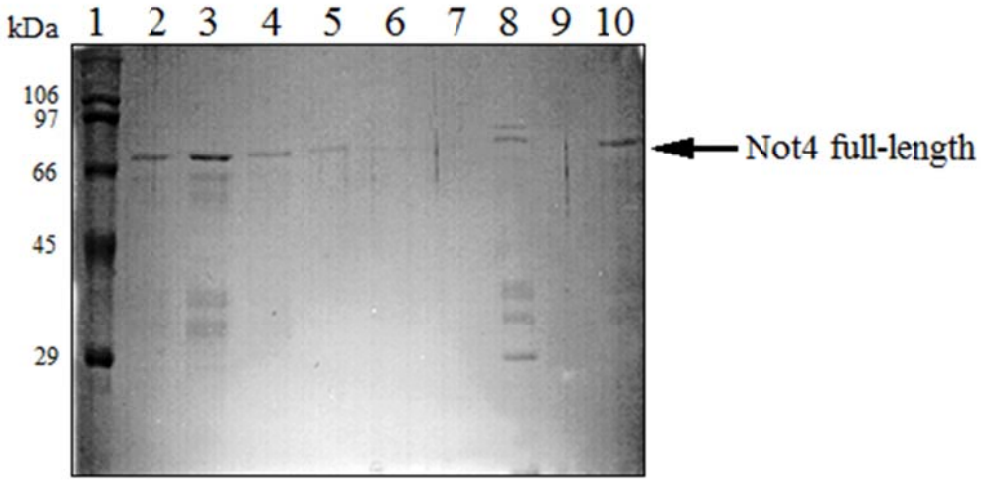
C. PreScission protease digestion of Not4 full-length. After 20 hours of digestion, the supernatant (lane 3) and one wash (lane 4) were collected. The remaining beads showed retention of the Not4 full-length (lane 2). A high molecular weight marker is shown in lane 1.

D.



D. Not4 full-length protein concentration. The protein supernatant (lane 2) and wash (lane 3) were pooled together (lane 5). Lane 4 shows remaining the Glutathione Sepharose4 Fast Flow beads after protein digestion. The pooled protein was concentrated using a Vivaspin 500 centrifugal concentrator (lane 7). No protein was detected in the flow-through (lane 6) from the Vivaspin 500 concentration procedure. A high molecular weight marker is shown in lane 1.

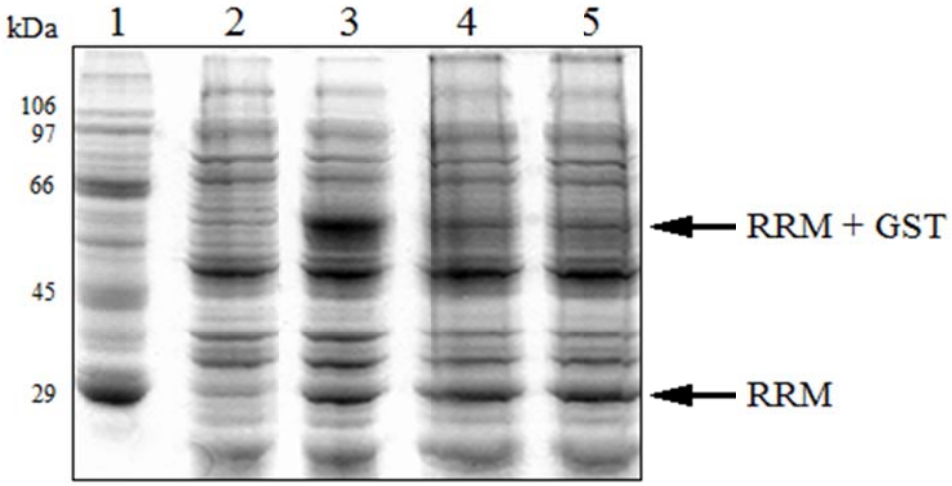
E.



E. Not4 full-length SP-sepharose column purification. The pooled and concentrated Not4 protein was passed over a 0.5 ml SP-sepharose column. The protein was eluted using buffer C. Six fractions were collected (lanes 2 through 7 respectively). The column was washed with buffer A (lane 9) followed by a wash with buffer B (lane 10). A sample of the flow-through from the Vivaspin 500 concentration procedure was run again: protein was detected in this gel (E, lane 8) whereas no protein was detected in the previous gel (D, lane 6) because of gel picture contrast and resolution. A high molecular weight marker is shown in lane 1.

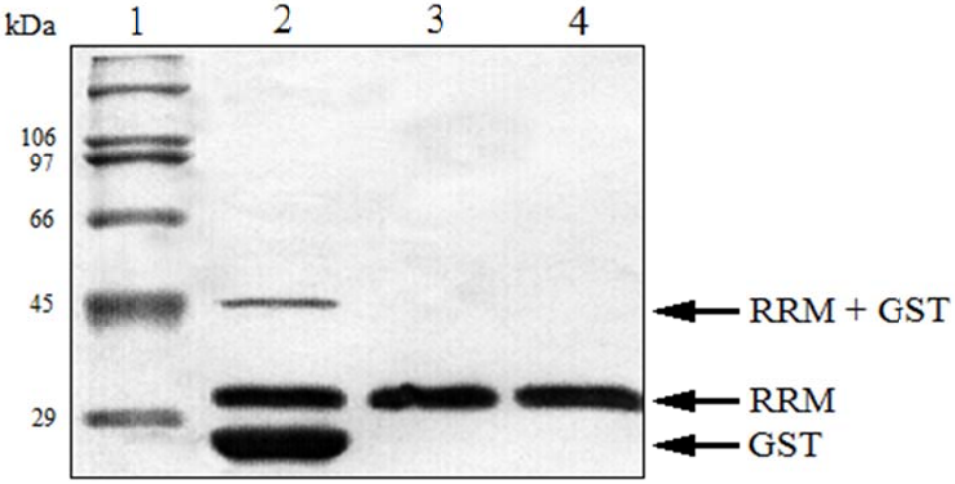
Figure 3.2 Not4 RRM protein expression and purification

A.



A. RRM expression in *E. coli* BL21-DE3 pLysS cells. Culture cells were sampled before IPTG induction (lane 2) and after IPTG induction (lane 3). GST protein purification and extraction was performed on the IPTG induced cell culture and the supernatant was sampled (lane 4). The protein was bound to Glutathione Sepharose4 Fast Flow beads (lane 6). The flow-through from the beads shows loss of protein (lane 5). A high molecular weight marker is shown in lane 1.

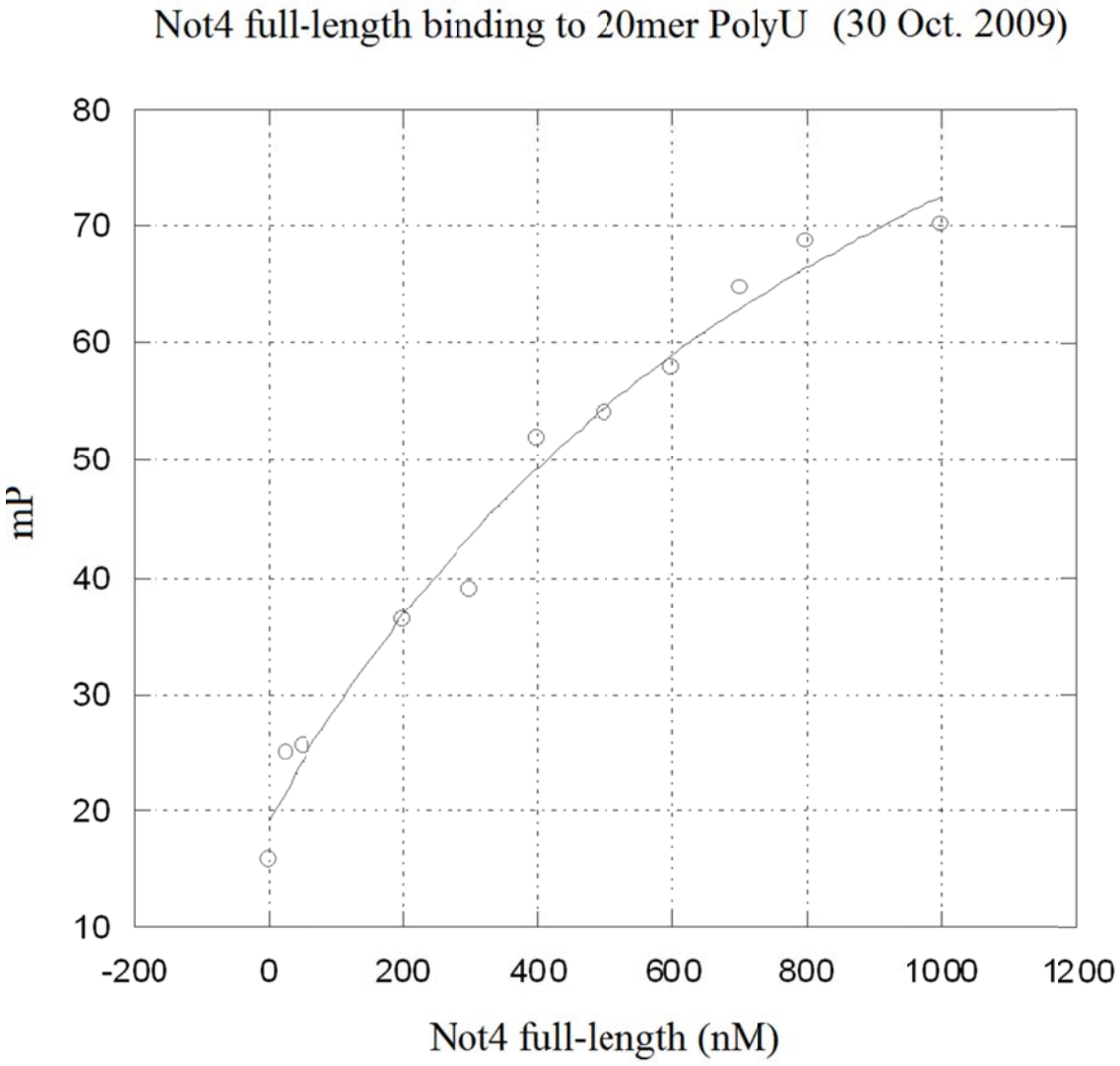
B.



