

CHARACTERIZING THE ROLE OF RNA-BINDING PROTEINS IN UBIQUITIN SIGNALING

Dissertation zur Erlangung des Grades
Doktor der Naturwissenschaften

am Fachbereich Biologie
der Johannes Gutenberg-Universität in Mainz

Andrea Hildebrandt, M.Sc.

Mainz, Juni 2019

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung:

PREFACE

Results of the presented work have partly been published previously in the following publications

- **Chapter I.3**
Sutandy FXR, Hildebrandt A, and König J. Profiling the Binding Sites of RNA-Binding Proteins with Nucleotide Resolution Using iCLIP. *Methods Mol Biol.* 2016.
- **Chapter III**
Hildebrandt A *et al.* Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system. *Sci Rep.* 2017.
- **Chapter IV**
Hildebrandt A *et al.* 2019. The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation. bioRxiv; doi: 10.1101/516005

The work included in this thesis was conducted within a joint research project in the laboratories of Dr. Julian König and Dr. Petra Beli at the IMB in Mainz. Dr. Julian König and Dr. Petra Beli conceived and supervised this study. Moreover, this work was supported by research collaborations with the laboratories of Dr. Kathi Zarnack (BMLS, Frankfurt), Prof. Miguel A. Andrade-Navarro (JGU, Mainz), Dr. Jean-Yves Roignant, and Prof. René Ketting (both IMB, Mainz). I was involved in the design and evaluation of all experiments and analyses. All experimental work and data analyses were performed by myself, with the following exceptions:

Experiments in the context of the publication Sutandy FXR, Hildebrandt, and König, 2016 (**Chapter I.3**), were performed together with Dr. FX Reymond Sutandy. The manuscript was written together with Dr. Julian König and Dr. FX Reymond Sutandy.

In the context of the publication Hildebrandt *et al.* 2017 (**Chapter III**), Prof. Miguel A. Andrade-Navarro and Dr. Gregorio Alanis-Lobato (JGU, Mainz) developed a method for the evaluation of the adapted AP. Dr. Gregorio Alanis-Lobato performed computational functional analyses, including GO term analyses, WI determination, Precision-Recall analyses, and comparisons of detected interactions to the HIPPIE database. Dr. Kathi Zarnack was involved in the design and analysis of the study, contributed ideas and wrote the manuscript together with me.

In the context of data presented in Hildebrandt *et al.* 2019 (**Chapter IV**), Dr. Anke Busch (IMB, Mainz), performed initial analyses of ubiquitylome and interactome data as well as initial iCLIP data processing and analysis steps. Susan Boerner (BMLS, Frankfurt) also performed initial analyses of iCLIP data. Mirko Brüggemann (BMLS, Frankfurt) performed bioinformatics analyses of the MKRN1 iCLIP data. Cornelia Rücklée (BMLS, Frankfurt) analyzed MKRN1 binding at polyadenylation sites and poly(A) tails. Moreover, she contributed to the evolutionary characterization of Makorin proteins under the supervision of Prof. Ingo Ebersberger (Goethe Universität, Frankfurt). Jan Heidelberger (IMB, Mainz) performed replicate ubiquitin remnant profiling experiments. Andrea Voigt (IMB, Mainz) performed AP-Western blot experiments and Western blot validations. Heike Hänel (IMB, Mainz) performed replicate iCLIP and replicate AP-Western blot experiments. Dr. Stefanie Ebersberger (IMB, Mainz) and Dr. Kathi Zarnack supervised the bioinformatics analyses. Moreover, Dr. Kathi Zarnack was involved in the design and analysis of the study. Annabelle Dold and Dr. Jean-Yves Roignant (both IMB, Mainz) performed complementary studies in *D. melanogaster*. The manuscript was written by myself, Dr. Kathi Zarnack, Dr. Julian König, and Dr. Petra Beli. Prof. René Ketting, Dr. Nadine Wittkopp, and Miguel V. Almeida participated in fruitful discussions.

For **Chapter V**, Dr. Anke Busch analyzed the RNA-Seq and Pulsed SILAC data. Dr. Kathi Zarnack, Dr. Stefanie Ebersberger, and Susan Boerner were involved in the RNA-Seq analyses. Sonja Debeç (Summer student at IMB, Mainz) performed the final steps of the TOPO cloning protocol for pGL3pro-YWHAB and set up the experimental framework for the Luciferase reporter assay. Heike Hänel helped with Western blots to characterize the cell lines stably expressing RPS10 (mutants). The IMB Genomics Core Facility prepared RNA-Seq libraries from submitted mRNA and sequenced all RNA-Seq and iCLIP libraries.

Supervisor confirmation

CONTENT

PREFACE	III
CONTENT	V
ZUSAMMENFASSUNG	VIII
SUMMARY	IX
I. INTRODUCTION	12
1. BIRTH OF A FUNCTIONAL PROTEIN	12
1.1 RNA-BINDING PROTEINS	12
1.2 mRNA 3' END PROCESSING AND THE POLY(A) TAIL	16
1.3 POLY(A)-BINDING PROTEINS	18
1.4 RIBOSOMES	18
1.5 TRANSLATION	19
1.6 OBSTACLES IN TRANSLATION AND CO-TRANSLATIONAL PROTEIN QUALITY CONTROL	20
1.7 UBIQUITYLATION	25
2. RELEVANCE IN DISEASE	31
2.1 ERRORS IN GENE EXPRESSION CAN LEAD TO DISEASE	31
2.2 MISREGULATION OF UBIQUITYLATION RESULTS IN DISEASE	31
3. PROFILING THE BINDING SITES OF RNA-BINDING PROTEINS WITH NUCLEOTIDE RESOLUTION USING ICLIP	33
4. REFERENCES	56
II. AIMS OF THE PROJECT	62
III. INTERACTION PROFILING OF RNA-BINDING UBIQUITIN LIGASES REVEALS A LINK BETWEEN POSTTRANSCRIPTIONAL REGULATION AND THE UBIQUITIN SYSTEM	65
ABSTRACT	66
INTRODUCTION	66
RESULTS	68
STRATEGY FOR RECOVERING STABLE INTERACTION PARTNERS OF RNA-BINDING UBIQUITIN LIGASES	68
FUNCTIONAL COHERENCE OF THE IDENTIFIED INTERACTION PROFILE OF PRPF19	73
DEFINING THE INTERACTION PROFILES OF SELECTED RBULS	74

DISCUSSION	81
MATERIAL AND METHODS	85
REFERENCES	91
SUPPLEMENTARY FIGURES	95
SUPPLEMENTARY TABLES	103

IV. THE RNA-BINDING UBIQUITIN LIGASE MKRN1 FUNCTIONS IN RIBOSOME-ASSOCIATED QUALITY CONTROL OF POLY(A)-TRANSLATION **111**

ABSTRACT	112
INTRODUCTION	113
RESULTS	114
MKRN1 INTERACTS WITH PABPC1 AND OTHER RBPs	114
MKRN1 BINDS TO POLY(A) TAILS AND AT INTERNAL A-STRETCHES	117
MKRN1 PROMOTES RIBOSOME STALLING AT POLY(A) SEQUENCES	119
MKRN1 MEDIATES THE UBIQUITYLATION OF RIBOSOME-ASSOCIATED PROTEINS	120
DISCUSSION	123
PABP RECRUITS MKRN1 UPSTREAM OF A-STRETCHES AND POLY(A) TAILS	123
MKRN1 UBIQUITYLATES RPS10 AND TRANSLATIONAL REGULATORS TO STALL RIBOSOMES	123
DIFFERENCES BETWEEN HUMAN AND YEAST RQC EXPLAIN THE REQUIREMENT FOR MKRN1	125
MATERIALS AND METHODS	126
REFERENCES	138
SUPPLEMENTARY TABLES AND FIGURES	142
SUPPLEMENTARY TABLES	142
SUPPLEMENTARY FIGURES	144

V. FURTHER CHARACTERIZATION OF MKRN1 AND PRELIMINARY EXPERIMENTS **164**

1. MATERIAL AND METHODS	165
1.1 MATERIALS	165
1.2 METHODS	177
2. RESULTS	185
2.1 <i>MKRN1</i> CAN BE EFFICIENTLY DEPLETED IN HEK293T CELLS	185
2.2 MKRN1 PROTEINS ARE ECTOPICALLY OR STABLY EXPRESSED IN HEK293T CELLS	186
2.3 MKRN1 BINDS TO MRNA WITH MORE THAN ONE ZNF DOMAIN	187
2.4 CHANGES IN MKRN1 LEVELS REDOUND TO GLOBAL ALTERATIONS OF MRNA LEVELS	188
2.5 MKRN1 AFFECTS TRANSLATIONAL REGULATION	193

2.6	MKRN1 UBIQUITYLATES PROTEINS FOUND WITHIN THE TRANSLATIONAL ELONGATING COMPLEX	198
2.7	MKRN1 LEVELS ARE REDUCED DURING STRESS CONDITIONS	202
3.	REFERENCES	204
<hr/>		
VI.	DISCUSSION AND OUTLOOK	207
1.	FUNCTIONAL INTERACTION NETWORK OF RBULS	207
2.	MKRN1 IS INVOLVED IN DETECTING ABERRANT MRNAs, IN TRANSLATIONAL REGULATION AND HAS AN EFFECT ON MRNA STABILITY	208
2.1	MKRN1 IS A MRNP MEMBER	208
2.2	POLY(A) TAILS AND 3' UTRS OF MRNAs ARE BOUND BY MKRN1	209
2.3	MKRN1 MEDIATES RIBOSOME STALLING	211
2.4	MKRN1 IS INVOLVED IN RQC AND UBIQUITYLATES RPS10	212
2.5	THE ROLE OF MKRN1 IN RQC MIGHT BE CONSERVED IN OTHER ORGANISMS	216
2.6	MKRN1 UBIQUITYLATES REGULATORS OF MRNA STABILITY AND TRANSLATION	218
2.7	MKRN1 MIGHT REGULATE TRANSLATION VIA THE 3' UTR OR THE POLY(A) TAIL	220
2.8	MKRN1 LEVELS AFFECT MRNA ABUNDANCES	222
2.9	MKRN1 LEVELS ARE DOWNREGULATED IN RESPONSE TO STRESS	223
2.10	WHY MKRN1 IS IMPORTANT	224
2.11	SUMMARY	225
3.	REFERENCES	226
<hr/>		
VII.	APPENDIX	230
	ABBREVIATIONS	230
	LIST OF FIGURES	234
	LIST OF TABLES	235

ZUSAMMENFASSUNG

Um ein funktionelles Protein herzustellen, wird DNA in Boten-RNA (mRNA) transkribiert. Diese wird dann posttranskriptionell verarbeitet und schließlich in ein Protein translatiert. Während der Genexpression wird die mRNA von RNA-bindenden Proteinen (RBPs), die das Schicksal der mRNA bestimmen, gebunden und verarbeitet. Eine besondere Klasse von RBPs sind RNA-bindende Ubiquitin-Ligasen (RBULs), welche Substratproteine ubiquitylieren können. Die zelluläre und molekulare Funktion der meisten RBULs ist kaum untersucht. Durch die Implementierung des adaptierten Affinitäts-Aufreinigung (AP)-Ansatzes werden in dieser Arbeit die Interaktionsnetzwerke von sechs RBULs bestimmt. Mit Hilfe der Genfunktions-Ähnlichkeits-Analyse werden die stabilen Bindungspartner von jedem der Köderproteine definiert. Die interagierenden Proteine geben Hinweise auf physiologische Funktionen dieser RBULs. Die identifizierten Interaktionspartner verbinden die Ubiquitylierungs-Maschinerie mit den verschiedenen Schritten des RNA-Metabolismus'. RBULs verbinden also die molekularen Signalwege der post-transkriptionellen Regulation mit dem Ubiquitin-System.

Das Ubiquitin-System ist am Abbau von defekten Proteinen beteiligt, um die Protein-Homöostase sicherzustellen. Schadhafte mRNAs, werden co-translationell erkannt. Im Rahmen dieser Arbeit habe ich herausgefunden, dass eine RBUL, das Makorin-Ringfinger-Protein 1 (MKRN1), an der Blockierung der Ribosomen an Poly(A)-Segmenten während der Ribosom-assoziierten Qualitäts-Kontrolle (RQC) beteiligt ist. RQC ermöglicht den Abbau anomaler mRNAs und derer neu translatierten Polypeptide. Dies vermeidet die Ansammlung toxischer Proteine in der Zelle. MKRN1 wird, wahrscheinlich mit der Hilfe seines Interaktionspartners PABPC1/4, vor (vorzeitigen) Poly(A)-Schwänzen und direkt vor A-Segmenten innerhalb von 3' UTRs platziert. So kann MKRN1 defekte mRNAs erkennen. Ribosomen werden durch eine Blockade an der Translation dieser Poly(A)-Schwänze gehindert. MKRN1 ubiquityliert RPS10, was darauf hindeutet, dass MKRN1 Ribosomen durch Ubiquitylierung dieses ribosomalen Proteins stoppt. Auf eine globalere Funktion von MKRN1 in der Regulation der post-transkriptionellen Genexpression weist seine Zugehörigkeit zu einem ‚Boten‘-Ribonukleoprotein-Partikel hin, welches aus mehreren RBPs wie PABPC1, LARP1 und IGF2BP1 besteht.

SUMMARY

In order to produce a functional protein, DNA is transcribed into mRNA, which is subsequently post-transcriptionally processed and finally translated into a protein. During gene expression, mRNA is bound and processed by RNA-binding proteins (RBPs), which determine the fate of the mRNA. A special class of RNA-binding proteins are RNA-binding ubiquitin ligases (RBULs), which are capable of ubiquitylating substrate proteins. The cellular and molecular function of most RBULs is hardly understood. By implementing the adapted affinity purification (AP) approach, the interaction networks of six RBULs is determined here. With the help of gene function similarity, a set of stable binding partners for each bait protein is defined. The interacting proteins provide hints towards physiological functions of those RBULs. The identified preys link the ubiquitylation machinery to different steps of RNA metabolism. Hence, RBULs are able to bridge molecular pathways of posttranscriptional regulation and the ubiquitin system.

The ubiquitylation system is involved in the degradation of aberrant proteins to ensure protein homeostasis. Aberrant mRNAs can be detected in a co-translational manner. Within this project, I found that one RBUL, namely Makorin Ring Finger Protein 1 (MKRN1), is involved in ribosome stalling at poly(A) stretches during ribosome-associated quality control (RQC). This process enables the degradation of aberrant mRNAs and nascent polypeptides and avoids the accumulation of toxic proteins in the cell. In more detail, MKRN1 is situated in front of (premature) poly(A) tails and upstream of A-segments within 3' UTRs, probably with the help of its interaction partner PABPC1/4. In this way, MKRN1 is capable of recognizing aberrant mRNAs and, by stalling ribosomes, of hindering these from translating into poly(A) tails. MKRN1 ubiquitylates RPS10, which suggests that MKRN1 stalls ribosomes by ubiquitylating a ribosomal protein. Moreover, the finding that MKRN1 is a member of a messenger ribonucleoprotein particle (mRNP), which consists of several RBPs, such as PABPC1, LARP1, and IGF2BP1, hints towards a more global function of MKRN1 in the regulation of post-transcriptional gene expression.



