Posttranscriptional Gene Regulation: From Global Models to Functional Mechanisms

by

Marshall Aaron Thompson

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Jack Keene, Supervisor

_____
Christopher Nicchitta

_____
Douglas Marchuk

_____
Fred Dietrich

_____
Lingchong You

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy in the University Program in
Genetics and Genomics in the Graduate School
of Duke University
2014

ABSTRACT

Posttranscriptional Gene Regulation: From Global Models to Functional Mechanisms

by

Marshall Aaron Thompson

University Program in Genetics and Genomics
Duke University

Date:_____

Approved:

_____
Jack Keene, Supervisor

_____
Christopher Nicchitta

_____
Douglas Marchuk

_____
Fred Dietrich

_____
Lingchong You

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the University Program in Genetics and Genomics in the
Graduate School of
Duke University
2014

Abstract

Precise regulation of the complex process of gene expression is essential for all aspects of life, and a large degree of this precision is mediated at the posttranscriptional level. The global and individual mechanisms by which posttranscriptional control is coordinated to maintain or alter levels of gene expression as necessary are not fully understood. Identification of the mRNA target sets of individual RNA binding proteins (RBPs) and characterization of the mechanisms by which RBPs regulate expression of individual mRNAs provide some insight into the global structure of the posttranscriptional environment. However, few studies have integrated these findings into a global model of posttranscriptional control. I have explored the structure and function of the posttranscriptional regulatory system through a combination of global modeling approaches, global studies of mRNA translation and decay, and mechanistic studies of the function of individual RBPs, specifically HuR and Pum1.

By combining RBP-mRNA association data and transcription factor (TF) target data from separate global studies in yeast, I developed an integrated model of gene expression regulation. Evaluation of this model indicates that posttranscriptional regulation may be responsible for substantially greater contributions to the overall gene expression program than transcriptional regulation. Further, I identified a self-regulatory feature of the posttranscriptional network that suggests a 'regulators of

regulators' structure may be a defining feature of posttranscriptional control of gene expression.

Additionally, I explored the mechanisms and functional consequences of dynamic association between an RBP, HuR, and its target RNAs through a combination of modeling and experimental approaches, including polysome profile analysis and global measurement of RNA stability. The model indicates that changes in total mRNA abundance are insufficient to fully explain the dynamics of association between HuR and its targets, suggesting a role for competition and cooperation with other RBPs. I also determined that HuR may play a role in inhibition of translation in a dynamic immunological system (T cell activation).

Finally, I performed a mechanistic analysis of the function of the Pum1 RBP and characterized the role of this protein in the translational regulation of several important target mRNAs through the use of luciferase reporter assays. I also provided the first in vivo evidence of a role for specific regions of the Pum1 protein in the mediation of gene expression. However, I was unable to verify previous in vitro reports of a role for Pum1 in control of translation elongation on verified in vivo mRNA targets, suggesting that Pum1's regulatory function may be context dependent.

Ultimately, the approaches and findings in this study will provide a framework for the development of a global integrated model of posttranscriptional control. Through iterative development of models and experimentation, hypotheses can be

generated by these integrated models, tested in the laboratory, and the results of these experiments will then further improve the development of the models to improve future hypothesis generation. An integrated approach of this type will be necessary to fully understand the highly complex and interconnected nature of the gene expression regulatory system.

Dedication

This is for the two most important people in my world, Evelyn and Afton. Your support and encouragement get me through the hard times and your love and companionship make the good times perfect.

There is no way I could have done this without both of you.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank Kyle Mansfield, Adam Morris, Neel Mukherjee, Matt Friedersdorf, Laura Simone, Jeff Blackinton, and Cindo Nicholson for their contributions to my various projects throughout my tenure in the Keene lab. Thanks also go to David Reid and Shelton Bradrick for assistance and guidance with polysome analyses. My committee members have provided me with support and assistance throughout my graduate career. Jack Keene has been unfailingly supportive of my work, through the disappointments and the successes of science, for which I am very grateful. The support of my friends and colleagues in Durham has truly helped me become a better person and a better scientist. Finally, I'd like to thank my brother and my parents for always supporting me in everything I've done.

1. Introduction

1.1 Posttranscriptional Regulation of Gene Expression

Gene expression is a varied and complex set of processes by which, ultimately, the genotype of an organism is manifested as a phenotype. The importance of these processes has been apparent since Jacob and Monod introduced the first model for coordinated regulation of expression of a group of genes, the bacterial *lac* operon (Jacob and Monod 1961), a discovery which was ultimately rewarded with a Nobel prize. Highly specific regulation of the myriad complex processes involved in gene expression is required, both temporally and spatially, to enable all aspects of life, evolution, and survival (Orphanides and Reinberg 2002). Historically, studies of the regulation of these processes have predominantly focused on the specific regulation of transcription (Keene 2007), and the advent of microarray and deep sequencing technologies has enabled the rapid and inexpensive measurement of total RNA abundance in many biological systems. However, numerous studies have shown a poor correlation between the steady-state levels of mRNAs (the transcriptome) and the abundance of the encoded proteins (the proteome)(Anderson and Seilhamer 1997, Futcher, Latter et al. 1999, Gygi, Rochon et al. 1999), indicating that additional levels of regulation must be involved in the overall process of gene expression.

More recently, it has become apparent that posttranscriptional regulation (PTR), or the regulation of RNA during the steps between transcription up to and including

translation, is one of these important additional levels of control of gene expression (Tenenbaum, Carson et al. 2000, Hieronymus and Silver 2004, Cheadle, Fan et al. 2005, Hao and Baltimore 2009). The relative paucity of studies on posttranscriptional regulation, as compared to transcriptional regulation, is partly due to the fact that the field is much younger. Additionally, posttranscriptional studies often require development of novel methods and experimental approaches, many of which will be discussed elsewhere in this document. The recent growth of interest in microRNAs (miRNA) and their role in gene expression has provided a boost in the general scientific awareness of the posttranscriptional environment (Ambros, Bartel et al. 2003, Chen and Rajewsky 2007). However, many of these studies fail to recognize that all steps of posttranscriptional regulation, including miRNA-mediated regulation, are controlled and coordinated by the action of RNA binding proteins (RBPs). As illustrated in Figure 1, RBPs are responsible for mediating gene expression at all steps of an mRNAs journey from transcription to translation. In contrast, miRNA mediate two important cytoplasmic steps in the process of posttranscriptional regulation, mRNA stability and translation, however it is also important to note that this mediation requires the action of many RBPs, specifically Argonaute proteins and other members of the RISC complex (He and Hannon 2004, Bartel 2009).

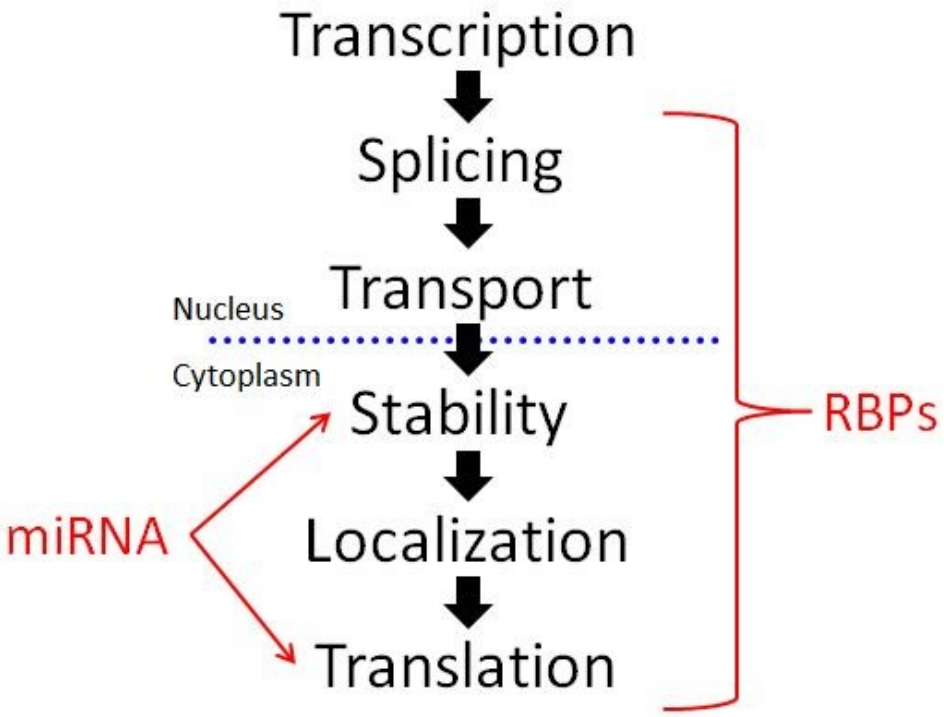


Figure 1: RBPs Coordinate and Control All Steps of an mRNAs Life

The life of an mRNA is regulated at multiple steps, both nuclear and cytoplasmic. RBPs are responsible for this regulation at all steps. It is also important to note that other important posttranscriptional regulatory factors, such as microRNAs, also require the action of RBPs to mediate their functions.

Although all steps of mRNA biogenesis and regulation are mediated by RBPs, this study focuses on the roles of two specific RBPs in the regulation of mRNA stability and translation. However, a general overview of the cellular processes that regulate mRNA translation and decay will provide necessary background prior to exploring the roles of these particular RBPs. The balance between mRNA decay and translation is largely responsible for determining the quantity of protein that is translated from a

particular mRNA. Following export from the nucleus, polyadenylated mRNA is recruited to or away from the translation and decay machinery by the action of a multitude of RBPs that recognize both general and specific sequence elements in the untranslated regions (UTRs) of the message.

The actions of numerous proteins are required for a message to engage the translation machinery (Gebauer and Hentze 2004). Generally, however, the 3′ Poly-A tail of an mRNA must be bound by the poly(A)-binding protein (PABP) and the 5′ cap of the RNA must be bound by the initiation complex, composed of several proteins, including eukaryotic initiation factors 4E and 4G (eIF4E and eIF4G). The bound PABP and eIF4G can directly interact, which creates a circular RNA/protein complex. The closely associated initiation factors and PABP recruit the small ribosomal subunit and additional eIFs to the 5′ end of the mRNA. Further action of several eIFs allows the small ribosomal subunit to 'scan' for the start codon where it can then recruit the large ribosomal subunit and begin translation of the mRNA. This process of translational initiation is generally believed to be the rate limiting step in the translation of most cellular RNAs (Gebauer and Hentze 2004, Kapp and Lorsch 2004), and numerous cellular factors have been identified as direct and indirect regulators of the process. For example, a class of proteins called eukaryotic translation initiation factor 4E-binding proteins (4E-BPs) associate with eIF4E and prevent its interaction with eIF4G, inhibiting the formation of the initiation complex. The binding of 4E-BPs to eIF4E is reversible,

through phosphorylation, and is altered in response to many cellular stimuli, including insulin, hypoxia, and oxidative stress (Fukuglta, Hogan et al. 1994, Patel, McLeod et al. 2002).

In addition to translational silencing through mediators such as 4E-BPs, the cell can reduce the protein expression of an mRNA by targeting the message for degradation. Similarly to translation, the pathways of mRNA decay require the coordinated action of many regulatory proteins, most of which are RBPs. In general, the first step in degradation of a message is the removal of the Poly-A tail through the action of a deadenylase enzyme complex, the primary of which is the Ccr4p/Pop2p/Not complex in eukaryotes (Tucker, Staples et al. 2002, Parker and Sheth 2007, Goldstrohm and Wickens 2008). Following deadenylation, mRNA can be digested through the 3'->5' exonuclease action of a protein complex called the exosome, or, more frequently, the 5' cap can be removed by the action of the Dcp1/Dcp2 decapping enzymes and the mRNA can then be degraded in the 5'->3' direction by the Xrn1 nuclease (Parker and Sheth 2007, Goldstrohm and Wickens 2008).

Both deadenylation and decay function at a basal level on all mRNA within a cell. However, the action of RBPs that recognize and bind specific sequences within their target messages can influence the recruitment of these enzymatic factors. For example, PABP, in addition to its role in translation initiation, is believed to shield the Poly-A tail from the action of the deadenylase complex, thereby preventing degradation

of the mRNAs to which it is bound (Wormington, Searfoss et al. 1996, MILONE, WILUSZ et al. 2004). HuR, an RBP that will be discussed in greater detail elsewhere in this document, has also been shown to promote the stability of its target messages (Fan and Steitz 1998, Levy, Chung et al. 1998, Peng, Chen et al. 1998), potentially by binding deadenylated mRNAs and preventing the action of the exonuclease machinery (Ford, Watson et al. 1999). Other RBPs have been shown to promote the activity of the deadenylase and decay machinery. For example, TTP, an RBP that binds similar RNA sequences as HuR, has been shown to recruit deadenylases (Ccr4), decapping enzymes (Dcp1), and nucleases (Xrn1) to its target mRNAs to trigger their rapid destruction (Lykke-Andersen and Wagner 2005). Similarly, members of the PUF family of RBPs, which will also be discussed later in this document, recruit deadenylase complexes to trigger the decay of target mRNAs (Goldstrohm, Hook et al. 2006, Hook, Goldstrohm et al. 2007).

1.2 Methods Used to Study Posttranscriptional Regulation

The study of posttranscriptional regulation has necessitated the development of a variety of new methods to measure previously unknown aspects of gene regulation. An important first question in many posttranscriptional studies is the identification of the subset of mRNAs that can be regulated by a given RBP. The majority of studies performing this analysis have utilized RNA ImmunoPrecipitation followed by microarray or deep sequencing analysis (RIP-Chip or RIP-Seq) (Tenenbaum, Carson et

al. 2000, Tenenbaum, Lager et al. 2002, Keene, Komisarow et al. 2006). This procedure involves the biochemical isolation of an RBP of interest under conditions that preserve its association with its target mRNAs, which are subsequently purified and characterized by microarray or sequencing. Numerous studies have utilized RIP and modified RIP procedures and identified sets of functionally related mRNAs that are co-regulated by one or more RBPs (Mansfield and Keene 2009).

Importantly, while RIP identifies the mRNAs associated with a given RBP, it does so at the level of the message. A complimentary experimental approach, Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation (PAR-CLIP), has recently been developed as a method to characterize the specific site within an RNA to which an RBP is bound (Hafner, Landthaler et al. 2010, Markus, Markus et al. 2010). PAR-CLIP involves the incorporation of 4-thiouridine (4SU) into cellular transcripts and the subsequent UV 365nm cross linking of these incorporated nucleotides to the adjacent RBPs. Upon immunoprecipitation of the RBP, the cross-linked RNA can then be nuclease digested to a small fragment, the specific site to which the RBP is bound, which is then identified through deep sequencing. A recent study performed both RIP and PAR-CLIP on HuR and found that while the results of both experiments agree, PAR-CLIP is well suited to identifying all transient and possible RBP-mRNA interactions, while RIP identifies the stable interactions within the cellular environment (Mukherjee, Corcoran et al. 2011).

In addition to identifying RBP-mRNA interactions, posttranscriptional studies frequently rely on techniques specifically designed to evaluate functional outcomes of posttranscriptional regulation. The characterization of the translation status of an mRNA has traditionally been performed with a technique known as polysome gradient analysis (Kuhn, DeRisi et al. 2001), in which cells are lysed under conditions that preserve ribosome occupancy on mRNA. The lysates are separated through a density gradient (typically ~10-50% sucrose) by ultracentrifugation and fractions are subsequently collected. Those messages that are engaged with multiple ribosomes (polysomes) migrate further in the gradient than free mRNA or mRNA associated with only one ribosome. Thus, the abundance of an mRNA in a given region of the gradient is indicative of the translational status of that message; mRNAs undergoing active translation are engaged with many ribosomes, while translationally silenced messages are located in the upper (less dense) regions of the gradient. Polysome analyses have yielded many important findings in a multitude of experimental systems. For example, the Parker lab utilized polysome gradients to demonstrate that translationally silent mRNA stored in cytoplasmic structures known as P-bodies (which were previously believed to be sites of RNA decay from which messages could not exit) can be returned to active translation (Brengues, Teixeira et al. 2005). Their findings highlight the dynamic nature of the posttranscriptional environment and have since been extended to

include a role for HuR and miRNAs in the recruitment of mRNA to and away from P-bodies and translating polysomes (Bhattacharyya, Habermacher et al. 2006).

As discussed previously, the regulation of mRNA degradation plays a major role in the outcome of a gene expression program. Analysis of the stability of mRNA has played a major role in the field of posttranscriptional regulation. These studies primarily utilize a small molecule inhibitor of transcription, Actinomycin D, to block RNA synthesis while taking a time course of samples following treatment to evaluate the rate at which an mRNA decays. Actinomycin D (ActD) binds directly to DNA within the transcription complex and prevents elongation of the nascent RNA chain (Sobell 1985). Although this treatment is toxic to the cell, short time course experiments (typically less than 8 hours) are quite feasible, as the cell is able to survive for some time without additional transcription. Evaluation of the amount of a given transcript remaining at several time points following ActD treatment, typically through PCR or array-based methods, allows for the calculation of a half-life of survival for that mRNA (the time required for the amount of that mRNA present to be reduced by half). Studies frequently utilize this method to evaluate the impact of mutations, overexpressions, or other cellular perturbations on the half-life of one or a few mRNAs. However, some studies have characterized mRNA decay kinetics globally. For instance, a study of activated T lymphocytes identified numerous groups of mRNAs that display similar rates of mRNA decay (Raghavan, Ogilvie et al. 2002, Raghavan, Dhalla et al. 2004).

Further characterization of these groups identified shared known regulatory elements in groups that displayed similar decay kinetics. Similarly, studies in yeast demonstrated that the mRNAs encoding proteins with shared or similar functions displayed similar decay kinetics (Wang, Liu et al. 2002), suggesting that co-regulation of RNA decay of functionally related groups of messages is a highly conserved feature of posttranscriptional regulation.

It is worthwhile to reemphasize the fact that these complex and interconnected steps of posttranscriptional regulation at the levels of translational control and regulation of mRNA decay are all mediated and controlled by the action of RBPs (Mansfield and Keene 2009). These complex processes are functionally coordinated by the combinatorial control of many regulatory RBP-mRNA interactions (Parker and Sheth 2007, Goldstrohm and Wickens 2008). The specificity of recognition of an mRNA by an RBP can range from very broad, as in the case of PABP which binds most mRNAs (Görlach, Burd et al. 1994), to very specific, for instance SLBP, an RPB that specifically binds only the mRNAs encoding histone proteins (Wang, Whitfield et al. 1996). This spectrum of binding specificity results from the interaction between specific protein domains and complementary sequence elements in the groups of target RNAs to which these protein domains can bind. These RNA sequences (and structures) have been termed untranslated sequence elements for regulation (USERs) and enable the

coordinated regulation of a group of functionally related messages by one or more RBPs, a model termed the posttranscriptional operon (Keene and Tenenbaum 2002).

1.3 HuR and Posttranscriptional Regulation

The first description of the posttranscriptional operon model was based on analysis of the sets of mRNA targets of a Hu-family RBP, HuB (Tenenbaum, Carson et al. 2000). Hu/ELAV RBPs are an evolutionarily conserved family of proteins with essential roles in development, cellular response to stress, and immune function (Abdelmohsen, Lal et al. 2007, Hinman and Lou 2008). In vertebrates, three Hu proteins, HuB, HuC, and HuD, are expressed solely in neurons, display a predominantly cytoplasmic subcellular localization, and have been shown to play vital roles in memory, neuronal plasticity, and neuronal development (Hinman and Lou 2008). The fourth vertebrate Hu family member, HuR, is ubiquitously expressed, primarily nuclear in localization, although it is known to shuttle between the nucleus and cytoplasm, and has been shown to play a multitude of roles in cellular response to stress, apoptosis, differentiation, and inflammation. Further, dysregulation of HuR has been demonstrated in a variety of human cancers, indicating an important role in the control of cellular proliferation (Simone and Keene 2013).

The functions of Hu family proteins are mediated through their ability to recognize and bind to sets of mRNAs and effect changes in the posttranscriptional regulation of these messages (Hinman and Lou 2008). Hu proteins are members of a

large class of RBPs that contain a protein domain called the RNA recognition motif (RRM), which in the case of Hu proteins is the protein domain that binds to AU-rich sequences (AREs) in the untranslated regions of target mRNAs (Query, Bentley et al. 1989, Maris, Dominguez et al. 2005). Generally, binding of Hu proteins to target mRNAs results in increased mRNA stability, translation, or both. However, the specific mechanisms by which this effect is mediated remain unclear. The predominant hypothesis suggests that by occupying AREs, which are traditionally considered destabilizing elements, Hu proteins prevent the binding of other ARE-binding proteins whose functions involve negative regulation of stability or translation (Hinman and Lou 2008). As discussed previously, several known ARE-binding proteins, including TTP and AUF1, have been shown to directly recruit members of the deadenylation and decay pathways (Goldstrohm and Wickens 2008). Thus, it is reasonable to hypothesize that steric hindrance by Hu proteins binding to these ARE sequences simply prevents the pro-decay mediators from interacting with target messages. Similarly, several recent studies have demonstrated that Hu proteins differentially influence the polyadenylation of target mRNAs in the nucleus, potentially by interfering with the ability of specific splicing and polyadenylation regulators from binding to the pre-mRNA (Zhu, Zhou et al. 2007, Mansfield and Keene 2012).

Although the binding of HuR to its target messages has been characterized as occurring at AREs, the specific features of an ARE that enable HuR binding remain

controversial. Early work identified numerous regions to which Hu family proteins bound with high affinity (Gao, Carson et al. 1994), and further explorations led to the proposition that the sequence recognized by Hu proteins might be a form of an ARE pentamer or nonamer (AUUUA or AUUUAUUUA). However, numerous additional studies indicated that the site recognized by Hu proteins could not be classified as a short simple motif. To this end, a study from the Gorospe lab endeavored to characterize a secondary structural RNA motif in HuR target mRNAs (Lopez de Silanes, Zhan et al. 2004) utilizing a covariance model approach originally designed to identify novel tRNA genes (Eddy and Durbin 1994). Interestingly, their study identified a structural motif that is highly enriched in the 3′ UTRs of known HuR target mRNAs. Following their work, our laboratory performed RIP-Chip analysis of HuR in a T cell activation system and identified a dynamic set of mRNA targets of HuR (Mukherjee, Lager et al. 2009). As shown in Figure 2, the Gorospe structural HuR binding motif is highly enriched in our set of mRNA targets as well (red stars).

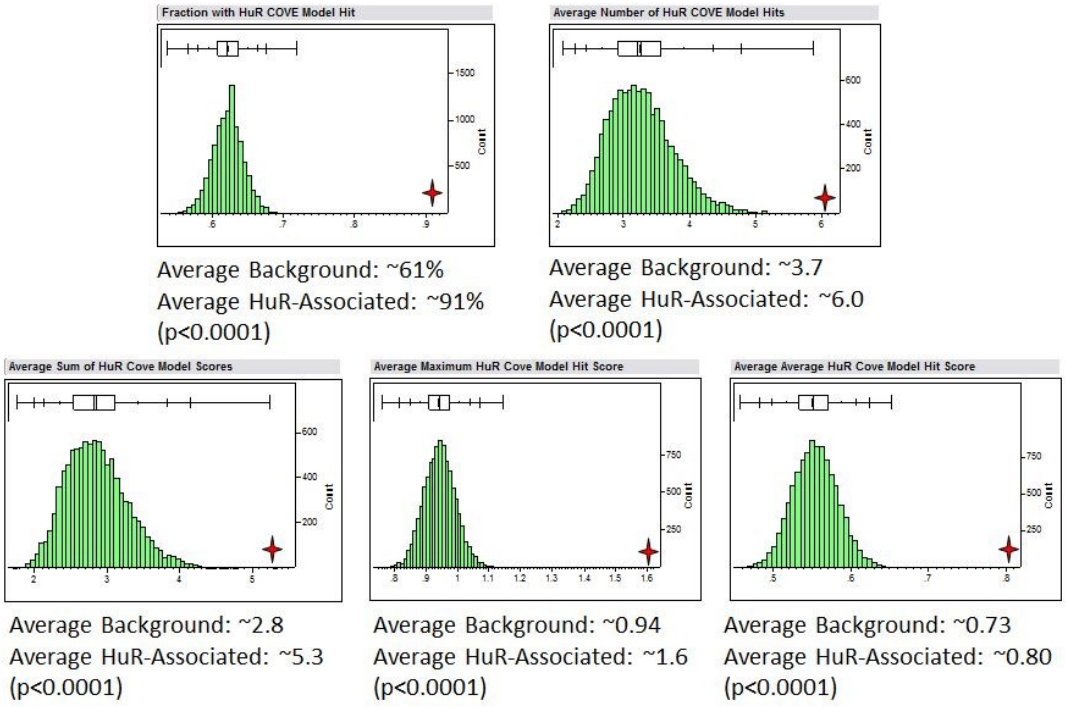


Figure 2: Enrichment of COVE Model Statistics in HuR Associated mRNAs

The set of HuR associated mRNAs was searched for the Gorospe HuR structural binding motif (COVE model). The prevalence of the motif and associated statistics were compared by hypergeometric testing to randomly generated sets of RNA sequence as background. The tested model statistics include: Fraction with HuR COVE Model Hit, which indicates the number of individual mRNAs in the dataset with a COVE model match; Average Number of HuR COVE Model Hits, which represents the arithmetic mean of the number of COVE model matches per mRNA; Average Sum, Maximum, and Average HuR COVE Model Hit Score indicate the sum, maximal, or mean of all scores generated by the COVE model matches across all mRNAs in the dataset. Red stars indicate the enrichment of the various statistics in the HuR target set and the green bars represent the background distribution generated by the hypergeometric sampling.

However, we were somewhat skeptical of the ability of HuR to bind a structural motif, as many studies have clearly characterized the RRM as a single-stranded RNA binding domain (Maris, Dominguez et al. 2005). Thus, I reevaluated the enrichment of

the covariance model in a set of sequences generated by randomly shuffling the sequences of the HuR target UTRs while preserving the frequency of all dinucleotide pairs. The preservation of dinucleotide frequencies in a sequence randomization maintains the degree of secondary structure in an RNA while removing any specific sequence or structural motifs (Workman and Krogh 1999). As shown in Figure 3, the enrichment of the covariance model statistics remains significant in the shuffled sequences, indicating that HuR does not bind a specific structure, but instead can bind to a group of messages that display similar folding free energy distributions (Mukherjee, Lager et al. 2009).