B. PreScission protease digestion of RRM. After 20 hours of digestion, the supernatant (lane 3) and one wash (lane 4) were collected. The remaining beads showed retention of the RRM (lane 2). A high molecular weight marker is shown in lane 1.

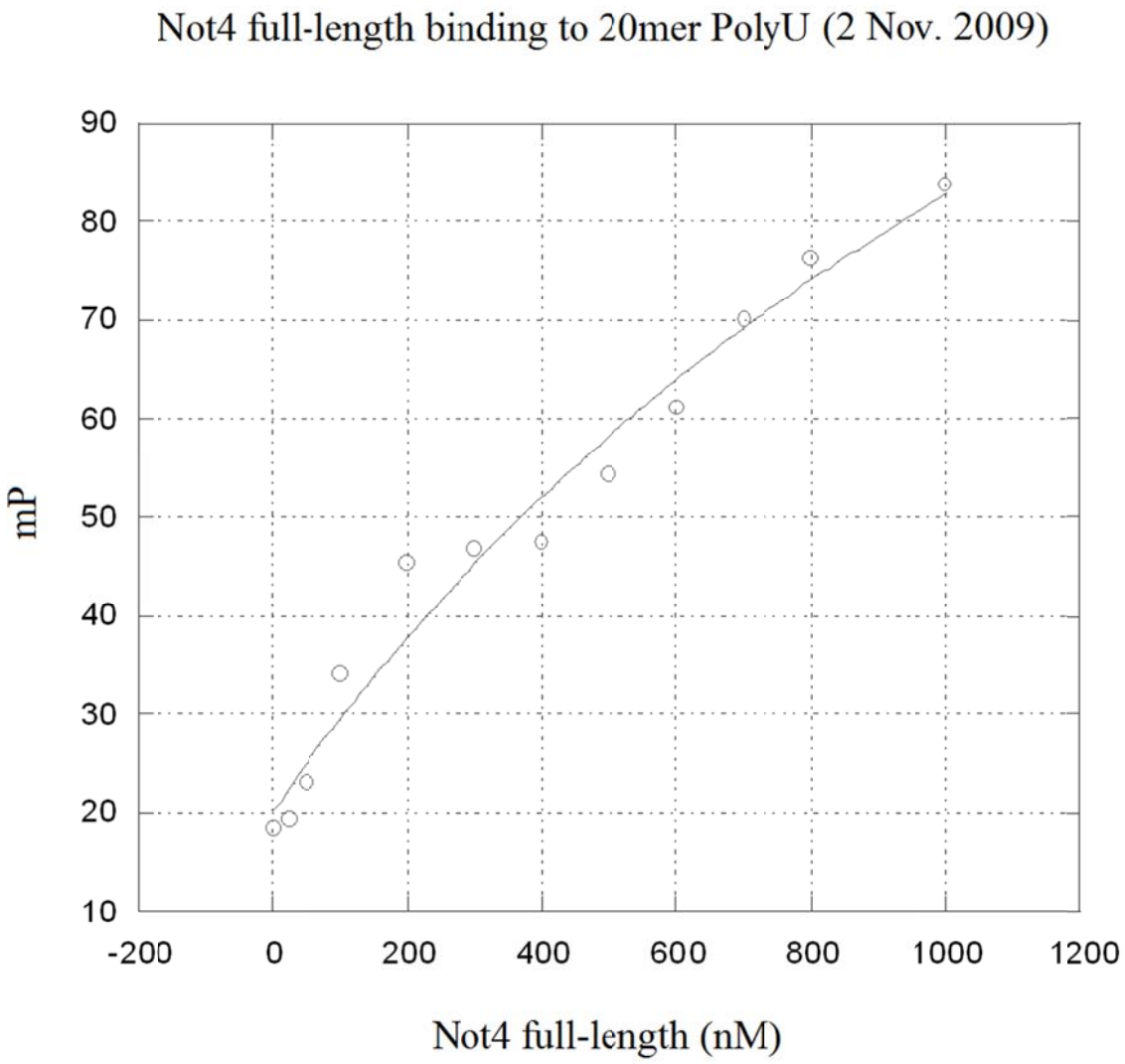
Figure 3.3 Not4 full-length RNA binding assay graphs

A.



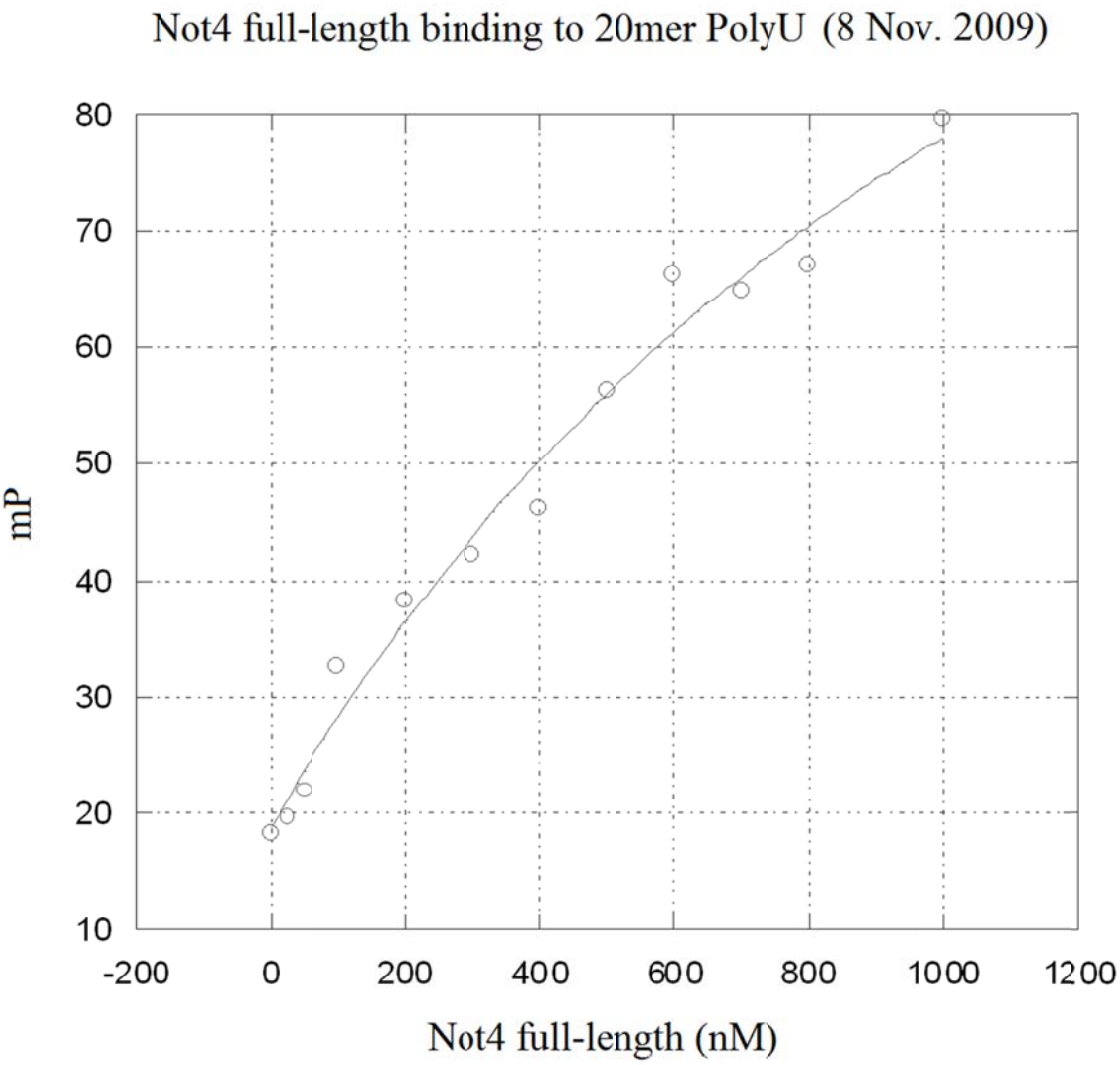
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	109.04	22.492
m2	1044.6	389.99
m3	19.172	1.9058
Chisq	66.612	NA
R	0.99071	NA

B.