CHAPTER I

INTRODUCTION

I. INTRODUCTION

1. Birth of a functional protein

In gene expression, DNA is transcribed by RNA-Polymerase (Pol) II into messenger RNA (mRNA), which is co-transcriptionally processed. During these processes, RNA-binding proteins (RBPs) form messenger ribonucleoprotein particle (mRNPs) with the nascent pre-mRNA in order to package the mRNA and avoid 'naked' mRNA in a co-transcriptional manner. Those nuclear mRNPs control the nuclear (pre-) mRNA processing. The 5' end of the mRNA is directly modified by the addition of a 7-methylguanosine (m⁷G) 'cap' to the first nucleotide of the mRNA by the CPB20/80 complex. Then, most pre-mRNAs are co-transcriptionally spliced to remove intronic sequences. When Pol II reads through the functional polyadenylation site (PAS), cleavage and polyadenylation complexes are recruited¹²⁻⁴. During the export of the mRNA to the cytoplasm, several components of the mRNP are interchanged⁴. In the cytoplasm, the exported mRNA is engaged by different RBP complexes that determine its destiny. mRNAs can be transported to a specific subcellular localization, degraded or translated in the cytoplasm. For mRNAs that are meant to be translated directly, the translation machinery encompasses mRNA when it emerges from the nuclear pore. Subsequently, the mRNA is translated into a protein⁴.

1.1 RNA-binding proteins

From transcription to translation, every mRNA processing step is tightly controlled. This task is fulfilled by RNA-binding proteins (RBPs), their interacting proteins and RNAs. Thus, while being processed, localized, degraded, or translated, mRNA is always occupied by RBPs. Hence, these proteins are involved in all steps of gene expression. They frequently act as connectors for other proteins of the cellular machineries that determine mRNA fate. Fifty years have passed since the discovery of RNA-protein interactions⁵. Screens for proteins capable of RNA-binding in yeast and 'RNA capture' studies in several organisms have revealed a wide range of RBPs in several organisms⁶⁻¹¹. To date, there is knowledge of more than 1300 human proteins that are part of the RBP family and bind RNA via RNA-binding domains (RBDs)¹¹.

1.1.1 Functions of RNA-binding proteins

RBPs exert different functions on RNAs. Within the cell, RBPs regulate processing, nuclear export, stability, folding, subcellular localization, and translation of mRNAs^{3,4,11-13}. Moreover, RBPs act in gene regulation mediated by non-coding RNAs and harbor RNA nuclease functions. Additionally, they are found in ribosomes and are involved in the biogenesis of transfer RNAs (tRNAs), small nucleolar RNAs, and ribosomal RNAs (rRNAs) and are constituents of ribosomes. Moreover, RBPs act in gene regulation mediated by long non-coding (lnc) RNAs or piwi RNAs and harbor RNA nuclease functions¹³.

Opposed to RBPs regulating RNA destiny, RNAs can also exert functions on RBPs. By scaffolding, RNAs can bring different RBPs in close proximity (Figure I-1A). For example the lncRNA HOTAIR simultaneously binds the proteins MEX3B and Snurportin-1, catalyzing the ubiquitylation of the latter by MEX3B¹⁴. Moreover, binding to RNA might alter protein conformation and enable the protein to exert enzymatic functions (Figure I-1B)¹¹⁻¹³.

RBPs bound to RNA form dynamic RNP complexes. Depending on their determined functions, those compounds remodel with help of post-translational modifications (PTMs) on RBPs, RNA helicases, and the joining or departing of RBPs. In some cases, RBPs or RNPs can act as 'RNA chaperons' that hinder mRNA from misfolding. Moreover, RBPs within RNPs help to avoid incomplete processing of mRNAs, enable transport, and control mRNA abundances^{11,13}.

1.1.2 Modular structure of RNA-binding domains

The regulation of mRNA fate is controlled by the interplay of RBPs. The proteins collaborate by interacting with each other or by contacting the same mRNA. In contrast, RBPs can act on mRNA in an antagonistic way by competing for binding sites or by binding to different *cis*-regulatory elements¹⁵. RBDs represent contact points between RNA and proteins. They enable proteins to bind RNA in a sequence- or structure-specific manner. RNA recognition motifs (RRM) are the most widespread RBDs, with other RBDs being K homology (KH) domains, Tudor domains, WD40 domains, or C₃H Zinc fingers (ZNFs), for instance¹².

Within an RRM, the β -sheet recognizes four RNA nucleotides. Additionally, four to eight nucleotides can be contacted by a single RRM using exposed loops and secondary

structure elements of the peptide chains¹². In contrast, C₃H ZNF proteins do not recognize RNA bases via the amino acid side chains but via hydrogen bonds between the protein backbone and the Watson-Crick edges of RNA bases.

Opposed to RRMs, that recognize RNA with the surface of β -sheet by hydrogen bonding and stacking as well as electrostatic interactions, CCCH ZNFs mainly use their protein backbone to contact RNA. The CCCH ZNF is composed of the motif C-X₅₋₁₄-C-X₅-C-X₃-H with C representing cysteine and H histidine. 'X' stands for any amino acid. The three cysteines and one histidine chelate a zinc ion, thereby forming a finger-shaped fold. Besides the binding of the zinc ion, which leads to the formation of a fold, the CCCH ZNF domain of this protein family shows little secondary protein structure. For example, in Tis11d, there is only a short three-amino acid helix present after the first cysteine. RNA can be bound in a sequence- or structure-specific manner through hydrophobic packing, stacking interactions and hydrogen bonds, which are established between the Watson-Crick edges of the RNA bases and the protein main chain. Hence, the protein shape provides a framework for hydrogen bonds and consequently is responsible for the RNA-binding specificity of the protein. Additionally, the amino acid composition of the ZNF domain has an effect on the RNA-binding behavior and -specificity of the protein. Within the ZNF domain, some amino acid side chains contact atoms of the main chain by hydrogen bonds and help to provide a structural template for RNA contacts. The side chains of the aromatic amino acids within the ZNF domain support the RNA recognition, for example by contributing to the formation of hydrophobic binding pockets for RNA bases, which are contacted by hydrogen bonds or stacking interactions. A single CCCH ZNF domain has only little RNA-binding capability. For this reason, CCCH ZNF proteins often contain several ZNF domains to strengthen the RNA affinity of the protein. Alternatively, proteins that only contain one ZNF domain act as dimers^{12,16-20}. As RBDs only recognize few nucleotides within the RNA or bind to small structural RNA motifs, many RBPs contain several RBDs, which are arranged in a modular manner and are connected by linkers. This modular RBD setup within one protein augments binding diversity, specificity and affinity (Figure I-1C and D). Hence, multiple domains binding to a single RNA form bigger contact areas that enable the protein to attach to longer nucleotide sequences.

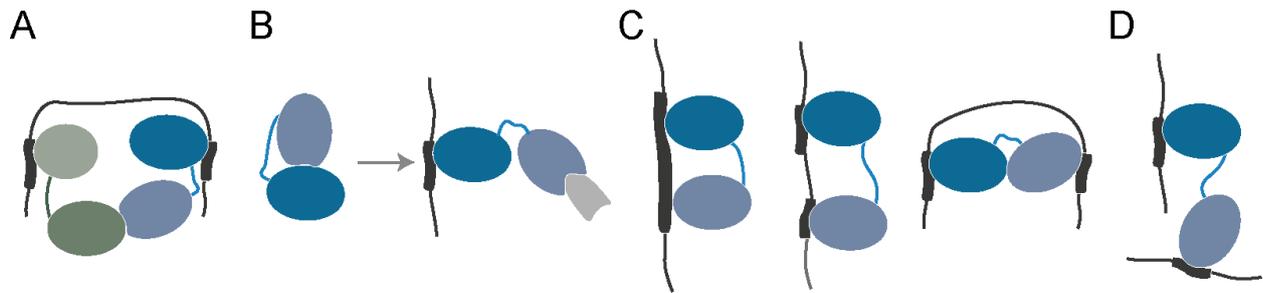


Figure I-1, RNA binding proteins often harbor several RBDs that collaboratively mediate RNA binding. **A**, RNA can act as a scaffold, bringing two different RBPs together and mediating their interaction. **B**, For bifunctional RBPs that harbor enzymatic activity, RNA-binding might enable enzymatic activity. **C**, Binding of several RBDs to sequences within one RNA (long segments (left) or shorter sequences that are further apart (middle, right)) allows binding with high affinity and specificity. **D**, With different RBDs, one RBP is capable of recognizing different RNAs. Figure adapted from Lunde, Moore, and Varani, 2017⁴.

As RBDs only recognize few nucleotides within the RNA or bind to small structural RNA motifs, many RBPs contain several RBDs, which are arranged in a modular manner and are connected by linkers. This modular RBD setup within one protein augments binding diversity, specificity and affinity (Figure I-1C and D). Hence, multiple domains binding to a single RNA form bigger contact areas that enable the protein to attach to longer nucleotide sequences.

1.1.3 Unstructured domains of RBPs in RNA- and protein-binding and their role in granule and hydrogel formation

To enlarge their RNA-binding surface, RBPs can utilize their interdomain regions, such as extended loops, for RNA-binding, in addition to 'conventional' RBDs. For example, while a single RRM recognizes four nucleotides in a sequence-specific manner, up to four additional nucleotides can be contacted by a single RRM employing exposed loops and secondary structure elements of the peptide chain, such as α -helices. This has been shown for La protein and PABP, for instance¹². Similarly, the linker connecting RBDs is involved in the RNA-binding behavior of a specific RBP. Due to the secondary structure of the peptide chain, those linkers provide RBPs with increased RNA-binding specificity and larger contact interfaces. Furthermore, inter-domain arrangements, protein-protein interactions, and unstructured domains influence RNA-binding^{12,21}.

Intrinsically disordered regions (IDRs) or unstructured low complexity domains are commonly present in RBPs and are rich in small, polar or charged amino acids, often found in repeats^{9,21,22}. When hydrophobic peptides are exposed on the surface of a protein, they facilitate binding of the unstructured domain to other proteins or RNA,

which results in the formation of membrane free granules, e.g. P bodies, by reversible liquid-liquid phase separation^{9,11,22–29}. Strikingly, a high convergence of IDRs and maturation of phase separated droplets can lead to the formation of hydrogels and even amyloid-like structures^{26–30}. In case of misregulation of granule formation or mutations that increase fiber predisposition, extreme fiber formation and protein aggregates can occur, culminating in disease^{21,28}.

1.1.4 Enzymatic functions in RNA-binding proteins

Several RBPs do not only harbor RNA-binding capability but additional enzymatic functions. In RNA-interactome capture experiments, several metabolic enzymes have been found to potentially bind to RNA^{8,9,11,21,31}. For example, phosphoglycerate kinase and thioredoxin are novel members of the RBP class³¹. In addition to metabolic enzymes, proteins involved in protein folding or actin-binding have been identified as RBPs³¹. Nonetheless, the role of many enzymatically functional RBPs in post-transcriptional regulation remains elusive. Interestingly, enzymes conferring PTMs, such as kinases and ubiquitin ligases, have also been shown to bind to RNA^{9,31,32}. RNA-binding ubiquitin ligases will further be discussed below (Chapter I.1.7.1).

1.2 mRNA 3'end processing and the poly(A) tail

The end of an mRNA's 3' UTR is defined by polyadenylation. In a co-transcriptional manner, 3' end processing factors are recruited to the mRNA's PAS with the help of the carboxy terminal domain (CTD) of Pol II. mRNA 3' end processing is a two-step process composed of endonucleolytic cleavage and poly(A) tail synthesis (polyadenylation). The consensus PAS within the 3' UTR contains several regulatory *cis* elements that are localized upstream and downstream of the cleavage site. Regulatory *cis* elements are, for instance, the hexameric motif A[A/U]UAAA (in short: AAUAAA), U-rich upstream auxiliary elements (USEs) and enhancing elements downstream (DSEs) (Figure I-2A). To induce the cleavage and polyadenylation reactions, the PAS is identified with the help of the multimeric complexes CPSF (cleavage and polyadenylation specificity factor), CstF (Cleavage stimulation factor), and CFI (cleavage factor I). After the PAS is defined, the mRNA is cleaved, which initiates polyadenylation of the mRNA. Poly(A) polymerase (PAP) adds adenosine (A) nucleotides (nt) to the newly generated mRNA 3' end, thereby forming the mRNA's poly(A) tail (Figure I-2A). While on average, the poly(A) tail consists of about 80 nt in yeast, the human poly(A) tail is 200 nt long^{2,33–36}.

The poly(A) tail of mRNAs is involved in regulating mRNA export and nuclear mRNA decay, as well as in mRNA translational regulation, especially during development. By cytoplasmic shortening and extending the poly(A) tail, the production of specific proteins can be regulated in a timely manner during early embryonic development and in neurons^{33,36–38}. Moreover, the poly(A) tail constitutes a binding platform for Poly(A) binding proteins (PABPs), which are key regulators of mRNA fate^{36,39}.

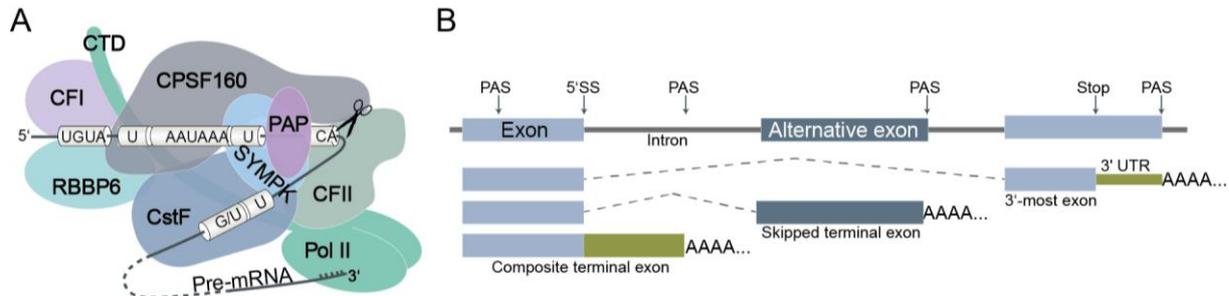


Figure I-2, mRNA (alternative) polyadenylation. **A**, Cleavage and polyadenylation machinery assembled at polyadenylation signal (PAS). Regulatory cis elements of pre-mRNA that signal for cleavage and polyadenylation are shown. The cleavage and polyadenylation machinery, consistent of the CPSF complex, the CstF complex, the CF I complex, the CF II complex, SYMPK, RBBP6, and PAP, assembles around the cis regulatory elements on the mRNA as well as on the Pol II CTD. The cleavage event is indicated by the scissors pictogram. **B**, Alternative polyadenylation (APA) of mRNA. Different APA events within one mRNA can lead to different mRNA isoforms that can code for proteins with specific functions or result in aberrant mRNAs. Splicing events resulting in three different mRNA isoforms are shown in dashed lines. On top, the ‘coding’ pattern is depicted with arrows indicates PASs, one 5’ splice site (SS) and the stop codon. The first ‘canonical’ isoform results from skipping of the alternative exon and translation termination at the stop codon. The 3’ UTR (in green) is followed by the ‘canonical’ poly(A) tail. The usage of the PAS 3’ of the alternative exon results in a transcript that contains the alternative exon and is polyadenylated but does not possess a 3’ UTR (middle). The third transcript depicts a transcript containing a composite terminal exon, which results from the usage of an internal PAS and the skipping repression of the 5’ SS. This leads to the partial inclusion of the intron and the polyadenylation of this transcript (bottom). Figure adapted from Tian and Manley, 2017³⁴.