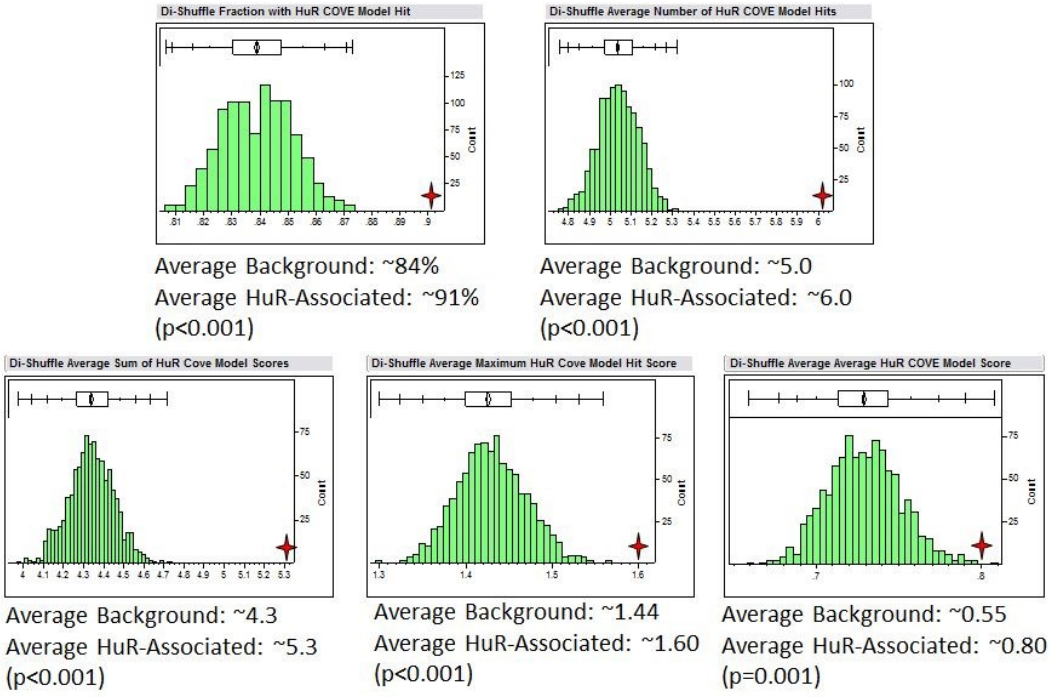


Figure 3: Enrichment of COVE Model Statistics in HuR Associated mRNAs, Dinucleotide Shuffled Background

The set of HuR associated mRNAs was searched for the Gorospe HuR structural binding motif (COVE model). The prevalence of the motif and associated statistics were compared by hypergeometric testing to randomly generated sets of RNA sequence with the distribution of dinucleotides in each random set of sequences maintained. The tested model statistics include: Fraction with HuR COVE Model Hit, which indicates the number of individual mRNAs in the dataset with a COVE model match; Average Number of HuR COVE Model Hits, which represents the arithmetic mean of the number of COVE model matches per mRNA; Average Sum, Maximum, and Average HuR COVE Model Hit Score indicate the sum, maximal, or mean of all scores generated by the COVE model matches across all mRNAs in the dataset. Red stars indicate the enrichment of the various statistics in the HuR target set and the green bars represent the background distribution generated by the hypergeometric sampling.

Recently, a combined PAR-CLIP and RIP-Chip study and a background

corrected PAR-CLIP study from our lab have further characterized the specifics of HuR-

mRNA binding (Mukherjee, Corcoran et al. 2011, Friedersdorf and Keene 2014). Together, these studies have demonstrated that HuR may bind a large proportion of the mRNAs within a cell, but only a subset of the bound messages represent stable interactions. Additionally, a comprehensive biochemical and computational analysis of HuR-mRNA recognition has indicated that the structural conformation of potential ARE sequences, which can be altered by the action of other RBPs or short RNAs, plays a major role in the ability of HuR to bind its target messages (Meisner, Hackermüller et al. 2004). Together, the studies of HuR-mRNA binding and mechanisms of action indicate that the complex and important role of HuR in posttranscriptional regulation remains only partially understood, and further application of both mechanistic and global studies will be required to fully characterize this well-studied RBP.

1.4 Pum and Posttranscriptional Regulation

Perhaps the most studied family of RBPs, particularly in non-mammalian eukaryotic systems, is the PUF family, a highly conserved family of RBPs, with members in most eukaryotic systems. The defining characteristic of a PUF protein is the presence of a PUF homolog domain (HD, also referred to as the PUM-HD) (Zamore, Williamson et al. 1997), a domain composed of eight repeated 36 amino acid motifs (the PUM repeat) (Zamore, Williamson et al. 1997, Zhang, Gallegos et al. 1997). Interestingly, despite the high degree of conservation of the HD, other regions of PUF family proteins are not highly conserved between species. Functional and structural studies have clearly

demonstrated the binding of the HD to RNA targets of PUF proteins (Wang, Zamore et al. 2001), many features of which will be discussed later in this section.

PUF family proteins display important roles in a variety of cellular functions. The ancestral function of the protein family is believed to be the maintenance of stem cell self-renewal and proliferation (Wickens, Bernstein et al. 2002, Spassov and Jurecic 2003). In addition to a role in stem cell maintenance, the single PUF family member in *Drosophila*, Pumilio, is required for proper development of germ cells (Forbes and Lehmann 1998) and for anterior/posterior patterning in the developing embryo (Murata and Wharton 1995). Similarly, the *C. elegans* PUF proteins FBF1 and FBF2 are required to maintain germ line stem cell self-renewal (Crittenden, Bernstein et al. 2002). Yeast studies have provided further evidence for the role of PUF proteins in stem cell maintenance. Specifically, yeast strains possessing deletions of the Puf5 gene are unable to divide as many times as normal strains before dying (Wickens, Bernstein et al. 2002).

As discussed in the first section of this introduction, many regulatory RBPs bind functionally related subsets of cellular mRNAs (Keene and Lager 2005). A particularly striking example of such a posttranscriptional operon is yeast Puf3. RIP-Chip analysis of Pufs 1-5 in yeast identified distinct and only partially overlapping subsets of target messages for each of the five related proteins (Gerber, Herschlag et al. 2004). Strikingly, Puf3 was found to bind nuclear-encoded mitochondrial messages almost exclusively,

and further studies (García-Rodríguez, Gay et al. 2007) demonstrated that Puf3 directly regulates mitochondrial biogenesis.

Additional characterization of the groups of mRNAs bound by various PUF family proteins (Gerber, Herschlag et al. 2004, Gerber, Luschnig et al. 2006, Galgano, Forrer et al. 2008, Morris, Mukherjee et al. 2008, Jiang, Guo et al. 2012) has illuminated an interesting feature of this family of RBPs, the posttranscriptional rewiring of PUF operons. Although the interactions between the PUM-HD and the RNA sequence it recognizes, the Pum binding element (PBE), have remained conserved throughout evolution, the specific groups of mRNAs containing PBEs have changed. Thus, *Drosophila* Pumilio can bind and regulate a group of mRNAs with completely different functions than a yeast Puf protein or a human Pum protein. Similar evolutionary mechanisms have been described for reuse of protein domains and transposable DNA elements (Jiang, Guo et al. 2012), thus it is unsurprising that the posttranscriptional level of gene regulation also displays the ability to evolve new regulatory functions and mechanisms.

In contrast to the pro-stability and translation mechanisms of the previously discussed Hu proteins, PUF family proteins are predominantly negative regulators of both mRNA stability and translation (Wickens, Bernstein et al. 2002, Spassov and Jurecic 2003, Quenault, Lithgow et al. 2011). *Drosophila* Pumilio and *C. elegans* FBF bind to the hunchback and fem-3 mRNAs, respectively, and prevent their translation. In *Drosophila*,

this repression is mediated through the additional protein factors Nanos and Brat, which interact directly with the Pumilio protein and bind the 5′ cap of the hunchback mRNA, thereby preventing eIF4E binding and translational initiation (Vardy and Orr-Weaver 2007). In *C. elegans*, FBF represses the translation of fem-3 through interaction with the *C. elegans* Nanos protein (Zhang, Gallegos et al. 1997, Kraemer, Crittenden et al. 1999), and further studies indicate that FBF can also recruit the Ccr-Not deadenylation complex to trigger mRNA decay (Suh, Crittenden et al. 2009). Several studies have demonstrated that both yeast Puf proteins and human Pums can directly interact with members of the Ccr4/Not complex of deadenylase enzymes and trigger deadenylation of target mRNAs (Goldstrohm, Hook et al. 2006, Van Etten, Schagat et al. 2012). In addition, it was recently shown that human Pum2 can influence translational elongation in vitro through interaction with Ago2 and a member of the translation initiation complex, eIF1A (Friend, Campbell et al. 2012).

Another distinct feature of PUF family proteins is the high degree of sequence specificity of the Pum USER code. Yeast Puf3, human Pum1 and Pum2, and fly Pumilio all bind the same RNA element, UGUAHAUA (Gerber, Herschlag et al. 2004, Gerber, Luschnig et al. 2006, Morris, Mukherjee et al. 2008). Other PUF family proteins recognize similar, but distinct, RNA elements. Figure 4 shows the RNA motif (the Pum USER code) recognized by multiple PUF proteins from human, *Drosophila*, and yeast.

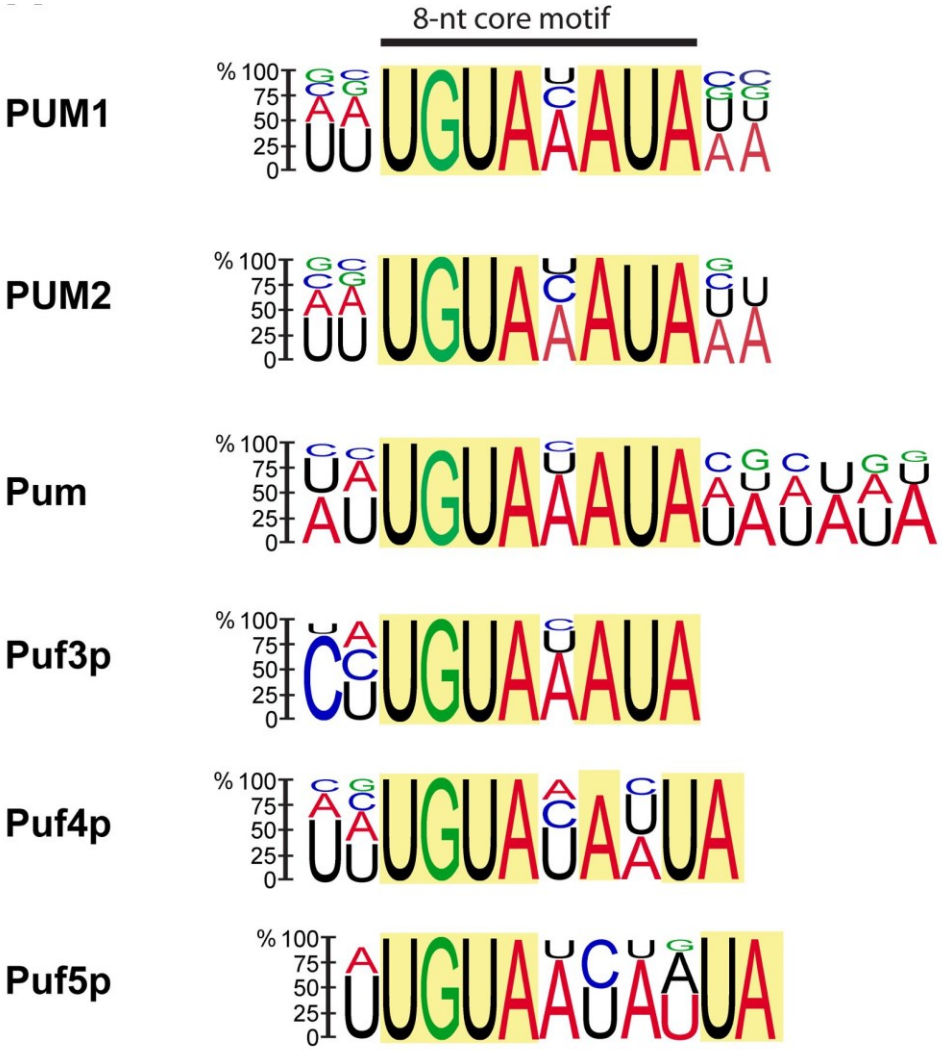


Figure 4: Conservation of the Pum USER Code.

Human (PUM1, PUM2), *Drosophila* (Pum), and yeast (Puf3p) recognize the same Pum USER Code. Other yeast Puf proteins (Puf3p, Puf4p, Puf5p) recognize similar but distinct sequences. Figure from (Galgano, Forrer et al. 2008).

Careful structural characterization of the Pum HD from multiple species has demonstrated a unique modular structure to the Pum domain (Wang, Zamore et al. 2001, Wang, McLachlan et al. 2002, Miller, Higgin et al. 2008). The HD of all studied

PUF family proteins forms a crescent shape, and each of the PUM repeats is structured as an alpha helix within the overall crescent. RNA interaction occurs on the inner concave surface of the crescent where each PUM repeat makes multiple contacts with a single base in the PBE. Figure 5 shows the crystal structure of the Pum-HD, both in complex with a target RNA and with the individual PUM repeats colored separately.

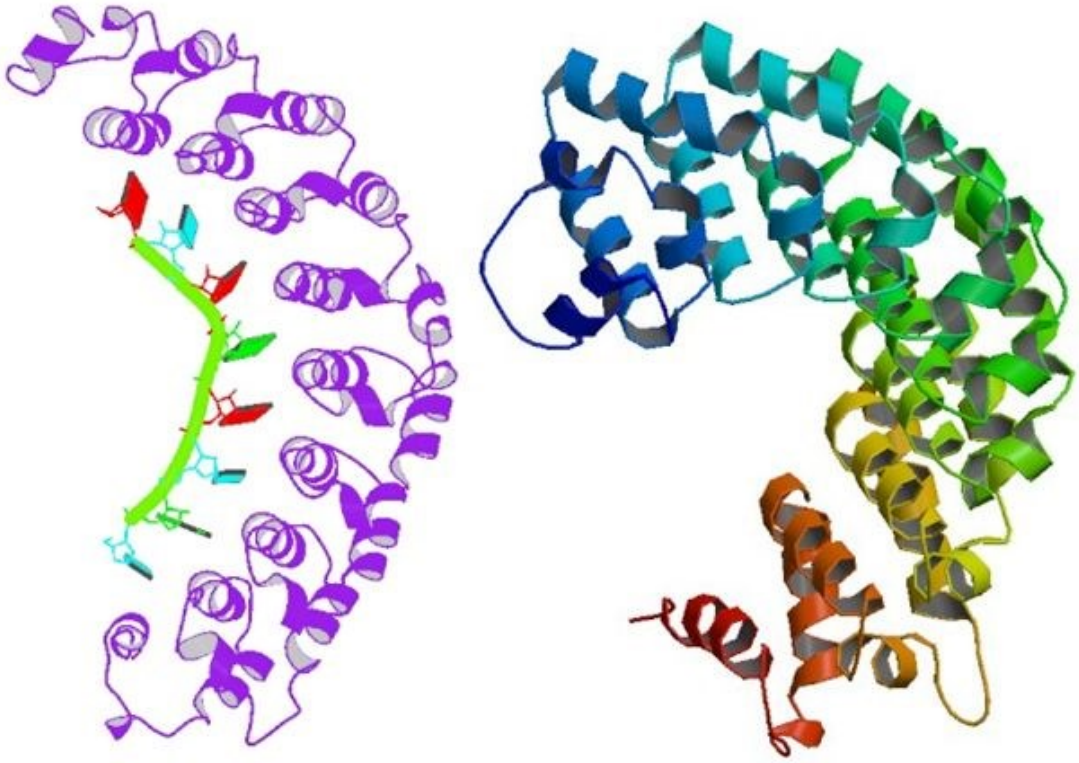


Figure 5: Structure of the PUM-HD.

The left image shows the Pum-HD of human Pum1 in complex with a Pum binding element from the p38alpha RNA, shown in green. The right image shows the PUM-HD of mouse Pum2 with the individual PUM repeats displayed in different colors. Images from the RCSB PDB (www.rcsb.org). Structures 3Q0M and 3GVT from (Jenkins, Baker-Wilding et al. 2009, Lu and Hall 2011).

Recent work has selected and engineered PUM repeats capable of binding any nucleotide (Filipovska, Razif et al. 2011), enabling the construction of a PUM-HD based RNA binding structure than can effectively recognize any eight nucleotide RNA sequence (Miller, Higgin et al. 2008, Lu, Dolgner et al. 2009, Cooke, Prigge et al. 2011).

The binding specificity and modularity of the PUM-HD make PUF family proteins are quite distinct from the RNA recognition characteristics of most other RBPs. As discussed above, for instance, the Hu family (and most other known ARE-binding proteins) display a much broader range of USER codes, which makes specific characterization of their RNA binding motif much more challenging. Interestingly, a comparison of the RNA targets of Pum1 and HuR identified by studies in our laboratory (Morris, Mukherjee et al. 2008, Mukherjee, Lager et al. 2009) show a substantial overlap. Given the contrast in established cellular functions of these two RBPs (HuR as an activator of expression and Pum1 as a repressor) and the distinct differences in their RNA binding specificity, I anticipate that comparisons between the activities of the two proteins will illuminate general features of posttranscriptional regulation. These features will provide a basis for developing rules and definitions of the functional consequences of RBP-mRNA interaction that will ultimately be required to evaluate and understand posttranscriptional regulation on a global scale.

2. A Global View of Posttranscriptional Regulation of Gene Expression

2.1 Introduction

Since the publication of the completed human genome sequence, it has become abundantly clear that the fundamental principles regulating gene expression are encoded, to a large degree, in the interactions between genes and their products (Remenyi, Scholer et al. 2004). This realization has led to the rapid expansion of the field of systems biology, which aims to understand the principles of biological systems in a holistic manner. Clearly, study of the complex network of interactions responsible for cellular behavior necessitates systems-level approaches to experimentation and data analysis.

Interestingly, studies in the field of systems biology have largely ignored a substantial and critical aspect of gene expression, namely, the posttranscriptional regulatory environment (Keene 2007). The recent explosion of microRNA (miRNA) studies has sparked some interest in the field of systems biology, promoting some exploration of the relationship between small RNAs and the overall gene expression network. However, earlier developments in the field of posttranscriptional control of gene expression have demonstrated that RNA binding proteins (RBPs) impart a significant amount of regulation to gene expression (Cheadle, Fan et al. 2005).

Until recently, the majority of studies on the posttranscriptional environment have focused on one or a few RBPs and their targets, making it very difficult to apply

any global systems-level analysis to this important aspect of gene expression. Fortunately, the Brown laboratory has published a study characterizing the RNAs associated with 47 different RBPs in the yeast *S. cerevisiae* (Hogan, Riordan et al. 2008). This groundbreaking study has provided a dataset that will enable global examination of the regulatory structure of the yeast posttranscriptional environment. Perhaps more importantly this data, when combined with the abundant yeast transcriptional datasets available, will illuminate the extent and structure of the relationship between transcriptional and posttranscriptional control of gene expression.

I have combined the RBP-mRNA association datasets from the Brown lab work with a transcription factor targeting dataset from a seminal yeast systems biology work from the Young laboratory (Harbison, Gordon et al. 2004) to generate and evaluate an integrated network of transcriptional and posttranscriptional control of gene expression in yeast.

2.2 Results

2.2.1 Network Definition

Edges were defined based on RBP-mRNA or transcription factor (TF)-DNA association, as measured by RIP-Chip or ChIP-Chip, respectively. RBP-mRNA edges were included in the network for all RBP interactions with mRNAs detected in the Brown study. TF-gene edges were included for all interactions identified in the Young lab study. To explore the relationships between the transcriptional and

posttranscriptional regulatory environments, the network was filtered to include only those nodes for which interactions were measured in one of these two studies. A representation of the filtered network is shown in Figure 6.

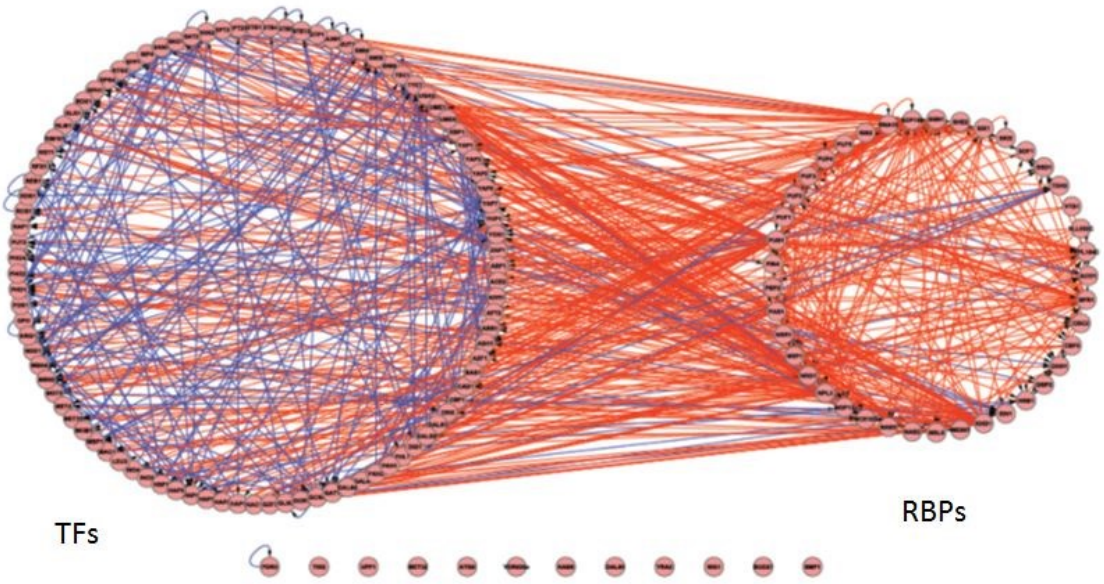


Figure 6: Combined RBP/TF Network

Each node in the left circle represents a TF, each node in the right circle represents an RBP. Each edge represents a targeting relationship. Edges originating from TF nodes indicate that the target node was identified as a TF target by ChIP, while edges originating from RBP nodes indicate that the target node was identified as an RBP target by RIP. Edges originating from TF nodes are blue, while edges originating from RBPs are red.

Although the network contains more than twice as many TF nodes as RBP nodes (102 TF, 46 RBP), the majority of network interactions are of the RBP-mRNA type (472 RBP-mRNA edges, 244 TF-DNA edges). This finding suggests that, at least within the subset of genes represented in this network, posttranscriptional interactions may be 2 to

4 times more common than transcriptional relationships. Examination of the relative number of interactions per RBP or TF indicates that this trend is supported on the scale of all genes. The mean number of interactions identified for RBPs and TFs by these two studies were 358 and 63, respectively. This data supports the argument that posttranscriptional interactions are substantially more prevalent than transcriptional interactions in the sets of RBPs and TFs tested.

2.2.2 Regulators of Regulators

Interestingly, RBPs and TFs appear to differ in their target gene-type preference. Specifically, the filtered network indicates that while RBPs display the expected frequency of interactions with both other RBP nodes and with TF nodes (68.92% of nodes are TF and 63.77% of the edges from RBP nodes target TF mRNAs), TFs show interaction with other TFs with much higher preference than interaction with RBPs (31.08% of nodes are RBPs, but only 13.52% of the edges from TFs target RBPs). Table 1 shows the relative frequency of interactions between the two types of nodes. This finding suggests that RBPs have the potential to regulate messages encoding both other RBPs and messages encoding TFs. In contrast, TFs appear to preferentially target other TFs, suggesting a self-regulating sub-network that may process the regulatory inputs provided by the RBP component of the overall regulatory network. Furthermore, the relative dearth of connections leading from TF nodes to RBP nodes suggests that RBPs

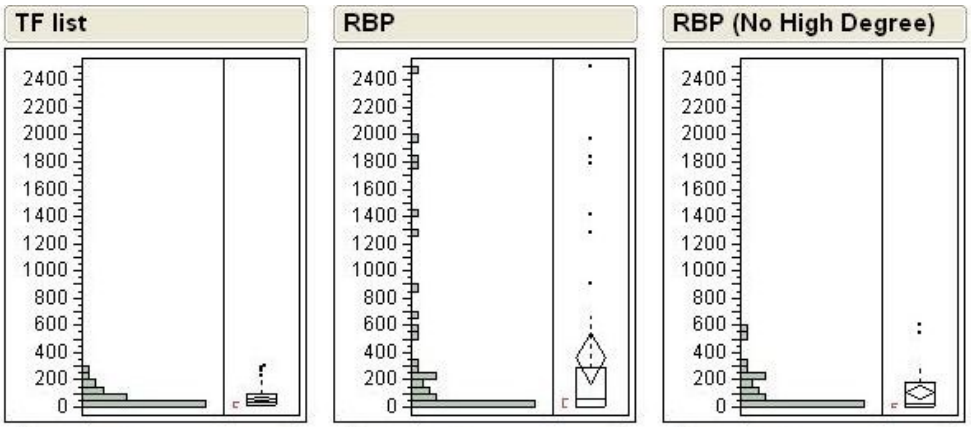
may compose a self-regulating sub-network responsible for controlling their own expression as well as the expression of key members of the TF network.

Table 1: Frequency of Interaction Types in RBP/TF Network

	RBP	TF
Nodes	46	102
Edges	472	244
% Nodes	**31.08%**	68.92%
Edges to TFs	301	211
Edges to RBPs	171	33
% of Edges To TFs	63.77%	86.48%
% of Edges to RBPs	36.23%	**13.52%**

In order to determine whether the network characteristics observed were simply a result of the large number of RBP-mRNA interactions, compared to TF-DNA interactions, I removed the 8 most highly connected RBP nodes and examined the resulting network. As shown in Figure 7, the distributions of number of interactions for RBPs and TFs were much more similar after removal of the high degree RBP nodes (right panel). Strikingly, removal of these RBP nodes did not impact the overall regulatory structure observed in the full network.

Table 2 illustrates that the relative frequency of interactions for RBP and TF nodes remains largely unchanged after removal of these nodes.



Quantiles

		TF list	RBP	RBP (No High Degree)
100.0%	maximum	285.00	2488.0	588.00
99.5%		285.00	2488.0	588.00
97.5%		268.70	2408.1	588.00
90.0%		164.00	1552.0	266.40
75.0%	quartile	93.00	289.5	185.50
50.0%	median	39.00	60.0	29.00
25.0%	quartile	14.00	4.0	3.00
10.0%		4.00	0.6	0.00
2.5%		1.55	0.0	0.00
0.5%		1.00	0.0	0.00
0.0%	minimum	1.00	0.0	0.00

Moments

	TF list	RBP	RBP (No High Degree)
Mean	63.554455	358.02222	103.56757
Std Dev	66.489319	616.28626	145.72492
Std Err Mean	6.6159345	91.870531	23.957029
Upper 95% Mean	76.680281	543.17511	152.15467
Lower 95% Mean	50.42863	172.86933	54.98046
N	101	45	37

Figure 7: Target Number Distributions of RBP, TF, and High-Degree-Removed RPB Networks.