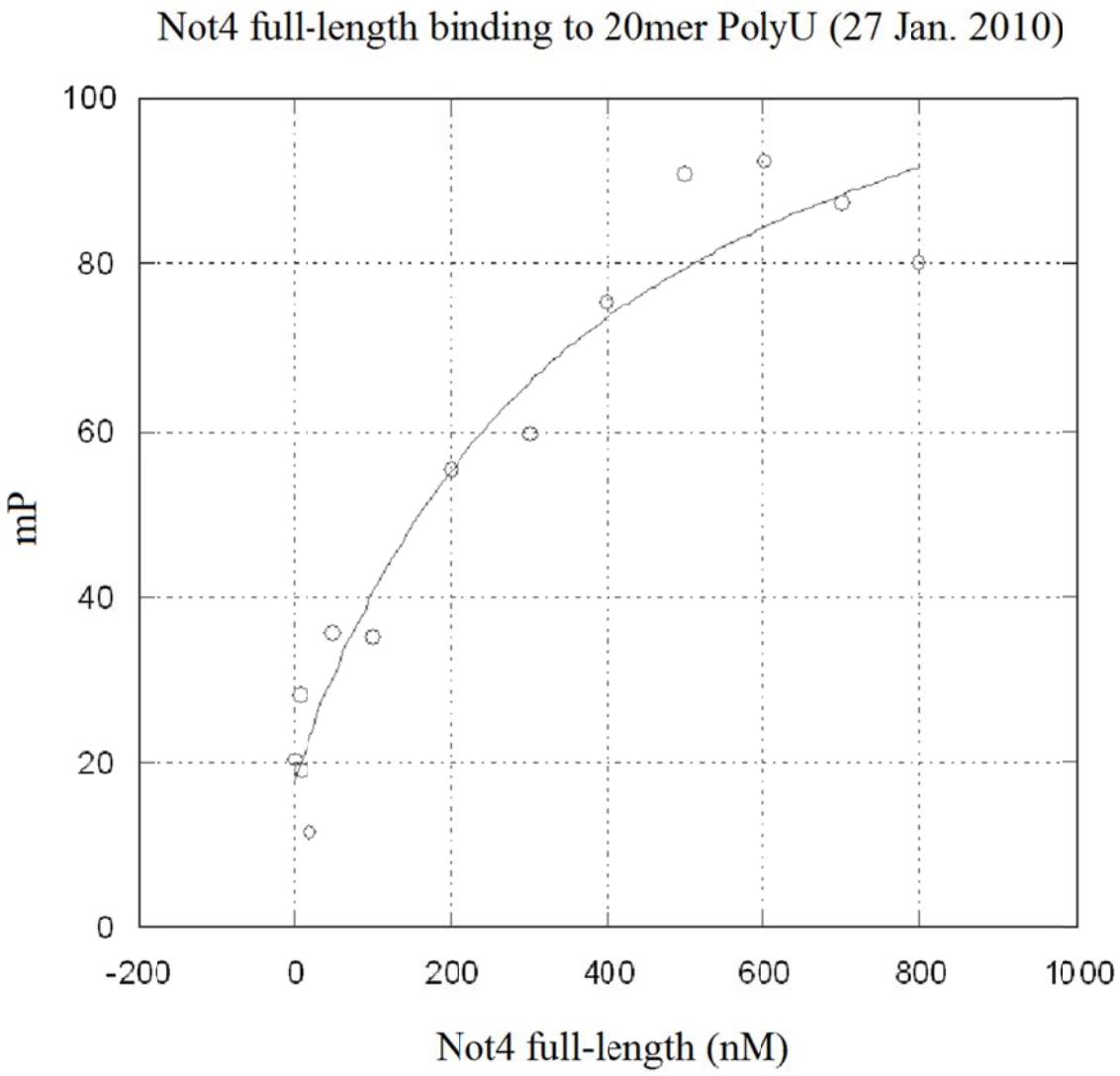
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	174.35	71.158
m2	1787.1	1087.4
m3	20.162	2.5054
Chisq	152.78	NA
R	0.98542	NA

C.



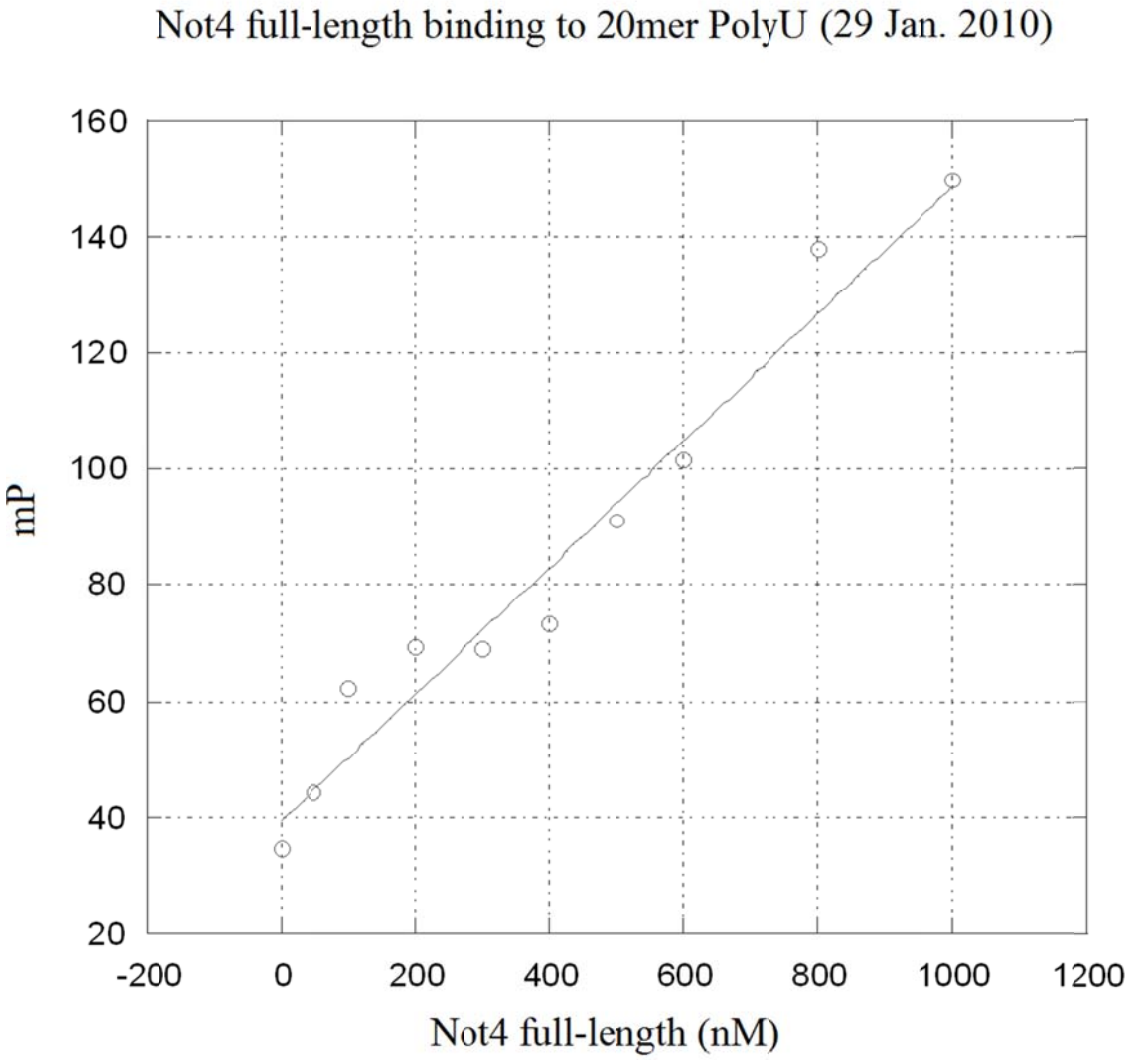
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	144.43	38.126
m2	1433.4	608.25
m3	18.7	1.9209
Chisq	86.84	NA
R	0.99082	NA

D.



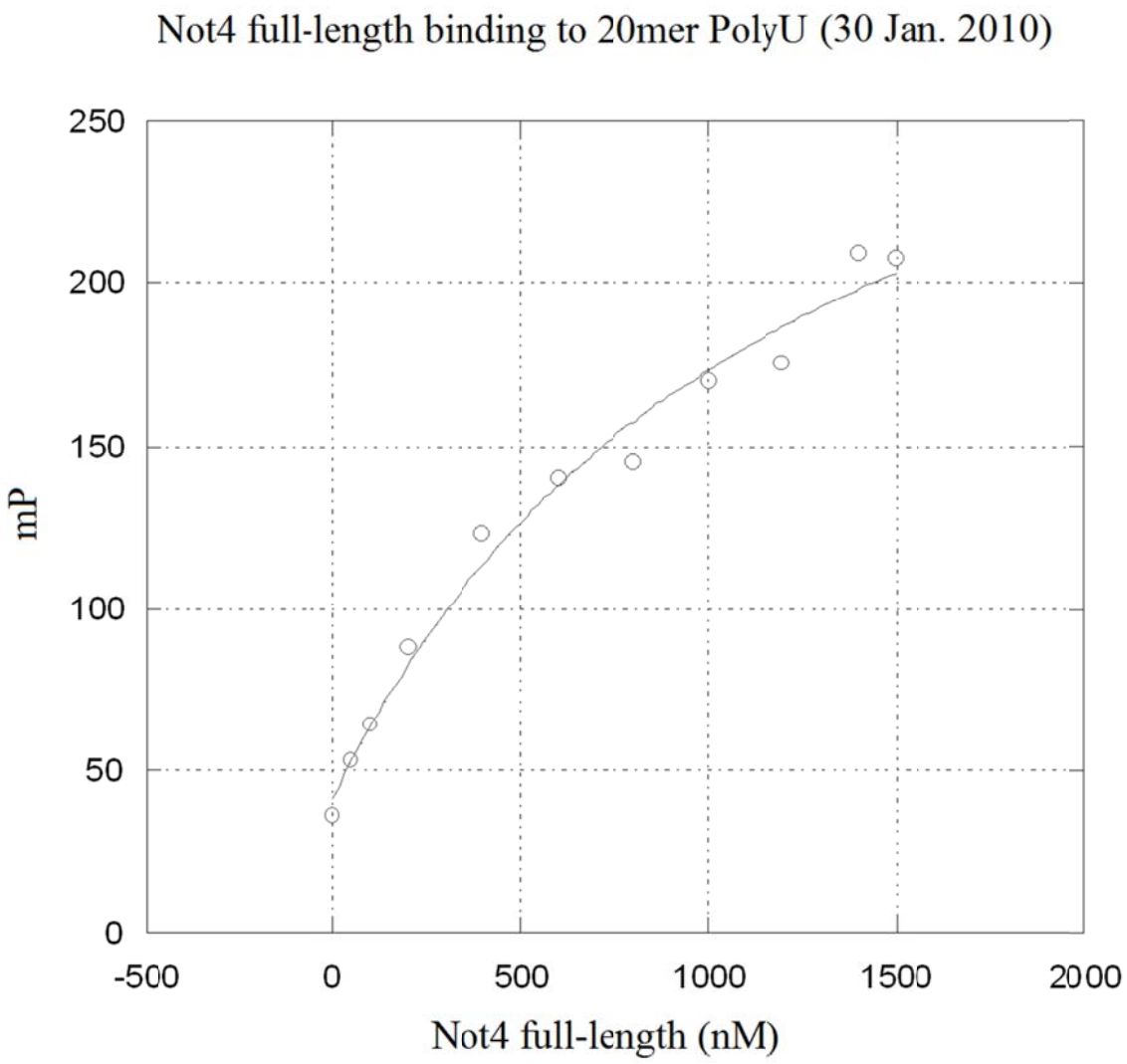
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	109.34	22.062
m2	384.62	202.8
m3	17.768	4.2963
Chisq	643.55	NA
R	0.96935	NA