Alternative polyadenylation (APA) gives rise to several mRNA isoforms that possess different characteristics concerning half-life, cellular localization, or translational behavior, aiding to protein diversity. Most pre-mRNAs harbor two polyadenylation sites within their 3’ UTR, culminating in diverse mRNA 3’ UTR lengths. Interestingly, PASs can be found within the open reading frame (ORF). Consequently, the usage of an internal PAS results in mRNAs with alternative termination exons and 3’ UTRs (Figure I-2B). In some cases, the mRNA created by the activation of an internal PAS does not contain a stop codon and gives rise to a truncated mRNA harboring the poly(A) tail within the ORF (prematurely polyadenylated mRNA; Figure I-2B). Of note, premature polyadenylation has been reported for 5% and 1% of mRNAs in yeast and human, respectively^{40–42}. Translation of this aberrant mRNA would culminate in the ribosome translating the polyA tail. For this reason, these kinds of mRNAs are targeted for decay (see Chapter I.1.6.1)^{2,34,35,40,41}.

1.3 Poly(A)-binding proteins

As mentioned above, the poly(A) tail is bound by Poly(A) binding proteins (PABPs). The human genome encodes one nuclear and four cytoplasmic PABPs. PABPs are highly conserved in eukaryotes and bind to poly(A) sequences via RRM. Opposed to the nuclear PABPN1, which only contains one RRM, cytoplasmic PABPs (PABPC1, PABPC3, PABPC4, and PABPC5) contain four RRMs that are functioning in RNA and protein binding. In some PABPs, a MLLE (methionine-leucine-leucine-glutamic acid) domain is connected to the RRMs via a proline- and glutamine-rich linker. The MLLE domain is not involved in contacting RNA but offers a platform for protein-protein interactions, via a PABP-interacting motif 2 (PAM2) that is present in PABP-binding proteins^{36,39,43,44}.

PABPs, especially PABPC1, play a role in many cellular pathways, such as regulation of polyadenylation, nuclear export, mRNA surveillance, mRNA stability, and regulation of translation during development^{36,39,45}. Exemplarily, PABPN1 has been implicated in polyadenylation, by stimulating PAP activity and by acting as a 'molecular ruler' for A nucleotides. It also acts in nuclear export, as well as in PAS recognition and APA regulation^{36,39}. PABPC1 stimulates cap-dependent translation initiation by binding to the poly(A) tail and EIF4G. By doing so, it helps to stabilize the binding of translation initiation factors to the 5' cap and mRNA. This enhanced binding aids the recruitment of the 43 Svedberg (S) ribosomal subunit (SU)^{36,39,45}.

1.4 Ribosomes

The ribosome is the heart of the translational process and its function is to translate mRNA into proteins by decoding nucleotide-triplets into amino acid chains. In eukaryotes, the cellular localization of the 80S ribosome is cytosolic, where they can be found in a 'free' state or attached to the endoplasmic reticulum. The ribosome consists of the ribosomal 40S and 60S SUs. Those SUs are composed of ribosomal RNA (rRNA), ribosomal proteins and associated proteins^{46,47}. The 40S SU contains the 18S rRNA and 33 ribosomal proteins (RPSs, proteins of the small ribosomal SU). The 60S SU consists of the 28S rRNA, the 5.8S rRNA, and the 5S rRNA, as well as of 47 ribosomal proteins (RPLs, proteins of the large SU). Within the ribosomal interface of the 40S and 60S SUs, the peptide bond synthesis takes place. Here, the rRNAs catalyze the joining of amino acids into a polypeptide. Ribosomal proteins form the framework of this ribozyme complex⁴⁷⁻⁴⁹. The ribosome harbors three RNA-binding

sites, which expand over the 40S and the 60S SUs and host the tRNAs during translation: the aminoacyl site (A site), the peptidyl site (P site), and the exit site (E site). Amino acid-charged tRNAs (aminoacyl-tRNAs) arrive and bind at the A site to enable peptide-bond formation. tRNAs that are attached to the nascent peptide chain are bound at the P site (peptidyl-tRNAs). When the tRNA is released from its amino acid (de-acylated or uncharged), it is present at the E site before detaching from the ribosome⁵⁰. Several translating ribosomes found on one mRNA are referred to as a polysome and represent the molecular machinery of protein synthesis.

1.5 Translation

Upon nuclear export, mRNAs can be directly translated into proteins by ribosomes. The translational process is divided into three steps: initiation, elongation and termination^{49–53}. During cap-dependent translation initiation, the 5' m⁷G cap is bound by the eukaryotic initiation factor (EIF) 4F complex. This complex consists of the proteins (EIF) 4A, EIF4E, and EIF4G. EIF4E binds the 5' cap and the helicase EIF4A dissolves secondary mRNA structures within the 5' UTR. EIF4G bridges the cap to the poly(A) tail by binding to PABP, which enables the mRNA to form a loop. The 43S preinitiation complex (PIC) consists of the 40S SU bound to EIF1, EIF1A, EIF3, EIF5, and the ternary complex (initiator tRNA^{Met} and guanosine-5'-triphosphate (GTP)-EIF2). After binding to the 5' cap, the PIC starts scanning the 5' UTR for the start codon (AUG). At the start codon, the PIC stops and forms the 48S initiation complex. Subsequently, the 60S ribosomal SU joins (subunit joining). With this, the 80S ribosome is formed with initiator tRNA^{Met} positioned at the P site on the AUG start codon^{49,54}.

In translation elongation, the ribosome slides along the mRNA decoding it in a three-nucleotide window and produces a polypeptide until it reaches one of the three stop codons (UAA, UAG, UGA) at the end of the ORF. In detail, the eukaryotic elongation factor (EEF) 1A delivers amino acid-charged tRNAs to the ribosomal A site, where tRNAs sample the mRNA codons. When the mRNA codon is specifically recognized by the tRNA anti-codon, the respective tRNA is inserted into the A site, in a process that is called tRNA decoding. Subsequently, a peptide bond is formed between the incoming amino acid and the amino acid of the peptidyl-tRNA. This is accompanied by the transfer of the peptide chain attached to the peptidyl-tRNA to the aminoacyl-tRNA in the A site. EEF2 assists the shifting of the 'new' aminoacyl-tRNA, now bearing the peptide chain, from the A site to the P site. Simultaneously, the empty tRNA is moved from the

P site to the E site. After moving to the E site, the deacylated tRNA is set free. This cycle is repeated until the termination codon at the end of the ORF is reached^{50,51,53}.

Eukaryotic release factor (ERF) 1 recognizes the stop codon within the A site and by reaching into the P site, it catalyzes the release of the nascent peptide. ERF3 positions ERF1 at the ribosomal A site at a stop codon. Moreover, ERF3 serves the peptide release and termination efficiency by hydrolyzing GTP^{50-53,55}. After termination, ribosomes are recycled. This means that the 80S ribosome is dissociated into the 40S and the 60S SU with the help of ABCE1. After splitting, the mRNA and tRNA are removed from the SUs. Due to these recycling steps, the ribosome can commence a new round of translation^{50,53}.

1.6 Obstacles in translation and co-translational protein quality control

Errors in translation can result in aberrant proteins, which might display reduced or no activity or are misfolded. Faulty proteins are prone to aggregation and thus pose negative impacts on cellular functions⁵⁶⁻⁵⁸. In order to avoid the production of incomplete and harmful peptides, several quality control (QC) mechanisms constantly survey translation. On the mRNA side, for example, mRNAs harboring premature stop codons are recognized by non-sense mediated decay (NMD). Those mRNAs are ultimately degraded by XRN1 (5' → 3') and the exosome (3' → 5')⁵⁹. mRNA surveillance pathways are accompanied by protein-centered QC mechanisms to avoid the production of faulty proteins, whose accumulation is linked to neurodegenerative diseases. Prematurely polyadenylated mRNAs, for instance, stall translating ribosomes and elicit the activation of the ribosome-associated quality control (RQC). RQC triggers the degradation of the nascent peptide and fosters ribosome recycling and simultaneously, non-stop decay (NSD) and no-go decay (NGD) degrade the aberrant mRNA (see Chapters I.1.6.1 and I.1.6.2). All together the QC systems take care of the detection of aberrant translation, thereby maintaining the cellular homeostasis. Thus, these mechanisms help to avoid stress responses, like the unfolded protein response (UPR), and monitor translational accuracy^{57,58,60-64}.

In order to maintain cellular homeostasis, translation rates of proteins are tightly regulated. As translation is one of the most energy-consuming processes within the cell, the decision whether translation takes place or not is taken as early as possible, this means at the translation initiation step. However, in case of stress, also translation elongation is inhibited to rapidly respond to environmental changes. Hence, by

regulating protein synthesis via translational control and protein degradation, the amount of cellular proteins is kept in balance^{56–58}. Errors in protein production can emerge in every aspect of protein synthesis; from the formation of aminoacyl-tRNAs to protein folding. Inaccuracies in translation can result in ribosome stalling and subsequently in premature translation termination^{55,57}. Temporarily paused ribosomes are capable of resuming translation after impediments have been removed. These ribosomes can be utilized to fine-tune translation, while terminally stalled ribosomes lead to translation abort⁵⁰. Subsequently, premature translation termination leads to downregulation of translation initiation. Moreover, the currently translated mRNA and the nascent chain are decayed, while the ribosome SUs are recycled or dismantled^{55,65,66}.

Ribosomes stall due to changes in tRNA modifications, the incorrect incorporation of amino acids, and the unavailability of specific aminoacyl tRNAs. Moreover, aberrant mRNAs that, for instance, arise due to chemicals, endonucleases, or faulty gene expression, negatively affect translating ribosomes^{55,67–69}. In addition, rare or problematic codons that are hard to decode (e.g. Pro) have been linked to translation speed reduction. In line, specific amino acids of the nascent chain might form bulky arrangements within the exit tunnel, slowing down or stalling ribosomes⁵⁰. In yeast, translation of six to twelve basic amino acids (Arg or Lys), are sufficient to stall ribosomes and induce nascent peptide decay^{63,64,70,71}. Interestingly, premature polyadenylation within the coding sequence of mRNAs, resulting in poly(A) tails within ORFs, as well as long A stretches, have the capability to retard translating ribosomes^{41,65,66,72,73}. Contrary to yeast, where basic amino acids are the reason for stalling, the nucleotide sequence, specifically A nucleotides, is the cause of terminal ribosome pausing in mammals^{70–74}. In summary, stalled ribosomes signal for erroneous translation. Thus, the ribosome presents a platform for surveillance and QC pathway factors to co-translationally sample the fidelity of the mRNA and the nascent chain. Below, I want to focus on the mRNA surveillance pathways no-go and non-stop decay as well as on ribosome-associated quality control (RQC, see Chapter I.1.6.2).

1.6.1 No-go and Non-stop mRNA decay

As mentioned above, mRNA surveillance pathways break down aberrant mRNAs. mRNAs that harbor sequences hindering ribosome elongation are subject to no-go decay (NGD). Those mRNAs contain stem loops, pseudo knots, or rare codons, for

instance^{67–69,71}. Non-stop decay (NSD) is responsible for the degradation of mRNAs that do not contain stop codons, truncated mRNAs, prematurely polyadenylated mRNAs or mRNAs that have a poly(A) tail but miss a stop codon^{40,41,66,75}. Ribosome stalling initiates NGD and NSD. However, the signals that actually activate the decay pathways stalling remain elusive. Opposed to this, the downstream events upon NGD and NSD initiation are known and quite similar⁶⁶. In both pathways a so far unknown endonuclease is recruited to stalled ribosomes. Subsequently, the mRNA is cleaved 5' of the stalled ribosome^{65,68,69,76,77}.

The Dom34/Pelota-Hbs1 complex as well as Asc1/RACK1 have been implied in fostering endonucleolytic cleavage^{67–69,76,78}. Upon cleavage, Dom34-Hbs1 mediates splitting of ribosomes stalled at 3' ends of mRNAs, enabling exonucleases to dismantle the mRNA^{65,66,69,79,80}. In the context of the NSD pathway, cleaved fragments of mRNAs with poly(A) stretches or without stop codons are decayed by the exosome (Figure I-3). In yeast, the exosome complex is recruited by Ski7, a relative of Hbs1^{41,81}. Interestingly, the exosome also harbors endonucleolytic activity and might be solely responsible for degradation of non-stop mRNA targets⁸². In NGD, the 5' fragment resulting from endonucleolytic cleavage is decayed by the exosome as well. Furthermore, it was proposed that the 3' cleavage fragment is targeted by Xrn1-mediated decay⁶⁸. An additional way to degrade non-stop and no-go transcript 5' fragments are continuous endonucleolytic cleavages. In detail, translating ribosomes following the first, already stalled ribosome will translate the 5' fragment and reach the newly generated 3' end of the mRNA. This truncation stalls the trailing ribosome resulting in endonucleolytic cleavage of the mRNA. With recurrent cleavages, the mRNA is degraded (Figure I-3)^{65,66,69,82}.

1.6.2 Ribosome-associated quality control

As discussed above, the ribosome can encounter problematic sequences, such as poly(A) tracts in ORFs, that lead to terminal ribosome pausing. The translation of faulty non-stop or no-go mRNA can lead to the (partial) translation of potentially deleterious polypeptides. In addition to mRNA decay, the defective polypeptide also needs to be degraded and the stalled ribosome desire recycling. Stalled ribosomes evict RQC activation in addition to NGD/NSD initiation. One the one hand, RQC provokes the decay of aberrant mRNA and the nascent peptide chain. On the other hand, ribosome

SUs and RQC components are degraded or recycled. Successful RQC supports the cell in avoiding the activation of cellular stress responses^{55,58,64}.

In yeast, ribosomes stall on stretches of poly-basic amino acids. Furthermore, it has been suggested that basic amino acids are impeded within the negatively charged ribosomal exit channel due to electrostatic interactions^{70,71,83}.