Histograms show the distribution of number of targets for all RBPs, all TFs, and the subset of RBPs with the high-degree RBP nodes removed. Quantiles and moments for the three distributions are provided below the histograms.

Table 2: Frequency of Interaction Types in High-Degree-RBP-Removed Network

	RBP	TF
Nodes	38	102
Edges	109	240
% Nodes	**27.14%**	72.86%
Edges to TFs	71	211
Edges to RBPs	38	29
% of Edges To TFs	65.14%	87.92%
% of Edges to RBPs	34.86%	**12.08%**

2.2.3 Reachability Analysis

To further explore the regulatory potential of the RBP and TF sub-networks, I extended the network model to include targeting data for all yeast genes, not just those encoding RBPs and TFs. I performed a 'reachability' analysis on this extended network by following all paths through the directed network and evaluating the number of targets (nodes) that can be reached through any path beginning at each RBP and TF node. As shown in Figure 8, a major cluster of 5331 genes (the majority of genes in the dataset) can be reached from ~45% of the RBP nodes, while only ~25% of the TF nodes have paths that lead to this cluster. This observation suggests that the indirect regulatory potential (that is regulation through one or more intermediate nodes) in the RBP sub-network is substantially greater than the potential in the TF network.

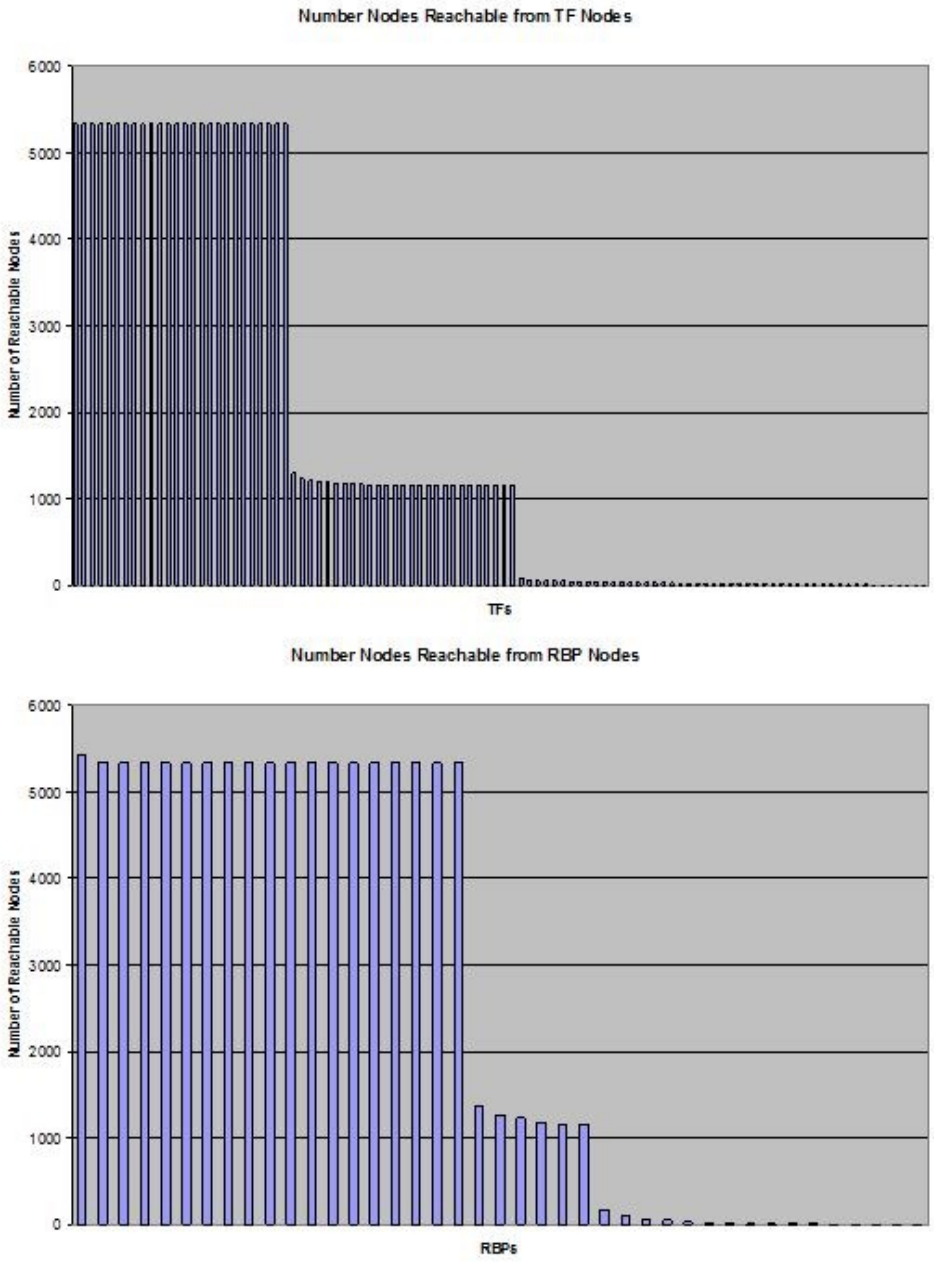


Figure 8: Reachability Analysis of RBP/TF Network

The number of nodes 'reachable' from individual TF or RBP nodes is indicated on the y axes. The x axis represents individual RBPs and TFs. Upper graph shows reachability analysis for TFs, lower graph shows analysis for RBPs.

2.2.4 Network Structure/Function Analysis

The relationship between regulatory network structure and functional regulation of gene expression was further explored through integration of Gene Ontology categories. The combined RBP/TF network was divided into groups according to the Yeast slim ontology (Eilbeck, Lewis et al. 2005) and the structural features of the resulting sub-networks were evaluated. Two structural metrics, modularity and conductance, were chosen as measures of network structure and information flow. Networks with high modularity display a large number of connections between nodes of the same 'module' while containing few connections between modules (Newman 2006). Many biological systems display high modularity, suggesting that local community interactions may underlie many biological functions. Several studies have determined that high modularity is frequently related to robustness and stability in biological networks (Pradhan, Dasgupta et al. 2011, Wang and Zhou 2012). We evaluated the modularity of the sub-networks resulting from the division of the RBP/TF network according to the GO-Slim ontology. Additionally, we evaluated the conductance of various sub-networks of the RBP/TF system. Conductance provides a quantification of the information flow across a cut of a graph. In this case, we evaluated the conductance in (or out) of a given sub-network of the combined RBP/TF network as a representation of the potential for that sub-network to send or receive information to or from the rest of the network.

Table 3 provides an overview of the results of the GO-Slim divisions and modularity calculations. Interstingly, no GO-Slim-based divisions resulted in sub-networks with conductance value z scores > 2. However, several of the GO-Slim divisions resulted in high modularity z scores. Both the "RNA binding" and "Transcription Regulator Activity" sub-networks displayed high modularity scores, further supporting the previous observation that many RBPs and TFs target other RBP and TF nodes in the network. Similarly, the "DNA Binding" and "Transcription" GO sub-networks likely represent a similar subset of TF nodes, as many GO-Slim categories contain overlapping sets of genes. It is interesting to note the sub-networks defined by the "Cell Cycle" and two metabolic GO categories ("cellular amino acid and derative metabolic process" and "generation of precursor metabolites and energy") also display high modularity. Previous work has shown the yeast cell cycle regulatory network to be highly robust to perturbation (Li, Long et al. 2004). Similarly, other studies have identified a relationship between robustness and modularity in metabolic networks (Holme 2011). It is important to note that previous studies have not included posttranscriptional regulatory data in the construction of their network models. Given the greater level of RBP-mediated regulatory potential in the combined RBP/TF network, it is likely that the high modularity values identified in these functional groups are, at least partially, due to the 'Regulators of Regulators' structure observed in the overall network.

Table 3: Network Strucural Analysis of GO Category Modules

GOID	Category	Size	Modularity z Score
GO:0003723	RNA binding	23	14.37
GO:0030528	transcription regulator activity	91	13.13
GO:0003677	DNA binding	78	12.61
GO:0007049	cell cycle	20	9.84
GO:0006519	cellular amino acid and derivative metabolic process	11	8.86
GO:0007124	pseudohyphal growth	9	8.12
GO:0005634	Nucleus	107	7.53
GO:0006350	Transcription	90	5.75
GO:0000746	conjugation	3	4.49
GO:0005694	chromosome	16	4.31
GO:0006091	generation of precursor metabolites and energy	7	3.09

2.2.5 Integration of Functional Datasets

I further evaluated the functional implications of the combined regulatory network by integrating data from several other yeast studies, including global measurements of mRNA abundance and transcription rate (Holstege, Jennings et al. 1998), mRNA decay rates (Holstege, Jennings et al. 1998, Wang, Liu et al. 2002), and protein abundance (Newman, Ghaemmaghami et al. 2006).

Table 4 shows the average values for the various datasets evaluated, separated into groups based on whether they are targets of RBP nodes only, TF nodes only, both RBP and TF nodes, or neither. Red entries in the table indicate $p < 0.05$ for a Welch's T-test vs. the 'Neither' group. Interestingly, those genes targeted by RBPs, either alone or

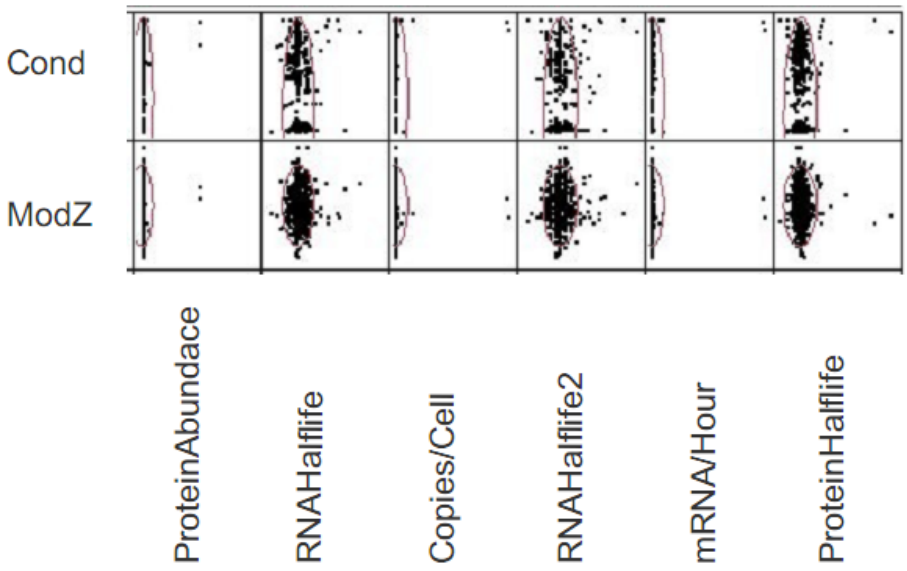
in combination with TFs, show significantly shorter halflives than the genes targed by TFs only or by neither TFs or RBPs. These findings may reflect the role of many RBPs in the destabilization of target mRNAs. It is also interesting to note that the transcription rate of those genes targeted by both RBPs and TFs is significantly higher than the rate of non-target genes. This observation agrees with the fact that this group of genes also displays a higher average mRNA abundance than the other groups. It is striking to note that the only group that displays a difference in average protein abundance is the TF-Only group. The highly elevated levels of protein abundance observed for this group may reflect a lack of translational control, as the RBP & TF group show a much lower average protein abundance, despite having nearly double the average mRNA abundance of the TF Only group. Together, these data support the previous argument that the posttranscriptional regulatory environment may be largely responsible for maintaining the steady state expression level of many of the expressed genes in the yeast system.

Table 4: Functional Dataset Integration Averages

	Protein Abundance	mRNA Half-life	mRNA Abundance	mRNA Half-life	Transcription Rate
Neither	1204.30	35.23	1.16	20.52	3.18
RBP Only	489.05	23.20	1.97	18.22	5.06
TF Only	5210.38	36.00	2.28	20.72	5.90
RBP & TF	1153.47	24.54	4.35	18.56	12.05

I also explored the relationship between these functional datasets and the previously discussed network structural metrics. In this case, the network was

partitioned by spectral clustering (Gleich 2006) and the clusters were evaluated for modularity and conductance. However, as shown in Figure 9, no significant relationship was identified between the various functional datasets examined and the conductance or modularity of the sub-networks.



Conductance	0.110	0.028	0.106	0.062	0.107	0.098
ModularityZ	-0.002	0.021	-0.008	0.005	-0.010	0.001

Figure 9: Relationship Between Functional Datasets and Structrual Metrics

The combined RBP/TF network was partitioned by spectral clustering and the conductance and modularity of the resulting clusters were evaluated. These values were then compared with data gathered from various global yeast studies. The upper panel shows the pairwise correlation plots between the conductance and modularity values and the functional datasets. The lower table shows the conductance and modularity values for the network partitions based on the functional datasets.

2.3 Materials and Methods

Data sources: Yeast RBP target data included in our dataset were obtained from the supplemental data of the Brown lab study (Hogan, Riordan et al. 2008). Interactions were included only if they were identified with an FDR < 1% in the published supplemental data. TF target data were obtained from the Frankel lab webpage (http://fraenkel.mit.edu/Harbison/) and only interactions identified with $p < 0.005$ were included in our dataset. All functional datasets were obtained from the published supplemental data of their respective publications.

'Reachability' Analysis was performed as an iterative (stack based) depth first search. The 'reachability' search was run beginning on every node with outgoing edges in the RBP/TF network and the results were compiled and analyzed in Microsoft Excel.

Conductance and Modularity calculations were performed in R, based on the definitions provided in (Newman 2006) and implemented with assistance from Anirudh Natarajan. The Spectral clustering method was based on algorithm provided in (Gleich 2006) and implemented in R by Scott Schmidler and modified with assistance from Anirudh Natarajan.

Statistical analysis of the functional datasets was performed in Microsoft Excel.

2.4 Discussion

Our analysis of the integrated yeast network suggests that posttranscriptional regulation of gene expression, as represented by RBP-mRNA association, composes a

substantial component of the overall gene expression regulatory network. In addition, I believe that the inclusion of RBP-mRNA interaction data to the gene expression network will enable discoveries that would not otherwise be possible (i.e. in the transcriptional network alone). Specifically, modeling of gene expression based on transcription networks alone lacks a whole level of regulation that can, at least partially, be captured by including posttranscriptional data in the network.

Further, comparisons between the TF and RBP sub-networks and their respective connections indicate a strong 'regulators of regulators' structure inherent to the posttranscriptional environment. While similar structures have been suggested in higher eukaryotic systems (Keene 2007, Keene 2010), large-scale RBP-mRNA interaction data for numerous individual RBPs was unavailable until the publication of the Brown work used to generate the RBP regulatory network. Since the conclusion of this work in our laboratory, several additionally studies have confirmed and extended many of the findings through additional reanalysis of the same data sets.

Strikingly, two studies from the Janga laboratory (Mittal, Roy et al. 2009, Mittal, Scherrer et al. 2011) determined that, in general, mRNAs encoding RBPs are expressed at higher levels than non-RBP mRNAs and that a large number (nearly 1/3) of RBPs autoregulate their own expression. Further, their study found RBPs with a large number of targets (node out degree in the combined RBP/TF network) tended to be expressed at higher levels and exhibited greater message and protein stability,

indicating that these highly connected RBPs are important and tightly regulated cellular factors.

Another group has also further explored the relationship between the transcriptional and posttranscriptional networks in yeast (Joshi, Van de Peer et al. 2011). Specifically, they found that clustering genes by similarity in expression level yields more functionally related groups when clustered based on translational abundance than when using RNA expression level. Further, they found that these groups of coordinately translated genes tended to be expressed at higher levels and with lower noise than clusters of genes co-expressed at the RNA level. By integrating RBP and TF target data in the same manner as I have, they also found that specific network structural motifs (feedback loops and bifans) enriched in the integrated network tend to target, and presumably regulate, these translationally co-expressed clusters. A later study (Joshi, Beck et al. 2012) from the same group confirmed that the RBP regulatory system in bacteria, yeast, mice, and humans tends to display less noise and contains more self-regulatory structural motifs than the corresponding TF regulatory system.

Although, I was unable to identify specific structural network attributes that correlate with the functional data evaluated, a previously discussed study (Mittal, Scherrer et al. 2011) found that the protein-protein interaction network displays a higher degree of modularity than either the TF-mRNA or the RBP-mRNA networks. Their findings suggest that co-regulation at the posttranscriptional level may be responsible

for establishing functional modularity in the resulting protein expression network. It is possible that the conductance and modularity metrics utilized (which were chosen for their potential to illuminate the relationships between 'regulators of regulators') are not ideal metrics to identify functional correlations between groups of nodes in the posttranscriptional network. It is also possible that identification of the true functional relationships encoded by the posttranscriptional RBP-mRNA interaction network will require characterization of the targets of a larger number of RBPs (although the Brown study examined 46 RBPs, there are believe to be over 550 RBPs in the yeast genome).

Together, our data and the findings provided by several other groups strongly support the conclusion that integration of posttranscriptional regulatory information into models of gene expression is critical to fully understand the complex process of regulation of gene expression.

3. Global Analysis of HuR Posttranscriptional Regulatory Dynamics

3.1 Introduction

Although the global network-based analysis of gene regulatory systems has been recognized as an important field of study for some time (Lee, Rinaldi et al. 2002), the understanding of the importance of dynamic data in the analysis of regulatory networks is a more recent development (Luscombe, Babu et al. 2004, Alexander, Kim et al. 2009, Blonder, Wey et al. 2012). As indicated in the previous chapter, the study of gene expression has largely focused on transcriptional regulation, while little attention has been paid to the posttranscriptional regulatory environment. Similarly, the integration of dynamic data in the development and analysis of regulatory networks has been almost exclusively focused on studies of transcriptional network dynamics. However, a few studies have strongly indicated a need for dynamic models of posttranscriptional regulation.

A study from our lab (Mukherjee, Lager et al. 2009) was the first to demonstrate dynamic remodeling of the posttranscriptional landscape. Specifically, the RBP HuR was found to dynamically associate with its target messages throughout the process of T cell activation. A novel application of probabilistic evaluation of the likelihood of RBP-mRNA association allowed us to quantify these association dynamics and identify correlated groups of functionally related mRNAs that are co-regulated by HuR. Additionally, the use of dynamic association data as an input to the Broad Institute's

Connectivity Map (CMAP) (Lamb, Crawford et al. 2006) enabled us to identify small molecule effectors that modulate the activity of an RBP, HuR in this case.

Although prior studies had characterized targets of HuR (Lopez de Silanes, Fan et al. 2004, Lopez de Silanes, Zhan et al. 2004, Abdelmohsen, Pullmann et al. 2007, Mazan-Mamczarz, Hagner et al. 2008), this study provided the first global picture of a dynamic RBP-mRNA landscape. Despite the advances provided by these studies, the underlying mechanisms involved in the coordination of these dynamic changes, as well as their functional consequences remain largely unknown. I explored these mechanisms and consequences through a combination of modeling and experimental approaches.

3.2 Results

3.2.1 Thermodynamic Model

Previous studies in our lab investigated the interaction between HuR (an RBP) and cellular mRNAs throughout the process of PMA/PHA activation of Jurkat T-cells (Mukherjee, Lager et al. 2009). We found that HuR dynamically associates with sets of mRNAs across the activation process. Based on these data, I generated a 'thermodynamic model' of RBP-mRNA association that describes the relationship between the relative concentrations of RBP, mRNAs, and RNP complexes by means of thermodynamic rate equations for reversible association in dynamic equilibrium:

Equation 1: mRNA/RBP Association Equilibrium

$$Ka = \frac{[AB]}{[A][B]}$$

where [B] indicates RBP concentration, [A] represents mRNA concentration, and the RBP/mRNA complex concentration is represented by [AB].

For each mRNA, i, Ka at time 0 is given by:

Equation 2: mRNA/RBP Equilibrium For Time 0

$$Ka_i^{\,0} = \frac{f_i^{0+IP} - f_i^{0-IP}}{f_i^{0Total} - (f_i^{0+IP} - f_i^{0-IP})} \bullet \frac{1}{[HuR]}$$

where the relative concentration of each message, fiTotal, is determined by total RNA microarray analysis. For each message, i, the relative concentration of [AB] is calculated from the difference of the positive and negative (control) IP mean microarray signals (f+IP – f-IP).

Array data from our previous study on Jurkat T-cell activation (Mukherjee, Lager et al. 2009) provide values for fiTotal, f+IP, and f-IP for multiple time points. The simple model presented in Equation 1 aims to explain dynamics in mRNA/RBP association as the result of temporal differences in RBP concentration ([B]) and mRNA concentration ([A]). If other cellular factors contribute to association dynamics, then the model will

not explain the observed changes in relative concentration of the mRNA/RBP complex ([AB]). Ideally, experimental values for [A], [B], and [AB] at each time point would be utilized to determine the degree to which the model predicts these observed changes. However, I was unable to obtain quantitative data on the change in HuR concentration ([B]), as the study did not analyze protein concentrations in a quantitative manner. Thus, I calculate an idealized change in [B], represented by c, assuming no factors other than RBP and mRNA concentrations are responsible for the observed changes in dynamics:

Equation 3: Estimate of Change in RBP Concentration

$$\hat{r}_0 = \frac{f_i^{0+IP} - f_i^{0-IP}}{f_i^{0\,Total} - (f_i^{0+IP} - f_i^{0-IP})}$$

$$c = \frac{\hat{r}_4}{\hat{r}_0} \qquad \hat{c} = median\left(\frac{\hat{r}_4}{\hat{r}_0}\right)$$

This estimate allows us to ask to what extent the observed changes in [A] and the calculated changes in [B] (represented by c) explain the observed dynamic changes in HuR RNP composition [AB] from 0 to 4 hours post activation:

Equation 4: Estimate of Dynamic Change in Association

$$\hat{r}_0 \overset{?}{\hat{c}} = \hat{r}_4 + \varepsilon$$

I found that the dynamics of HuR association for many messages are well explained by changes in mRNA and HuR abundance alone (Figure 10). However, for many other messages, further experimental data, such as competition with other RBPs in the cell or differential phosphorylation of the RBP, will be required to determine the factors underlying the dynamics of RBP association. Additionally, the changes in association observed for a number of messages fall within the experimental noise and I was unable to draw conclusions regarding the possible mechanisms responsible for the observed changes in their association with HuR.

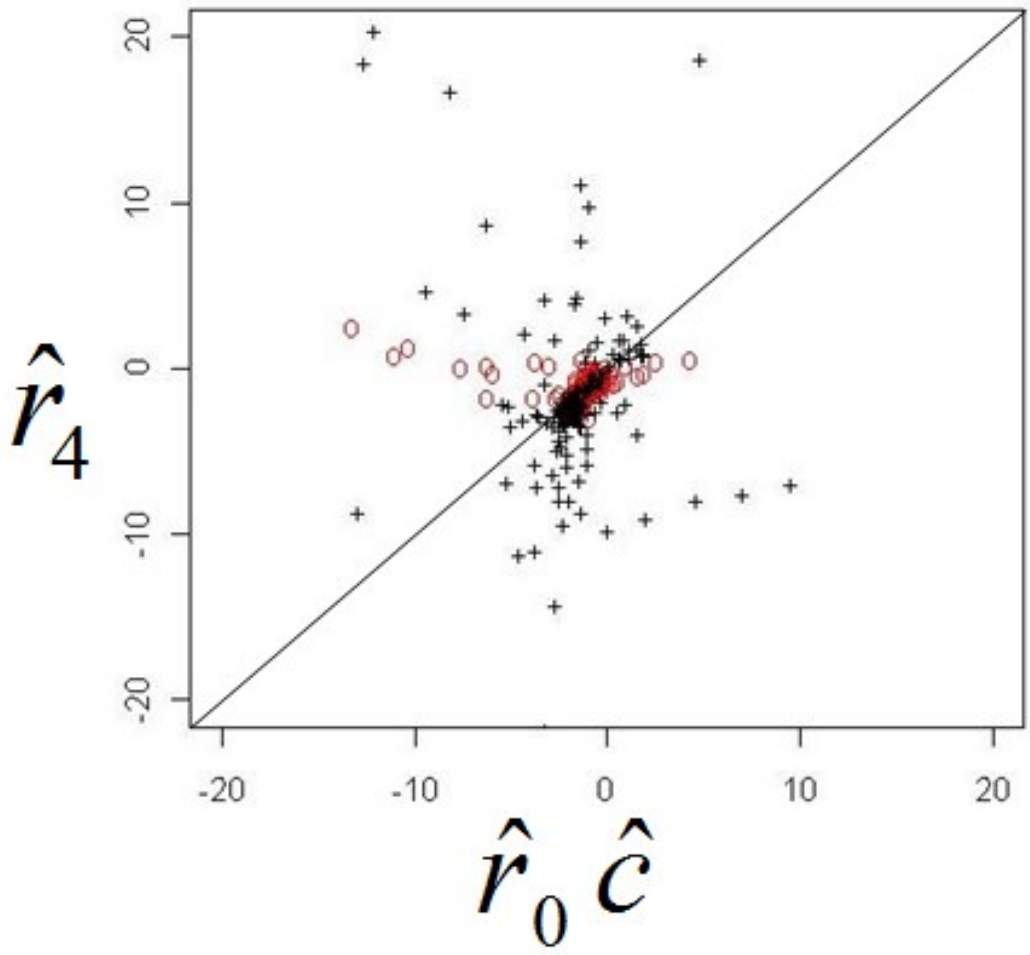


Figure 10: Values of r4 are largely predicted by r0c.

The points whose confidence intervals do not cross the y=x line are denoted by red circles. These points represent messages whose dynamic association with HuR is not explained by a change in [HuR] from 0 to 4 hours post-activation. This suggests that posttranslational modification or the state of other RNPs may be responsible for these dynamics. The confidence intervals for points denoted by (+) cross the y=x line, indicating that the experimental noise for these measurements makes it impossible to draw conclusions regarding possible mechanisms for the observed changes.

3.2.2 Functional Dynamics

3.2.2.1 Translational Regulation

Generally, RIP-Chip based posttranscriptional studies evaluate global RNP composition by microarray or deep sequencing. However, functional characterization of RNP composition has been limited to studying a handful of mRNA regulatory elements using reporter assays. Understanding of posttranscriptional gene expression coordination from a global perspective has been greatly limited by the absence of global approaches to evaluate functional outcomes of RNP association. To this end, I examined an additional level of global posttranscriptional gene expression regulation based on polysome profiling followed by microarray analysis to examine the functional "translatomic" outcomes of dynamic RNP association. Many studies have employed polysome profiling microarrays to evaluate global changes in translation following a perturbation, and such information is closely related to RNP analysis by RIP-Chip (Lu, de la Pena et al. 2006, Mamane, Petroulakis et al. 2007).