E.



y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	3.0083e+13	9.3877e+19
m2	2.7527e+14	8.5945e+20
m3	39.376	4.1773
Chisq	474.84	NA
R	0.98104	NA

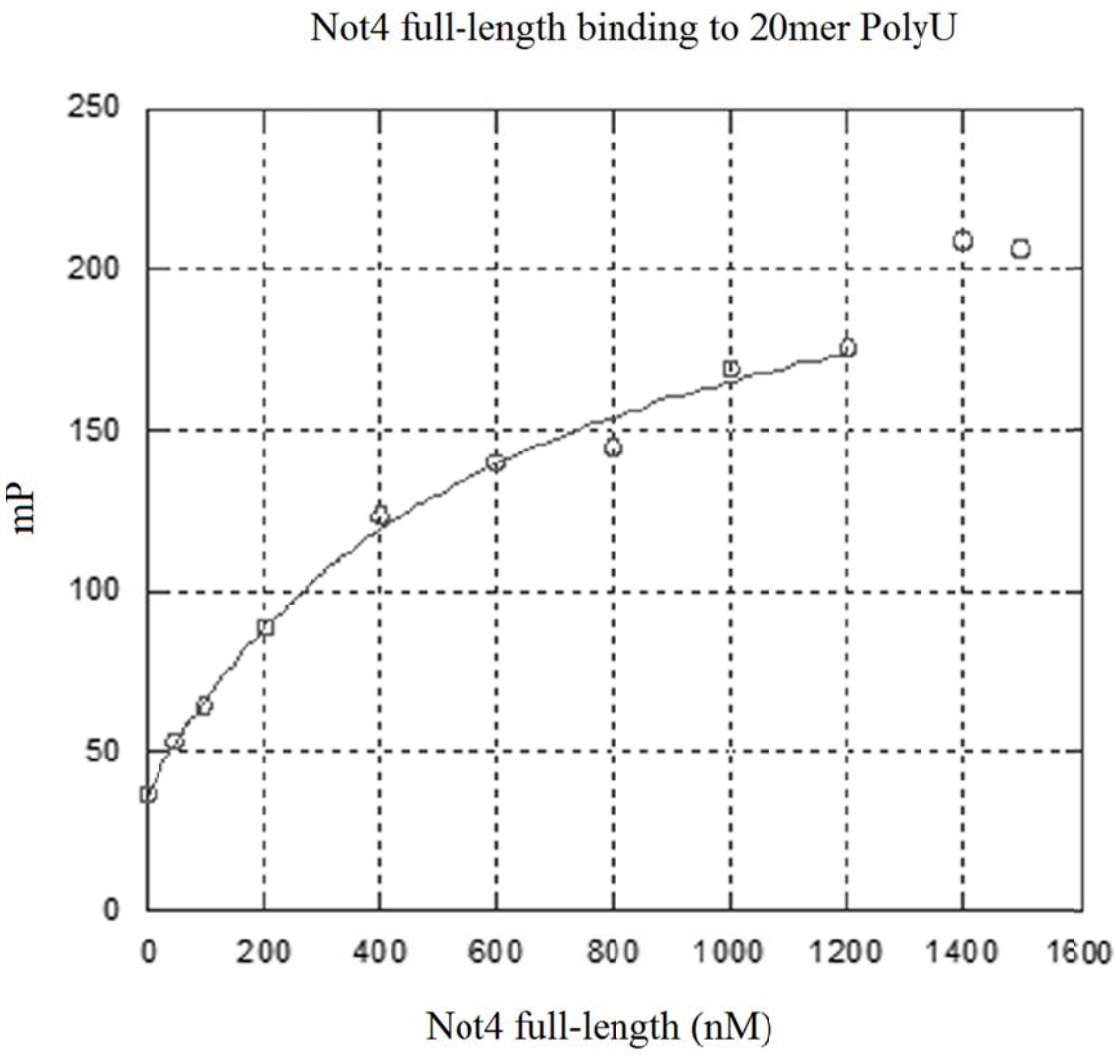
F.



y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	292.86	45.748
m2	1221.2	412.35
m3	41.474	6.0895
Chisq	604.09	NA
R	0.99182	NA

Figure 3.3 Not4 full-length RNA binding assay graphs. [A] Fluorescence polarization was used to determine the RNA binding activity of Not4. Assays were performed by combining 0.1 nM 3'-fluorescein-labeled polyU RNA with increasing concentrations of the Not4 full-length protein and monitoring the change in the fluorescence polarization mP value. The K_d dissociation constant for Not4 full-length binding to RNA was determined by plotting fluorescence polarization as a function of protein concentration and fitting the experimental data to a rectangular hyperbola with the following equation: y=(m1*m0)/(m2+m0)+m3, where m0 is the protein concentration independent variable, m1 is the maximum fluorescence polarization mP value, m2 is the K_d dissociation constant, m3 is the y-axis shift, and y is the fluorescence polarization mP dependent variable. [B] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [C] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [D] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [E] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [F] Same as A except that the binding assay was performed again on a different date with a separate protein preparation.

Figure 3.4 Re-plotted Not4 full-length binding to 20mer polyU

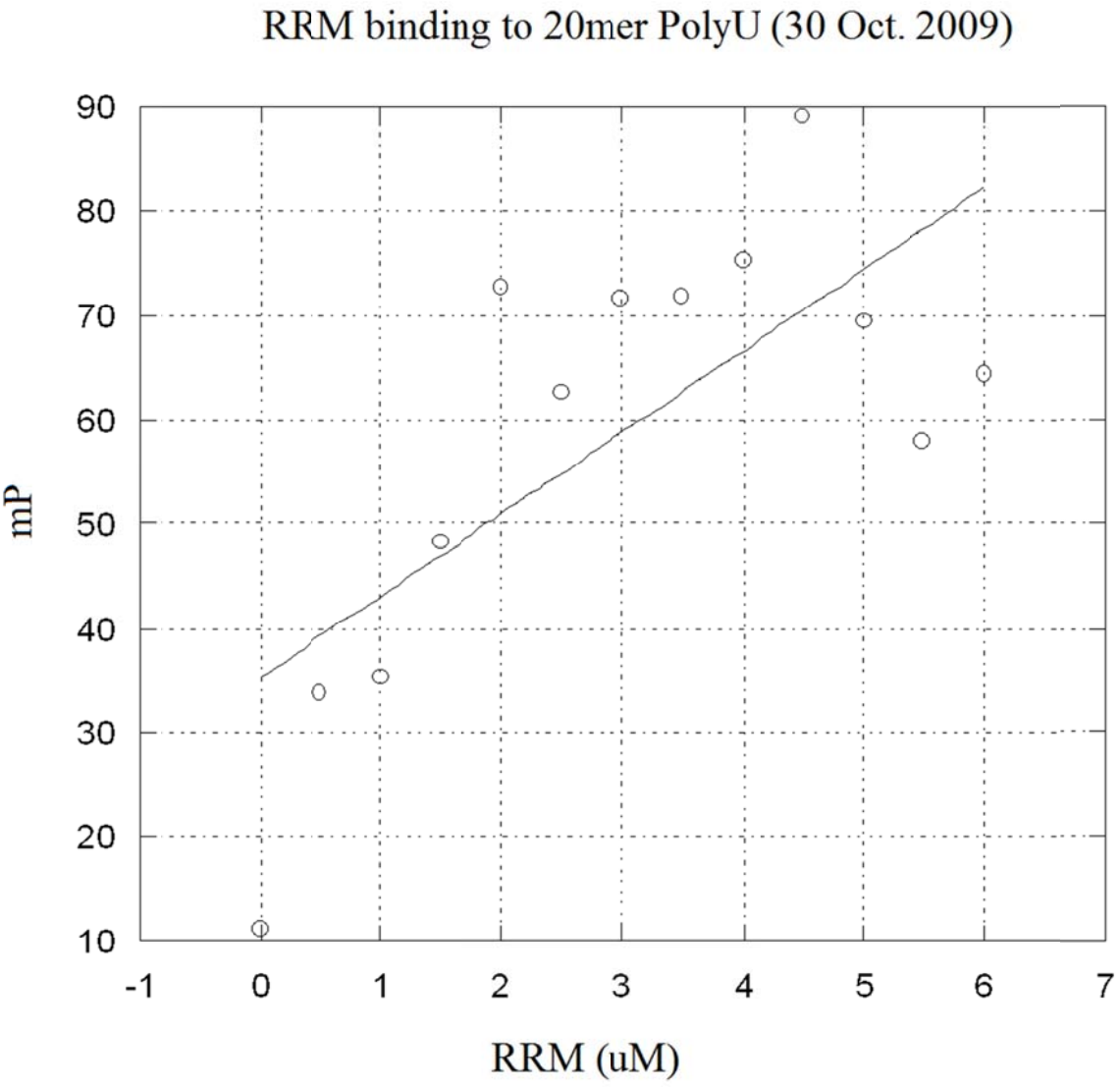


y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	205.06	14.132
m2	594.46	114.28
m3	36.329	3.6336
Chisq	124.46	NA
R	0.99707	NA

Figure 3.4 Figure 3.3 [F] re-plotted for Not4 full-length binding to 20mer polyU. The binding assay graph shown in Figure 3.3 F with a K_d value of 1221.2 ± 412.35 was re-plotted with the two terminal data points removed. The K_d value determined from this re-adjusted plot is 594.46 ± 114.28.