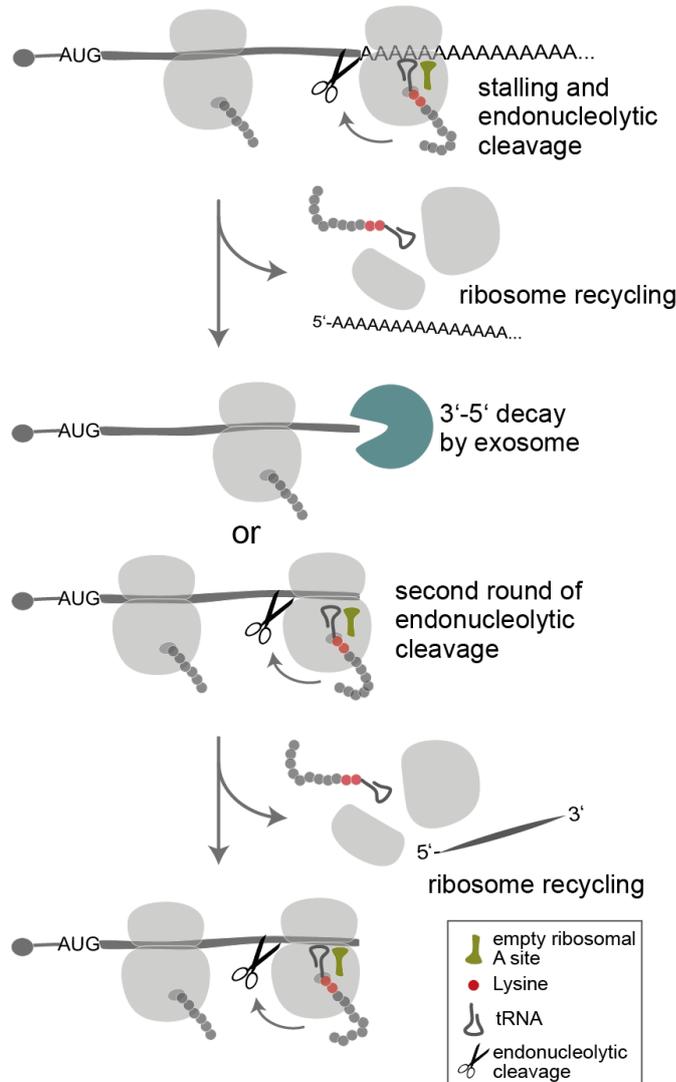


Figure I-3, Co-translational non-stop decay. In human, translating ribosomes are stalled when translating into the poly(A) tail. Endonucleolytic cleavage takes place 5' to the ribosome. The ribosome is recycled and the polypeptide chain, as well as the mRNA are degraded by RQC and exosomal decay, respectively. The remaining mRNA is degraded 3' – 5' by the exosome. Alternatively, reoccurring rounds of endonucleolytic cleavages upon ribosome stalling at the 3' end of the transcript can occur and result in small mRNA fragments. Figure adapted from Guydosh and Green, 2017⁶⁴.

Those terminally stalled ribosomes are recognized by Hel2 and Asc1. Together with the RQC triggering complex (consistent of Slh1, Cue3, and KYR023W), Hel2 and Asc1 are capable of inducing downstream RQC signaling at poly-basic stretches in yeast^{67,71,84–90}. Asc1 acts upstream of Hel2 and might stabilize the binding of Hel2 to the interface of ribosomal 40S SUs^{72,84,85,89}. Hel2 is an E3 ligase, which ubiquitylates the ribosomal proteins RPS20 and RPS3^{84,86,89,91}. PTMs on proteins of the translational machinery provide a level of translational regulation. An example is the phosphorylation of EIF2 α upon stress. Also ribosomal proteins are ubiquitylated upon stress, like the UPR. Ribosome ubiquitylation has previously been shown to influence RQC^{56,73,92}.

In mammals, lysine-coding poly(A) nucleotide stretches are causative of ribosome stalling, independently of interactions with the exit channel^{72–74,93}. Those ribosomes are recognized by ZNF598 and RACK1 (Hel2 and Asc1 in yeast, respectively). ZNF598 has been suggested to mediate ribosome stalling by ubiquitylation of ribosomal proteins^{72,73,93}. This E3 ligase ubiquitylates the ribosomal proteins RPS20, RPS10, RPS3A, and RPS3. Notably, the main ubiquitylation substrate of ZNF598 involved in eliciting RQC is RPS10, with ubiquitylation sites K138 and K139^{72,73,93}. It was proposed that ZNF598 travels with the ribosome and thereby recognizes the translation of poly(A) sequences due to its crosslinking to tRNA^{Lys}(UUU) coding for lysine in mammalian cells⁹³. Whether and how poly(A) translation is actually recognized from an 'mRNA point of view' is still under debate. So far, it has been shown that active translation is required to identify those aberrant mRNAs⁴¹.

Upon detection of the stall, the ribosome subunits are split and recycled. The molecular mechanism of ribosome stalling differ between yeast and mammalian cells. However, downstream signaling events of RQC, such as the recognition of stalled ribosomes and eviction of the polypeptide chain, are highly similar between yeast and mammalian RQCs, which hints towards a conserved QC mechanism^{55,64,79,94}. As most research has been conducted in yeast, I will mainly refer to the yeast homologs of the proteins mentioned below. Dom34, Hbs1, and Rli1 (Pelota, HBS1, and ABCE1 in mammals; Figure I-4) are responsible for splitting of the ribosome. Interestingly, Hbs1 is a homolog of ERF3 and EEF1A1, while Dom34 is homologous to ERF1 and resembles tRNAs^{95–97}. It is hypothesized that, under normal conditions, the Dom34-Hbs1 complex is outcompeted by EEF1A1-tRNA or ERF3-ERF1 at translating or terminating ribosomes, respectively. In slow or stalled ribosomes, however, Dom34 can be accommodated into the ribosomal A site with the help of Hbs1 as a result of missing competition. Subsequently Rli1 is recruited and the ribosomal SUs are then irrevocably split^{53,55,60,64,79,94}. Upon ribosomal splitting, the 40S ribosomal SU is recycled and the mRNA is degraded by NSD or NGD (see above)^{41,64,65,68,81,94}.

Opposed to the free 40S SU, the 60S SU is still attached to the tRNA and the nascent chain. This composite is bound by Ltn1/LTN1. Rqc2/Tae2 (NEMF in human) recognizes the residual tRNA and stabilizes binding of Ltn1 to the 60S SU. In addition, Rqc2 prevents the rejoining of the 40S SU by covering the majority of the 60S SU interface. With the RING (really interesting new gene) domain of Ltn1 sited at the ribosomal exit tunnel, the E3 ligase is capable of polyubiquitylating lysines, produced during translation

of non-stop or prematurely polyadenylated mRNAs (Figure I-4)^{55,58,102,60,63,64,88,98–101}. Subsequently, the Cdc48/VCP, together with Udf1 and Npl4, is recruited by Rqc2/Tae2, Ltn1, and Rqc1, disassembles the 60S composite by evicting the polyubiquitylated nascent peptide chain. The latter is guided towards the proteasome, where it is degraded (Figure I-4). By cleaving the tRNA bound to the polypeptide within the 60S SU, ANKZF1 assists the detachment of the polyubiquitylated nascent chain^{55,58,105–108,63,64,70,88,101–104}.

Besides Ltn1 and Hel2/ZNF598, one other RING ubiquitin ligase has been linked to RQC. The E3 ligase Not4 is involved in translation inhibition during RQC⁸⁸. Firstly, Not4 interacts with polysomes. Secondly, Not4 was shown to be involved in repressing translation of poly-lysine-containing proteins, seemingly independent of its ubiquitylation function. Furthermore, it seems to regulate the levels and stability of nonstop mRNAs^{55,63,88,109–113}.

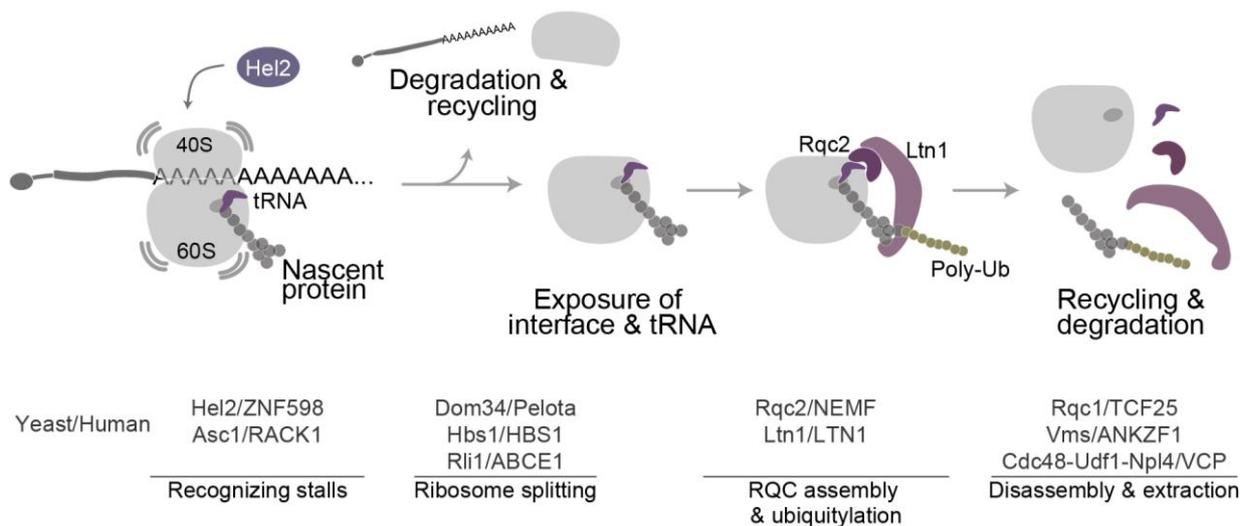


Figure I-4, Steps of the ribosome-associated quality control pathway. A ribosome stalled at poly-basic/ poly(A) segments is recognized by Hel2/ZNF598. Ribosome splitting factors disassemble the translational complex and send the 40S SU for recycling. The mRNA is degraded by NGD/NSD. The exposed tRNA and 60S SU interface are recognized by Rqc2/NEMF, which stabilizes Ltn1/LTN1 binding. The E3 ligase Ltn1 ubiquitylates the nascent polypeptide chain, which is then extracted with the help of the Cdc48/VCP and targeted for proteasomal degradation. The other components of the RQC are subsequently recycled. Protein factors from yeast and human involved in the respective step are mentioned below. Figure adapted from Brandman and Hegde, 2016⁶⁴.

1.7 Ubiquitylation

Protein PTMs provide increased functional diversity to proteins and the ability to fine-tune molecular interactions. PTMs are covalently attached to proteins and can be either chemical residues (e.g. phosphate-, methyl-, or acetyl groups) or small proteins (e.g.

ubiquitin or SUMO (small ubiquitin-related modifier)). PTMs can regulate protein stability and localization, for instance.

Ubiquitin is reversibly attached to proteins, including itself, in a three step reaction, named ubiquitylation. With the help of adenosine triphosphate (ATP) hydrolysis, the E1 enzyme activates ubiquitin and forms an E1~ubiquitin thioester. The ubiquitin conjugating enzyme E2, is linked to ubiquitin in a trans-thioesterification reaction and guides it to the E3 ligase. Finally, the E3 mediates the ubiquitylation of the substrate, either by binding to the substrate and mediating the transfer of ubiquitin from the E2 to the target protein (RING ligases) or by first binding to ubiquitin itself and then transferring it to the substrate (HECT ligases). With this, the C'-term end glycine of ubiquitin is attached to a substrate's lysine (Figure I-5A)¹¹⁴⁻¹¹⁷.

Ubiquitin is a small protein of 8.5 kDa. It is produced by deubiquitylases (DUBs) that cleave of single ubiquitin moieties from either a fusion peptide of ubiquitin and a ribosomal protein (encoded by UBA52 and RPS27A) or from polyubiquitin cassettes (encoded by UBB and UBC)^{118,119}. Ubiquitin can impact on protein stability or confer signaling functions. For example, K48-modified proteins are degraded by the proteasome. Furthermore, ubiquitin signaling is involved in DNA repair, immune signaling, endocytosis, and cell cycle. Ubiquitin also plays a role in the chromatin architecture, as well as in the DNA damage response and immune signaling. Consequently, the effects of ubiquitin signaling in the cell are versatile and widespread^{115,117}.

While there are only two E1 ligases (UBA1 and UBA6) and approximately 40 E2 enzymes in human, more than 700 E3 ligases are known. The group of E3 ligases can be divided into RING, HECT (homologous to the E6-AP carboxyl terminus), and RBR (RING between RING) ubiquitin ligases. The catalytic domains of RING E3 ligases consist of either RING C₃HC₄ zinc fingers or Ubox domains. To transfer ubiquitin onto a substrate, HECT ligases perform a two step process; First, a thioester between the HECT's active domain and ubiquitin is formed. Second, ubiquitin is transferred from the E3 onto the substrate by forming an isopeptide bond (Figure I-5B). In contrast, RING E3 ligases catalyze the transfer of ubiquitin from the E2 directly to the substrate without covalently binding ubiquitin themselves (Figure I-5B). Interestingly, RING E3 ligases can act as monomers, as dimers (e.g. BRCA1/BARD1), or in multimeric complexes (like Cullin E3 ligases).

The ubiquitylation activity of RING E3 ligases can be regulated on different levels. For example, PTMs on E3 ligases regulate the interactions of the E3 ligases with proteins within E3 complexes, E2s, substrates, and regulatory binding partners. Moreover, PTMs can regulate E3 ligase activity and have an effect on the conformation of the modified E3 ligase. Moreover, autoubiquitylation of E3 ligases has been reported to regulate the ligase activity or lead to proteasomal degradation, posing a self-regulatory mechanism^{114–117,120–123}.