As shown in Figure 11, HuR protein is partially localized in the heavy polysome fractions during all three time points in the Jurkat activation system.

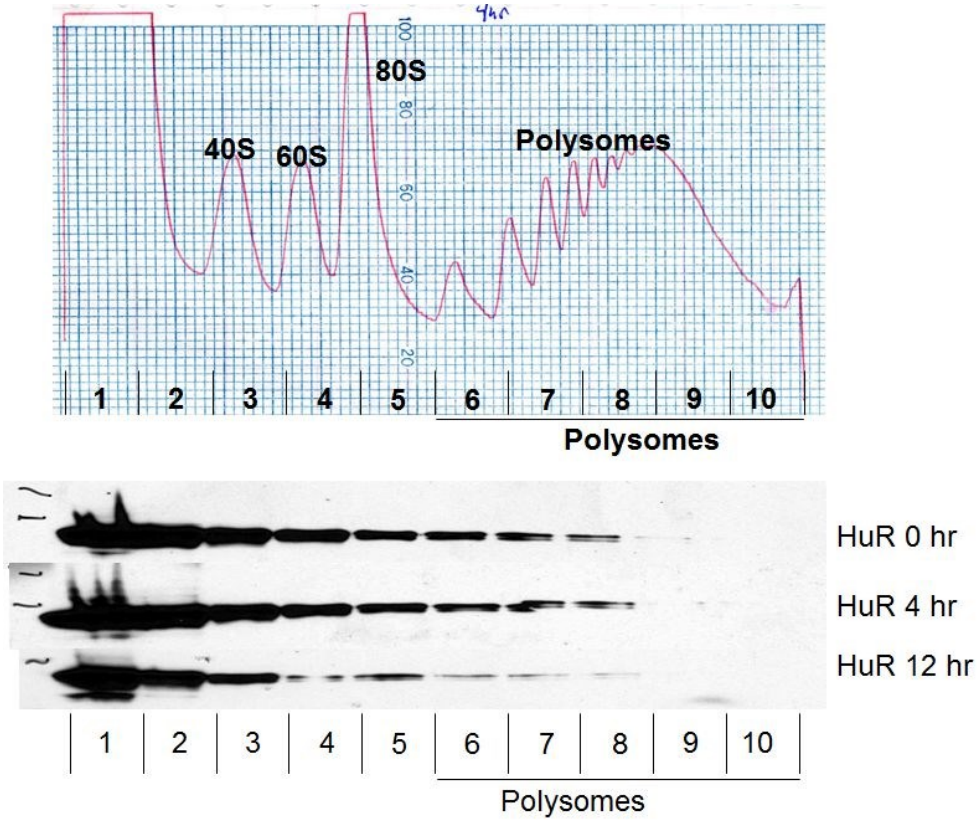


Figure 11: HuR Protein is localized to heavy polysome fractions during Jurkat T cell activation.

Upper panel shows a representative UV absorbance plot for a polysome gradient. Fraction numbers are indicated below the trace. Lower panel shows western blots of HuR protein localization across polysome gradients at 0, 4, and 12 hours post Jurkat activation.

Given our previous findings that the set of mRNAs associated with HuR changes dynamically throughout the process of Jurkat activation, I sought to explore the functional consequences of HuR association, as reflected in translation status of HuR target mRNAs. Thus, I performed microarray analysis of total mRNA and heavy

polysome fraction mRNA from Jurkat T-cells at 0hr, 4hr, and 12hr post activation (Figure 12).

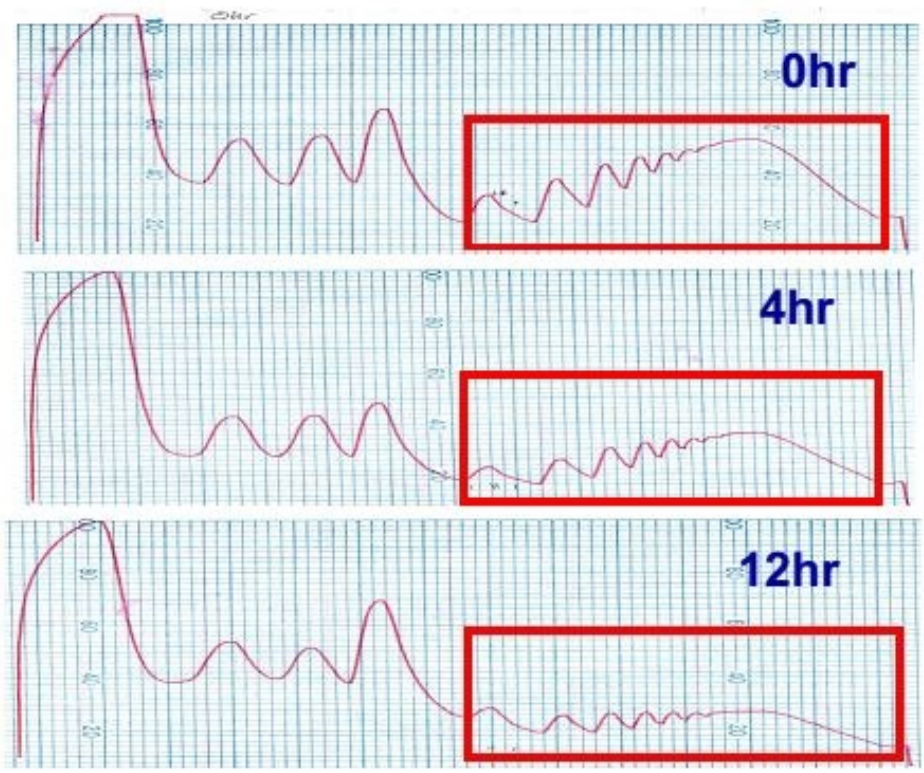


Figure 12: Polysome profiling of activated Jurkat T-Cells

Polysome-associated mRNA (red boxes) was pooled and analyzed by microarray analysis from Jurkat T cells at 0, 4, and 12 hours post activation.

Following array normalization, T-scores of change in polysome-association and change in total mRNA levels between time points were calculated for each message. These T-scores were then compared to one another (pairwise correlation) to examine the differences in total mRNA levels or polysome association from 0-4hr and from 4-12hr following Jurkat T-cell activation. The data, presented in Figure 13, indicate that the

total message levels and polysome association dynamics differ across the two time intervals (red circles). Further, the dynamic changes observed in the total mRNA levels differ from the changes observed in polysome association (blue circles), suggesting the activity of specific translational control mechanisms during this process.

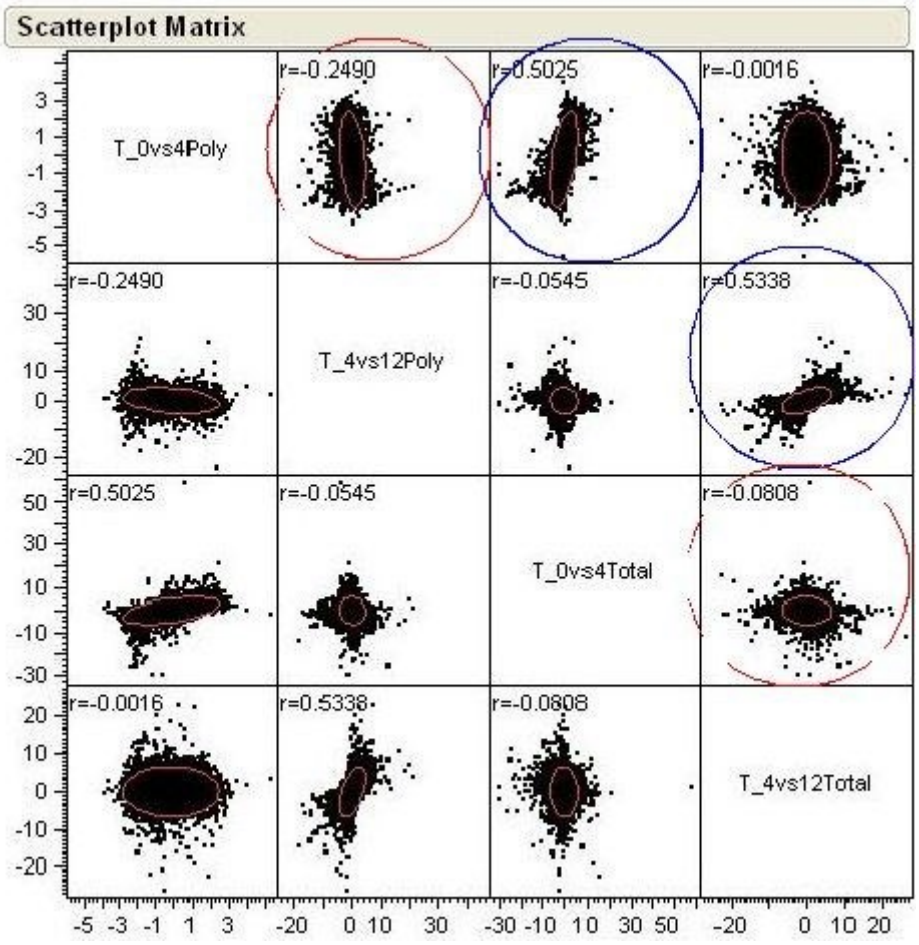


Figure 13: Correlations between T-scores measuring change in total mRNA levels or polysome association from 0-4hr and from 4-12hr post Jurkat T-cell activation

Changes in total RNA and polysome-associated RNA abundance were evaluated by microarray analysis. T-scores representing the differences between timepoints were calculated: T_0vs4Poly indicates the T-score of change in polysome association from 0 to 4 hours post-activation; T_4vs12Poly indicates the T-score of change in polysome association from 4 to 12 hours post-activation; T_0vs4Total indicates the T-score of change in total RNA abundance from 0 to 4 hours post-activation; T_4vs12Total indicates the T-score of change in total RNA abundance from 4 to 12 hours post-activation. Red circles indicate comparisons between the changes from 0 to 4 hours and 4 to 12 hours. Blue circles indicate comparisons between changes in total RNA abundance and changes in polysome association.

3.2.2.2 Integration with RBP Targeting Data

To further explore the functional consequences of RBP-mRNA interaction, I compared dynamic changes in HuR-mRNA association with the dynamic changes of mRNA in the polysome gradient experiments. Because I am interested in specific alterations in translational control, I calculated changes in polysome association relative to changes in total mRNA abundance. Thus, differences observed would not be due to increases or decreases in the absolute amount of a given mRNA, but instead should represent message-specific recruitments to or away from the translation machinery. Data from our previously published study (Mukherjee, Lager et al. 2009) on dynamic changes in HuR-mRNA association between 0 and 4 hours post Jurkat T-Cell activation were classified as 1) HuR associated at both 0 and 4 hours, 2) loss of HuR association from 0 to 4 hours, 3) gain of HuR association from 0 to 4 hours, and 4) not HuR associated at either time point for every mRNA. Figure 14 compares the cumulative probability of an mRNA increasing in association with the heavy polysome fractions between messages classified into these four groups based on their HuR association dynamics. Surprisingly, the comparison shows that those messages that lose HuR association from 0 to 4 hours (green line) are significantly ($p = 0.0148$) more likely to increase in their association with polysomes than those messages that are always HuR associated (red line).

Although HuR is generally believed to increase the stability and translation of its target messages, numerous studies (Katsanou, Papadaki et al. 2005, Leandersson, Riesbeck et al. 2006, Kim, Kuwano et al. 2009) have identified systems in which HuR acts as a repressor of translation. Also, it is interesting to note that in these data, those messages that gain association with HuR are not differentially localized in the polysome gradient, compared to those that are never HuR associated or those that are HuR associated at both 0 and 4 hours post-activation. However, the data do show a significant difference in the likelihood of polysome association between those messages that are HuR associated at both 0 and 4 hours (red line) and those that are never HuR associated (orange line, $p = 0.0006$). Similar to those messages that lose HuR association from 0 to 4 hours post-activation, the messages that are not found associated with HuR at either time point show a greater tendency to be localized in the heavy polysome fractions of the gradient after activation.

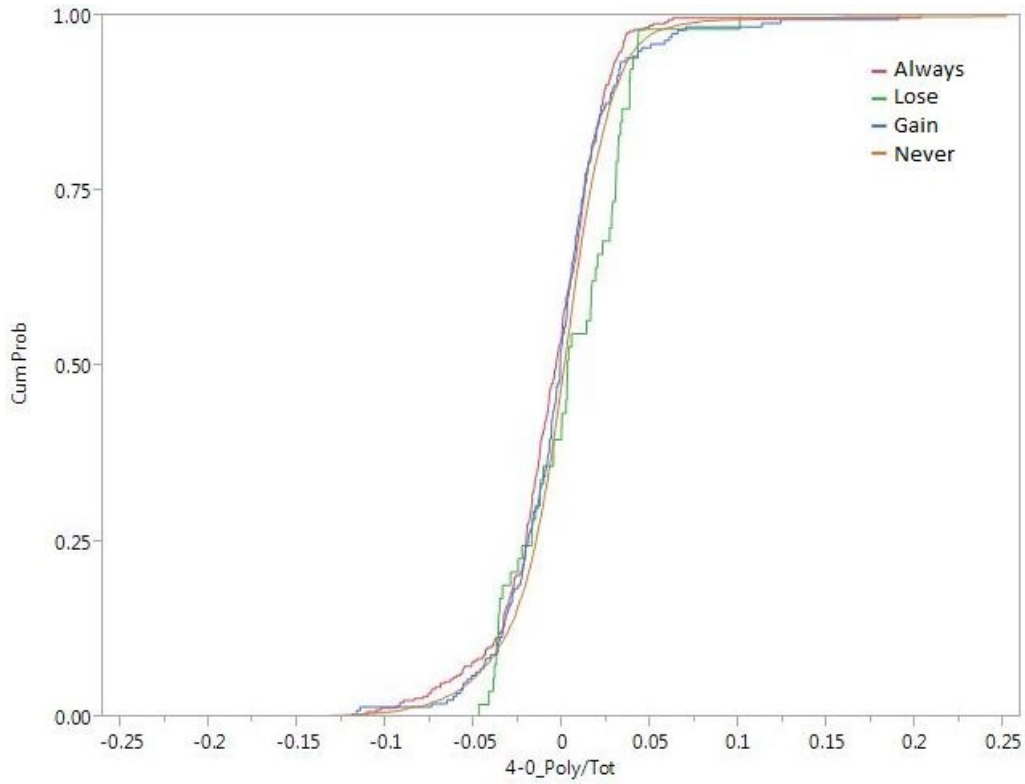


Figure 14: Relationship Between Change in HuR Association and Change in Polysome Association

Cumulative probability function comparing change in polysome association to change in HuR target status. The horizontal axis represents change in polysome association (normalized to change in total RNA level) from 0 to 4 hours post-activation. The red line represents those messages that are HuR associated at both 0 and 4 hours; the green line represents those messages that are HuR associated at 0 hours but not 4 hours; the blue line represents those messages that are HuR associated at 4 hours but not at 0 hours; the orange line represents those messages that are not HuR associated at 0 or 4 hours.

In addition to comparing the polysome association data with RBP targeting data from the Jurkat T-Cell activation system, I also examined the relationships between this data and previously published RBP targeting data sets. Specifically, I utilized PAR-Clip binding site data for numerous RBPs available in the literature that was collected in a

previous study from our lab (Friedersdorf and Keene 2014). Using the binding site information from these datasets, we categorized each message as a target or non-target (and in some cases, strong target) of each RBP. These lists were then compared to the dynamic polysome association data. Figure 15 shows several representative comparisons for target sets based on PAR-Clip experiments on the RBPs Ago2, Pum2, and HuR. In agreement with the data presented previously, this analysis indicates that association with HuR negatively correlates with association with the heavy polysome fractions at 4 hours post-activation in Jurkat T-Cells ($p < 0.0001$). Indeed, the group of 'strong targets' (those with more than 4 PAR-Clip sites, blue line) show a larger negative correlation than the less strong targets.

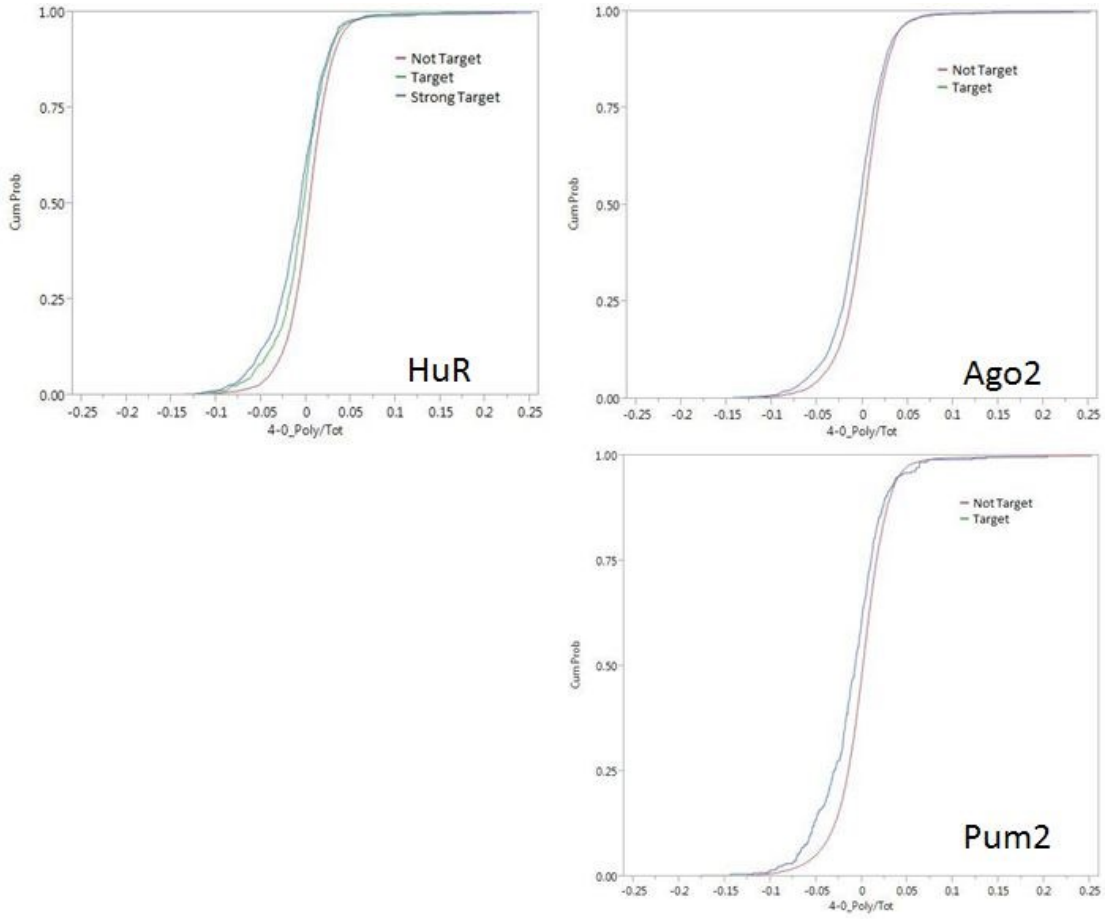


Figure 15: Relationship Between RBP Targeting and Change in Polysome Association

Cumulative probability functions comparing change in polysome association to RBP targeting databases. The horizontal axis represents change in polysome association (normalized to change in total RNA level) from 0 to 4 hours post-activation. The red lines represent messages that are not targets of the specific RBP tested. Blue and green lines represent messages that are target or strong targets (green) of the specific RBP tested.

Similarly, those messages identified as containing PAR-Clip sites for both Pum2 and Ago2 (blue lines) show significant ($p < 0.0001$) negative correlation with association with the heavy polysome fractions at 4 hours post-activation. Table 5 shows the p

values of association between all PAR-CLIP target sets and the total-normalized change in polysome association from 0 hours to 4 hours post activation.

Table 5: p-values of Relationship Between RBP Targeting and Change in Polysome Association

RBP	p value
Ago2	< 0.0001
ALKBH5	< 0.0001
C17ORF85	0.0873
C22ORF28	0.0003
CAPRIN1	0.0001
FMR1	0.5747
FMR1 iso1	< 0.0001
FXR1	< 0.0001
FXR2	< 0.0001
HuR	< 0.001
IGF2BP2	< 0.0001
LIN28B	< 0.0001
Pum2	< 0.0001
TTP	0.38
ZC3H7B	< 0.0001

In all cases of significant differences between groups, the 'Target' group showed a negative correlation with the change in polysome association compared to the 'Non-Target' group. This trend is unsurprising for several well-established mRNA repressors (Ago2, Pum2, FMR1, and FXR1), and may further suggest that much of the transcriptome is, in fact, translationally repressed during the process of T-Cell activation. It is also possible that this negative correlation trend is, at least partially, due to the fact that the majority of the RBPs for which PAR-CLIP data are available are predominantly negative regulators of expression.

3.2.2.3 Integration with Global Stability Data

A currently ongoing collaborative study in our lab is exploring the role of changes in mRNA stability during the process of Jurkat T-Cell activation. RNA stability is tightly linked to the regulation of translation. Many stability-regulating factors interact with and directly regulate translation complexes, as well as the stability of other target mRNAs. Utilizing 4SU labeling and separation, Jeffrey Blackinton performed global stability analysis on Jurkat T-Cells at 0 and 4 hours post-activation. Data from these experiments were analyzed to determine mRNA half-lives during this activation process. Messages were then separated into groups according to their change in half-life (HL) from 0 to 4 hours post-activation (Decreased HL, No change, Increased HL).

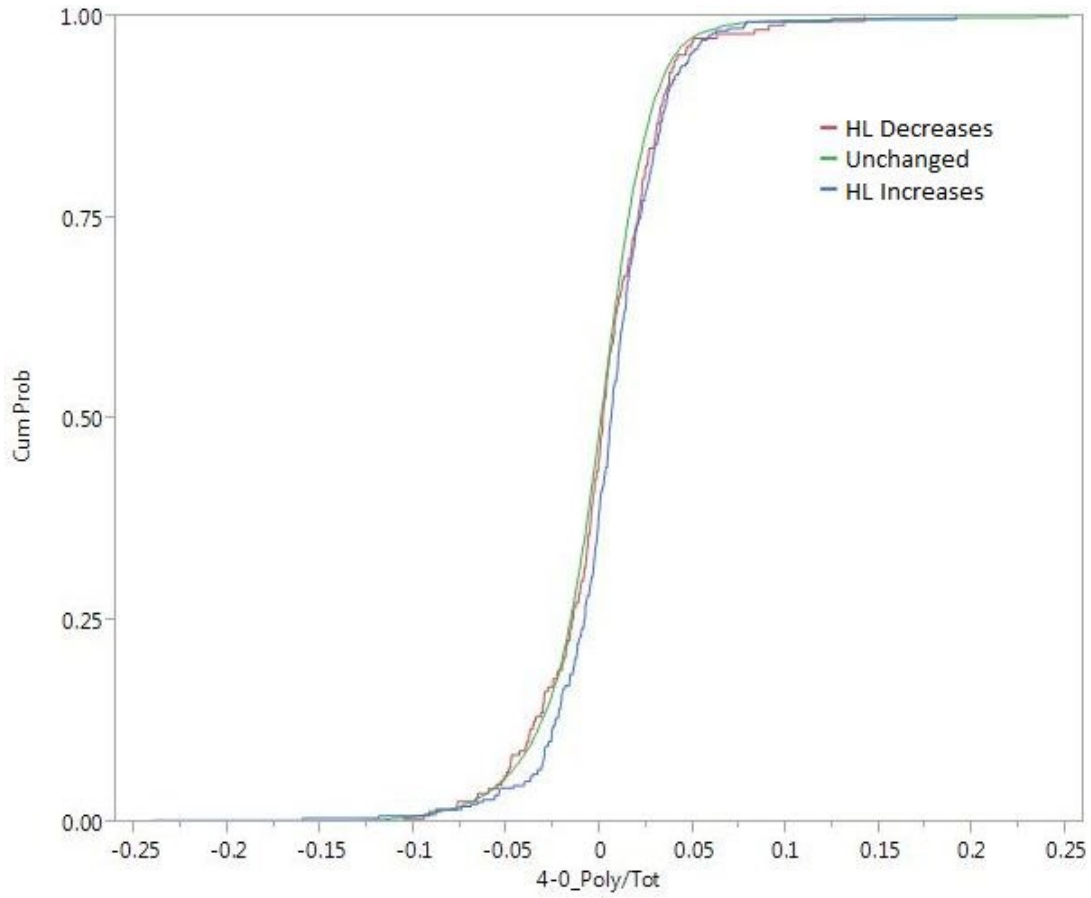


Figure 16: Relationship Between Change in mRNA Half-life and Change in Polysome Association

Cumulative probability functions comparing change in polysome association to change in mRNA half-life. The horizontal axis represents change in polysome association (normalized to change in total RNA level) from 0 to 4 hours post-activation. The red line represents messages with RNA half-lives that decrease from 0 to 4 hours post-activation; the green line represents messages with RNA half-lives that do not change from 0 to 4 hours; and the blue line represents messages with RNA half-lives that increase from 0 to 4 hours post-activation.

Figure 16 shows the comparison between change in half-life and dynamic association with the heavy polysomes from 0 to 4 hours following Jurkat activation. The group of messages with half-lives that increase from 0 to 4 hours (blue line) show a

significant positive association ($p = 0.0003$) with the likelihood of associating with the heavy polysomes after activation. Given the tight mechanistic linkages between translation and mRNA stability, it is unsurprising to find that stabilized messages are also more likely to increase in translation. However, the previously described findings that none of the tested RBP targeting datasets positively correlate with the change in polysome association suggest that this increase in stability and translation may be due to the actions of a yet unidentified RBP or RBPs that stimulate the translation and prevent decay of mRNAs important for the process of T-Cell activation.

3.3 Materials and Methods

Data sources and analysis: Data used for the thermodynamic model and HuR target data used for comparison with polysome association were obtained from the supplemental information of (Mukherjee, Lager et al. 2009). PAR-CLIP target data were obtained from the supplemental information of (Friedersdorf and Keene 2014). mRNA half-life data were kindly provided by Jeffrey Blackinton. All calculations for the thermodynamic model were performed in R, and correlation and cumulative distribution plots were generated with JMP 11 (SAS). Statistical significance was considered $p < 0.05$ via all pairs Student's T tests.

Cell culture: Jurkat cells were cultured in RPMI 1640 with the addition of 10% FBS and penicillin/streptomycin at 37 C and 5% CO2. Activation was performed by addition of 50ng/mL PMA and 2ug/mL PHA for the indicated times.