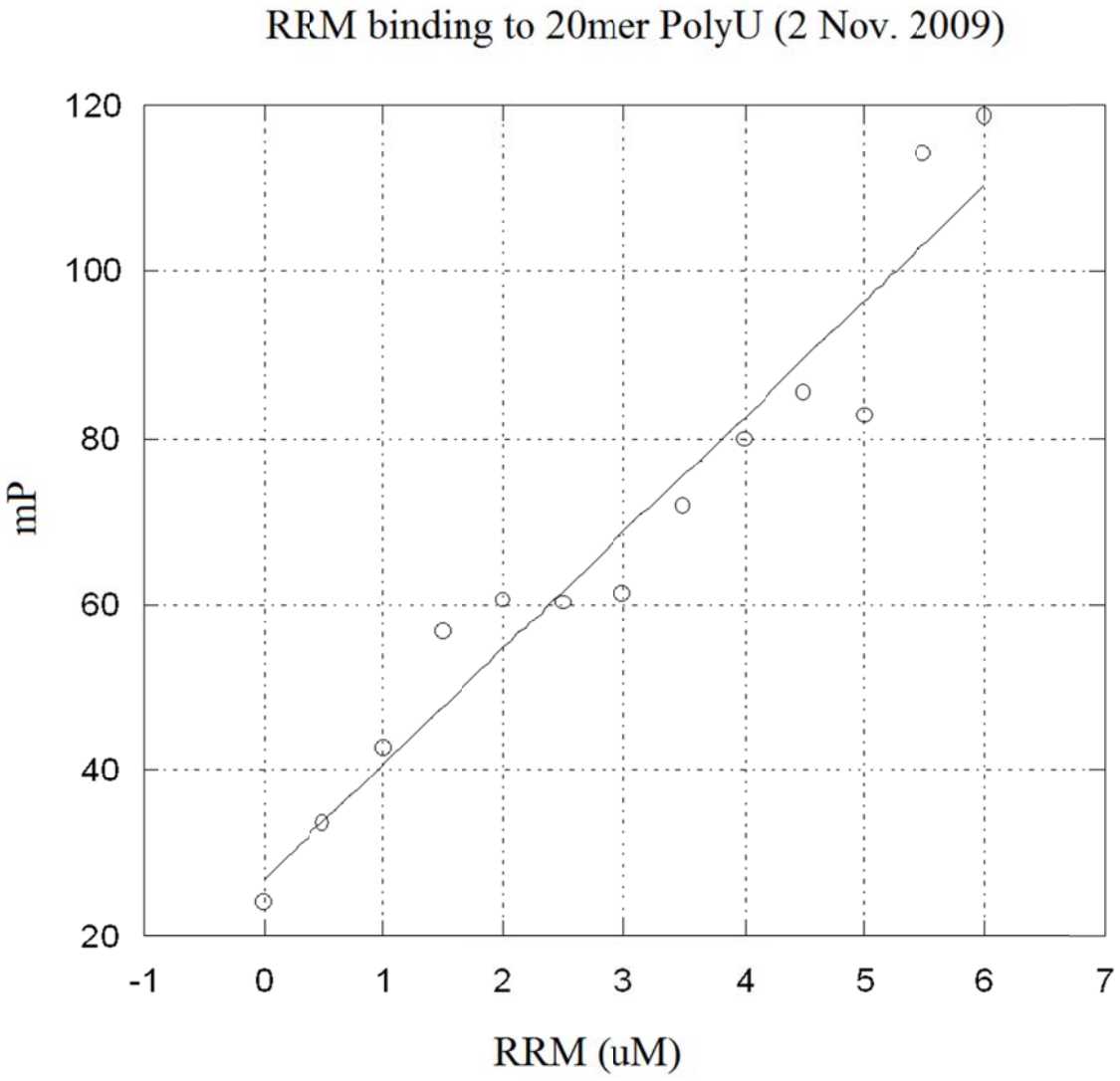
Figure 3.5 RRM RNA binding assay graphs

A.



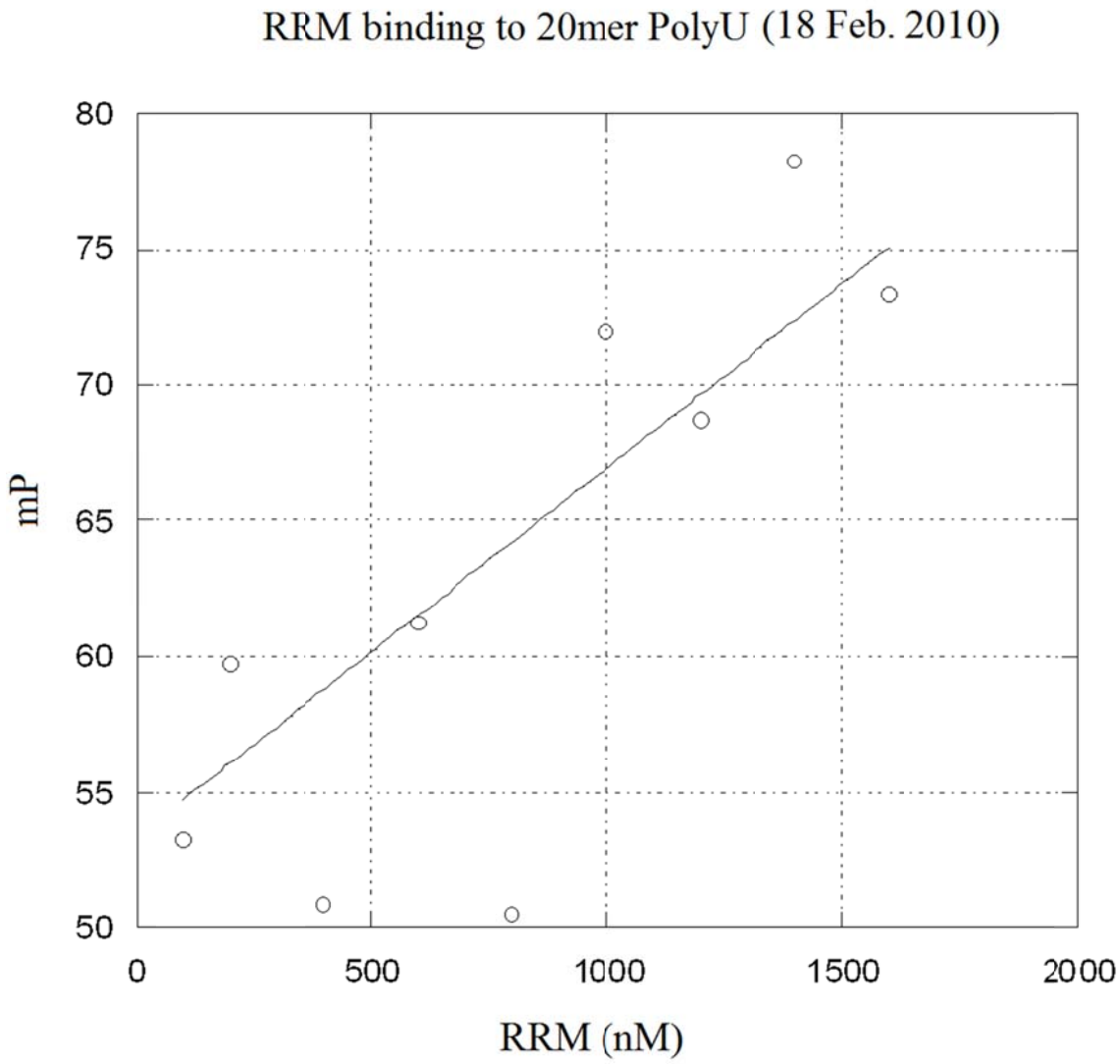
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	-3.1803e+40	3.6319e+47
m2	-4.0659e+39	4.6455e+46
m3	35.303	8.53
Chisq	2648.5	NA
R	0.71586	NA

B.