The process of modifying a substrate with one ubiquitin is referred to as monoubiquitylation. The addition of several single ubiquitins to different lysines on one substrate is named multi-monoubiquitylation. Monoubiquitylation usually has an impact on protein-protein interactions. Ubiquitin consists of 76 amino acids, with seven of these being lysines. Each lysine, in addition to the N' term Met, can be ubiquitylated, thereby forming ubiquitin chains. Homotypic chains consist of ubiquitins linked via the same amino acid. Examples are K48- or K63-linked chains. Additionally, in mixed chains different linkages within an ubiquitin chain occur. While in mixed chains, each ubiquitin only is modified with one ubiquitin, ubiquitins in branched chains can bear two or more ubiquitins with the same or different linkages (Figure I-5C). As the modus operandi of ubiquitylation

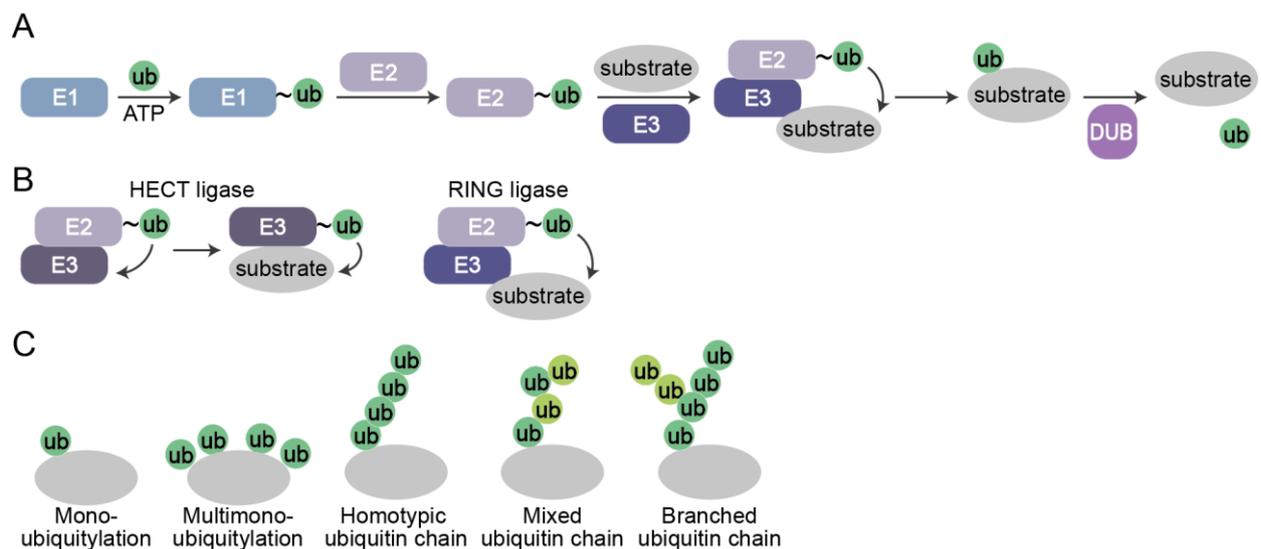


Figure I-5, Ubiquitylation. **A**, Attachment of ubiquitin onto a substrate. By ATP hydrolysis, ubiquitin is activated and forms a thioester with the activating E1 enzyme. In a trans-thioesterification reaction, ubiquitin is transferred from the E1 to the E2 enzyme. The latter brings ubiquitin to the E3 ligase, which catalyzes the transfer of ubiquitin onto a substrate. DUBs are able to remove ubiquitin from substrates. **B**, HECT E3 ligases ubiquitylate substrates in a two-step reaction. First they take over ubiquitin from the E3. In a second step, the ubiquitin, bound at the active site of the HECT ligase, is then passed onto the substrate. RING E3 ligases do not bind to ubiquitin themselves but only catalyze the transfer of ubiquitin from the E2 onto the substrate. **C**, The attachment of single moieties of ubiquitin to substrates is referred to as monoubiquitylation of multimono-ubiquitylation. Polyubiquitin chains can be composed of homotypic

chains with a single linkage. Moreover, there are heterotypic or mixed ubiquitin chains with multiple linkages and branched ubiquitin chains.

is different for HECT and RING E3 ligases, chain formation by the two E3 families also has different modes. As HECT ligases bind the donor ubiquitin directly, they are also responsible for the chain linkage. Opposed to this, the linkage specificity of RING ligases is conferred by their interacting E2 enzymes. Some E2s only transfer a single ubiquitin, mediating (multi-) monoubiquitylation. Other E2s only transfer ubiquitin onto another ubiquitin, thereby mediating chain formation. Lastly, a third group of E2s is capable of mono- and polyubiquitylation. The ubiquitin code is read by proteins harboring ubiquitin-binding domains. Thus, they decipher the ubiquitin signal by specifically recognizing ubiquitin and ubiquitin linkages and transfer it into downstream signals. The recognition of the respective ubiquitin chain has very specific signaling outcomes. K48, K29, and K11 chains as well as K48/K11 heterotypic chains are involved in protein degradation. But also non-proteolytic functions have been attributed to ubiquitin chains. For example, K63 and M1 chains are active in DNA damage and immune response signaling, respectively^{114–117,120,122–125}.

As mentioned above, ubiquitylation is reversible. In human, there are approximately 100 deubiquitylases (DUBs), which have the capacity to completely or partially remove ubiquitin from a substrate. An important function of DUBs is to ensure the availability of monoubiquitin within the cell by protecting ubiquitin from degradation together with ubiquitylated proteins. Notably, trimming ubiquitin chains can be involved in changing cellular signaling. Leaving a single ubiquitin attached or debranching of chains can lead to altered signal outcomes, among other events. Additionally, by cleaving ubiquitin off a substrate, DUBs have the ability to rapidly terminate signaling events^{115,117}.

1.7.1 RNA-binding ubiquitin ligases

As mentioned above, some E3 ubiquitin ligases have been reported to bind to RNA (RNA-binding ubiquitin ligases, RBULs; Table I)^{9,31,32}. The known RBULs harbor different RBDs, such as RRM (CNOT4), KH domains (MEX3 proteins), or WD repeats (PRPF19)^{9,32}. Globally, those proteins have been linked to post-transcriptional regulation and ubiquitylation³². However, the general functional relevance of the RNA-binding capacity of this subclass of E3 ligases is poorly understood and only single RBULs, like MDM2, have been characterized in detail. Another example is MEX3C, which has been found to bind to *HLA-2A* mRNA. By sole RNA-binding, translation of

HLA-2A mRNA is inhibited. In addition, the ubiquitylation of the deadenylase subunit CNOT7 by MEX3C results in *HLA-2A* mRNA degradation, linking ubiquitylation to the regulation of mRNA stability^{32,126}.

1.7.2 Makorin1

Another RBUL is Makorin1 (MKRN1), which is a member of the Makorin ring finger protein family. As these proteins are highly conserved, orthologous proteins can be found in chicken, fruit flies, nematodes, plants, and fungi. In vertebrates, up to four MKRN paralogs have been identified^{127,128}.

Table I, RNA-binding ubiquitin ligases

Gene	RNA binding domain	E3 ligase domain	Role
ARIH2	unkown	RING between RING	DNA damage response
CNOT4	RRM	RING	DNA damage response, Deadenylation
DZIP3	Lys-rich region	RING	Splicing
HUWE1	unknown	HECT	DNA damage response
MDM2	RING ZNF	RING	DNA damage response
MEX3A	KH	RING	RNA processing
MEX3B	KH	RING	RNA processing
MEX3C	KH	RING	mRNA decay, immune response
MEX3D	KH	RING	mRNA decay, apoptosis
MKRN1	CCCH ZNF	RING	DNA damage response
MKRN2	CCCH ZNF	RING	DNA damage response
MKRN3	CCCH ZNF	RING	DNA damage response
PRPF19	WD repeat	RING	DNA damage response, splicing
RBBP6	CCHC ZNF	RING	DNA damage response, splicing
Roquin1	CCCH ZNF	RING	RNA processing, decay
Roquin2	CCCH ZNF	RING	RNA processing
RNF113A	CCCH ZNF	RING	DNA damage response
RNF113B	CCCH ZNF	RING	DNA damage response
RNF17	TUDOR	RING	Not known
RNF31	unknown	RING between RING	Immune response
TRIM25	PRY/SPRY domain	RING	Immune response
TRIM33	unknown	RING	Transcription
TRIM56	unknown	RING	immune response
TRIM71	C-terminal NHL-domain	RING	Proliferation, miRNA binding, translation
unk	CCCH ZNF	RING	Differentiation,

Human *MKRN1* codes for a 482 amino acid protein of 53.3 kDa. Mouse and human *MKRN1* share 92% protein identity. The proteins of the *MKRN* family encode two to four C₃H ZNF domains, which confer RNA-binding capability, a C₃HC₄ RING ZNF and an unusual C₂H₂CH ZNF of unknown function¹²⁷. Additionally, *MKRN1* harbors a PABP-interacting motif 2 (PAM2) that facilitates protein–protein interactions^{44,129,130}. Due to its RING domain, *MKRN1* is assigned to the group of E3 ligases. Indeed, ubiquitylation targets of *MKRN1*, such as p53, p21, p14ARF, PTEN, and FADD, have been previously identified. The subsequent degradation of those substrates link *MKRN1* to the regulation of cell growth, cellular senescence, and apoptosis, as well as to tumor development^{131–134}. Moreover, a role for *MKRN1* in the regulation of telomere length and RNA Pol II transcription has been suggested^{135–137}. While several studies have shed light on putative ubiquitylation targets of *MKRN1*, its ability to bind to RNA remains inadequately understood. In mouse embryonic stem cells (mESCs), *MKRN1* binds to hundreds of RNAs and interacts with several RBPs, such as YBX1, IGF2BP1, ELAVL1, and PABPC1, linking this RBUL to post-transcriptional and translational regulation¹³⁸. Importantly, the interaction of a short *MKRN1* variant with PABPC1 and IGF2BP1 has been confirmed in human HEK293 cells. Moreover, in rat neurons, *MKRN1* co-localizes with PABP and together, these proteins are involved in the regulation of local translation¹²⁹. In line with this, *MKRN1* associates with ribosomal proteins in mESCs in a RNA-dependent manner⁴⁷. Physiologically, *MKRN* proteins seem to play a role in germ cell formation, embryonic development, and neurogenesis. In mice, *MKRN1* is highly expressed in early embryos and in the developing central nervous system^{127,139}. Accordingly, early human and zebrafish embryos express high *MKRN1* and *mkrn4* transcript levels, respectively^{128,139}. In adult tissues and gonads, *MKRN1* is ubiquitously expressed at low to medium levels^{127,140}. Furthermore, a role of *MKRN* proteins in oogenesis and spermatogenesis is implicated by the high *mkrn1*, *mkrn2*, and *mkrn4* transcript levels in ovaries of zebrafish, clawed frogs, and medaka¹⁴⁰. Due to the strong conservation of this protein family, a general vital cellular function has been suggested for *MKRN1*¹²⁷. This role, however, still needs to be elucidated.

2. Relevance in disease

2.1 Errors in gene expression can lead to disease

Errors in gene expression can result in diverse pathologies, such as cardiovascular diseases, neurodegeneration, and cancer. Premature polyadenylation, for example, has been reported to frequently occur in tumor suppressor genes, resulting in truncated proteins. In chronic lymphocytic leukemia, those proteins either lose their cancer-constraining function or gain opposing capacities¹⁴¹. When the clearance of aberrant mRNAs cannot be ensured, defects in mRNA surveillance generally entail the accumulation of faulty proteins that harbor mutations or do not fold correctly and hence, aggregate^{57,118,142,143}. Furthermore, upon identification by mRNA surveillance pathways, nonstop mRNAs are usually degraded by the exosome. However, they accumulate in case of nonfunctional exosomal decay. Additionally, nonstop mutations resulting in reduced transcript and thus protein levels cause different pathologies. This is the case for stop codon-mutated *DEFB126* (reduced fertility) or *DYSF* (muscular dystrophy), for instance¹⁴⁴. Moreover, translational inaccuracy or mutations in translation machinery factors or in their regulatory elements have been linked to such as neurodegenerative diseases and developmental disorders^{145–147}.

2.2 Misregulation of ubiquitylation results in disease

Deleterious aggregated proteins are usually targeted by E3 ubiquitin ligases for proteasomal degradation. In cystic fibrosis and lysosomal storage disorder, it can be observed that the pool of functional, properly folded proteins is depleted due to aggregation and enhanced decay of proteins^{60,148}. Consequently, it is vital for the cell to only remove faulty proteins with the help of protein QC pathways¹⁴³.

However, when QC pathways are dysregulated, nonfunctional proteins accumulate and lead to neurodegenerative diseases, cardiovascular diseases, metabolic disorders, or cancers^{60,148}. Moreover, the shortage of QC machinery components can be embryonically lethal, while mutations in those constituents are again involved in pathologies like neurodegeneration, immune system function, cancers, and developmental defects^{60,118,149}. The ubiquitin system is a major component of QC surveillance systems. As ubiquitin is involved in all cellular pathways, it is not surprising that the depletion of ubiquitin encoding genes, e.g. *Ubb*, has negative effects on embryonic neuronal and on hepatocyte differentiation and results in neurodegeneration

in adulthood^{118,150,151}. In neurons, the ubiquitin system is involved in synaptic plasticity and in ensuring the general well-being of the cell. Thus, disturbed ubiquitylation resulting in deviant translation, trafficking, and protein interactions, for instance, is prevalent in neurodegenerative disorders like Alzheimer's disease, Parkinson's disease, and Amyotrophic lateral sclerosis¹⁵². Along these lines, neurodegeneration and birth defects are also found in developmental processes that are affected by malfunctioning E3 ubiquitin ligases and DUBs. Moreover, E3 ligases play a role in cancer development. For example, mutations in *BRCA1* are linked to increased risks for ovarian and breast cancer¹¹⁸. In conclusion, functional proteins, high translational fidelity, protein QC pathways and correct ubiquitylation are indispensable for cell survival, embryonic development and health in general.

3. Profiling the Binding Sites of RNA-binding Proteins with Nucleotide Resolution using iCLIP

Methods Mol Biol. 2016;1358:175-95. doi: 10.1007/978-1-4939-3067-8_11.

Profiling the Binding Sites of RNA-binding Proteins with Nucleotide Resolution using iCLIP

FX Reymond Sutandy¹, Andrea Hildebrandt¹, and Julian König^{1,§}

¹Institute of Molecular Biology gGmbH, Ackermannweg 4, 55128, Mainz, Germany.

§Address correspondence to: Julian König, j.koenig@imb-mainz.de

Summary

The importance of posttranscriptional regulation in cellular metabolism has recently gone beyond what was previously appreciated. The regulatory mechanisms are controlled by RNA-binding proteins (RBPs), which form complexes with RNA and regulate RNA processing, stability and localization, among others. Consistently, mutations in RBPs result in defects in developmental processes, diseases and cancer. Gaining deeper insights into the biology of RNA-RBP interactions will lead to a better understanding of regulatory processes and disease development. Several techniques have been developed to capture the properties of RNA-RBP interactions. Furthermore, the development of high-throughput sequencing has broadened the capability of these methods. Here, we summarize individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP), a powerful technique that provides genome-wide information on RNA-RBP interactions at nucleotide resolution. In this chapter, we outline the iCLIP protocol and list possible controls that allow a targeted and cost-minimizing optimization of the protocol for an RBP-of-interest. Moreover, we provide notes on experimental design and a troubleshooting guideline for common problems that can occur during iCLIP library preparation.

Keywords: RNA, RNA-binding protein, iCLIP, RBP binding site, protein-RNA interactions

1. Introduction

RNA processing steps, such as splicing and polyadenylation, are major regulatory mechanisms in posttranscriptional gene regulation. Other points of action include RNA stability and localization by which the cell can coordinate RNA turnover and local protein synthesis. During all these processes, RNA-binding proteins (RBPs) act as the minders and controllers.