Polysome gradients: Activated Jurkat cells were treated with cycloheximide (CHX) at a concentration of 200uM for 15 minutes then harvested in lysis buffer (400mM KOAc, 24mM HEPES, 5mM MGOAc2, 0.5% Na-deoxycholate, 0.5% Triton-X-100, 1mM DTT, 1mM PMSF, 200uM CHX, 50U/mL RNaseOut) and incubated for 15 minutes on ice. Lysates were precleared by centrifugation at 3,000 x g for 8 minutes and loaded on 15-50% sucrose density gradients. Gradients were spun for 3hr at 35,000 rpm in a SW41 swinging bucket rotor. Gradients were analyzed and collected as 10 fractions each on an ISCO density gradient fractionator. RNA was extracted from 250uL of each fraction with Trizol LS (Invitrogen) according to the manufacturer's protocol.

Western blots: HuR was detected using the mouse monoclonal 3A2 antibody and an HRP conjugated anti-mouse secondary antibody (GE Healthcare). Visualization was performed using Pierce ECL reagent.

Microarrays: Trizol-LS extracted RNA from the heavy polysome fractions (indicated by red boxes in Figure 12) was pooled and equal quantities of polysome RNA and total RNA were submitted for array analysis. All samples were run in biological triplicate on HO36K spotted oligo microarrays (Operon). Data were quantile normalized and all further analysis was performed in JMP 11 (SAS).

3.4 Discussion

Our Thermodynamic Model of RBP-mRNA association provides, to my knowledge, the first predictive statistical model of dynamics in a posttranscriptional

regulatory system. My analysis of dynamic HuR association data using this model highlights the combinatorial nature of the posttranscriptional regulatory environment. In agreement with the conclusions from the study of the yeast posttranscriptional network, this modeling work indicates that while some of the dynamics of association observed in the T-Cell system may be simply explained by changes in mRNA abundance (which are likely the result of transcriptional dynamics), many of the changes observed can't be attributed to such alterations. Instead, I suggest that competition and/or cooperation between other posttranscriptional regulatory factors, including RBPs and small RNAs, is likely responsible for driving these changes.

Several studies have identified specific examples of both competition and cooperation between HuR and another ARE-binding protein, AUF1 (Lal, Mazan-Mamczarz et al. 2004, Zou, Rao et al. 2010). Additionally, HuR has been found to interact with the miRNA regulatory system to both counteract (Kundu, Fabian et al. 2012) and promote (Kim, Kuwano et al. 2009) miRNA-mediated repression of target mRNAs. A recent review (Ho and Marsden 2014) summarized many of the complex and somewhat conflicting relationships that have been characterized between HuR, the miRNA regulatory system, and other RBPs (including AUF and TTP). Strikingly, a major conclusion of their summary is that global characterization of the posttranscriptional regulatory environment, and particularly the myriad interactions within the system, will be a critical step in understanding the various mechanisms of

gene expression. The modeling approaches presented in Chapter 2 of this document are well suited to integration and analysis of this type of complex regulatory data.

Our thermodynamic model may be extended and adapted to answer a number of important biological questions. For example, are RBPs or mRNAs the limiting factor in RNP formation? Similarly, future integration of miRNA target association data and RBP-mRNA association data will permit the model to estimate the relative contributions of each individual element (RBP or miRNA) to the experimentally observed dynamic changes in mRNP composition.

To further explore the relationship between RBP association dynamics and functional outcomes, I endeavored to globally evaluate functional consequences of changes in mRNA-RBP associations. I quantified global changes in polysome occupancy of mRNAs following Jurkat T-Cell activation via microarray analysis and correlated those data with previously discussed data on HuR-mRNA association dynamics in the same system. Although I was somewhat surprised by the observation that loss of HuR association correlates with increased localization to the polysome fractions, several studies have identified similar non-canonical behaviors of HuR (Katsanou, Papadaki et al. 2005, Leandersson, Riesbeck et al. 2006, Kim, Kuwano et al. 2009). In particular, the Katsanou study demonstrated that exogenous expression of HuR resulted in the shift of target messages encoding inflammatory response proteins away from the polysome fractions in response to inflammatory stimulation (LPS). Similarly to this study, they

found that treatment with pro-inflammatory mitogens, LPS in their case and PMA/PHA in my case, resulted in the reduction of polysome association of HuR target messages. They further concluded that HuR likely acts through multiple mechanisms and co-effectors that enable multifunctional control of gene expression, a conclusion that further supports the argument of the importance of global modeling and integrated data analysis.

Unfortunately, the distribution of FDRs in the polysome dataset indicates a very high level of noise in quantifying the changes in polysome association, making more fine-grained modeling and prediction somewhat intractable. I believe that this suggests the need for additional replicates or new analytical techniques to characterize the dynamics of polysome association in this system. Further, the array technology available when these experiments were performed is vastly inferior to modern sequencing-based approaches.

Additional characterization of the links between the functional translation data and previously established RBP-mRNA interactions was performed by correlating the dynamic translation data with published lists of RBP target sets identified through PAR-CLIP (Friedersdorf and Keene 2014). My observations for many of the examined RBP target lists agreed with my findings on HuR targets. Those mRNAs with putative target sites for many known repressive RBPs, such as Ago2, Pum2, and FXR, are less likely to increase in their association with the heavy polysome fractions following Jurkat

activation. Ideally, dynamic targeting data for additional RBPs would be available in the Jurkat activation system, allowing more thoroughly quantification and exploration of the combinatorial contributions of multiple RBPs on the dynamic regulation of the translational state of mRNAs on a global scale.

4. Mechanisms of Control of Gene Expression Mediated by Pum1

4.1 Introduction

Although global studies of regulatory networks will ultimately be required to understand and predict the overall behaviors of complex biological gene expression processes, the development of in-depth models requires knowledge of specific mechanistic relationships between members of the regulatory networks. The previously discussed observation that RBP targeting correlates with changes in translation status suggests a mechanistic linkage between association with many RBPs and recruitment to or away from the translation machinery. Further, the tight linkage between regulation of mRNA translation and stability leads us to ask to what degree the dynamic associations with RBPs affects the regulation of mRNA decay.

Evaluation of the contributions of RBPs to either or both of these facets of posttranscriptional control has proved somewhat difficult, as direct influences on RNA stability are often responsible for secondary effects on translation and vice versa. While some RBPs may be directly responsible for the recruitment to or away from the translation machinery (for example, eIF4E) or the direct catalysis of RNA (for example ADAR2 and Ago) (Lunde, Moore et al. 2007), most require the action of other RBPs and cellular co-factors. I investigated the specific mechanisms of posttranscriptional control mediated by a well-characterized RBP, Pum1. Pum1 is closely related to one of the RBPs previously described as displaying a likely effect on translational control based on the

dynamic polysome analysis (Pum2). Additionally, abundant literature on PUF family proteins indicates the potential for a dual role in translational control and regulation of mRNA stability, although the details of the mechanistic interactions required for Pum1's regulatory functions remain largely uncharacterized. The mechanistic study of Pum1 will provide important functional linkage data that can be incorporated into network models of posttranscriptional regulation to further enhance our ability to evaluate and predict behaviors of these complex regulatory systems.

4.2 Results

4.2.1 Pum1 Binds SIRT1 and PCNA mRNAs

Previous studies have demonstrated that Pum1 protein associates with a subset of cellular mRNAs (Galgano, Forrer et al. 2008, Morris, Mukherjee et al. 2008). Although several of these studies identified SIRT1 as a putative mRNA target of Pum1 protein, further characterization of this interaction or its functional consequences has not been presented. To evaluate the ability of Pum1 to bind the SIRT1 mRNA, I performed RNA Immunoprecipitation (RIP) analysis of FLAG-tagged Pum1 protein from HEK-293T cells. Enrichments in FLAG IPs compared to negative IPs (RIPs performed on lysates lacking FLAG-Pum) and to total RNA levels were calculated using GAPDH for normalization. As shown in Figure 17, SIRT1 mRNA is highly (~200 fold) enriched in the Pum1 IPs, while the mRNAs for HuR, B2M, and Histone 1 do not show substantial enrichment compared to negative IPs. Additionally, PCNA mRNA, a known Pum1

target message (Morris, Mukherjee et al. 2008), also shows strong enrichment in the Pum1 IPs (~50 fold).

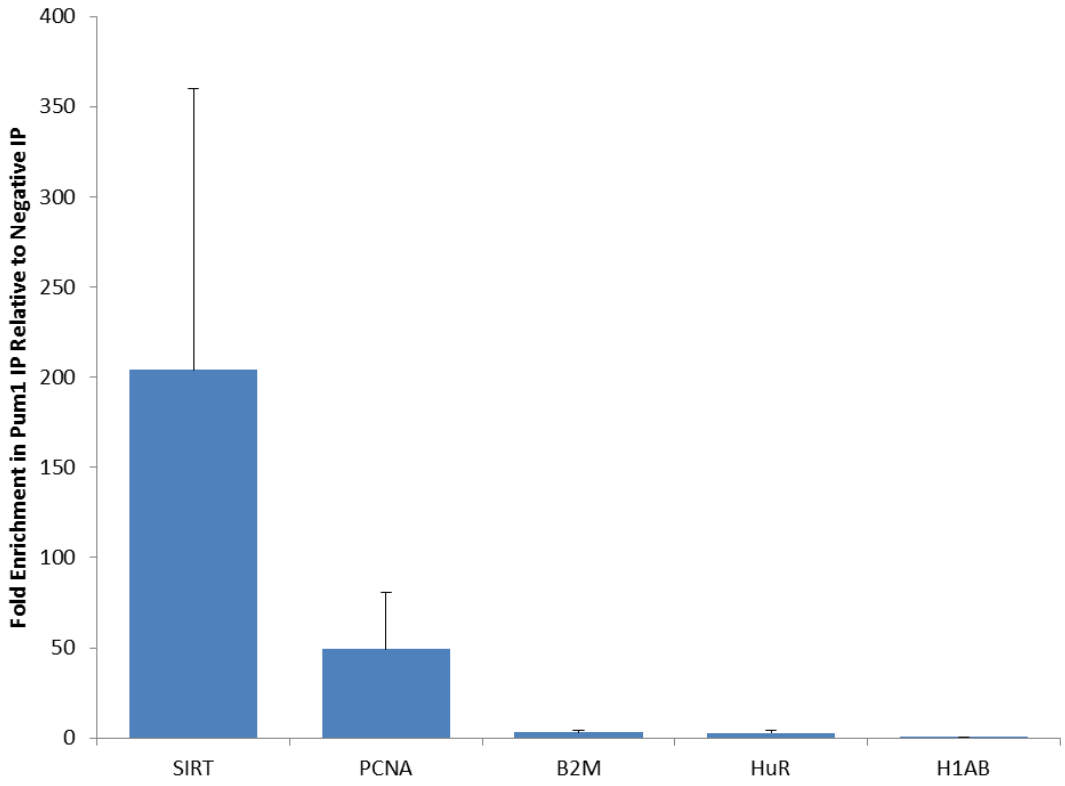


Figure 17: Pum1 Binds SIRT1 and PCNA mRNAs

Fold enrichments of mRNAs in Pum1 RIP experiments were calculated relative to a negative RIP and normalized to the GAPDH mRNA. SIRT1 and PCNA are putative and known RNA targets of Pum1, respectively, while B2M, HuR, and H1AB (Histone 1) are not known targets of Pum1.

4.2.2 Pum1 represses expression through the SIRT1 and PCNA UTRs

To evaluate the functional consequences of the interaction between Pum1 and the SIRT1 mRNA, I generated luciferase reporters containing the full length 3′ UTRs of SIRT1, PCNA, and TTP (as a non-target control). Cells were co-transfected with

luciferase reporters and a FLAG-Pum overexpression plasmid. The presence of a target UTR (either PCNA or SIRT1) and full length Pum1 protein resulted in a decrease in luciferase expression of between 40 and 60% (Figure 18). In contrast, luciferase expression was unchanged when the reporter contained a non-Pum1 target UTR (TTP).

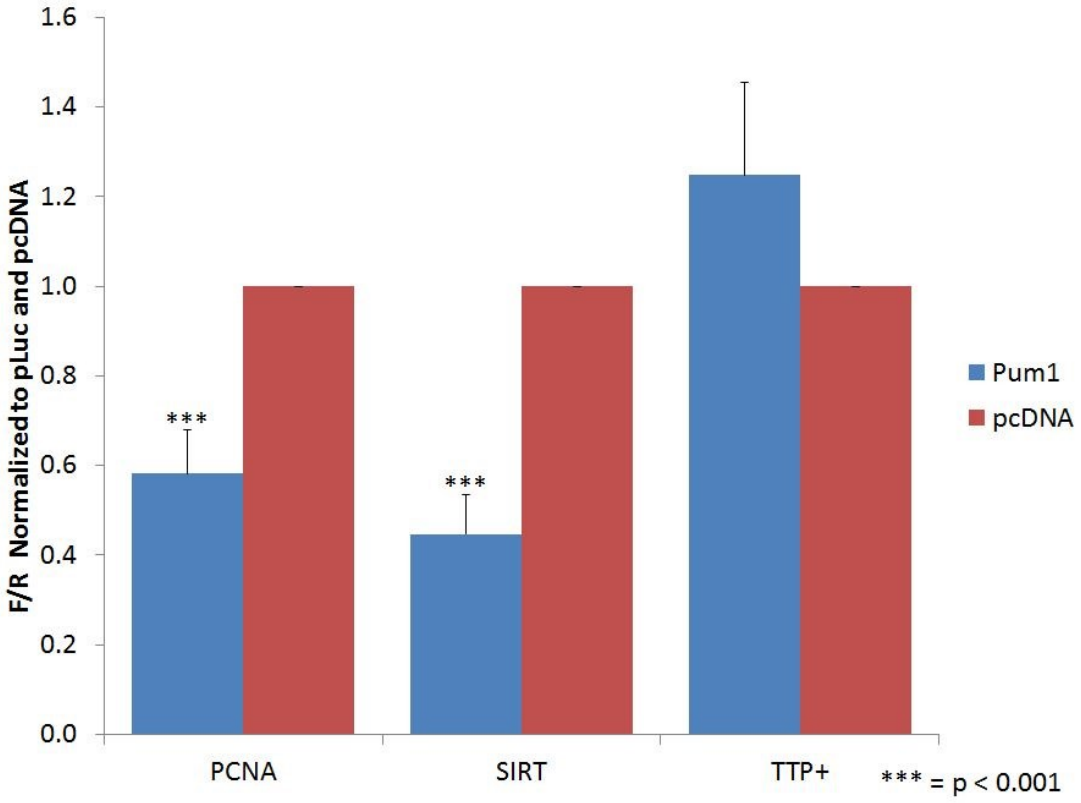


Figure 18: Pum1 Represses Expression of its Targets

Full length FLAG-Pum1 was expressed together with luciferase constructs containing the 3′ UTRs of Pum1 target (PCNA, SIRT1) and non-target (TTP) mRNAs. Luciferase activity was normalized to the luciferase vector lacking a UTR (pLuc) and cells transfected with an empty expression vector (pcDNA). * indicates $p < 0.001$ vs. pcDNA control.**

4.2.3 The region upstream of the Pum1 RNA binding domain is required for repression

Pum1 contains a single Pumilio-homology domain (HD) near its C-terminal region. Previous studies have demonstrated that the HD is responsible for the RNA binding activity of Pum1 (Zamore, Williamson et al. 1997), as well as the repressive function of Pum1 through recruitment of deadenylase complexes (Goldstrohm, Hook et al. 2006, Goldstrohm, Seay et al. 2007, Van Etten, Schagat et al. 2012). However, another study identified regions upstream of the HD as potentially important for the regulatory function of *Drosophila* Pumilio and human Pum1 in an in vitro system (Weidmann and Goldstrohm 2012). To determine whether these upstream regions are required for Pum1-mediated repression of SIRT1, I co-transfected luciferase reporters with constructs expressing Pum1 truncations. Figure 18 provides an overview of the Pum1 protein truncations tested. The FLAG-HD construct contains only the defined Pumilio-homology domain of Pum1. The FLAG-HD+ construct includes the homology domain with an additional 239 upstream amino acids and includes the putative upstream regulatory regions identified by (Weidmann and Goldstrohm 2012).

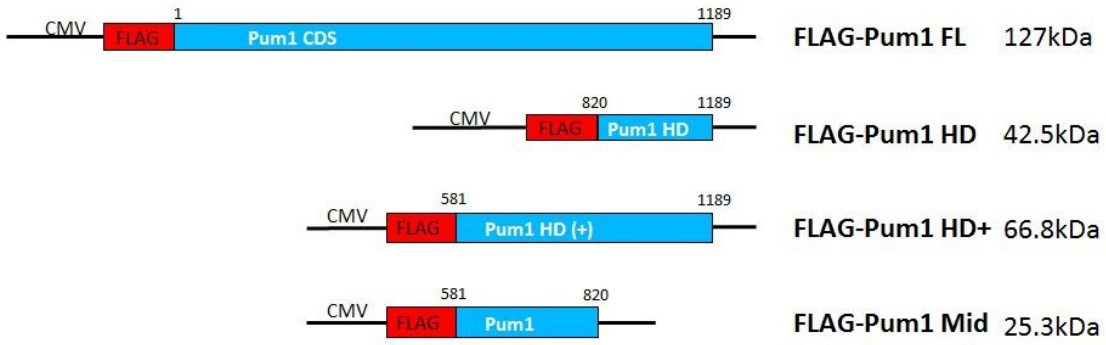


Figure 19: Pum1 Truncation Constructs

Diagram of the Pum1 RBP truncation vectors and their respective protein product sizes. FLAG-Pum FL encodes the full length Pum1 protein with an N-terminal FLAG tag. FLAG-Pum1 HD encodes the C-terminal 369 amino acids of the Pum1 protein, which includes the entire PUM homology domain (HD). FLAG-Pum1 HD+ encodes the C-terminal 608 amino acids of the Pum1 protein, including the PUM homology domain and an additional 239 upstream amino acids.

As shown in Figure 20, the expression of the Pum1 homology domain alone does not affect the expression of any luciferase vectors, independent of the presence of a Pum1 target UTR. However, the expression of both the SIRT1 and PCNA UTR constructs is dramatically repressed when co-transfected with the FLAG-HD+ construct.

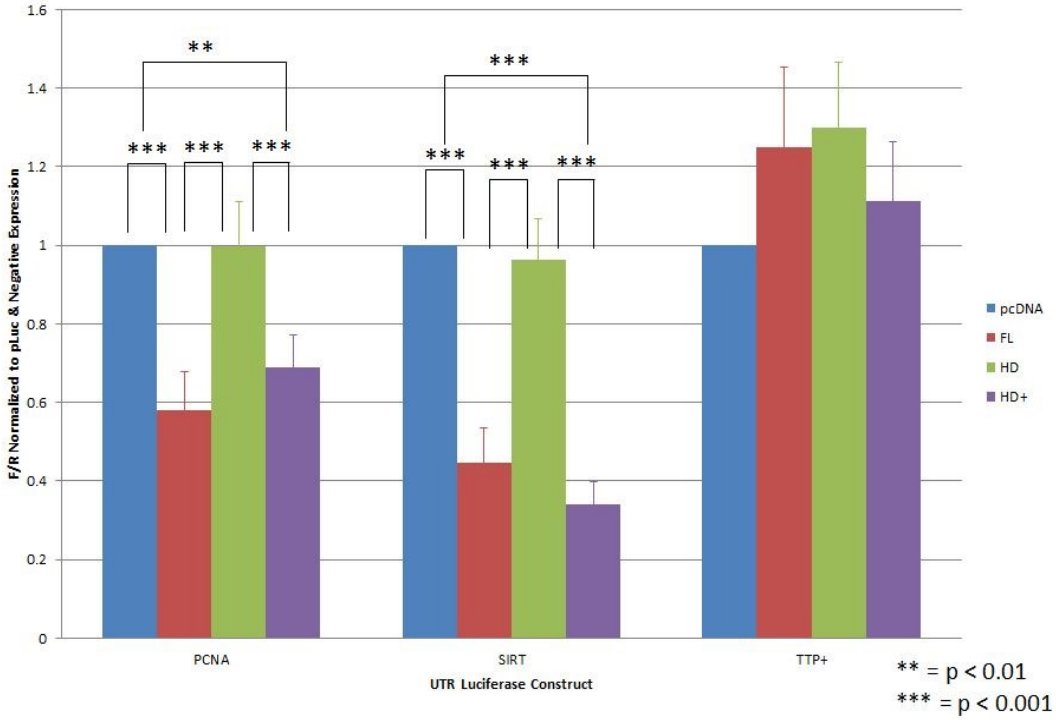


Figure 20: The Region Upstream of the Pum1 HD is Required for Translational Repression

The full length, HD, and HD+ Pum1 truncations were expressed together with luciferase constructs containing the 3′ UTRs of Pum1 target (PCNA, SIRT1) and non-target (TTP) mRNAs. Luciferase activity was normalized to the luciferase vector lacking a UTR (pLuc) and cells transfected with an empty expression vector (pcDNA). ** indicates p < 0.01 vs pcDNA control, * indicates p < 0.001 vs. pcDNA control.**

Figure 21 shows that this loss of regulation is not due to the loss of RNA binding, as SIRT1 and PCNA mRNAs are highly enriched in IPs of Pum-HD truncation constructs compared to negative IPs.

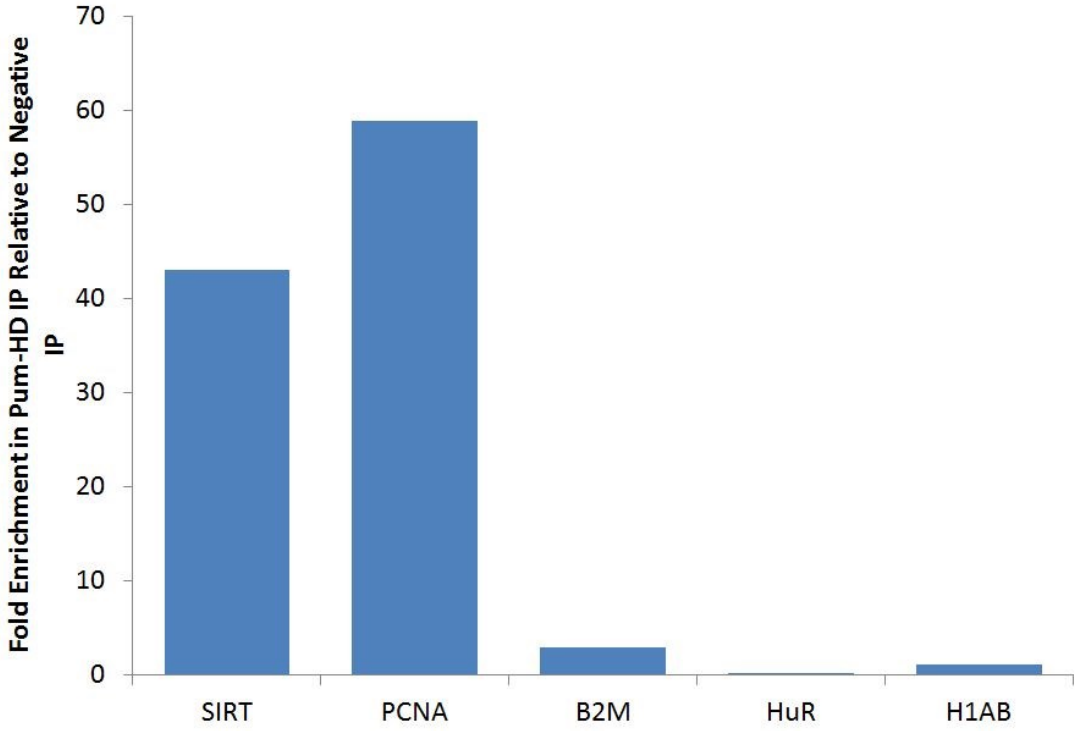


Figure 21: Pum HD Truncation Binds Target mRNAs

Fold enrichments of mRNAs in Pum1-HD RIP experiments were calculated relative to a negative RIP and normalized to the GAPDH mRNA to ensure the Pum1-HD construct retained its RNA binding function. Enrichment of Pum1 target mRNAs (PCNA and SIRT1) in the RIP relative to non-target mRNAs (B2M, HuR, Histone) indicate that the HD truncation binds Pum1 target mRNAs.

Because the homology domain has been previously shown to interact with the cellular deadenylase system to trigger destabilization of its target messages, I was surprised to find that the HD region alone is unable to repress expression in this system. These data suggest that the observed repression may be due to a mechanism other than recruitment of the deadenylase complex.

4.2.4 Role of Pum1 in Regulation of Stability

PUF family proteins have been characterized as negative regulators of mRNA stability in yeast (Olivas and Parker 2000, Hook, Goldstrohm et al. 2007). In addition, in vitro and in vivo studies have demonstrated functional interactions between PUF family proteins and enzymes responsible for mRNA deadenylation (Goldstrohm, Hook et al. 2006, Goldstrohm, Seay et al. 2007). To investigate the role of Pum1 in the regulation of mRNA stability, I treated cells with actinomycin D to inhibit transcription and evaluated the rate at which specific mRNAs decayed over a 3 hour time course both in the presence and absence of siRNA-mediated knockdown of Pum1. As shown in Figure 22, the rates at which the endogenous SIRT1 and PCNA mRNAs decayed was unaffected by Pum1 knockdown.

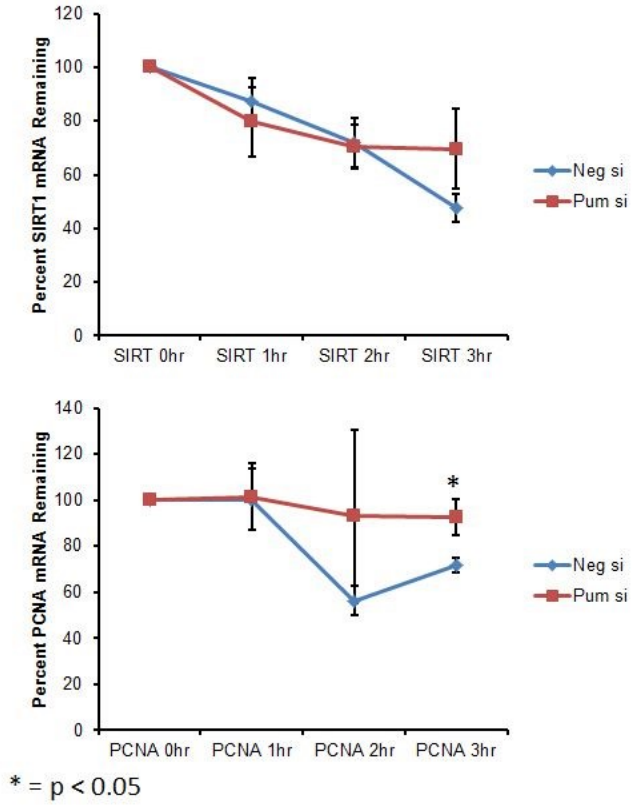


Figure 22: Effect of Pum1 Knockdown on Decay of Target mRNAs

Actinomycin D decay assays were performed on cells treated with siRNA against Pum1 mRNA (Pum si) or a control siRNA (Neg si). Abundance of mRNAs was evaluated at 1-hour timepoints following treatment via real time PCR, normalized to the initial amount of mRNA at time 0. The upper panel shows the decay curve for the SIRT1 mRNA, and the lower panel shows the decay curve for the PCNA mRNA.