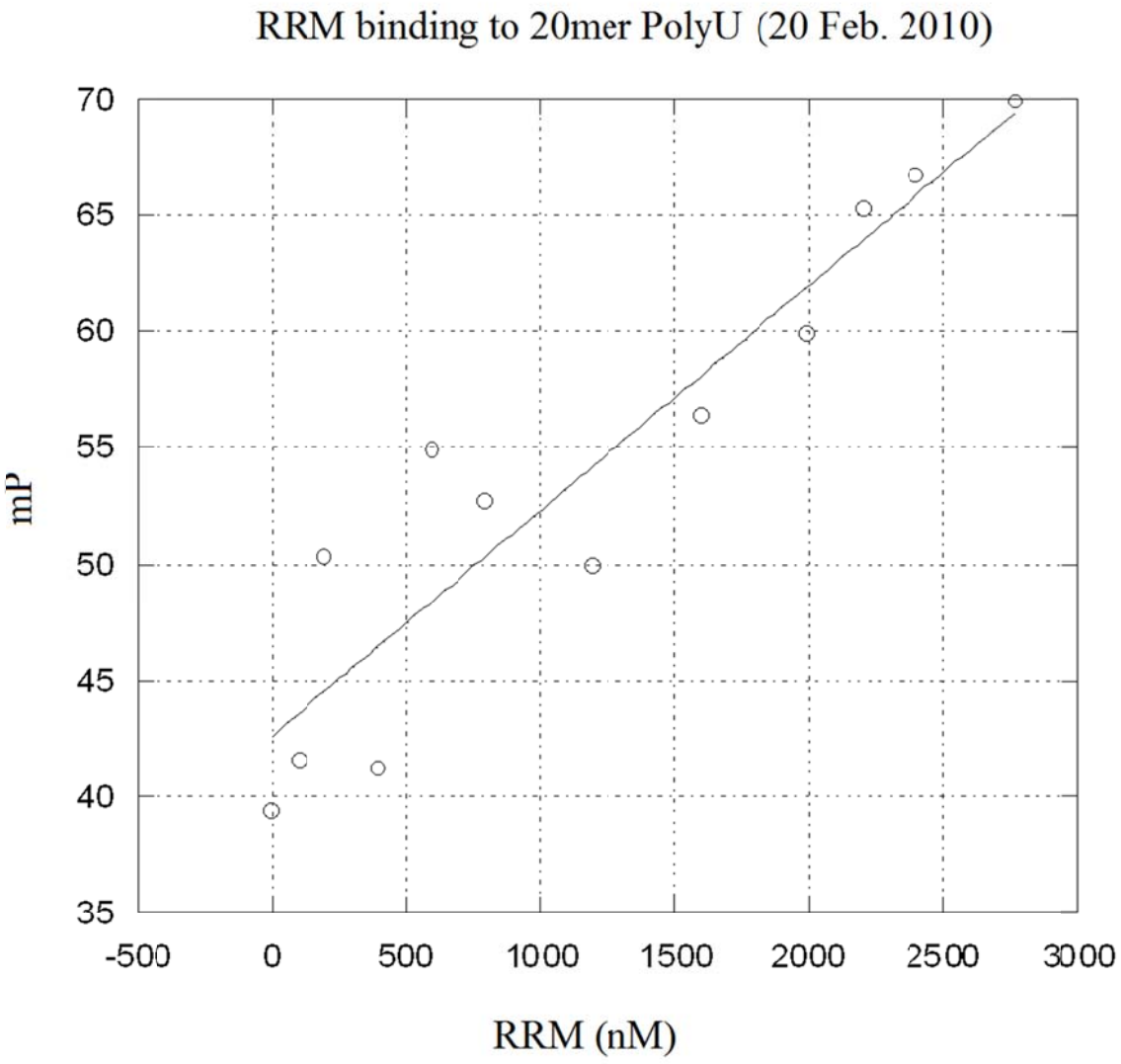
y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	3.1335e+12	3.6008e+22
m2	2.2517e+11	2.5887e+21
m3	26.797	0
Chisq	594.97	NA
R	0.96786	NA

C.



y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	19416	2.4231e+07
m2	1.4281e+06	1.7852e+09
m3	53.356	7.0533
Chisq	330.48	NA
R	0.78583	NA

D.



y = (m1*m0)/(m2+m0)+m3		
	Value	Error
m1	3.6462e+10	1.5016e+17
m2	3.7627e+12	1.5504e+19
m3	42.597	1.9217
Chisq	152.58	NA
R	0.9314	NA

Figure 3.5 RRM protein RNA binding assay graphs. [A] Fluorescence polarization was used to determine the RNA binding activity of the RRM of Not4. Assays were performed by combining 0.1 nM 3'-fluorescein-labeled polyU RNA with increasing concentrations of the RRM derivative protein and monitoring the change in the fluorescence polarization mP value. The K_d dissociation constant for RRM binding to RNA was determined by plotting fluorescence polarization as a function of protein concentration and fitting the experimental data to a rectangular hyperbola with the following equation: y=(m1*m0)/(m2+m0)+m3, where m0 is the protein concentration independent variable, m1 is the maximum fluorescence polarization mP value, m2 is the K_d dissociation constant, m3 is the y-axis shift, and y is the fluorescence polarization mP dependent variable. [B] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [C] Same as A except that the binding assay was performed again on a different date with a separate protein preparation. [D] Same as A except that the binding assay was performed again on a different date with a separate protein preparation.

Figure 3.6

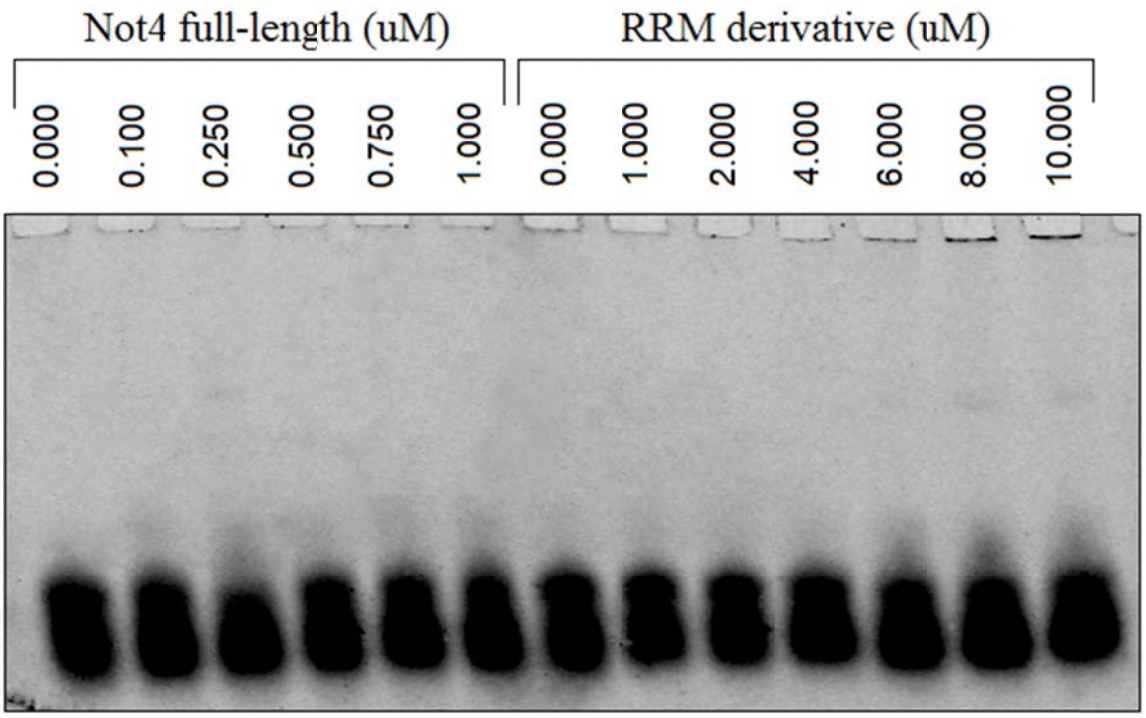


Figure 3.6 Electrophoretic mobility shift assay with 20mer polyU RNA. The electrophoretic mobility shift assay (EMSA) was performed with polyU RNA and increasing concentrations of either the Not4 full-length or the Not4 RRM derivative proteins. The samples were run on a 5% gel at 100 V for 1 hour. The gel scanned with the Molecular Dynamics Typhoon system and a Fluorescence 526 SP emission filter. ImageQuant was used to view the scan.

Table 3.1 Summary of dissociation constants for Not4 full-length binding to RNA

	K_d dissociation constant (nM)	R-value
Figure 3.3		
A	1044.6 ± 389.99	0.99071
B	1787.1 ± 1087.4	0.98542
C	1433.4 ± 608.25	0.99082
D	385.62 ± 202.8	0.96935
E	no binding	
F	1221.2 ± 412.35	0.99182
Figure 3.4		
Figure 3.3 [F] re-plotted	594.46 ± 114.28	0.99707

Table 3.1 Summary of dissociation constants for Not4 full-length binding to RNA. RNA binding assay results for the Not4 full-length K_d dissociation constant is listed. The curve fit R-value for each graph is listed as well.

CHAPTER 4

DISCUSSION

4.1 Role of RNA binding in Not4 E3 ubiquitin ligase activity

In this project, we report that the Not4 E3 ubiquitin ligase has RNA binding activity. Previous studies have demonstrated the existence of RNA-binding E3 ubiquitin ligases that bind to and regulate RNA stability (Cano et al., 2010). Comparative genomics have also confirmed many proteins in animals, plants and fungi containing both ubiquitination and RNA-binding domains (Lucas et al., 2005). Although it is not clear what role RNA binding plays in E3 ubiquitin ligase activity, it is interesting to note that the regulation of protein and mRNA stability may be linked through the ubiquitin system. In light of the role that Ccr4-Not complex plays as a global regulator of gene expression (Figure 1.5), the result that the Not4 subunit of the complex is an E3 ubiquitin ligase with RNA binding activity does not come as a contradiction to the current understanding of the function of the subunit or the protein complex.

4.2 Not4 full-length shows RNA binding activity

Our results from *in vitro* fluorescence polarization binding assays show that Not4 binds RNA at a K_d of approximately 594.46 ± 114.28 nM (Figure 3.4). In comparison to TcUBP1, a well-characterized RRM protein, this K_d value is demonstrative of RNA binding. TcUBP1, a trypanosome cytoplasmic RNA-binding protein involved in the

selective destabilization of U-rich mRNAs, contains a single, conserved RRM domain (Volpon et al., 2005) and binds RNA at a K_d of 14300 ± 0.1 nM

We suggest that the RNA binding activity is important to the function of Not4 and the Ccr4-Not complex. On one hand, Not4 contains an RRM, which is highly conserved throughout evolution and is known to bind RNA sequences. On the other hand, Not4 is a part of the larger Ccr4-Not complex, which is involved in nuclear transcription and cytoplasmic mRNA degradation (Liu et al., 1998; Collart and Timmers, 2004; Maillet and Collart, 2002; Tucker et al., 2001). The RNA binding activity of Not4 may very well support the activity of the Ccr4-Not complex, although it is not yet clear what such a role may be.

The results of our fluorescence polarization assays, although indicative of RNA binding, show variable Not4-RNA binding strengths. K_d dissociation constants range from 385.62 ± 202.8 (Figure 3.3 E) to 1787.1 ± 1087.4 nM (Figure 3.3 B). The K_d determined from the selected representative binding assay (Figure 3.4) is 594.46 nM. This K_d varies 2 – 3-fold with different protein preparations, similar to the variability noted in other RNA – protein systems (Hall and Stump, 1992). We believe these discrepancies may be due to variations in the quality of the protein used for the RNA binding assays and to imprecision in the estimation of the protein concentration due to break-down products. Protein quality may be affected in the process of protein induction, purification and concentration and in protein storage time and temperature. Additionally, binding assays

conditions may affect both protein stability as well as RNA binding affinity, leading to variable K_d dissociation constants.