RNA-protein interactions have been subject to basic research for more than 40 years and are still investigated to date (1, 2). As early as 1979, Jay R. Greenberg found that crosslinking of proteins to RNA is feasible at a wavelength of 254 nm. This led to the conclusion that RNA and proteins can be found in close proximity in the cell (3). Many RBPs were originally described by their conserved RNA binding domains, which exist in single or multiple copies and enable the RBPs to associate with RNA in a sequence- or structure-dependent manner (4). More recently, approaches using protein microarrays and proteomics identified several hundred novel RBPs in yeast (5, 6), and further screens for human RBPs based on oligo(dT)-pulldowns coupled to mass spectrometry added hundreds of novel RBPs to the list (7–9). Notably, in these screens, an RNA binding function was often assigned to proteins that had previously only been known to be involved in other cellular processes, such as metabolic functions.

Following from their abundance and widespread function, mutations in RBPs can disturb fundamental biological processes and are often linked to human diseases and cancer (10). For example, loss of function of the fragile X mental retardation protein (FMRP) leads to the fragile X syndrome, a form of mental retardation (11). Therefore, identifying RNA targets and especially disease-associated RBP-RNA interactions will be beneficial to understand the molecular mechanisms of disease and to ultimately develop treatments.

In order to identify the RNA targets of an RBP-of-interest, RNA can be crosslinked to directly interacting proteins *in vivo* using UV irradiation (3). Crosslinking is vital, since earlier approaches have shown that RNA-RBP complexes can re-associate during affinity purification (12). The method UV crosslinking and immunoprecipitation (CLIP) exploits UV crosslinking to obtain covalently bound RBP-RNA complexes, which are immunoprecipitated with specific antibodies against the RBP-of-interest (13, 14). It is then possible to extract the interacting RNAs in subsequent steps. In a related method called PAR-CLIP, the crosslinking is achieved through the use of photoreactive nucleosides, such as 4-thiouridine, which can be specifically crosslinked at 365 nm (15).

Coupled with high-throughput sequencing (HITS-CLIP or CLIP-Seq) (16), CLIP allows to map RNA-RBP interactions on a genomic scale, which can be used for instance to identify target sequences (17) or to compute RNA-maps (16).

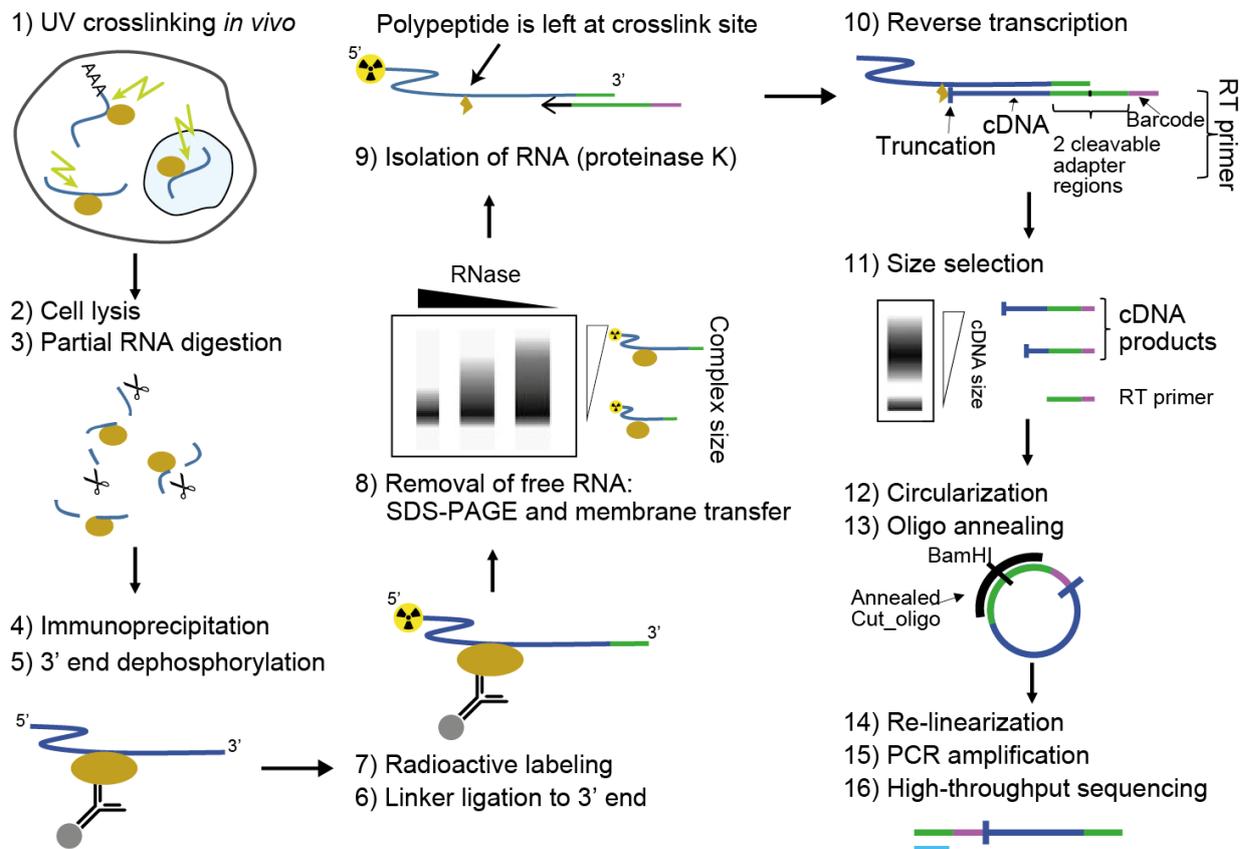


Figure I-3-1, Schematic overview of the iCLIP protocol. The experiment starts with UV irradiation of the samples, followed by cell lysis and partial RNase digestion. The specific protein–RNA complexes are then immunoprecipitated with a specific antibody, and an adaptor is ligated to the co-purified RNAs. The complexes are subsequently run on an SDS-PAGE and transferred to a nitrocellulose membrane, which will remove most unspecifically bound RNAs. Proteinase K digestion is applied to extract the cross-linked RNAs, which are then subjected to reverse transcription. The resulting cDNAs are size-selected and circularized. The following linearization step provides the basis for amplification of the iCLIP library, which is then ready for high-throughput sequencing. Figure adapted Sutandy, Hildebrandt, and König, 2016.

We recently introduced individual-nucleotide resolution CLIP (iCLIP), which allows to quantitatively map RNA-RBP interaction sites with high resolution and efficiency (18). Basis to the iCLIP procedure is the fact that during cDNA synthesis from the co-purified RNAs, the reverse transcriptase often truncates at the crosslink site, at which a small residual of the RBP is left after protein digestion (19). The iCLIP library preparation protocol allows to capture these truncated cDNAs, which provide a footprint of the RBP binding site (20). More precisely, the cDNA sequences start one nucleotide downstream of the crosslink sites, thereby providing nucleotide-resolution information on RBP binding. To date, iCLIP has been widely applied to investigate diverse functions of

RBPs in RNA metabolism, such as regulating alternative splicing and RNA stability (21, 22), mediating RNA modifications (23), or binding small RNAs (24, 25), among others.

In this chapter, we introduce the iCLIP protocol, including a description of essential controls and guidelines for troubleshooting. The iCLIP experiment starts with UV irradiation of the sample (**Fig. 1**), which introduces a covalent bond between the RBP and the interacting RNAs. The cells are then lysed, and a partial RNase digestion is applied to cut the interacting RNAs down to a specific size range, which is optimal for library preparation. An example of isolated RNAs under different RNase conditions is shown in **Fig. 2B**. This is followed by an immunoprecipitation step to capture the RBP-of-interest in complex with the crosslinked RNA fragments. Due to the covalent bond between the RBP and the RNA fragments, a stringent purification and washing scheme can be applied to remove unspecific interactions. Next, the 3' ends of the RNA fragments are dephosphorylated and then ligated to a DNA linker, while the 5' ends are radioactively labeled to enable their visualization on an autoradiograph upon SDS-PAGE and transfer to a nitrocellulose membrane (**Fig. 2A**). The latter is critical to further remove non-crosslinked RNAs from the sample. The crosslinked RNAs are then extracted from the membrane by proteinase K digestion and subjected to reverse transcription (RT) with RT primers complementary to the DNA linker sequence (**Fig. 2C**). During this step, most cDNAs will truncate at the crosslink site, where a small peptide residue of the RBP remains after proteinase K digestion. Importantly, the RT primers include a random barcode sequence, which facilitates removal of PCR duplicates during computational data analysis, allowing to ultimately keep only unique RT products. Gel purification and size selection of the cDNAs remove residual RT primers. **Fig. 3** provides a mask for selecting cDNAs in the optimal size range from these gels. Upon cDNA circularization, an oligonucleotide is annealed to the linker region to create a doublestranded cutting site for a restriction enzyme. This re-linearization renders cDNAs that are flanked on either side by adapter regions for PCR amplification. After another round of size selection to remove excess PCR primers, the iCLIP library is ready for high-throughput sequencing.

Positive and negative controls are critical during the iCLIP library preparation. **Table 1** summarizes important controls that should be used, especially when optimizing the procedure. Since iCLIP is an extensive protocol, these controls serve as early indicators for problems with library quality, which can reduce working time and costs and decrease

the risk of producing suboptimal iCLIP libraries. **Table 2** lists troubleshooting strategies for common problems in iCLIP library preparation.

2. Materials

Table I-3-1, Possible controls during iCLIP library preparation

Type of control	Purpose
Negative control without UV crosslink	Confirm that the radioactive signal comes from crosslinked molecules. This is to exclude that e.g. the protein itself can be radioactively labeled through autophosphorylation or a contaminating protein kinase.
Negative control without antibody	Detect possible contaminations. Unspecific proteins binding to the beads would still give a signal in this control. To prevent this, washing steps could be carried out more stringently.
Negative control upon knock-down of the target protein	Evaluate the specificity of the antibody used in the immunoprecipitation. A specific antibody would produce weaker or no protein-RNA signals in the autoradiograph image.
Positive control with antibody against a different protein	Crosslinking and IP efficiency are protein-dependent. This control ensures that all steps of the protocols are working and can help to estimate whether the amount of crosslinked RNA is sufficient to successfully prepare an iCLIP library for the protein-of-interest. Best for this control are antibodies against proteins, which were successfully used in previous iCLIP experiments.
Visualization of RNA sizes using different RNase conditions	The RNase concentration has to be optimized to gain suitable RNA sizes. In order to determine the RNA size range, the RNA fragments resulting from different RNase concentrations can be visualized on a gel after their isolation from the membrane using proteinase K digestion (Fig. 2B).

2.1 Buffers

Store all buffers in the fridge and keep them on ice during experiments.

1. Lysis buffer: 50 mM Tris-HCl, pH 7.4, 100 mM NaCl, 1 % Igepal CA-630 (Sigma), 0.1 % SDS, 0.5 % sodium deoxycholate. On the day of the experiment, add 1/100 volume of protease inhibitor cocktail III (Calbiochem/Merck) to the amount of buffer required for lysis (but not washing). See **Note 1**
2. High-salt wash buffer: 50 mM Tris-HCl, pH 7.4, 1 M NaCl, 1 mM EDTA, 1 % Igepal CA-630, 0.1 % SDS, 0.5 % sodium deoxycholate.
3. PNK buffer: 20 mM Tris-HCl, pH 7.4, 10 mM MgCl₂, 0.2 % Tween-20.
4. 5x PNK pH 6.5 buffer: 350 mM Tris-HCl, pH 6.5, 50 mM MgCl₂, 5 mM dithiothreitol. Freeze aliquots of the buffer, do not thaw and freeze again.
5. 4x ligation buffer: 200 mM Tris-HCl, pH 7.8, 40 mM MgCl₂, 4 mM dithiothreitol. Freeze aliquots of the buffer, do not thaw and freeze again.

6. PK buffer: 100 mM Tris-HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA.
7. PK buffer/7 M urea: 100 mM Tris-HCl, pH 7.4, 50 mM NaCl, 10 mM EDTA, 7 M urea.
- 8.

2.2 Linker and primer sequences

1. pre-adenylated linker L3-App, IDT, rApp/AGATCGGAAGAGCGGTTCAG/ddC/
2. The linker is modified with adenylation at the 5' end (rApp) and dideoxycytidine (ddC) at the 3' end.
3. Each RT primer (Rt#clip) has a different experimental barcode, which provides the possibility for sample multiplexing during high-throughput sequencing. In addition, all Rt#clip primers contain a random barcode sequence (NNxxxxNNN, where N is any base and x the bases of the defined experimental barcode) for the purpose of duplicate removal during sequencing data analysis. Here are examples of some Rt#clip primers:

Rt1clip X33/NNAACCN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt2clip X33/NNACAANN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt3clip X33/NNATTGN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt4clip X33/NNCGCCNN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt5clip X33/NNGCCANN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt6clip X33/NNGACTNN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt7clip X33/NNGTGGNN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt8clip X33/NNATTNN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC
Rt9clip X33/NNTTAANN NNAGATCGGAAGAGCGTCGTGgatcCTGAACCGC

X33 = 5' phosphate (see **Note 2**)

1. Cut_oligo, IDT, GTTCAGGATCCACGACGCTCTTCaaaa
2. For PCR amplification and high-throughput sequencing, we used the official Illumina P3 and P5 primers.

P5Solexa: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACAC
GACGCTCTTCCGATCT

P3Solexa: CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCTGC
TGAACCGCTCTTCCGATCT

3. Methods

All steps in this protocol should be performed on ice unless stated otherwise.

3.1 UV crosslinking

The following iCLIP protocol uses cell culture samples.

1. Remove media and add 6 ml cold PBS to cells growing in a 10 cm plate, remove lid and place on ice (usually enough for 3 immunoprecipitations).
2. Irradiate once with 150 mJ/cm² in a Stratalinker 2400 at 254 nm (see **Note 3**).
3. Harvest cells by scraping.
4. Add 2 ml suspension to each microtube (RNase-free, non-stick), spin at 0.4 g for 1 min at 4 °C, and remove supernatant.
5. Snap freeze pellets on dry ice and store at -80 °C until use (or use directly).

3.2 Bead preparation

See **Notes 4 & 5**

1. Add 50 - 100 µl of protein G Dynabeads (Life Technologies) per experiment to a microtube (see Note 6).
2. Wash beads 2x with lysis buffer.
3. Resuspend beads in 100 µl lysis buffer with 2 - 10 µg antibody per experiment (see **Note 7**).
4. Rotate tubes at room temperature for 30 - 60 min (until lysate is ready).
5. Wash 1x with high-salt wash buffer.
6. Wash 2x with lysis buffer and leave in 400 µl lysis buffer with added protease ^ inhibitors (see section 2.1) until ready to proceed.

3.3 Pellet resuspension

When ready to proceed, resuspend the cell pellets (from section 3.1) in 1 ml lysis buffer. This should result in a protein concentration of ~ 2 mg/ml (see **Note 8**). Do not forget to add the proteinase inhibitor to the lysis buffer before use (see section 2.1).

3.4 Sonication of samples (optional step)

See **Notes 9 & 10**

Variant A: Sonicate sample on ice. The probe should be approximately 0.5 cm from the bottom of the tube and not touching the tube sides in order to avoid foaming. Sonicate 2x with 10 s bursts at 5 decibels. Clean the probe by sonicating water before and after sample treatment.