Similarly, overexpression of the HD and HD+ Pum1 truncation constructs had no effect on the decay rate of the PCNA mRNA (Figure 23, bottom panel).

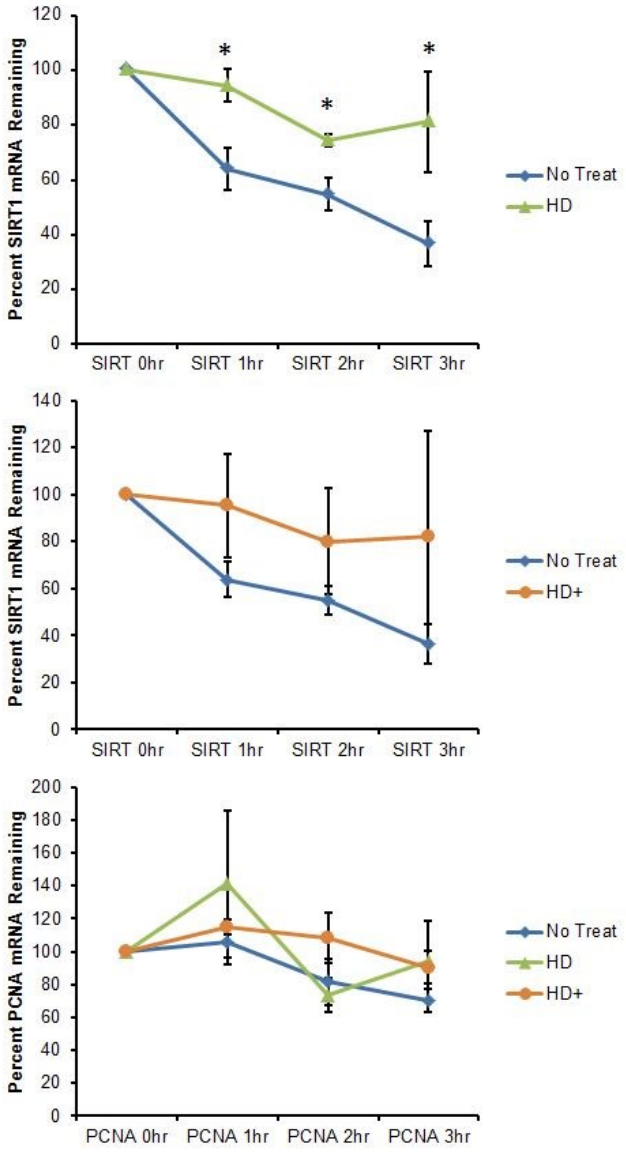


Figure 23: Effect of Pum1 Truncation Constructs on Decay of Target mRNAs

Actinomycin D decay assays were performed on cells expressing Pum1 truncation constructs. Abundance of mRNAs was evaluated at 1-hour timepoints following treatment via real time PCR, normalized to the initial amount of mRNA at time 0. The upper and middle panels show the decay curves for the SIRT1 mRNA in cells expressing the Pum-HD and Pum-HD+ truncations, respectively. The lower panel shows the decay curve for the PCNA mRNA in cells expressing the Pum-HD and Pum-HD+ truncations.

Interestingly, however, overexpression of the HD truncation construct stabilized the SIRT1 mRNA (Figure 23, top panel). Expression of the HD+ truncation displayed a similar effect on the SIRT1 mRNA, however the difference was not significant (Figure 23, middle panel). These observations further suggest that the repression observed in the luciferase assays may involve a pathway other than the recruitment of deadenylases and enhancement of mRNA decay.

4.2.5 Role of Pum1 in Translational Regulation

Numerous studies have demonstrated roles for PUF-family proteins in translational silencing (Wharton, Sonoda et al. 1998, Gu, Deng et al. 2004). For example, *Drosophila* Pumilio has been extensively studied as an important regulator of *hunchback* and Cyclin B mRNA translation (Asaoka-Taguchi, Yamada et al. 1999, Sonoda and Wharton 1999). A recent study also demonstrated a role for human Pum2 in regulation of translational elongation in vitro (Friend, Campbell et al. 2012). To investigate whether Pum1 affects the translation of its target mRNAs in vivo, I performed polysome gradient fractionation and evaluated the polysome distribution of Pum1 target mRNAs. As shown in Figure 24, siRNA-mediated knockdown of Pum1 dramatically reduced Pum1 protein levels in total cell lysate. However, the absence of Pum1 protein did not affect the overall distribution of the SIRT1 or PCNA mRNAs in the polysome gradient (Figure 25).

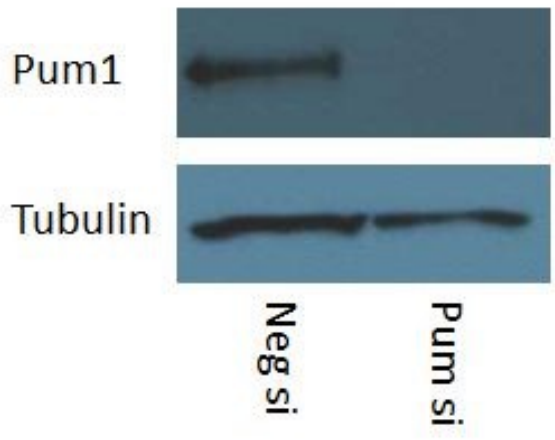


Figure 24: siRNA Knockdown of Pum1 Protein

Western blot showing the reduction in Pum1 protein level by siRNA treatment. Tubulin is shown as a loading control.

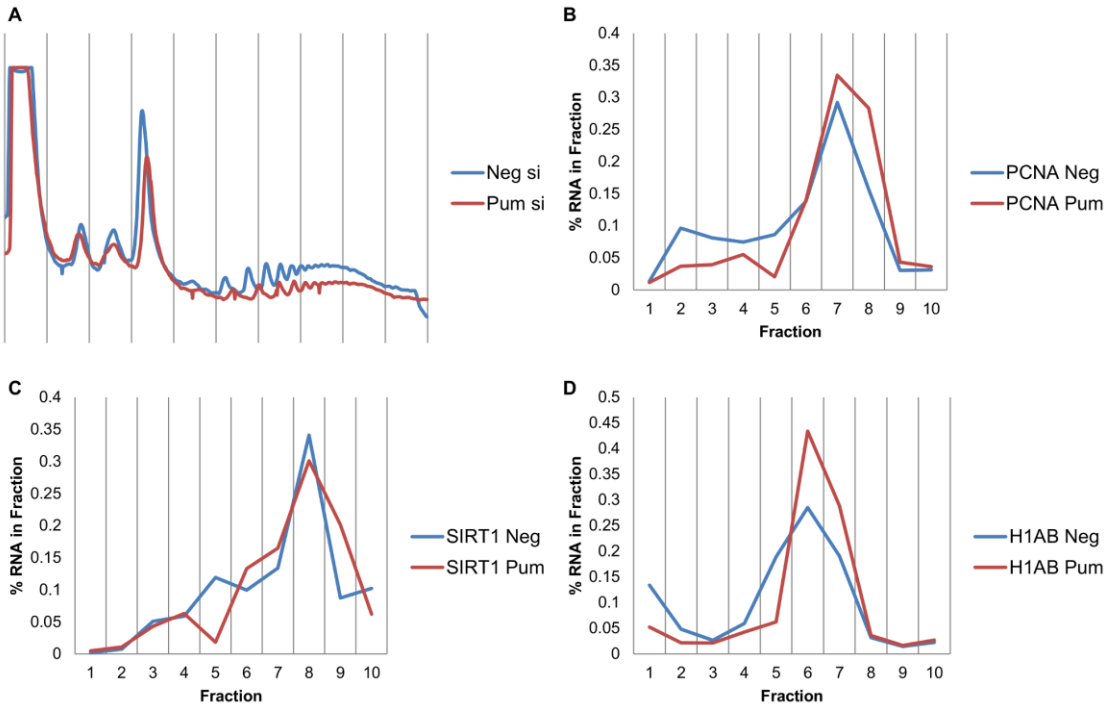


Figure 25: Effect of Pum1 Knockdown on Polysome Distribution of Target mRNAs

Polysome gradient analysis was performed on cells treated with siRNA against Pum1 mRNA. The proportion of mRNA in each fraction of the gradient was analyzed by real time PCR and normalized as a percentage of the total abundance of that mRNA across the gradient. Panel A is the UV 265 absorbance of the total RNA across the gradients. Panels B, C, and D show the distributions of the PCNA, SIRT1, and Histone mRNAs, respectively, in the polysome gradients from cells treated with either Pum1 siRNA or a negative control siRNA.

I performed a similar analysis in cells overexpressing the Pum1HD and Pum1HD+ truncation constructs. As shown in Figure 26, the expression of the various Pum1 truncations did not alter the distribution of Pum1 target mRNAs in the polysome gradient.

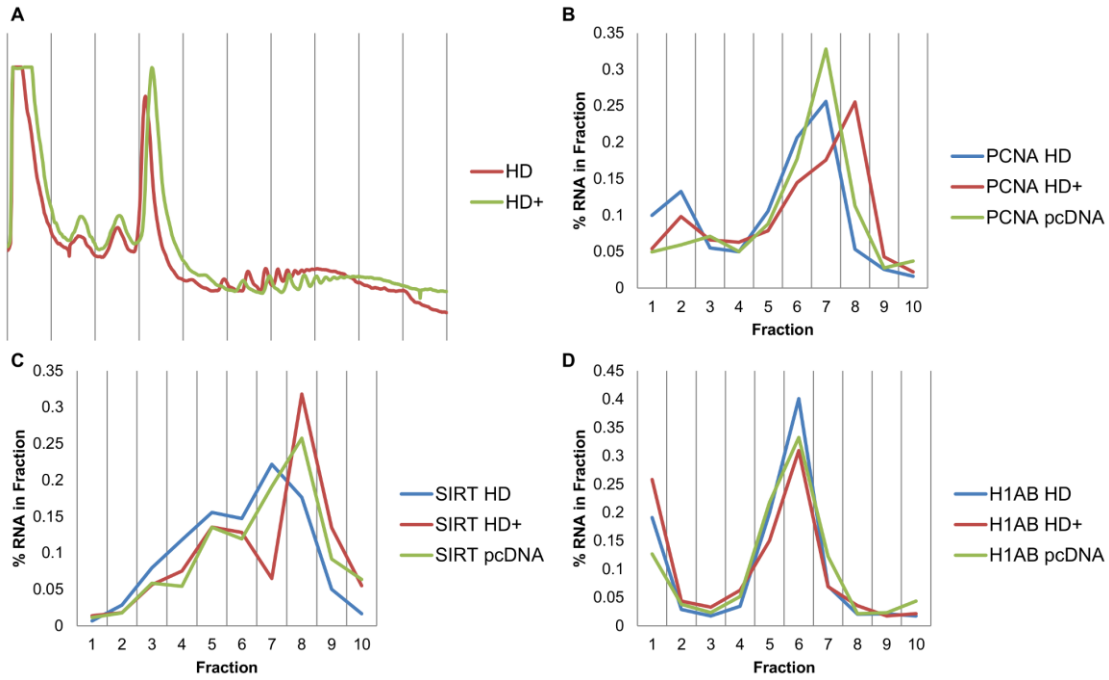


Figure 26: Effect of Pum1 Truncations on Polysome Distribution of Target mRNAs

Polysome gradient analysis was performed on cells expressing Pum1 truncation constructs. The proportion of mRNA in each fraction of the gradient was analyzed by real time PCR and normalized as a percentage of the total abundance of that mRNA across the gradient. Panel A is the UV 265 absorbance of the total RNA across the gradients. Panels B, C, and D show the distributions of the PCNA, SIRT1, and Histone mRNAs, respectively, in the polysome gradients from cells expressing Pum1-HD, Pum1-HD+, or an empty control expression vector (pcDNA).

Steady state polysome gradient analysis only reflects the average translational status of a given mRNA. This distribution is generally dominated by the rate of translation initiation, and the previous demonstration that Pum2 affects translational elongation suggested that a steady state analysis might not identify differences in elongation rate due to Pum1. To investigate these differences, I treated cells with

pactamycin, an inhibitor of translation initiation, for either 5 or 10 minutes prior to harvesting and polysome gradient analysis. As shown in Figure 27 and Figure 28 (top left panels), the addition of pactamycin permitted translational 'run off' while preventing additional initiation, resulting in a reduction in the heavy polysome fractions and an increase in the lighter fractions.

However, the distributions of Pum1 target messages across these gradients did not differ from the distributions of non-target messages. Additionally, neither the siRNA-mediated knockdown of Pum1 protein nor the overexpression of the Pum1 truncations or altered these distributions in the pactamycin-treated cells (Figure 27, Figure 28).

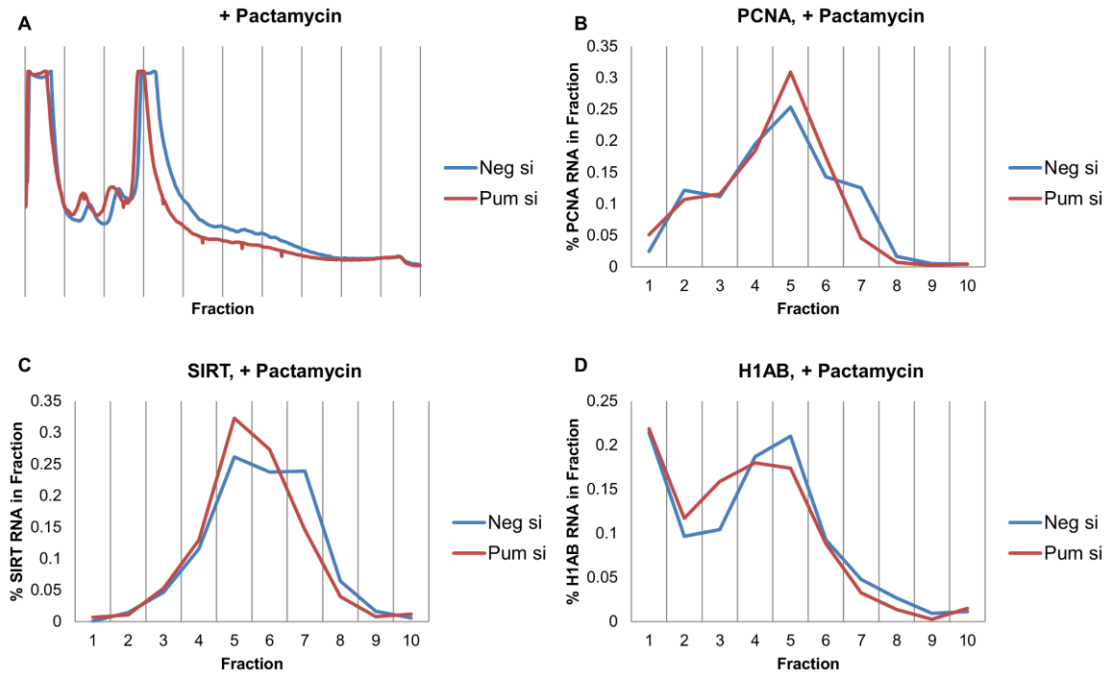


Figure 27: Effect of Pum1 Knockdown on Polysome Distribution of Target mRNAs After Pactamycin Treatment

Polysome gradient analysis was performed on cells treated with siRNA against Pum1 following treatment with pactamycin to halt translation initiation. The proportion of mRNA in each fraction of the gradient was analyzed by real time PCR and normalized as a percentage of the total abundance of that mRNA across the gradient. Panel A is the UV 265 absorbance of the total RNA across the gradients. Panels B, C, and D show the distributions of the PCNA, SIRT1, and Histone mRNAs, respectively, in the polysome gradients from cells treated with either Pum1 siRNA or a negative control siRNA following pactamycin treatment.

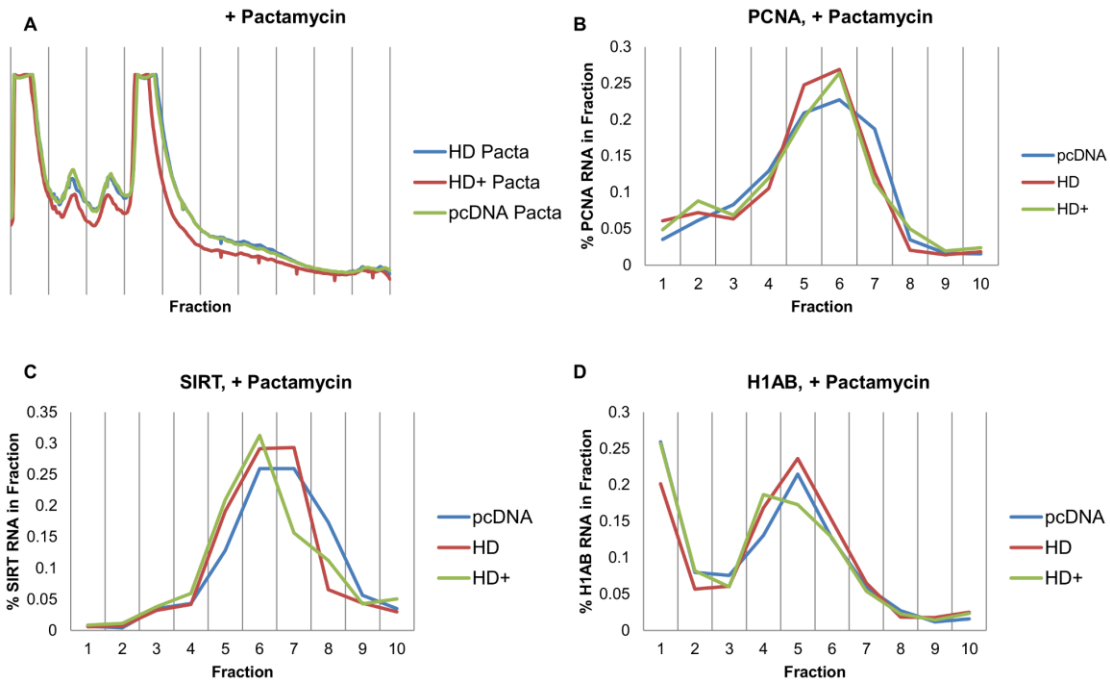


Figure 28: Effect of Pum1 Truncations on Polysome Distribution of Target mRNAs After Pactamycin Treatment

Polysome gradient analysis was performed on cells expressing Pum1 truncation constructs following treatment with pactamycin to halt translation initiation. The proportion of mRNA in each fraction of the gradient was analyzed by real time PCR and normalized as a percentage of the total abundance of that mRNA across the gradient. Panel A is the UV 265 absorbance of the total RNA across the gradients. Panels B, C, and D show the distributions of the PCNA, SIRT1, and Histone mRNAs, respectively, in the polysome gradients from cells expressing Pum1-HD, Pum1-HD+, or an empty control expression vector (pcDNA) following pactamycin treatment.

4.2.6 Functional interaction partners

Several studies have identified functional interactions involved in PUF-family mediated repression of gene expression. Specifically, interaction with members of the CCR4-NOT complex has been identified as a major mechanism by which PUF family

proteins repress, primarily through the recruitment and promotion of deadenylation (Goldstrohm, Hook et al. 2006, Goldstrohm, Seay et al. 2007). However, these studies utilized in vitro methods to identify PUF-interacting proteins. Further, the bulk of the characterization performed in these studies focused on yeast Puf proteins, and interactions with the mammalian Pum proteins were only briefly examined.

To identify proteins that interact with human Pum1 in vivo, I immunoprecipitated full length FLAG-Pum1 from 293T cells and subjected the resulting IP eluate to proteomic analysis. As shown in Figure 29, I performed both positive and negative (control) IPs and eluted from the IP resin with FLAG peptide.

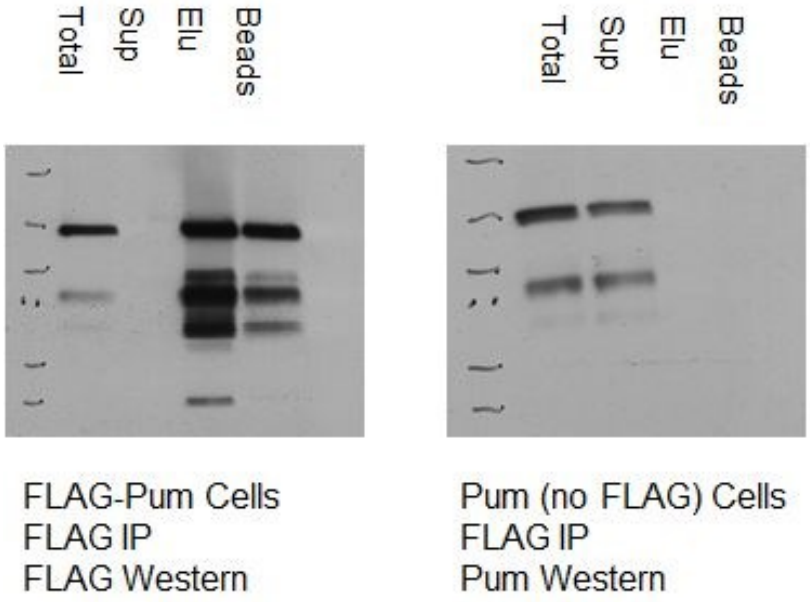


Figure 29: FLAG-Pum1 IP Western Blot

Western blots of FLAG-Pum (left) and control (right) IPs for proteomic analysis. Each blot contains total protein, IP supernatants (Sup), fractions eluted using the 3X FLAG peptide (Elu), and the beads following elution. The positive IP was performed on cells overexpressing full length FLAG-Pum1 (left blot), while the control IP was performed on untransfected cells (right blot). The positive blot was probed with an anti-FLAG primary antibody, while the negative blot was probed with an anti-Pum1 antibody.

As illustrated in Figure 30, protein was recovered from both the positive and negative IPs, but the positive IP showed substantially more protein signal than the background IP.

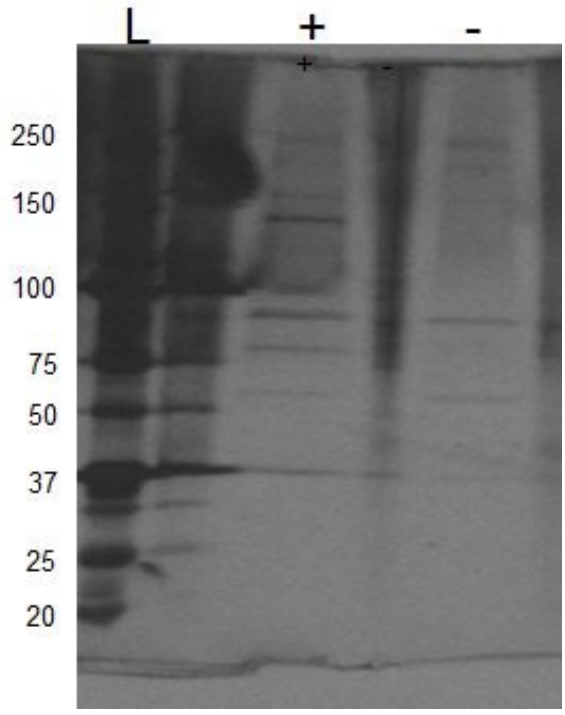


Figure 30: IP to Identify FLAG-Pum1 Interacting Proteins

Eluted fractions from positive (+) and control (-) IPs were separated by electrophoresis and visualized by silver stain. L indicates a molecular weight ruler, weights are indicated as kDa to the left of the blot.

Both samples (+ and – IPs) were submitted for proteomic analysis. Table 6 shows the list of co-immunoprecipitated proteins identified as significantly enriched in the positive IP compared to the negative IP. Encouragingly, a member of the CCR4-NOT complex, CNOT1, was identified as a Pum1-interacting protein by this analysis. Additionally, several other RBPs and known RBP-related proteins were identified (HNRNPC1, PABP1, HNRNPA1, RHA). It is important to note that our IP protocol did not involve any RNAse treatment or RNP-disrupting methods. Thus, I expected to

identify members of the overall Pum1-containing RNP complex, not only direct protein-protein interactors.

Table 6: FLAG-Pum1 Interacting Proteins

Name	Mol. Weight	Fisher's (pval)
Pumilio homolog 1	126 kDa	1.40E-45
Heat shock 70 kDa protein 4L	95 kDa	0.00064
Clathrin heavy chain 1	192 kDa	0.0012
Heat shock protein 105 kDa	97 kDa	0.0014
Heterogeneous nuclear ribonucleoproteins C1/C2	34 kDa	0.0014
Matrin-3	95 kDa	0.0014
Structural maintenance of chromosomes protein 4	147 kDa	0.0033
14-3-3 protein epsilon	29 kDa	0.0039
Polyadenylate-binding protein 1	71 kDa	0.0056
Heat shock cognate 71 kDa protein	71 kDa	0.0074
Heterogeneous nuclear ribonucleoprotein A1	39 kDa	0.0095
Nucleolar RNA helicase 2	87 kDa	0.02
Nuclear pore complex protein Nup205	228 kDa	0.02
ATP-dependent RNA helicase A	141 kDa	0.023
CCR4-NOT transcription complex subunit 1	**267 kDa**	**0.026**
Nucleoprotein TPR	267 kDa	0.026
Regulator of nonsense transcripts 1	124 kDa	0.038
Heat shock 70 kDa protein 1A/1B	70 kDa	0.04
Stress-70 protein, mitochondrial	74 kDa	0.047

As I began validation experiments and further characterization of the interactions between Pum1 and CNOT1, a study was published by the Goldstrohm laboratory identifying and carefully characterizing the RNA-independent interactions between Pum1 (and Pum2) and the CCR4-NOT complex (Van Etten, Schagat et al. 2012). This study clearly illustrated a strong role for the interaction between Pum and the CCR4-NOT complex in the promotion of mRNA deadenylation and repression of

reporter constructs. Interestingly, however, the results of this study also suggested that Pum proteins may exhibit a secondary repressive function that is not dependent on deadenylation of target mRNAs.

Importantly, previous studies on the interactions between PUF proteins and the CCR4-NOT complex demonstrated that the region of the human Pum1 protein responsible for the interaction is located in the Pumilio Homology Domain. Given the previously discussed findings that the repression of the SIRT1 and PCNA reporters requires a region of Pum1 upstream of the HD, it seems unlikely that Pum1's interaction with the CCR4-NOT complex is responsible for this repression. Thus, I sought to identify other proteins that interact with regions of Pum1 upstream of the HD in the hopes of identifying the factors responsible for the Pum1-mediated repression of SIRT1 and PCNA.