Of the binding assay, we analyzed the plot with the best hyperbolic fit, or highest R-value (Figure 3.3 F). Figure 3.3 F was re-fitted with the terminal two data points removed (Figure 3.4). We believe that the deviation of the terminal two data points at 1400 nM and 1500 nM Not4 protein concentrations mark the beginnings of protein aggregation and non-specific binding due to the increase in protein concentration. A comparison between the original fit and the adjusted fit shows a two-fold decrease in K_d from 1221.2 ± 412.35 to 594.46 ± 114.28 nM and an small increase in R-value from 0.99182 to 0.99707 (Table 3.1).

20mer 3'-fluorescein-labeled polyU RNA was used in the RNA binding assays because previous studies have shown that RNA-binding proteins show higher binding affinity for polypyrimidines than polypurines (Albo et al., 1995) and that RRMs can bind polyU RNA with an affinity ten-fold higher than nonspecific substrate polyA (Pérez et al., 1997).

4.3 Not4 RRM does not show RNA binding activity

The RRM is approximately 90 amino acids in length (Birney et al., 1993) and its domain structure is characterized by two α-helices packed on a four-stranded β-sheet with the β-sheets acting as the primary RNA binding surface (Figure 1.9) in a $\beta_1 - \alpha_1 - \beta_2 - \beta_3 - \alpha_2 - \beta_4$ sequence. The RRM is an evolutionarily conserved domain, but it has only a few well-

conserved residues. These conserved residues reside mostly in the RNP-1 (octomer) and RNP-2 (hexamer) submotifs with the RNP-1 being the most obvious signature for the typical RRM (Birney et al., 1993). The RNP-1 and RNP-2 submotifs reside in the central anti-parallel β_3 and β_1 sheets respectively (Birney et al., 1993). While there is no universal consensus sequence for the RRM, the RNP consensus sequences have been determined. The RNP-1 consensus sequence is defined as Lys / Arg-Gly-Phe / Tyr-Gly / Ala-Phe / Tyr-Val / Ile / Leu-X-Phe / Tyr and the RNP-2 consensus sequence is defined as Ile / Val / Leu-Phe / Tyr-Ile / Val / Leu-X-Asn-Leu, where X can be any amino acid (Maris et al., 2005).

While some RRMs are capable of forming a globular domain that is able to independently bind RNA, RNA – binding usually requires a synergy between two or more RRMs in a multi-domain protein (Birney et al., 1993). In eukaryotic proteins, RRMs will appear as multiple copies 44% of the time and together with other domains 21% of the time (Maris et al., 2005). Neighboring RRM domains and other protein domains influence the RNA-binding ability of the RRM by modulating its affinity and specificity (Maris et al., 2005).

The presence of an RRM in protein is strongly suggestive of a functional role in RNA binding. Furthermore, because the RRM domain of the human counterpart of the Not4 protein, hNot4, has been shown to be involved in the binding of single-stranded nucleic acids (reviewed in Nagai et al., 1995), we hypothesized that the RNA binding activity of Not4 occurs through the RRM region in the protein. After fluorescence polarization

confirmed the RNA binding activity of Not4, we further analyzed the Not4 RRM in *in vitro* RNA binding assays. However, despite greater protein purity in RRM expression as compared to the Not4 protein, no RNA binding was detected through either fluorescence polarization (Figure 3.5) or EMSA for the Not4 RRM derivative (Figure 3.6).

Absence of Not4 RRM RNA binding activity may be due to several factors. It is possible that the K_d value for the RRM domain had not been reached at the protein concentration range used (0.0 – 10.0 μM). This possibility is unlikely when the compared to the K_d dissociation constants of the polypyrimidine tract binding protein (PTB) RRMs, which are in the 0.010 – 0.020 μM range (Liu et al., 2002). However, PTB contains four RRMs. The multiple RRMs may synergistically increase the binding affinity of PTB for RNA and account for the lower K_d, suggesting the possibility that additional Not4 domains are needed for proper RRM domain folding, stabilization and RNA binding.

It is known that a combination of two or more RRM domains will often notably increase RNA binding affinity by lengthening the RNA binding platform of the protein and that most RNA binding proteins do not solely rely on one RRM domain for RNA binding (Maris, 2005). Full-length PTB protein binds RNA with dissociation constants in the range of 0.03 – 0.06 μM. Truncation of PTB to include just the third and fourth RRM domains in the C-terminal reduces RNA binding by approximately seven-fold (Conte et al., 2000). Thus, RRM deletion constructs of PTB retain RNA binding activity at a lower binding affinity (Pérez et al., 1997).

Not4, however, only contains one RRM domain. Not4 also contains one RING-type zinc finger and one C3H1-type zinc finger. Interestingly, in eukaryotic proteins, RRMs are found in 21% of all cases together with other protein domains, the most common of which are the C3H1 and C2HC zinc finger (Maris et al., 2005). Hence, it may be that additional Not4 domains and secondary structures are needed for proper RRM domain folding, stabilization and RNA binding.

A third possibility is that Not4 RNA binding occurs in a region of the protein other than the RRM domain. A variety of domains other than the RRM enable proteins to bind RNA. Such domains include the K Homology (KH) domain, Zinc finger (mainly C-x8-C-x5-C-x3-H type), RGG box, DEAD/DEAH box, Pumilio/FBF (PUF) domain, double-stranded RNA binding domain (DS-RBD), Piwi/Argonaute/Zwille (PAZ) domain, and the Sm domain (Lee and Schedl, 2006). In a recent study of the translocated in liposarcoma (TLS) RNA-binding protein, it was shown that the zinc finger domain of TLS binds RNA with a dissociation constant of approximately 10 μM (Iko et al., 2004). TLS, like Not4, contains both an RRM and a zinc finger domain. Interestingly, the RRM domain of TLS showed no observable interaction with RNA, suggesting that the zinc finger domain is more important for RNA recognition and binding than the RRM domain (Iko et al., 2004). Thus, Not4 may very well rely on a protein domain other than the RRM to bind RNA.

4.4 Summary

Not4 is an E3 ubiquitin ligase with RNA binding activity. It has been suggested that the RNA binding activity of E3 ubiquitin ligases helps to regulate RNA stability. While it is known that Not4 contains an RRM domain, we were not able to show that Not4 binding to RNA occurs through the RRM domain. Further research will answer these questions and open up venues towards a better understanding of the properties and functions of Not4.

BIBLIOGRAPHY

Albert, T.K., Lemaire, M., Berkum, N.L., Gentz, R., Collart, M.A., and Timmers, H.T.M. (2000). Nucleic Acids Res *28*, 809-817.

Alberts, B., Bray, D., Hopkin, K., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. Essential Cell Biology. 2nd edition (2003) New York, Garland Science Publishing.

Albo, C., Valencia, A., and Portela, A. (1995). Identification of an RNA binding region within the N-terminal third of the influenza A virus nucleoprotein. J Virol *69*, 3799-3806.

Bai, Y., Salvadore, C., Chiang, Y.C., Collart, M.A., Liu, H.Y., and Denis, C.L. (1999). The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. Mol Cell Biol *19*, 6642-6651.

Birney, E., Kumar, S., and Krainer, A.R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res *21*, 5803-5816.

Campbell, N.A., Reece, J.B., Molles, M., Urry, L.A., and Heyden, R. Biology. 7th edition (2005) New York, Pearson Publishing.

Cano, F., Miranda-Saavedra, D., and Lehner, P.J. (2010). RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. Biochem Soc Trans *38*, 1621-1626.

Chen, J., Rappsilber, J., Chiang, Y.C., Russell, P., Mann, M., and Denis, C.L. (2001). Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex. J Mol Biol *314*, 683-694.

Collart, M.A. (2003). Global control of gene expression in yeast by the Ccr4-Not complex. Gene *313*, 1-16.

Collart, M.A., and Timmers, H.T. (2004). The eukaryotic Ccr4-not complex: a regulatory platform integrating mRNA metabolism with cellular signaling pathways? Prog Nucleic Acid Res Mol Biol *77*, 289-322.

Conte, M.R., Grüne, T., Ghuman, J., Kelly, G., Ladas, A., Matthews, S., and Curry, S. (2000). Structure of tandem RNA recognition motifs from polypyrimidine tract binding protein reveals novel features of the RRM fold. The EMBO J *19*, 3132-3141.

Denis, C.L., and Chen, J. (2003). The CCR4-NOT complex plays diverse roles in mRNA metabolism. Prog Nucleic Acid Res Mol Biol *73*, 221-250.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol *8*, 9-22.

Fox, B.G. and Blommel, P.G. (2009). Autoinduction of protein expression. John Wiley & Sons, Inc. Curr Protoc Protein Sci *56*, 5.23.1-5.23.18.

Gene: MOT2 (YER068W). Ensembl Fungi. 8 Feb. 2011. EMBL-EBI. 5 Apr. 2011 <http://www.ebi.ac.uk/s4/redirect/aHR0cDovL3d3dy5lYmkuYWMudWsvczQvZXllcmVzdWx0Lz9pbmRleD0zJnNwZWNpZXM9U2FjY2hhcm9teWNlc19jZXJldmlzaWFlJnRlcm09Tm90NCZ0aWQ9c3luRU5TTVVTRzAwMDAwMDM4Nzg0/56390?url=http://fungi.ensembl.org/Saccharomyces_cerevisiae/Gene/Summary?g%3DYER068W%3Bdb%3Dcore>.

Gingras, A.C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu Rev Biochem *68*, 913-963.