Variant B: Transfer sample to 1.5 ml microtubes and use Bioruptor for 5 cycles with alternating 30 sec on/off at low intensity (needs to be switched on in advance to allow for cooling down of the water bath). Six samples can be sonicated at the same time.

Table I-3-2, Troubleshooting for common problems during iCLIP library preparation.

Step	Problem	Possible reason	Solution
Autoradiograph imaging (section 3.10)	No/faint band	Failure in immunoprecipitation	Optimize the antibody used in the immunoprecipitation step.
		Poor crosslinking efficiency	Low crosslinking efficiency would cause huge loss of RNA-protein complexes, which leads to poorly detectable bound RNAs in the autoradiograph. Since the crosslinking efficiency is protein-dependent, you could either increase irradiation time or try 4-thiouridine crosslinking at 365 nm.
	No/little smear	Too much RNase digestion	Adjust RNase concentration.
	Strong band in the no-UV crosslinking control	The protein itself is radioactively labeled (see also Table 1)	In principle, this does not interfere with the isolation of bound RNA fragments. However, the presence of this signal has to be considered when interpreting the signal strength of the crosslinked samples.
	Multiple bands	Protein contaminations	Apply more washing steps.
		Protein dimerization	Some proteins might form a dimer, either by themselves or as a result of the purification tag that is fused to the proteins. In this case, RNA from all bands should be extracted.
		Two proteins are crosslinked in close proximity on the same RNA	This phenomenon could result in a 'supershift' at the added size of both proteins. In this case, RNA from all bands should be extracted.
RNA precipitation (section 3.11)	No RNA pellet	Some residues, such as high salt, gel or phenol	Emergency action: Do not take off the supernatant, but split the sample into two tubes. Add 200 µl TE buffer followed by 20 µl sodium acetate and 500 µl 100% ethanol to each tube to further dilute the sample. Incubate overnight at -20 °C and spin the samples again on the next day.
Post-PCR (section 3.16)	Unspecific bands above the expected size	Overamplification	Secondary PCR products can be removed by reducing the number of amplification cycles. The post-PCR gel size selection also helps to remove undesired PCR products before high-throughput sequencing.
	Broader band distribution	Suboptimal cDNA size selection	Optimize the cDNA gel size selection step (section 3.13) or perform post-PCR gel size selection.

3.5 Partial RNase digestion

1. Make 1/500 RNase I (Ambion) dilution in lysis buffer and add 10 μ l to the lysate together with 2 μ l Turbo DNase (Ambion) (see **Note 11**).

High-RNase control (**optional**, recommended for initial optimizations): Treat one sample with high RNase: prepare a 1/50 RNase I dilution in lysis buffer and add 10 μ l to the lysate together with 2 μ l Turbo DNase. To minimize the use of reagents, it is possible to use only 1/5 of the cell lysate and all other reagents for this experiment (see **Notes 12 & 13**). Since this RNA is too short for DNA linker ligation, this control can skip 3' end dephosphorylation and linker ligation and go straight from section 3.6 to section 3.9.

2. Incubate for 3 min at 37 °C shaking at 1100 rpm in thermomixer (Eppendorf). After incubation transfer to ice for >3 min (see **Note 14**).
3. Spin at 4 °C at top speed for 10 min and transfer the supernatant to a new 1.5 ml microtube.
4. Load 500 μ l of the lysate onto a Proteus Clarification Mini Spin Column (Generon). Spin at 4 °C at 16000 x g for 1 min. Transfer flow-through to a new microtube. Repeat with second half of the lysate and combine both (see **Note 15**).

3.6 Immunoprecipitation

1. Add the cell extract to the beads.
2. Rotate bead/lysate mix for 1 h (max. 2 h) at 4 °C (see **Notes 16 & 17**).
3. Discard the supernatant and wash 2x with high-salt wash buffer (rotate the second wash for at least 1 min in the cold room).
4. Wash 2x with PNK buffer and then resuspend in 1 ml PNK buffer (samples can be left like this at 4 °C until you are ready to proceed to the next step).

3.7 RNA 3' end dephosphorylation

See **Note 18**

1. Discard supernatant. Resuspend the beads in 20 μ l of the following mixture:
4 μ l 5x PNK pH 6.5 buffer, 0.5 μ l PNK (with 3' phosphatase activity, NEB), 0.5 μ l RNasin (Promega), 15 μ l water.
2. Incubate for 20 min at 37 °C in a thermomixer at 1100 rpm.
3. Wash 1x with PNK buffer.
4. Wash 1x with high-salt wash buffer (rotate wash for at least 1 min in cold room).
5. Wash 2x with PNK buffer.

3.8 L3 linker ligation

- Carefully remove the supernatant and resuspend the beads in 20 μ l of the following mix:
8 μ l water, 5 μ l 4x ligation buffer, 1 μ l T4 RNA ligase (NEB), 0.5 μ l RNasin, 1.5 μ l pre-adenylated linker L3-App (20 μ M), 4 μ l PEG400 (Sigma).
- Incubate overnight at 16 $^{\circ}$ C in a thermomixer at 1100 rpm.
- Add 500 μ l PNK buffer.
- Wash 2x with 1 ml high-salt wash buffer, rotating the wash for 5 min in the cold room.
- Wash 2x with 1 ml PNK buffer. Transfer to new microtubes after the first wash and leave in 1 ml of the second wash.

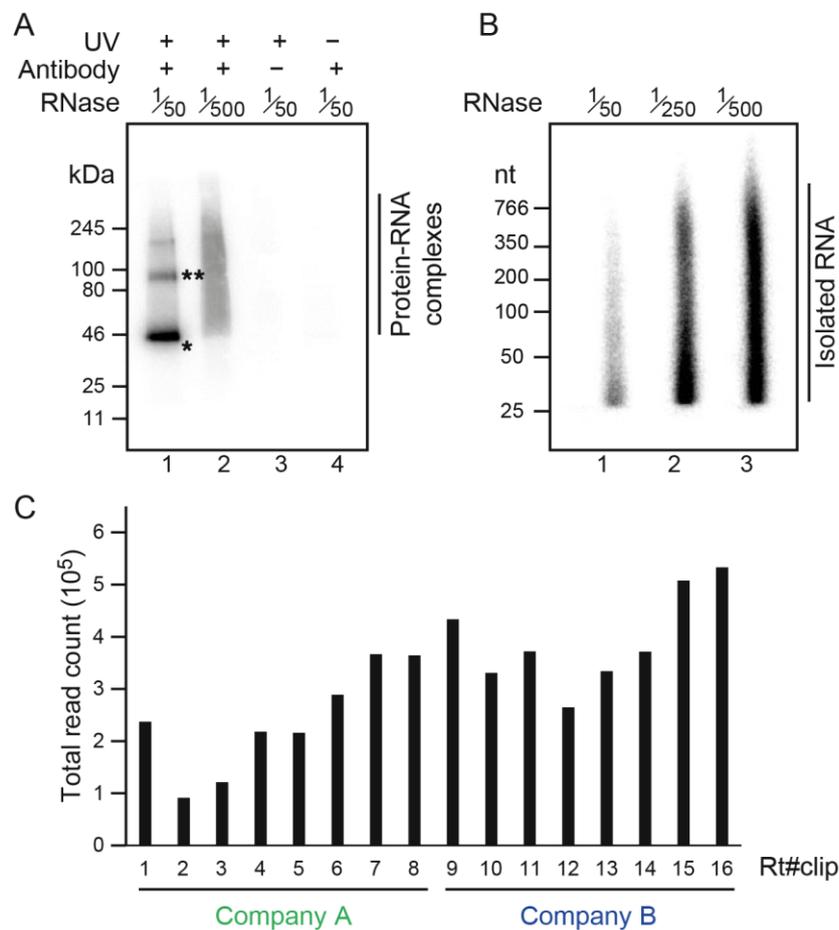


Figure I-3-2, Visualization of RNA-RBP complexes and isolated RNAs. (a) Radioactively labeled RNA-RBP complexes upon membrane transfer. Lane 1, high RNase concentration resulting in short RNA fragments on the RBP. The radioactive signal appears immediately upstream of the estimated size of the RBP (labeled by *). **Higher molecular weight band, indicating two proteins cross-linked to the same RNA molecule. Lane 2, low RNase concentration producing a smear of RNA-RBP complexes with RNA fragments of varying sizes. Lane 3, noantibody control. Lane 4, no-UV cross-linking control. (b) Gel separation of RNA fragments that were isolated from the nitrocellulose membrane after digestion with different RNase concentrations. In this experiment, the 1:500 dilution shows the highest amount of RNAs in the optimal size range of 30–150 nt. (c) Comparison of read numbers obtained from the same sample that was processed with Rt#clip primers with different barcodes and synthesized by two different companies. The variability in library preparation efficiency needs to be considered when comparing between samples. Figure adapted Sutandy, Hildebrandt, and König 2016.

3.9 5' end labeling

1. Collect 200 μl (20 %) of beads from the previous step (section 3.8) and remove the supernatant.
2. Add 4 μl of hot PNK mix:
0.2 μl PNK, 0.4 μl ^{32}P - γ -ATP (Perkin Elmer), 0.4 μl 10x PNK buffer, 3 μl water.
3. Incubate for 5 min at 37 °C in a thermomixer at 1100 rpm.
4. Remove the supernatant and add 20 μl of 1x NuPAGE loading buffer prepared by mixing 4x stock with water). Incubate at 70 °C for 5 min in a thermomixer at 1100 rpm.

Optional: Use reducing agent and antioxidant to avoid potential interference of antibodies with the protein-RNA complexes during SDS-PAGE and nitrocellulose transfer.

6. Place on a magnet to precipitate the beads, transfer the supernatant to a new microtube, place it again on a magnet and load the supernatant on the gel.

3.10 SDS-PAGE and nitrocellulose transfer

1. Load the samples on a 4 - 12 % NuPAGE Bis-Tris gel according to the manufacturer's instructions (see **Note 19**). Use 0.5 l 1x NuPAGE MOPS SDS running buffer. Also load 5 μl of a pre-stained protein size marker.
2. Run the gel for 50 min at 180 V.
3. Remove the dye front and discard it as solid radioactive waste (it contains the radioactive ATP).
4. Transfer the protein-RNA complexes from the gel to a Protan BA85 nitrocellulose membrane using the Western Blot wet transfer apparatus according to the manufacturer's instructions (transfer for 1 h at 30 V; do not forget to add 10 % methanol to the transfer buffer) (see **Note 20**).
5. After the transfer, rinse the membrane in PBS buffer, then wrap it in Saran Wrap and expose it to a film at 4 °C for ~1 h (place a fluorescent sticker next to the membrane to later align the film and the membrane).
6. Visualize the film on a phosphorimager (see **Note 21**).

3.11 RNA isolation

1. Use the high-RNase condition to examine the specificity of the protein-RNA complexes (see **Note 22**).
2. Isolate the protein-RNA complexes from the low-RNase experiment using your autoradiograph as a mask for cutting the respective region out of the nitrocellulose membrane. Place the membrane fragments into 1.5 ml microtubes. If a piece of membrane is too large to fit down to the bottom of the tube, cut it into several pieces before placing it into the microtube.

Optional: Re-expose the membrane after excising the bands to confirm accuracy of cutting.

3. Add 10 μ l proteinase K (Roche) in 200 μ l PK buffer to the nitrocellulose pieces (all should be submerged). Incubate in a thermomixer at 1100 rpm for 20 min at 37 °C.
4. Add 200 μ l of PK buffer/7 M urea and incubate for further 20 min at 37 °C and 1100 rpm.
5. Collect the solution and add it together with 400 μ l phenol/chloroform to a 2 ml Phase Lock Gel Heavy tube (VWR) (see **Note 23**).
6. Incubate for 5 min at 30 °C shaking at 1100 rpm (do not vortex). Separate the phases by spinning for 5 min at 16000 x g at room temperature.
7. Transfer the aqueous layer into a new microtube (be careful not to touch the gel matrix with the pipette).

Optional: Spin again for 1 min and transfer into a new microtube.

8. Precipitate by addition of 0.75 μ l GlycoBlue (Life Technologies) and 40 μ l 3 M sodium acetate pH 5.5. Then mix and add 1 ml 100 % ethanol, mix again and place over night at -20 °C (see **Note 24**).
9. Spin for 20 min at 21.000 x g at 4 °C. Remove the supernatant, wash the pellet with 0.9 ml 80 % ethanol and spin again for 5 min. Resuspend the pellet in 5 μ l water and transfer to a PCR tube (see **Note 25**).

3.12 Reverse transcription

1. Add the following reagents to the resuspended pellet (see **Notes 26 & 27**):
1 μ l Rt#clip (0.5 pmol/ μ l), 1 μ l dNTP mix (10 mM).
2. Denature by running the following thermal program:
70 °C for 5 min, then hold at 25 °C until the RT mix is added (see below).
3. Add 13 μ l RT mix per tube:

7 μ l H₂O, 4 μ l 5x RT buffer (Life Technologies), 1 μ l 0.1 M DTT, 0.5 μ l RNasin, 0.5 μ l Superscript III (Life Technologies).

Mix by pipetting.

Perform the reverse transcription with the following thermal program:

25 °C for 5 min, 42 °C for 20 min, 50 °C for 40 min, 80 °C for 5 min, then hold at 4 °C.

4. Add 1.65 μ l 1 M NaOH and incubate at 98 °C for 20 min.
5. Add 20 μ l 1 M HEPES pH 7.3 (see **Note 28**).
6. Add 350 μ l TE buffer, 0.75 μ l GlycoBlue and 40 μ l 3 M sodium acetate pH 5.5. Mix, then add 1 ml 100 % ethanol. Mix again and precipitate overnight at -20 °C.

3.13 Gel purification

1. Spin down for 15 min at 21000 x *g* at 4 °C. Remove the supernatant and wash the pellet with 0.5 ml 80 % ethanol. Spin down again, remove the supernatant and resuspend the pellet in 6 μ l water.
2. Add 6 μ l 2x TBE-urea loading buffer to the cDNA. It is recommended, at least in initial experiments, to add loading buffer also to 6 μ l DNA size marker (dilution 1/30). Heat samples to 80 °C for 5 min directly before loading.
3. Prepare 0.8 l 1x TBE running buffer and fill the upper chamber with 0.2 l and the lower chamber with 0.6 l. Use a p1000 tip to flush precipitated urea out of the wells before loading 12 μ l of each sample. Leave one lane after each sample to facilitate cutting. Load the marker into the last lane.
4. Run the 6 % TBE-urea gel (Life Technologies) for 40 min at 180 V until the lower (dark blue) dye is close to the bottom.
5. Cut off the last lane containing the size marker and stain it by incubation for 10 min shaking in 10 ml TBE buffer with 2 μ l SYBR Green II (Life Technologies) stock. Wash 1x with TBE and visualize by UV transillumination. Produce a mask to guide band excision by printing the marker image scaled to 100 % (**Fig. 3**).
6. Together with the L3-App linker, the primer sequence accounts for 52 nt of the cDNA. The upper (light blue) dye runs at \pm 110 - 130 nt, and the lowest rim of the plastic gel cassette is at \pm 50 nt; these marks can be used to guide excision together with the size marker. Cut three bands at 70 - 80 nt, 80 - 150 nt and 150 - 200 nt. Use the schematic given in **Fig. 3** as template where to cut the bands. Place each gel piece into a 1.5 ml microtube (see **Note 29**).
7. Add 400 μ l TE and crush the gel piece into small pieces with a 1 ml syringe plunger,

OR

Prepare 0.5 ml microtubes by piercing a hole in the bottom using a 21G needle. Place a gel fragment inside and then place the microtubes into a 2 ml collection tube. Spin at 16000 x *g* for 2 min.