To identify these interactors, I again performed proteomic analysis of Pum1 IPs. However, in this case, I expressed FLAG-tagged truncation constructs containing either the region upstream of the HD previously identified as necessary for repression (FLAG-Mid), or the previously described FLAG-HD+ construct. A negative IP was also performed on lysates expressing an untagged HD+ construct. Figure 31 shows the expression of the truncation constructs (indicated with arrows) and the resulting protein isolated by IP (lower panel).

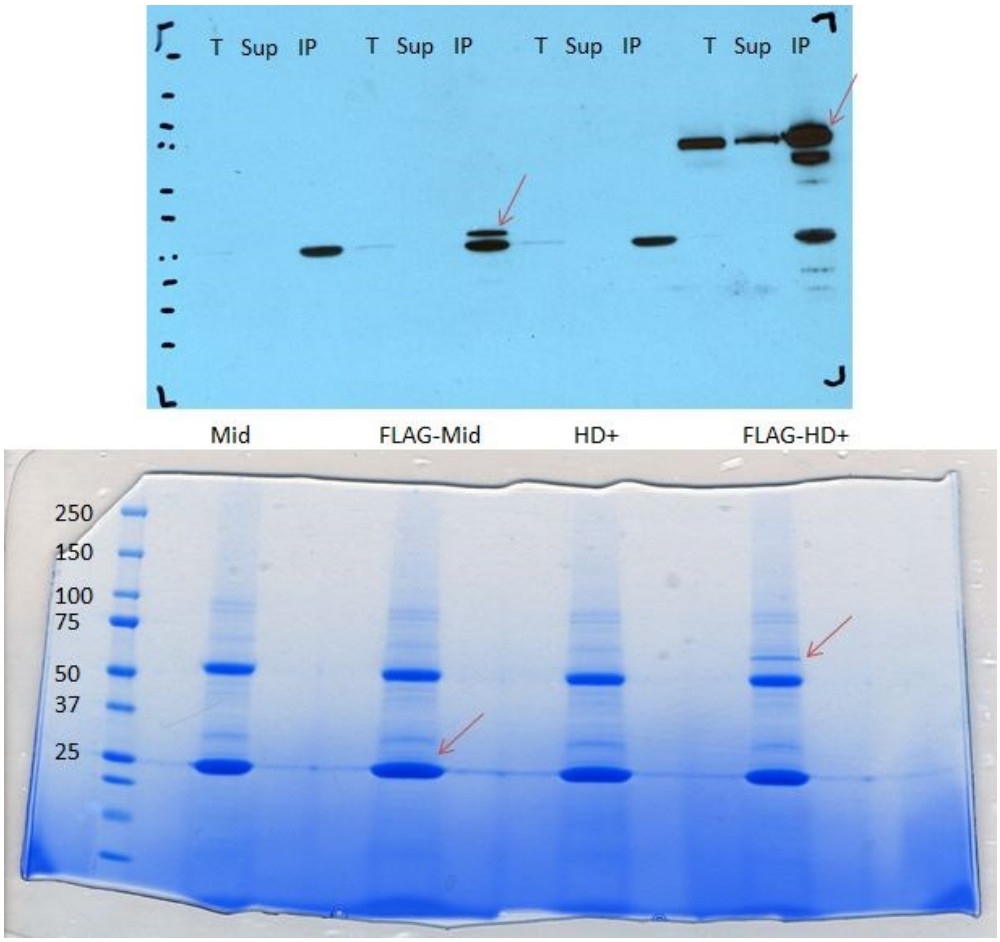


Figure 31: IP to Identify Protein Interaction Partners Upstream of the HD

The indicated constructs (Mid, FLAG-Mid, HD+, FLAG-HD+) were expressed in 293T and IPs were performed with anti-FLAG antibodies. Upper panel shows a western blot of total protein (T), IP supernatants (Sup), and immunoprecipitated samples (IP) probed with an anti-FLAG primary antibody. Lower panel shows a coomassie stain for total protein of the indicated IP samples. Red arrows indicate FLAG constructs.

Proteins identified through the proteomic assay were considered targets only if they were identified in both the FLAG-HD+ and FLAG-Mid IPs and were not identified in the control (HD+) IP sample.

Table 7 lists the proteins identified in this analysis as potential interactors upstream of the Pum1 HD. Strikingly, the top two identified proteins as ranked by number of peptide counts were also identified in the full length Pum1 proteomic analysis. However, one of these, Heat shock protein 70 4L, is a known chaperone protein, and may be present simply due to its general protein-binding function.

Table 7: FLAG-HD+ and FLAG-Mid Interacting Proteins

Name	Mol. Weight	FLAG-HD+ Peptide Counts	FLAG-Mid Peptide Counts
Pumilio homolog 1	126 kDa	31	8
Heat shock 70 kDa protein 4L	95 kDa	13	6
ATP-dependent RNA helicase A	141 kDa	9	2
Staphylococcal nuclease domain-containing protein 1	102 kDa	6	2
Serum albumin	69 kDa	2	3
Glutathione S-transferase P	23 kDa	4	4
40S ribosomal protein S20	13 kDa	3	2
40S ribosomal protein S11	18 kDa	4	1
Gem-associated protein 5	169 kDa	4	1
28S ribosomal protein S27, mitochondrial	48 kDa	2	2
Transmembrane protein 109	26 kDa	2	2
60S ribosomal protein L23a	18 kDa	3	1
DNA replication licensing factor MCM6	93 kDa	1	3
Calcyclin-binding protein	26 kDa	3	1
Ornithine aminotransferase, mitochondrial	49 kDa	1	3
40S ribosomal protein S26	13 kDa	2	1
60S ribosomal protein L22	15 kDa	1	2
Signal recognition particle 14 kDa protein	15 kDa	2	1
60S ribosomal protein L12	18 kDa	2	1
Plasminogen activator inhibitor 1 RNA-binding protein	45 kDa	2	1
Splicing factor U2AF 35 kDa subunit	28 kDa	1	2
28S ribosomal protein S29, mitochondrial	46 kDa	1	2
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	69 kDa	1	2
Annexin A5	36 kDa	2	1

ATP-dependent RNA helicase A (RHA or DHX9), the other potential interactor identified in both assays, is a known RBP, with roles in transcriptional and translational

regulation (Fuller-Pace 2006, Ranji, Shkriabai et al. 2011, Manojlovic and Stefanovic 2012). Interestingly, studies have shown a direct interaction between RHA and another protein identified by the assay, Staphylococcal nuclease domain-containing protein 1 (SND1 or p100) (Valineva, Yang et al. 2006). Given the potential of these two proteins as functional interactors with Pum1, I chose to validate their interaction through IP and western blot. Figure 32 shows a western blot of both negative (lysate containing no FLAG tag) and positive IPs of the same truncations utilized for the proteomic assay.

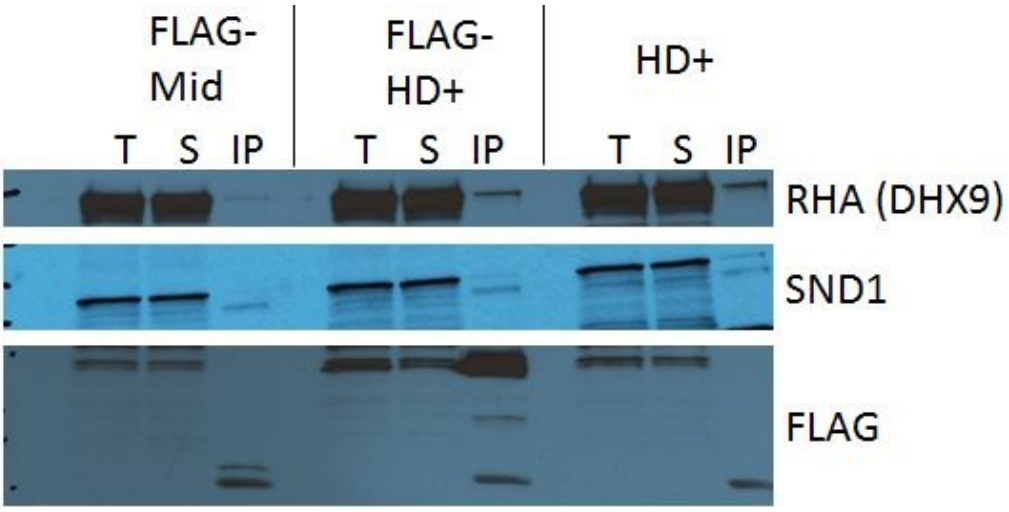


Figure 32: Putative Interaction Partners In Negative IPs

Total (T), supernatants (S), and immunoprecipitates (IP) from positive (FLAG-Mid and FLAG-HD+) and control (HD+) IP reactions were subjected to western blots probed with anti- RHA, SND1, and FLAG primary antibodies.

Unfortunately, as shown in the rightmost lanes of the figure, both RHA and SND1 were present in the eluate from the negative IP as well as the positive IPs. Additional optimization of wash conditions, IP conditions, and negative lysates did not

alter these results. Discussions are currently underway with the proteomics facility to resolve this uncharacteristic difference between the two detection methods (neither RHA nor SND1 were present in the negative sample by proteomic analysis). Further characterization of the interactions between these potential functional regulators and Pum1 will require a resolution of these discrepancies.

4.3 Materials and Methods

Plasmids: The FLAG-Pum1 plasmid was generated by inserting an N-terminal FLAG sequence upstream of the full length CDS of human Pum1 in a pcDNA3 backbone. The Pum1 truncation constructs, FLAG-HD and FLAG-HD+, were generated by PCR using primers containing N-terminal FLAG sequences and cloned into pcDNA3. The FLAG-HD construct encodes amino acids 820-1189 of human Pum1, and FLAG-HD+ encodes amino acids 581-1189. Luciferase reporter vectors were generated by inserting full length 3′ UTR PCR products of the SIRT1, PCNA, and TTP mRNAs downstream of the Firefly Luciferase ORF in the pcDNA3-Luc vector. The pcDNA3-Luc vector containing no UTR insert was used for normalization. All constructs were validated by sequencing.

Cell culture and transfection: HEK-293T cells (ATCC) were cultured in DMEM supplemented with 10% FBS and Penicillin/Streptomycin at 5% CO2 and 37 C. All plasmids and siRNA were transfected using Lipofectamine 2000 (Invitrogen) per the manufacturer's instructions. The day prior to transfection, cells were seeded in 12-well

plates to an approximate density of 1.5x10^5 cells/mL. Cells were transfected with 1ug of total DNA and the media was replaced after 24 hours. For siRNA experiments, 30 pmol of Pum1 siRNA or a non-silencing negative control RNA (Ambion Silencer siRNA) were co-transfected with 1ug of pcDNA3 as a carrier. All experiments were harvested 48 hours after transfection.

Immunoprecipitation: FLAG immunoprecipitation (IP) was performed using preconjugated anti-FLAG M2 magnetic beads (Sigma). Cell lysates used for IP were harvested in PLB (100mM KCl, 5mM MgCl2, 10mM HEPES (pH 7.0), 0.5% NP40, 1 mM DTT, 100 units/mL RNase Out, 400 μM VRC, 1X Protease inhibitor cocktail (Roche)) (Keene, Komisarow et al. 2006) and frozen at -80C. Lysates were precleared by centrifugation at 15,000 x g for 15 minutes and IP reactions were prepared in NT2 buffer (50mM Tris-HCl (pH 7.4), 150mM NaCl, 1mM MgCl2, 0.05% IGEPAL). Beads were washed 5X with cold NT2 buffer then tumbled with the IP reactions for 2 hours at 4 C. The IPs were then washed 3X with NT2 buffer containing 5M urea, followed by two washes in NT2 buffer. The beads were suspended in Trizol and stored at -80 C until RNA preparation. For RNA-coimmunoprecipitation (RIP), negative IPs were performed on lysates transfected with an untagged Pum1 expression vector. The enrichment in mRNA presence in the FLAG-containing-IP over the untagged IP was used to evaluate protein-RNA association.

Western Blots: Rabbit anti-Pum1 was purchased from GeneTex, rabbit anti-SIRT1 was purchased from Cell Signaling Technologies, mouse anti-PCNA and rat anti-tubulin were purchased from Thermo, rabbit anti-RHA and mouse anti-SND1 were purchased from Pierce.. Chemiluminescent secondary antibodies were obtained from GE Healthcare and detection was performed using Pierce ECL reagent.

Coomassie and Proteomics: Gels used for total protein visualization were stained with colloidal coomassie blue G-250 according to the protocol in (Candiano, Bruschi et al. 2004). LC-MS/MS samples were delivered in 1x LDS loading buffer (Invitrogen) and separated via SDS-PAGE on a 4-12% Bis/Tris gradient gel (Invitrogen) in MES buffer, for approximately 15 minutes. Seven distinct gel regions were dissected using a scalpel, as annotated in Figure 33. In-gel digestion was performed according to a standard protocol (http://www.genome.duke.edu/cores/proteomics/sample-preparation/documents/In-gelDigestionProtocolrevised.pdf). Briefly, gel pieces are repeatedly shrunken and swelled in MeCN and 50 mM ammonium bicarbonate (AmBic) respectively. Proteins are then reduced using a solution of 10 mM dithiothreitol, alkylated with 20 mM iodoacetamide. Supernatant was removed, then 300 ng Sequencing Grade Modified Trypsin (Promega) was added at 10ng/uL in 50 mM ammonium bicarbonate (pH 8). Digestion was allowed proceed overnight at 37°C, then quenched with the addition of 1% Formic Acid. Peptides are extracted and dried, then resuspended in 12 uL 0.2/2/97.8 v/v/v formic acid/MeCN/water for analysis.

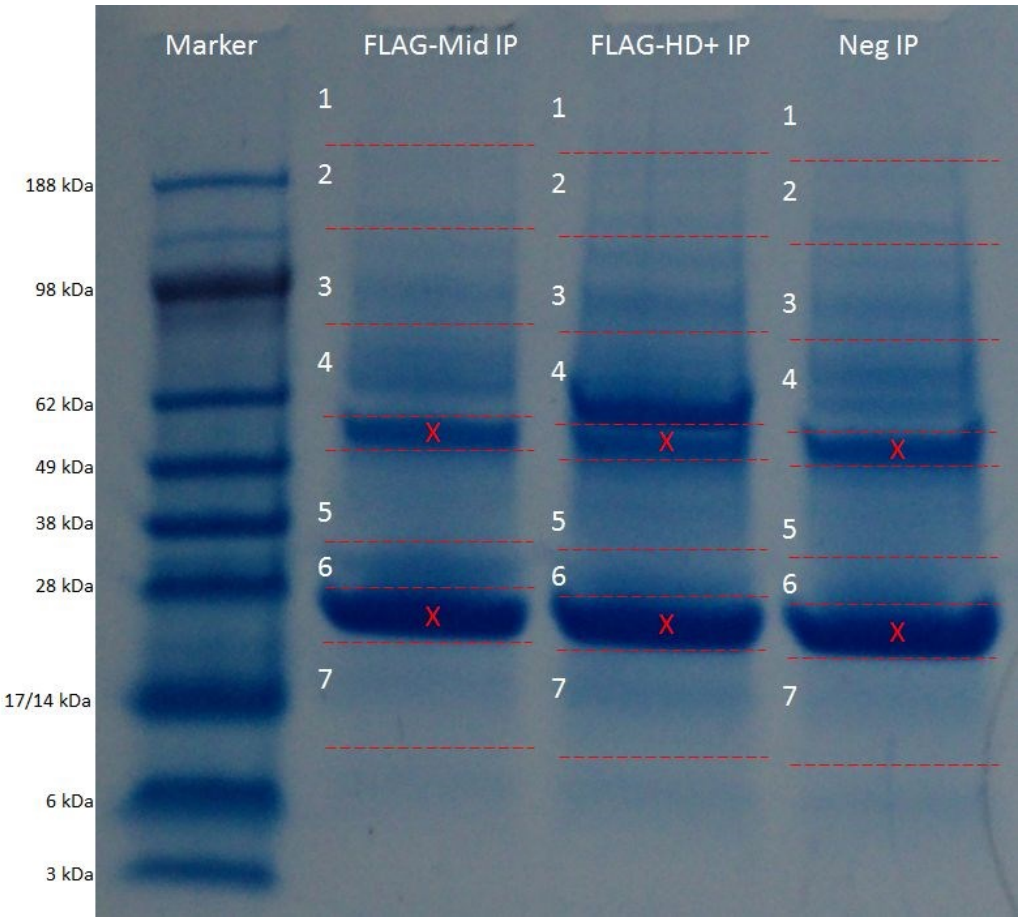


Figure 33: Gel Cuts for Proteomic Analysis

Coomassie stained gel of co-IP reactions submitted for proteomic analysis. Positive (FLAG-Mid IP and FLAG-HD+ IP) and control (Neg IP) samples were separated by electrophoresis and the lanes were cut into 7 samples each, as indicated in the figure. The regions marked with X indicate heavy and light chain antibody bands and were avoided.

The in-gel digested peptides from each band were analyzed via LC-MS/MS as follows: Approximately ½ of each digest (5 ul) was injected onto a 75µm x 250 mm BEH C18 column (Waters) and separated using a gradient of 5 to 40% acetonitrile with 0.1% formic acid, with a flow rate of 0.4 µL/min, in 30 minutes on a nanoAcquity liquid

chromatograph (Waters). Electrospray ionization was used to introduce the sample in real-time to a Q-Tof Synapt G2 mass spectrometer (Waters), collecting data for each sample in DDA mode with 0.6 second survey scans and three 0.6-second MS/MS scans in CID mode of the top three most abundant precursor ions, allowing MS/MS sequencing of ions with charge state 2 or greater. Raw data was processed in Mascot Distiller v2.3 (Matrix Sciences, Inc) and interrogated using automated searching in Mascot v2.2 (Matrix Science). Search tolerances 10 ppm on precursor and 0.04 Da product ions. Carbamidomethylation (C) was included as a fixed modification, and deamidation (N and Q) and oxidation (M) were allowed as variable modifications. Data curation was performed in Scaffold v4 (Proteome Software, Inc), and annotation performed at a false discovery rate of 0% at the protein level, using target-decoy database strategy.

RT and Real Time PCR: RNA was extracted using Trizol reagent (Invitrogen) per the manufacturer's instructions and cDNA synthesis was performed using the BioRad iScript kit. Real time PCR was carried out in a Roche LightCycler capillary qPCR instrument using SYBR-Supermix UDG (Invitrogen). Primer sequences used for qPCR are listed in

Table 8. Data were analyzed using the delta-delta Ct method, normalized to GAPDH as a 'housekeeping' control and a negative treatment sample.

Table 8: Real Time PCR Primer Sequences

mRNA	Forward Primer	Reverse Primer
SIRT	AGTGGCAAAGGAGCAGATTAG	CTGCCACAAGAACTAGAGGATAAG
PCNA	AGCTCTTCCCTTACGCAAGTCTCA	ACGAGTCCATGCTCTGCAGGTTTA
B2M	TGTCTGGGTTTCATCCATCCGACA	TCACACGGCAGGCATACTCATCTT
HuR	GGATGAGTTACGAAGCCTGTT	CAAGCTGTGTCCTGCTACTT
H1A	GGAGAAGAACAACAGCCGCAT	TTGAGCTTGAAGGAACCCGAG
GAPDH	TCGACAGTCAGCCGCATCTTCTTT	ACCAAATCCGTTGACTCCGACCTT

Luciferase Assays: Equimolar quantities of luciferase reporter vectors were co-transfected with a Renilla luciferase expression vector (pRL) and Pum1 expression constructs or empty pcDNA3 as a negative control. Cells were harvested using Passive Lysis Buffer (Promega) and diluted 100x prior to performing luciferase assays. Assays were performed on a Mithras LB 940 (Berthold Technologies) using Promega Dual Luciferase reporter reagents. The ratios of Firefly to Renilla luciferase signal (F/R) for experimental groups were normalized to the F/R ratio of the pLuc vectors and untreated samples. Data are presented as mean +- s.d. Data analysis was performed in Excel with Daniel's XL Toolbox (Microsoft; Daniel Kraus, Wurzburg, Germany).

Actinomycin D experiments: Cells were treated with 5ug/mL of Actinomycin D and harvested in Trizol at hourly time points. qPCR analysis of mRNA abundance at each time point relative to the untreated sample was used to generate decay curves.

Polysome Gradients: Cells were treated with cycloheximide (CHX) at a concentration of 200uM for 5 minutes then harvested in lysis buffer (400mM KOAc, 24mM HEPES, 5mM MGOAc2, 0.5% Na-deoxycholate, 0.5% Triton-X-100, 1mM DTT,

1mM PMSF, 200uM CHX, 50U/mL RNaseOut) and incubated for 15 minutes on ice. For translational runoff experiments, cells were treated with 200nM pactamycin for 5 or 10 minutes prior to the addition of CHX. Lysates were precleared by centrifugation at 3,000 x g for 8 minutes and loaded on 15-50% sucrose density gradients. Gradients were spun for 2.5hr at 35,000 rpm in a SW41 swinging bucket rotor. Gradients were analyzed and collected as 10 fractions each on an ISCO density gradient fractionator. RNA was extracted from 250uL of each fraction with Trizol LS (Invitrogen) according to the manufacturer's protocol. qPCR analysis was used to evaluate the distribution of specific mRNAs across the gradient. Equal volumes of RNA from each fraction were used for cDNA synthesis and raw C(t) values were used to calculate the relative percentage of each mRNA in each fraction of the gradient.

4.4 Discussion

Although SIRT1 has been identified in several genome-level studies as a potential mRNA target of the Pum1 RNA binding protein (RBP)(Galgano, Forrer et al. 2008, Morris, Mukherjee et al. 2008), I have provided the first evidence that this targeting may also result in the posttranscriptional regulation of SIRT1 expression. Additionally, I have demonstrated that the regulation of SIRT1 mRNA by Pum1 is mediated through the 3' UTR of the SIRT1 message. The only other reported RBP-mediated regulation of SIRT1 mRNA is the condition dependent stabilization of SIRT1 message through interaction with HuR (Abdelmohsen, Pullmann et al. 2007). Strikingly, their study

found that SIRT1 mRNA loses its interaction with HuR upon induction of oxidative stress and that this disassociation results in destabilization of the SIRT1 mRNA. A previous study from our group (Morris, Mukherjee et al. 2008) found that Pum1 localizes to stress granules during oxidative stress, however attempts to quantify differences in Pum1 mRNA targeting before and after induction of oxidative stress have not, as yet, been successful. It is intriguing to hypothesize that the destabilization of SIRT1 mRNA following oxidative stress may be partially mediated by Pum1. Further, I propose that the lack of regulation observed on the endogenous SIRT1 message may be a consequence of the dynamic and condition specific nature of posttranscriptional regulation.

Our observations that overexpression of Pum1 has little effect on stability or translation of its target messages is consistent with previously published findings (Van Etten, Schagat et al. 2012), suggesting that Pum1 protein is present in excess in 293T cells. Further, Pum2, which is also expressed in 293T cells, targets a very similar set of mRNAs as Pum1 (Galgano, Forrer et al. 2008), suggesting that regulatory compensation by Pum2 may be responsible for maintaining levels of SIRT1 expression when Pum1 levels are depleted.

I have also provided the first in vivo validation of the role for regions upstream of the Pum1 homology domain (HD) in regulation of target expression. A previous study (Weidmann and Goldstrohm 2012) identified several regions upstream of the HD

in multiple PUF family proteins that are independently capable of repressing expression in an in vitro tethering assay. My data indicate that the presence of this upstream regulatory region is required for Pum1-mediated repression through the SIRT1 and PCNA 3' UTRs, but this region is not required for mRNA binding. These findings suggest the presence of a functional domain upstream of the Pum1 HD. I hypothesize that this region is responsible for the recruitment of other effector proteins that mediate the mechanistic repression of mRNA expression, either through translational silencing or deadenylation and decay. It is interesting to note that previous studies have identified the interaction sites between Pum1 and the deadenylase complex in the Pum HD (Goldstrohm, Hook et al. 2006, Goldstrohm, Seay et al. 2007). This suggests that the repression mediated by the upstream regions of Pum1 may involve a different and novel set of regulatory mediators.

As indicated by the proteomic data, the region upstream of the HD likely interacts with several additional elements of the posttranscriptional system. The identification and characterization of these potential co-regulators has been a focus of numerous experiments in our laboratory. The previously described disagreement between western blot and proteomic detection methods has somewhat hampered progress in the specific identification of these co-regulators. However, it is possible that other candidate interactors identified through the proteomic experiments will prove

specific and functionally relevant. Experiments are currently underway to test and characterize these additional potential interactions.

Numerous studies have investigated the role of PUF family proteins in the regulation of translation (Cao, Padmanabhan et al. 2010, Chritton and Wickens 2010, Chritton and Wickens 2011) in multiple species. Recent work has implicated human Pum2 as a negative regulator of translational elongation in a cell free translation assay (Friend, Campbell et al. 2012). Although the luciferase reporter data support a role for Pum1 in the regulation of translation, I observed no effect on the distribution of endogenous Pum1 target messages in polysome gradients in response to overexpression or depletion of Pum1 protein. I further explored the potential role of Pum1 in translational elongation through a translation run-off experiment, but again failed to confirm previous reports of Pum1-mediated repression. A major difference between this study and the Friend et al. work is the in vivo vs in vitro natures of the respective systems. It is likely that the intact cell system utilizes multiple redundant pathways to ensure proper expression of important regulatory genes, such as SIRT1.

Additionally, although SIRT1 message shows strong enrichment in Pum1 IP experiments, I have not quantified the proportion of total SIRT1 mRNA that is bound by Pum1 protein in the cell. Subcellular localization and binding accessibility of a given mRNA may vary greatly within the cell, resulting in differential regulation and coordination of subsets of a particular mRNA species.

The specific contributions of the multiple regulatory mechanisms utilized by the PUF family proteins remain uncharacterized. It will be an important future step in the characterization of this family of RBPs to assess the degree to which they affect deadenylation, mRNA decay, and translation, particularly in the context of specific biological conditions and perturbations. As previously discussed, experimental designs utilizing dynamic conditions, such as oxidative stress or developmental perturbations, may provide insights into many of these remaining questions regarding the mechanisms of Pum1 mediated regulation of SIRT1 and other target mRNAs. Ultimately, the integration of these important mechanistic relationships into global-scale regulatory models will be necessary to fully characterize the complex system of posttranscriptional control of gene expression.