Goldstrohm, A.C., and Wickens, M. (2008). Multifunctional deadenylase complexes diversify mRNA control. Nature Reviews Molec Cell Biol *9*, 337-344.

Hall, K.B. and Stump, W.T. (1992). Interaction of N-terminal domain of U1A protein with an RNA stem/loop. Nucleic Acids Res *20*, 4283-4290.

Hanzawa, H., de Ruwe, M.J., Albert, T.K., van Der Vliet, P.C., Timmers, H.T., and Boelens, R. (2001). The structure of the C4C4 ring finger of human NOT4 reveals features distinct from those of C3HC4 RING fingers. J Biol Chem *276*, 10185-10190.

Iko, Y., Kodama, T.S., Kasai, N., Oyama, T., Morita, E.H., Muto, T., Okumura, M., Fujii, R., Takumi, T., Tate, S., and Morikawa, K. (2004). Domain architectures and characterization of an RNA-binding protein, TLS. J Biol Chem *279*, 44834-44840.

Lablife: pGEX-2T. addgene. 2011. Lablife. 5 Apr. 2011 <http://www.lablife.org/p?a=vdb_view&id=g2.fpzDkSJdEuqykKDjECRqRQzAQdI->.

Laribee, R.N., Shibata, Y., Mersman, D.P., Collins, S.R., Kemmeren, P., Roguev, A., Weissman, J.S., Briggs, S.D., Krogan, N.J., and Strahl, B.D. (2007). Ccr4/Not complex associates with the proteasome and regulates histone methylation. Proc Natl Acad Sci U S A *104*, 5836-5841.

Lee, M.H. and Schedl, T. (April 18, 2006). RNA-binding proteins. *WormBook*, ed. The *C. elegans* Research Community, WormBook, doi/10.1895/wormbook.1.79.1, http://www.wormbook.org.

Liu, H., Zhang, W., Reed, R.B., Liu, W., and Grabowski, P.J. (2002). Mutations in RRM4 uncouple the splicing repression and RNA-binding activities of polypyrimidine tract binding protein. RNA *8*, 137-149.

Liu, H.Y., Badarinarayana, V., Audino, D.C., Rappsilber, J., Mann, M., and Denis, C.L. (1998). The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. EMBO J *17*, 1096-1106.

Lucas, J.I., Arnau, V., and Marin, I. (2005). Comparative genomics and protein domain graph analysis link ubiquitination and RNA metabolism. J Molec Biol *357*, 9-17.

Maillet, L., and Collart, M.A. (2002). Interaction between Not1p, a component of the Ccr4-not complex, a global regulator of transcription, and Dhh1p, a putative RNA helicase. J Biol Chem *277*, 2835-2842.

Maillet, L., Tu, C., Hong, Y.K., Shuster, E.O., and Collart, M.A. (2000). The essential function of Not1 lies within the Ccr4-Not complex. J Mol Biol *303*, 131-143.

Maris, C., Dominguez, C., and Allian, F.H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J *272*, 218-2131.

Mersman, D.P., Du, H.N., Fingerman, I.M., South, P.F., and Briggs, S.D. (2009). Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression. Genes Dev *23*, 951-962.

Mulder, K.W., Brenkman, A.B., Inagaki, A., van den Broek, N.J., and Timmers, H.T. (2007). Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex. Nucleic Acids Res *35*, 2428-2439.

Nagai, K., Oubridge, C., Ito, N., Avis, J., and Evans, P. (1995) The RNP domain: a sequence-specific RNA-binding domain involved in processing and transport of RNA. Trends Biochem Sci *20*, 235-240.

Panasenko, O., and Collart, M.A. (2011). Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29. Mol Cell Biol. http://mcb.asm.org/cgi/reprint/MCB.01210-10v1.

Panasenko, O., Landrieux, E., Feuermann, M., Finka, A., Paquet, N., and Collart, M.A. (2006). The yeast Ccr4-Not complex controls ubiquitination of the nascent-associated polypeptide (NAC-EGD) complex. J Biol Chem *281*, 31389-31398.

Parker, R., and Sheth, U. (2007). P bodies and the control of mRNA translation and degradation. Mol Cell *25*, 635-646.

Pérez, I., McAfee, J.G., and Patton, J.G. (1997). Multiple RRMs contribute to RNA binding specificity and affinity for polypyrimidine tract binding protein. Biochemistry *36*, 11881-11890.

Sachs, A.B., Sarnow, P., and Hentze, M.W. (1997). Starting at the beginning, middle, and end: translation initiation in eukaryotes. Cell *89*, 831-838.

Sartorius Stedim Biotech. "Vivaspin 500 µl and 2 ml." Sep. 2008. Publication No.: SLU6093-e08093. 19 Mar. 2011 <http://www.sartorius-stedim.com.tw/Attachment/FCKeditor/Product/file/PDF/lab/viva/Manual_Vivaspin_500_2-SLU6093-e.pdf>.

Stefl, R., Skrisovska L., and Allain, F.H. (2005). RNA sequence- and shape-dependent recognition by proteins in the ribonucleoprotein particle. EMBO reports *6*, 33-38.

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

Volpon, L., D'Orso, I., Young, C.R., Frasch, A.C., and Gehring, K. (2005). NMR structural study of TcUBP1, a single RRM domain protein from *Trypanosoma cruzi*: contributino of a β hairpin to RNA binding. Biochem *44*, 3708-3717.

Woelk, T., Sigismund, S., Penengo, L., and Polo S. (2007). The ubiquitination code: a signalling problem. Cell Division *2*, 1-12.

Zlatanova, J. and Victor, J.M. (2009). How are nucleosomes disrupted during transcription elongation? HFSP J *3*, 373-378.

ACADEMIC VITA

Bonnie B. Lu
bxl210@psu.edu
814-769-0378

EDUCATION

The Pennsylvania State University, University Park, PA
B.Sc. in Biochemistry and Molecular Biology, May 2011
B.Sc. in Immunology and Infectious Disease, May 2011
B.Sc. in Toxicology, May 2011
Schreyer Honors College

RESEARCH EXPERIENCE

Honors Thesis Research
The Pennsylvania State University, University Park, PA, May 2009 – May 2011
Thesis advisor Dr. Joseph C. Reese
Thesis title "A biochemical analysis of the RNA binding activity of Not4 of the Ccr4-Not complex in yeast"

- Studied the RNA binding activity of the Not4 full-length protein and the Not4 RRM derivative
- Performed PCR, agarose gels, DNA purifications, mini-preps, transductions, cell cultures, fluorescence polarization assays, electrophoretic mobility shift assays

Immunology Research
The Pennsylvania State University, University Park, PA, January 2008 – August 2008
Research supervisor Dr. Na U. Xiong

- Genotyped mice ears and mice fetal tails through PCR and agarose gels
- Prepared stocks of cell culture plates and growth media

Fishery Sciences Research
The Pennsylvania State University, University Park, PA, May 2007 – September 2007
Research supervisor Dr. Paola C. Ferreri

- Maintained five 4000-gallon and over three hundred 50-gallon tanks of warm-water and African fish species
- Categorized freshwater macro-invertebrates according to genus and species for river soil sample analysis
- Electro-fished at local Centre County streams and gill-netted on the Ohio River to sample fish populations

WORK EXPERIENCE

Resident Assistant
The Pennsylvania State University, University Park, PA, August 2009 – May 2011
Supervisor Moses K. Davis

- Served as campus resource for students
- Provide monthly programs and community building activities
- Fostered sense of community among residents

Hospital Volunteer
Mount Nittany Medical Center, State College, PA, May 2009 – August 2010
Volunteer services supervisor Alice E. Clark

- Volunteered in Oncology Unit, Sterile Processing, Pharmacy, and Emergency Departments
- Trained new hospital volunteers
- Volunteered over 100 hours

Cerebral Palsy Rehabilitation Center Volunteer
Sunshine Homes, Shanghai, China, July 2010 – August 2010
Study abroad program director Justin O'Jack

- Led games and activities for 10 young adults
- Taught introductory English language classes
- Volunteered 15 hours

ORGANIZATIONS AND CERTIFICATIONS

Coordinating, Public Relations and Scheduling Chair
The Penn State Science Lions, The Pennsylvania State University, University Park, PA, August 2007 – May 2009

- Coordinated 38 Science Lions programs with 29 public schools and organizations
- Performed science demonstrations relating to electrostatic electricity and magnetism for grade-school children
- Created science-related activities for events

EMT-B certified
Pennsylvania Department of Health, December 2008 – December 2012

- CPR and AED certified by the American Heart Association
- Trained to assess patient condition, manage respiratory and cardiac emergencies, and provide first aid
- Shadowed Paramedics and EMTs for 30 hours on emergency calls

ACADEMIC AWARDS AND RECOGNITION

- **First Graduate,** highest graduating G.P.A., Immunology and Infectious Disease major, The College of Agricultural Sciences, The Pennsylvania State University (2011)

- **Standard Bearer,** Spring 2011 Commencement Ceremony, Immunology and Infectious Disease major, The College of Agricultural Sciences, The Pennsylvania State University (2011)

- **Braddock Scholarship,** four-year full tuition, room and board merit scholarship, The Eberly College of Science, The Pennsylvania State University (2007 – 2011)

- **Academic Excellence Scholarship,** $1750 per semester merit scholarship, The Schreyer Honors College, The Pennsylvania State University (2007 – 2011)

- **National Merit Finalist,** $2500 merit scholarship, National Merit Scholarship Corporation (2007)

- **Schreyer Honors College Travel Grant,** $750 education abroad scholarship, The Schreyer Honors College, The Pennsylvania State University (2010)

- **Whole World Scholarship,** $700 education abroad scholarship, The Pennsylvania State University (2010)

- **Dean's List,** maintained semester GPA of 3.50 or higher, The Pennsylvania State University (2008 – 2011)

LANGUAGE PROFICIENCY

English (native proficiency)
Mandarin Chinese (native proficiency)
Spanish (elementary proficiency)