5. Conclusions and Future Directions

I have explored the posttranscriptional regulatory environment from the global perspective, thorough the development of systems-level interaction models, from the perspective of global translational control in a dynamic system, and through mechanistic analysis of individual RBP-mRNA interactions. Global analyses indicate that RBPs compose a substantial proportion of the gene regulatory network, potentially contributing greater regulatory function than the transcription network. Further, the global and mechanistic analyses presented in this document emphasize the importance of intramolecular cooperation and interactions in the regulatory control of gene expression at all levels.

Global analysis of the combined yeast transcriptional and posttranscriptional gene regulatory network provided us with evidence for the extent and significance of the posttranscriptional regulatory system. This work, and subsequent studies by other groups, determined that those genes displaying regulation by RBPs show a greater degree of controlled expression at both the transcriptomic and proteomic levels than those genes that are only controlled by TFs. The analysis presented in Chapter 2 also illuminates a global 'regulators of regulators' structure to the posttranscriptional environment. This structure has been suggested by several small-scale studies previously, but this is the first global level demonstration of the tendency of posttranscriptional regulators to target and regulate one another in a combinatorial

fashion. Additional studies from other laboratories exploring similar relationships in the same large-scale datasets I studied, as well as further studies in mouse and human systems, also observed similar 'regulators of regulators' behaviors in both the yeast and mammalian regulatory networks. Together, these data suggest that the structure of interconnected regulatory networks of posttranscriptional mediators of gene expression may be an organizing principle that drives overall coordination of gene expression.

While these findings are quite intriguing, they only include posttranscriptional interaction data for 46 yeast RBPs. It has been suggested that yeast may express as many as 500 different RBPs, and integration of data from larger scale characterization of their target messages would provide a more complete picture of the highly complex network of posttranscriptional regulation in this system.

Another important feature lacking from the integrated network analysis is biological dynamics. In a first attempt to model the dynamic binding kinetics of an RBP, I explored the role of HuR during T cell activation. I successfully modeled the dynamics of interaction between an RBP and its target mRNAs and was able to partially attribute these dynamics to changes in mRNA abundance. However, I also determined that competition or cooperation with other RBPs, as well as potential posttranslational modification of HuR itself, are likely involved in the dynamic changes observed in HuR-mRNA association. I also analyzed HuR binding dynamics in the context of translational control and found that, in T cell activation, HuR may play a role in

translation inhibition, much as it does in other immune cells. Further characterization of the global dynamics of RNA behavior in the T cell activation system is currently underway in our laboratory.

Ultimately, I hope to integrate many levels of RNA regulation, including RBP targeting, RNA stability, and translation status into the first global model of PTR in a dynamic mammalian system. Inclusion of global miRNA targeting data would further enhance the scope and impact of a PTR model. To this end, I also hope to characterize the RNA targets of additional RBPs, either through direct experimentation or the integration of published large-scale interaction data, such as PAR-CLIP binding data sets and miRNA target databases. Such data sets are rapidly emerging, and integration of these large-scale targeting studies into a statistical model of the functional consequences of RBP-mRNA interaction would provide great insights into the dynamic and complex nature of the overall posttranscriptional regulatory environment.

Finally, I explored specific mechanisms of posttranscriptional regulation mediated by the Pum1 RBP. I demonstrated a role for Pum1 in the regulation of SIRT1 mRNA, an important regulator of cell aging and senescence. I also provided the first in vivo evidence for the role of a region upstream of the Pum-HD in the function of the Pum1 protein. The specific mechanisms of this repression function remain somewhat unclear. I described preliminary approaches to identifying additional factors that may be involved in mediating this repression; however I have not yet definitively

characterized these factors. Additional studies are currently underway to validate and characterize the factors that interact with Pum1 and mediate its ability to repress translation.

Importantly, these studies of Pum1 have been performed in a static system, unperturbed growing 293T or HeLa cells. Based on previous experience with HuR, I believe that exploration of Pum1 biology in a dynamic biological system, such as response to stress or cellular development and differentiation is an important next step in understanding the mechanisms and global functions of Pum1. Given the importance of the SIRT1 protein in regulation of aging and cell senescence, and the hypothesized ancestral role for Pum proteins in stem cell self-renewal, exploration of the regulatory dynamics of SIRT1, as mediated by Pum1 protein, in an aging or stem-like experimental system would provide key insight into the biological role of mammalian Pum proteins in development. Further, although the greater biological roles of PUF-family proteins have been studied extensively in yeast, flies, and worms, most mammalian studies of Pum1 and Pum2 have exclusively focused on cellular phenotypes and molecular mechanisms. The important biological roles of these highly evolutionarily conserved genes may become apparent through studies in dynamic and developmental mammalian systems.

I hope to study Pum1 and other RBPs in such a system, with the ultimate goal of integrating multiple levels of RBP-mRNA interaction data, global mRNA decay data,

and global translational profiling data into a dynamic systems-level model of the posttranscriptional regulatory environment.

References

Abdelmohsen, K., et al. (2007). "Posttranscriptional orchestration of an anti-apoptotic program by HuR." Cell Cycle **6**(11): 1288.

Abdelmohsen, K., et al. (2007). "Phosphorylation of HuR by Chk2 regulates SIRT1 expression." Mol Cell **25**(4): 543-557.

Alexander, R. P., et al. (2009). "Understanding modularity in molecular networks requires dynamics." Science signaling **2**(81): pe44.

Ambros, V., et al. (2003). "A uniform system for microRNA annotation." RNA **9**(3): 277-279.

Anderson, L. and J. Seilhamer (1997). "A comparison of selected mRNA and protein abundances in human liver." Electrophoresis **18**(3-4): 533-537.

Asaoka-Taguchi, M., et al. (1999). "Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos." Nat Cell Biol **1**(7): 431-437.

Bartel, D. P. (2009). "MicroRNAs: target recognition and regulatory functions." Cell **136**(2): 215-233.

Bhattacharyya, S. N., et al. (2006). "Relief of microRNA-mediated translational repression in human cells subjected to stress." Cell **125**(6): 1111-1124.

Blonder, B., et al. (2012). "Temporal dynamics and network analysis." Methods in Ecology and Evolution **3**(6): 958-972.

Brengues, M., et al. (2005). "Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies." Science **310**(5747): 486-489.

Candiano, G., et al. (2004). "Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis." Electrophoresis **25**(9): 1327-1333.

Cao, Q., et al. (2010). "Pumilio 2 controls translation by competing with eIF4E for 7-methyl guanosine cap recognition." RNA **16**(1): 221-227.

Cheadle, C., et al. (2005). "Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability." BMC Genomics **6**: 75.

Chen, K. and N. Rajewsky (2007). "The evolution of gene regulation by transcription factors and microRNAs." Nature Reviews Genetics **8**(2): 93-103.

Chritton, J. J. and M. Wickens (2010). "Translational repression by PUF proteins in vitro." RNA **16**(6): 1217-1225.

Chritton, J. J. and M. Wickens (2011). "A role for the poly(A)-binding protein Pab1p in PUF protein-mediated repression." J Biol Chem **286**(38): 33268-33278.

Cooke, A., et al. (2011). "Targeted translational regulation using the PUF protein family scaffold." Proc Natl Acad Sci U S A **108**(38): 15870-15875.

Crittenden, S. L., et al. (2002). "A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans." Nature **417**(6889): 660-663.

Eddy, S. R. and R. Durbin (1994). "RNA sequence analysis using covariance models." Nucleic Acids Res **22**(11): 2079-2088.

Eilbeck, K., et al. (2005). "The Sequence Ontology: a tool for the unification of genome annotations." Genome Biol **6**(5): R44.

Fan, X. C. and J. A. Steitz (1998). "Overexpression of HuR, a nuclear–cytoplasmic shuttling protein, increases the in vivo stability of ARE-containing mRNAs." EMBO J **17**(12): 3448-3460.

Filipovska, A., et al. (2011). "A universal code for RNA recognition by PUF proteins." Nature chemical biology **7**(7): 425-427.

Forbes, A. and R. Lehmann (1998). "Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells." Development **125**(4): 679-690.

Ford, L. P., et al. (1999). "ELAV proteins stabilize deadenylated intermediates in a novel in vitro mRNA deadenylation/degradation system." Genes Dev **13**(2): 188-201.

Friedersdorf, M. B. and J. D. Keene (2014). "Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs." Genome Biol **15**(1): R2.

Friend, K., et al. (2012). "A conserved PUF-Ago-eEF1A complex attenuates translation elongation." Nat Struct Mol Biol **19**(2): 176-183.

Fukuglta, M., et al. (1994). "Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5-cap function." Nature **371**: 27.

Fuller-Pace, F. V. (2006). "DExD/H box RNA helicases: multifunctional proteins with important roles in transcriptional regulation." Nucleic Acids Res **34**(15): 4206-4215.

Futcher, B., et al. (1999). "A sampling of the yeast proteome." Mol Cell Biol **19**(11): 7357-7368.

Galgano, A., et al. (2008). "Comparative analysis of mRNA targets for human PUF-family proteins suggests extensive interaction with the miRNA regulatory system." PLoS One **3**(9): e3164.

Gao, F.-B., et al. (1994). "Selection of a subset of mRNAs from combinatorial 3'untranslated region libraries using neuronal RNA-binding protein Hel-N1." Proc Natl Acad Sci U S A **91**(23): 11207-11211.

García-Rodríguez, L. J., et al. (2007). "Puf3p, a Pumilio family RNA binding protein, localizes to mitochondria and regulates mitochondrial biogenesis and motility in budding yeast." J Cell Biol **176**(2): 197-207.

Gebauer, F. and M. W. Hentze (2004). "Molecular mechanisms of translational control." Nature reviews Molecular cell biology **5**(10): 827-835.

Gerber, A. P., et al. (2004). "Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast." PLoS Biol **2**(3): e79.

Gerber, A. P., et al. (2006). "Genome-wide identification of mRNAs associated with the translational regulator PUMILIO in *Drosophila melanogaster*." Proc Natl Acad Sci U S A **103**(12): 4487-4492.

Gleich, D. (2006). Hierarchical directed spectral graph partitioning, Stanford University.

Goldstrohm, A. C., et al. (2006). "PUF proteins bind Pop2p to regulate messenger RNAs." Nat Struct Mol Biol **13**(6): 533-539.

Goldstrohm, A. C., et al. (2007). "PUF protein-mediated deadenylation is catalyzed by Ccr4p." J Biol Chem **282**(1): 109-114.

Goldstrohm, A. C. and M. Wickens (2008). "Multifunctional deadenylase complexes diversify mRNA control." Nature reviews Molecular cell biology **9**(4): 337-344.

Görlach, M., et al. (1994). "The mRNA poly (A)-binding protein: localization, abundance, and RNA-binding specificity." Experimental cell research **211**(2): 400-407.

Gu, W., et al. (2004). "A new yeast PUF family protein, Puf6p, represses ASH1 mRNA translation and is required for its localization." Genes Dev **18**(12): 1452-1465.

Gygi, S. P., et al. (1999). "Correlation between protein and mRNA abundance in yeast." Mol Cell Biol **19**(3): 1720-1730.

Hafner, M., et al. (2010). "Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP." Cell **141**(1): 129-141.

Hao, S. and D. Baltimore (2009). "The stability of mRNA influences the temporal order of the induction of genes encoding inflammatory molecules." Nature immunology **10**(3): 281-288.

Harbison, C. T., et al. (2004). "Transcriptional regulatory code of a eukaryotic genome." Nature **431**(7004): 99-104.

He, L. and G. J. Hannon (2004). "MicroRNAs: small RNAs with a big role in gene regulation." Nature Reviews Genetics **5**(7): 522-531.

Hieronymus, H. and P. A. Silver (2004). "A systems view of mRNP biology." Genes Dev **18**(23): 2845-2860.

Hinman, M. and H. Lou (2008). "Diverse molecular functions of Hu proteins." Cellular and molecular life sciences **65**(20): 3168-3181.

Ho, J. J. and P. A. Marsden (2014). "Competition and collaboration between RNA-binding proteins and microRNAs." Wiley Interdiscip Rev RNA **5**(1): 69-86.

Hogan, D. J., et al. (2008). "Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system." PLoS Biol **6**(10): e255.

Holme, P. (2011). "Metabolic robustness and network modularity: a model study." PLoS One **6**(2): e16605.

Holstege, F. C., et al. (1998). "Dissecting the regulatory circuitry of a eukaryotic genome." Cell **95**(5): 717-728.

Hook, B. A., et al. (2007). "Two yeast PUF proteins negatively regulate a single mRNA." J Biol Chem **282**(21): 15430-15438.

Jacob, F. and J. Monod (1961). "Genetic regulatory mechanisms in the synthesis of proteins." J Mol Biol **3**(3): 318-356.

Jenkins, H. T., et al. (2009). "Structure and RNA binding of the mouse Pumilio-2 Puf domain." J Struct Biol **167**(3): 271-276.

Jiang, H., et al. (2012). "Rewiring of posttranscriptional RNA regulons: Puf4p in fungi as an example." Molecular biology and evolution **29**(9): 2169-2176.

Joshi, A., et al. (2012). "Post-transcriptional regulatory networks play a key role in noise reduction that is conserved from micro-organisms to mammals." FEBS J **279**(18): 3501-3512.

Joshi, A., et al. (2011). "Structural and functional organization of RNA regulons in the post-transcriptional regulatory network of yeast." Nucleic Acids Res **39**(21): 9108-9117.

Kapp, L. D. and J. R. Lorsch (2004). "The molecular mechanics of eukaryotic translation." Annual review of biochemistry **73**(1): 657-704.

Katsanou, V., et al. (2005). "HuR as a negative posttranscriptional modulator in inflammation." Mol Cell **19**(6): 777-789.

Keene, J. D. (2007). "RNA regulons: coordination of post-transcriptional events." Nat Rev Genet **8**(7): 533-543.

Keene, J. D. (2010). "Minireview: global regulation and dynamics of ribonucleic Acid." Endocrinology **151**(4): 1391-1397.

Keene, J. D., et al. (2006). "RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts." Nat Protoc **1**(1): 302-307.

Keene, J. D. and P. J. Lager (2005). "Post-transcriptional operons and regulons co-ordinating gene expression." Chromosome Res **13**(3): 327-337.

Keene, J. D. and S. A. Tenenbaum (2002). "Eukaryotic mRNPs may represent posttranscriptional operons." Mol Cell **9**(6): 1161-1167.

Kim, H. H., et al. (2009). "HuR recruits let-7/RISC to repress c-Myc expression." Genes Dev **23**(15): 1743-1748.

Kraemer, B., et al. (1999). "NANOS-3 and FBF proteins physically interact to control the sperm–oocyte switch in *Caenorhabditis elegans*." Current biology **9**(18): 1009-1018.

Kuhn, K. M., et al. (2001). "Global and specific translational regulation in the genomic response of Saccharomyces cerevisiae to a rapid transfer from a fermentable to a nonfermentable carbon source." Mol Cell Biol **21**(3): 916-927.

Kundu, P., et al. (2012). "HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA." Nucleic Acids Res **40**(11): 5088-5100.

Lal, A., et al. (2004). "Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs." EMBO J **23**(15): 3092-3102.

Lamb, J., et al. (2006). "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease." Science **313**(5795): 1929-1935.

Leandersson, K., et al. (2006). "Wnt-5a mRNA translation is suppressed by the Elav-like protein HuR in human breast epithelial cells." Nucleic Acids Res **34**(14): 3988-3999.

Lee, T. I., et al. (2002). "Transcriptional regulatory networks in Saccharomyces cerevisiae." Science **298**(5594): 799-804.

Levy, N. S., et al. (1998). "Hypoxic stabilization of vascular endothelial growth factor mRNA by the RNA-binding protein HuR." J Biol Chem **273**(11): 6417-6423.

Li, F., et al. (2004). "The yeast cell-cycle network is robustly designed." Proc Natl Acad Sci U S A **101**(14): 4781-4786.

Lopez De Silanes, I., et al. (2004). "Global analysis of HuR-regulated gene expression in colon cancer systems of reducing complexity." Gene Expr **12**(1): 49-59.

Lopez De Silanes, I., et al. (2004). "Identification of a target RNA motif for RNA-binding protein HuR." Proc Natl Acad Sci U S A **101**(9): 2987-2992.

Lu, G., et al. (2009). "Understanding and engineering RNA sequence specificity of PUF proteins." Current opinion in structural biology **19**(1): 110-115.

Lu, G. and T. M. T. Hall (2011). "Alternate modes of cognate RNA recognition by human PUMILIO proteins." Structure **19**(3): 361-367.

Lu, X., et al. (2006). "Radiation-induced changes in gene expression involve recruitment of existing messenger RNAs to and away from polysomes." Cancer Res **66**(2): 1052-1061.

Lunde, B. M., et al. (2007). "RNA-binding proteins: modular design for efficient function." Nature reviews Molecular cell biology **8**(6): 479-490.

Luscombe, N. M., et al. (2004). "Genomic analysis of regulatory network dynamics reveals large topological changes." Nature **431**(7006): 308-312.

Lykke-Andersen, J. and E. Wagner (2005). "Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1." Genes Dev **19**(3): 351-361.

Mamane, Y., et al. (2007). "Epigenetic activation of a subset of mRNAs by eIF4E explains its effects on cell proliferation." PLoS One **2**(2): e242.

Manojlovic, Z. and B. Stefanovic (2012). "A novel role of RNA helicase A in regulation of translation of type I collagen mRNAs." RNA **18**(2): 321-334.

Mansfield, K. D. and J. D. Keene (2009). "The ribonome: a dominant force in co-ordinating gene expression." Biol Cell **101**(3): 169-181.

Mansfield, K. D. and J. D. Keene (2012). "Neuron-specific ELAV/Hu proteins suppress HuR mRNA during neuronal differentiation by alternative polyadenylation." Nucleic Acids Res **40**(6): 2734-2746.

Maris, C., et al. (2005). "The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression." FEBS J **272**(9): 2118-2131.

Markus, H., et al. (2010). "PAR-CliP-a method to identify transcriptome-wide the binding sites of RNA binding proteins." Journal of Visualized Experiments(41).

Mazan-Mamczarz, K., et al. (2008). "Post-transcriptional gene regulation by HuR promotes a more tumorigenic phenotype." Oncogene **27**(47): 6151-6163.

Meisner, N. C., et al. (2004). "mRNA openers and closers: modulating AU-rich element-controlled mRNA stability by a molecular switch in mRNA secondary structure." Chembiochem **5**(10): 1432-1447.

Miller, M. T., et al. (2008). "Basis of altered RNA-binding specificity by PUF proteins revealed by crystal structures of yeast Puf4p." Nat Struct Mol Biol **15**(4): 397-402.

Milone, J., et al. (2004). "Characterization of deadenylation in trypanosome extracts and its inhibition by poly (A)-binding protein Pab1p." RNA **10**(3): 448-457.

Mittal, N., et al. (2009). "Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks." Proc Natl Acad Sci U S A **106**(48): 20300-20305.

Mittal, N., et al. (2011). "Interplay between posttranscriptional and posttranslational interactions of RNA-binding proteins." J Mol Biol **409**(3): 466-479.

Morris, A. R., et al. (2008). "Ribonomic analysis of human Pum1 reveals cis-trans conservation across species despite evolution of diverse mRNA target sets." Mol Cell Biol **28**(12): 4093-4103.

Mukherjee, N., et al. (2011). "Integrative regulatory mapping indicates that the RNA-binding protein HuR couples pre-mRNA processing and mRNA stability." Mol Cell **43**(3): 327-339.

Mukherjee, N., et al. (2009). "Coordinated posttranscriptional mRNA population dynamics during T-cell activation." Mol Syst Biol **5**: 288.

Murata, Y. and R. P. Wharton (1995). "Binding of pumilio to maternal *hunchback* mRNA is required for posterior patterning in drosophila embryos." Cell **80**(5): 747-756.

Newman, J. R., et al. (2006). "Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise." Nature **441**(7095): 840-846.

Newman, M. E. (2006). "Modularity and community structure in networks." Proc Natl Acad Sci U S A **103**(23): 8577-8582.

Olivas, W. and R. Parker (2000). "The Puf3 protein is a transcript-specific regulator of mRNA degradation in yeast." EMBO J **19**(23): 6602-6611.

Orphanides, G. and D. Reinberg (2002). "A unified theory of gene expression." Cell **108**(4): 439-451.

Parker, R. and U. Sheth (2007). "P bodies and the control of mRNA translation and degradation." Mol Cell **25**(5): 635-646.

Patel, J., et al. (2002). "Cellular stresses profoundly inhibit protein synthesis and modulate the states of phosphorylation of multiple translation factors." European Journal of biochemistry **269**(12): 3076-3085.

Peng, S. S. Y., et al. (1998). "RNA stabilization by the AU-rich element binding protein, HuR, an ELAV protein." EMBO J **17**(12): 3461-3470.

Pradhan, N., et al. (2011). "Modular organization enhances the robustness of attractor network dynamics." EPL (Europhysics Letters) **94**(3): 38004.

Quenault, T., et al. (2011). "PUF proteins: repression, activation and mRNA localization." Trends in cell biology **21**(2): 104-112.

Query, C. C., et al. (1989). "A common RNA recognition motif identified within a defined U1 RNA binding domain of the 70K U1 snRNP protein." Cell **57**(1): 89-101.

Raghavan, A., et al. (2004). "Patterns of coordinate down-regulation of ARE-containing transcripts following immune cell activation." Genomics **84**(6): 1002-1013.

Raghavan, A., et al. (2002). "Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes." Nucleic Acids Res **30**(24): 5529-5538.

Ranji, A., et al. (2011). "Features of double-stranded RNA-binding domains of RNA helicase A are necessary for selective recognition and translation of complex mRNAs." J Biol Chem **286**(7): 5328-5337.

Remenyi, A., et al. (2004). "Combinatorial control of gene expression." Nat Struct Mol Biol **11**(9): 812-815.

Simone, L. E. and J. D. Keene (2013). "Mechanisms coordinating ELAV/Hu mRNA regulons." Curr Opin Genet Dev **23**(1): 35-43.

Sobell, H. M. (1985). "Actinomycin and DNA transcription." Proc Natl Acad Sci U S A **82**(16): 5328-5331.

Sonoda, J. and R. P. Wharton (1999). "Recruitment of Nanos to hunchback mRNA by Pumilio." Genes Dev **13**(20): 2704-2712.

Spassov, D. and R. Jurecic (2003). "The PUF Family of RNA-binding Proteins: Does Evolutionarily Conserved Structure Equal Conserved Function?" IUBMB life **55**(7): 359-366.

Suh, N., et al. (2009). "FBF and its dual control of gld-1 expression in the *Caenorhabditis elegans* germline." Genetics **181**(4): 1249-1260.

Tenenbaum, S. A., et al. (2000). "Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays." Proc Natl Acad Sci U S A **97**(26): 14085-14090.

Tenenbaum, S. A., et al. (2002). "Ribonomics: identifying mRNA subsets in mRNP complexes using antibodies to RNA-binding proteins and genomic arrays." Methods **26**(2): 191-198.

Tucker, M., et al. (2002). "Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *Saccharomyces cerevisiae*." EMBO J **21**(6): 1427-1436.

Valineva, T., et al. (2006). "Characterization of RNA helicase A as component of STAT6-dependent enhanceosome." Nucleic Acids Res **34**(14): 3938-3946.

Van Etten, J., et al. (2012). "Human Pumilio proteins recruit multiple deadenylases to efficiently repress messenger RNAs." J Biol Chem **287**(43): 36370-36383.

Vardy, L. and T. L. Orr-Weaver (2007). "Regulating translation of maternal messages: multiple repression mechanisms." Trends in cell biology **17**(11): 547-554.

Wang, S.-J. and C. Zhou (2012). "Hierarchical modular structure enhances the robustness of self-organized criticality in neural networks." New Journal of Physics **14**(2): 023005.

Wang, X., et al. (2002). "Modular recognition of RNA by a human pumilio-homology domain." Cell **110**(4): 501-512.

Wang, X., et al. (2001). "Crystal structure of a Pumilio homology domain." Molecular cell **7**(4): 855-865.

Wang, Y., et al. (2002). "Precision and functional specificity in mRNA decay." Proc Natl Acad Sci U S A **99**(9): 5860-5865.

Wang, Z.-F., et al. (1996). "The protein that binds the 3'end of histone mRNA: a novel RNA-binding protein required for histone pre-mRNA processing." Genes Dev **10**(23): 3028-3040.

Weidmann, C. A. and A. C. Goldstrohm (2012). "Drosophila Pumilio protein contains multiple autonomous repression domains that regulate mRNAs independently of Nanos and brain tumor." Mol Cell Biol **32**(2): 527-540.

Wharton, R. P., et al. (1998). "The Pumilio RNA-binding domain is also a translational regulator." Mol Cell **1**(6): 863-872.

Wickens, M., et al. (2002). "A PUF family portrait: 3′ UTR regulation as a way of life." TRENDS in Genetics **18**(3): 150-157.

Workman, C. and A. Krogh (1999). "No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution." Nucleic Acids Res **27**(24): 4816-4822.

Wormington, M., et al. (1996). "Overexpression of poly (A) binding protein prevents maturation-specific deadenylation and translational inactivation in *Xenopus* oocytes." EMBO J **15**(4): 900.

Zamore, P. D., et al. (1997). "The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins." RNA **3**(12): 1421-1433.

Zhang, B., et al. (1997). "A conserved RNA-binding protein that regulates sexual fates in the *C. elegans* hermaphrodite germ line." Nature **390**(6659): 477-484.

Zhu, H., et al. (2007). "Hu proteins regulate polyadenylation by blocking sites containing U-rich sequences." J Biol Chem **282**(4): 2203-2210.

Zou, T., et al. (2010). "Polyamines regulate the stability of JunD mRNA by modulating the competitive binding of its 3' untranslated region to HuR and AUF1." Mol Cell Biol **30**(21): 5021-5032.

Biography

Marshall Aaron Thompson was born November 10, 1983 in Las Cruces, New Mexico. Marshall graduated from the University of New Mexico in 2005 with a Bachelor of Science degree in Biology and a Bachelor of Science degree in Computer Science.

Publications:

A. Slepoy, M. A. Thompson. Sandia Report, SAND2004-4185, "Stability of Biological Networks as Represented in Random Boolean Nets."

Mukherjee, N., P. J. Lager, M. B. Friedersdorf, M. A. Thompson and J. D. Keene (2009). "Coordinated posttranscriptional mRNA population dynamics during T-cell activation." Mol Syst Biol 5: 288.

J. Blackinton, M. A. Thompson, and J. D. Keene (2014). "Functional integration of mRNA decay and translation reveals coordinated regulatory dynamics in T cell activation." Manuscript in preparation.

M. A. Thompson and J.D. Keene (2014). "SIRT1 mRNA is Posttranscriptionally Regulated by the Pum1 RNA-Binding Protein." Manuscript submitted for publication.

Awards:

Duke University Center For Systems Biology Seed Grant. "RNA Network Coordination Model for Posttranscriptional Gene Expression." July 2008

Sigma Xi