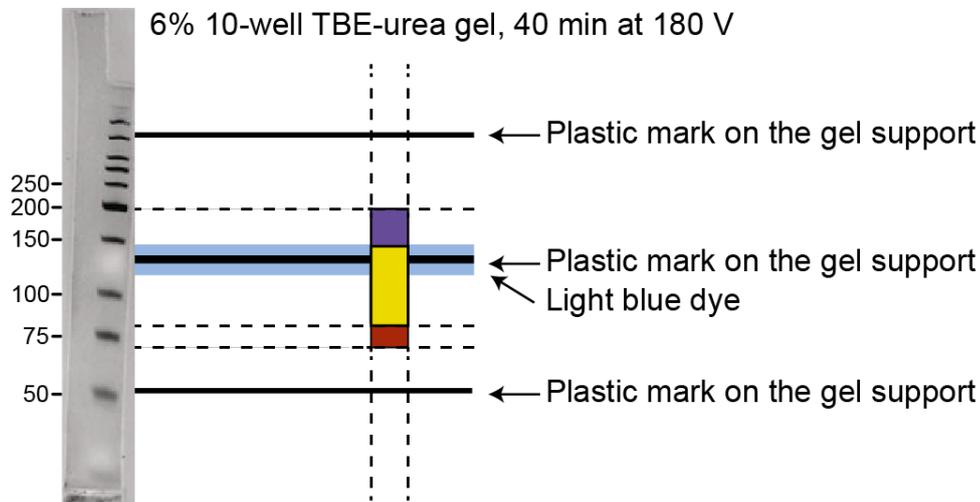


Figure I-3-3, Gel cutting mask for cDNA size selection. This mask can be used to align the cDNA gel via the blue dye as well as the rims on the plastic gel support on the 6 % TBE-urea pre-cast gel. The size marker should be used as an indicator that the gel has indeed run properly. The mask indicates three areas for cutting: Red, cDNA size $\pm 70\text{--}80$, insert size $\pm 18\text{--}28$: This band contains high cDNA complexity (i.e., high concentration of cDNA), but due to the short insert sizes, not all cDNAs can be mapped as unique hits to the genome. Also, this band has a tendency to isolate the primer contaminant. We recommend to use this band only when binding to short RNAs, such as miRNAs, is of interest. Green, cDNA size $\pm 80\text{--}150$ nt, insert size $\pm 28\text{--}98$: This band contains the best cDNAs that are long enough to map as unique hits to the genome. It should not have the primer contaminant and the complexity is generally high. Violet, DNA size $\pm 150\text{--}200$, insert size $\pm 98\text{--}148$: These cDNAs will all map to the genome, but cDNA complexity is usually limited. We nevertheless recommend to include this band initially for testing. Figure adapted from Sutandy, Hildebrandt, and König, 2016.

8. Incubate in a thermomixer at 1100 rpm for 1 h at 37 °C, then place on dry ice for 2 min, and place back at 1100 rpm for 1 h at 37 °C. Transfer the liquid portion of the supernatant into a Costar SpinX column (Corning Incorporated), into which you have placed two 1 cm glass pre-filters (Whatman).
9. Spin at 16000 x *g* for 1 min at room temperature. Collect the solution and add it together with 400 μ l phenol/chloroform into a 2 ml Phase Lock Gel Heavy tube.
10. Incubate for 5 min at 30 °C shaking at 1100 rpm (do not vortex). Separate the phases by spinning for 5 min at 16000 x *g* at room temperature.
11. Transfer the aqueous phase into a new microtube (be careful not to touch the gel matrix with the pipette). Spin again for 1 min and transfer into a new microtube.
12. Add 1 μ l GlycoBlue and 40 μ l 3 M sodium acetate, pH 5.5. Mix, then add 1 ml 100 % ethanol. Mix again and precipitate overnight at -20 °C.

3.14 Circularization

1. Spin down and wash with 80 % ethanol as described above and resuspend in 8 μ l ligation mix:
6.5 μ l water, 0.8 μ l 10x CircLigase Buffer II (Epicentre), 0.4 μ l 50 mM $MnCl_2$, 0.3 μ l CircLigase II (Epicentre). Transfer to PCR tubes and incubate for 1 h at 60 °C.
2. Add 30 μ l oligo annealing mix:
26 μ l H_2O , 3 μ l FastDigest Buffer (Fermentas), 1 μ l 10 μ M Cut_oligo.
3. Anneal the oligonucleotide with the following thermal program:
95 °C for 2 min; then successive cycles of 20 s, starting from 95 °C and decreasing the temperature by 1 °C each cycle down to 25 °C; then hold at 25 °C.
4. Add 2 μ l FastDigest BamHI (Fermentas) and incubate for 30 min at 37 °C, then incubate at 80 °C for 5 min.
5. Add 350 μ l TE, 0.75 μ l GlycoBlue, 40 μ l 3M sodium acetate, pH 5.5, and mix. Then add 1 ml 100 % ethanol. Mix again and precipitate overnight at -20 °C.

3.15 PCR amplification

Spin down and wash the cDNA with 80 % ethanol as described above, then resuspend it in 21 μ l water.

3.15.1 Optimize PCR amplification

See **Note 30**

1. Prepare the following PCR mix:
1 μ l cDNA, 0.25 μ l primer mix of P5Solexa and P3Solexa (10 μ M each), 5 μ l Accuprime Supermix 1 enzyme (Life Technologies), 3.75 μ l H_2O .
2. Run the following PCR:
94 °C for 2 min,
followed by as less as possible cycles of: (see **Notes 31 & 32**)
94 °C for 15 s, 65 °C for 30 s, 68 °C for 30 s, 68 °C for 3min,
then hold at 25 °C.
3. Run 2 μ l of the amplified library on capillary gel electrophoresis using the High Sensitivity D1000 Kit (Agilent Technologies) on a TapeStation system (Agilent Technologies) (see **Notes 33 & 34**).

3.15.2 Preparative PCR

1. From the results of the capillary gel electrophoresis, estimate the minimum number of PCR cycles to amplify 1/2 of the library (see **Note 35**).
2. Prepare the following PCR mix:
10 μ l cDNA (from step 3.15), 9 μ l H₂O, 1 μ l primer mix of P5Solexa and P3Solexa (10 μ M each), 20 μ l Accuprime Supermix 1 enzyme.
3. If everything is fine also amplify the second 1/2 of the library with the same cycle number and combine the two PCR reactions.

3.16 Post-PCR processing

1. Perform PCR clean-up protocol using the MinElute PCR Purification Kit (Qiagen) (see **Note 36**). Elute the library in 15 μ l water. Measure the concentration of eluted library with NanoDrop.
2. Take ca. 300 ng (according measured concentration from the previous step) of library and dilute them in water to a total volume of 10 μ l.
3. Perform size selection with the DNA 500 LabChip Kit (Perkin Elmer) on the LabChip system (see **Note 37**). Follow the manufacturer's guideline to perform the size selection. Cut the band in the range of 150 - 225 bp (the insert accounts for 20 - 95 bp).
4. Measure the concentration of size-selected library using the Qubit dsDNA HS Kit (Life Technologies). Use 2 μ l of the sample for Qubit measurement.
5. Run 2 μ l of size-selected library on capillary gel electrophoresis using the High Sensitivity D1000 Kit (Agilent Technologies) in a TapeStation system (Agilent Technologies).
6. Use the peak size from the capillary gel electrophoresis together with the Qubit measurement to estimate the molar concentration of the library (see **Note 38**).
7. Submit the library for high-throughput sequencing or store at -80 °C.

4. Notes

1. If you are working with a tissue with high RNase A activity, adding 1/1000 volume of ANTI-RNase (Ambion) will control the RNase conditions, without affecting the activity of RNase I.
2. The quality of the RT-clip primers depends strongly on the individual round of synthesis (**Fig. 2C**). Therefore it is important to compare all primers on the same input RNA, when reordering primers or when ordering new primers for the first time.
3. The length of crosslinking should be optimized for each protein, as each RNA-binding domain crosslinks with different efficiency depending on its content of aromatic amino acids and the nucleotide composition of the binding site. Try 100, 200 and 400 mJ/cm², then use the shortest condition that gives >70 % of the maximum signal.
4. Unless specified differently, all washes throughout the protocol are performed in a volume of 900 μ l.
5. You can prepare the beads with antibody in advance and store them on ice if preparation of the cell extract takes more than 60 min.
6. For rabbit antibodies, protein A Dynabeads (Life Technologies) can work better in some cases.
7. The required amount of antibody depends on its quality and purity. This should be optimized in preliminary experiments.
8. Cell culture pellets have ± 20 mg. Weighing pellets before freezing can help estimate the required volume of lysis buffer. An even better way is to determine RNA/protein concentration with NanoDrop/Bradford assay and normalize concentrations to the lowest sample. Then take 1 ml from each sample and proceed to the next step. Comparable RNA/protein concentrations should lead to more reproducible RNase digestions.
9. Sonication helps when using cell culture, as undigested viscous DNA can sometimes cause problems with the IP. It can also alleviate problems caused by mild lysis buffers or hard-to-lyse tissues. When the protein of interest is shown to have interaction with DNA, this step will be critical to reduce the protein loss. However in most of the cases, the sonication can be excluded from the protocol.
10. Optionally, the lysate can be pre-cleared with protein A sepharose (this doesn't hurt, but usually makes little difference; it may reduce background when using protein A Dynabeads with certain extracts). Prepare a 30 % protein A sepharose slurry in

water. Add 100 μ l protein A sepharose slurry to 1.5 ml lysate and rotate for 10 min in the cold room before spinning.

11. The optimal dilution factor for the low-RNase condition depends on the batch of RNase, so in the first experiment several dilutions should be tested (**Fig. 2B**). It is advisable to use two low-RNase concentrations that are close to the optimized range (this could be 1:500 and 1:2000, or any other combination that works). Since there is variability between experiments, this will ensure that at least one of the concentrations will be in the right range.
12. The high-RNase control is important to monitor the specificity of the IP. Other recommended controls include a control where the RNA-binding protein is absent from the original material (such as a knockout animal or knockdown cells), a control where crosslinking is omitted and a control where no antibody is used during IP.
13. Unlike other RNases, RNase I has little base preference, and therefore cleaves after all four nucleotides. Under high-RNase conditions, the size of the radioactive band in the SDS-PAGE has to change in comparison to low-RNase conditions, confirming that the band corresponds to a protein-RNA complex. Furthermore, this experiment helps to determine the size of the immunoprecipitated RNA-binding protein, as the protein will be bound to short RNAs and thus will migrate as a less diffuse band ~5 kDa above the expected molecular weight.
14. It is important to digest for exactly 3 min. Use 1.5 ml microtubes in a thermomixer for 1.5 ml microtubes to make the warming to 37 °C efficient and reproducible.
15. To test a new antibody, collect 15 μ l of lysate at this step for Western blot comparison of lysate before and after IP (to visualize depletion of the protein from the lysate).
16. If monitoring depletion efficiency, place on magnet and save 15 μ l supernatant for Western blot analysis (see **Note 15**).
17. The rotating/incubation time should always be the same to be comparable between experiments.
18. Sections 3.7 and 3.8 do not need to be carried out on no-UV and high-RNase controls. However, they need to be carried out on the no-antibody control to use it as a background estimate for the complete library preparation.
19. The NuPAGE gels are critical. A pour-your-own SDS-PAGE (Laemmli) changes its pH during the run, which can get to ~9.5 leading to alkaline hydrolysis of the RNA. The NuPAGE buffer system is close to pH 7. We use the NuPAGE MOPS SDS running buffer.

20. The pure nitrocellulose membrane is a little fragile, but it works better for the RNA/protein extraction step.
21. Use the shortest exposure that gives visible bands, visualize using the Image J software, and determine if the shift of the protein-RNA complex is reproducible in parallel experiments. If the signal in shortest exposure is strong enough, then expose for 15 min to a phosphorimager cassette and quantify with a Typhoon Phosphorimager (GE Healthcare) instead of Image J.
22. When performing iCLIP for the first time, use questions a) to c) to check that a specific UV crosslink and IP have been performed. Then follow points d) and e) as a guideline to properly cut the membrane and extract the complexes :
23. Is there a radioactive band ~5 kDa above the molecular weight of the protein in the high-RNase experiment?
24. Does the band disappear in the control experiments? These might include: no UV crosslink, pulldown without antibody (beads only or pre-immune serum), samples from a knockout organism or knockdown cells, or an appropriate control for overexpressed tagged proteins.
25. Does the band move up and become more diffuse under low-RNase conditions? Because the RNA digestion is random, the RNA sizes vary more in the low RNase condition and thus the RNA-protein complexes are more heterogeneous in size. On this basis, if you are convinced of the veracity of your results, proceed to RNA isolation and amplification. Note the following guidelines:
26. The average molecular weight of 70 nt RNA is ~20 kDa. As the tags contain a linker of 21 nt (L3-App), the ideal position of RNA-protein complexes that will generate iCLIP cDNAs of sufficient length is ~20 - 60 kDa above the expected molecular weight of the protein.
27. The width of the excised band depends on potential other RNA-protein complexes present in the vicinity as seen in the high-RNase experiment. If none are apparent, cut a wide band of ~20 - 60 kDa above the molecular weight of the protein. If, however, contaminant bands are present above the size of the protein, cut only up to the size of those bands. If the contaminating bands run below your RNA-protein complex, you might consider cutting an additional band between the contaminating band and your protein-RNA complex. The RNA sequences cloned from this band can later be used to compare with those purified with the protein-of-interest.

28. Over 90 % of the radioactive signal should be removed after proteinase K treatment. This can be monitored by a Geiger counter measurement of the membrane pieces before adding proteinase K and after removing it.
29. GlycoBlue is necessary to efficiently precipitate the small quantity of RNA.
30. Remove the wash first with a p1000 and then with a p20 or p10. Try not to disturb the pellet, but if you do, spin it down again. Leave on the bench for 3 min, but no longer, with the cap open to dry. When resuspending, make sure to pipette along the back area of the tube.
31. Use only the RT primers that have worked well in test experiments. You can find a list of nine RT primers that worked well in our experiments in chapter 2.2.
32. Do not forget a negative control. This can either be a reaction where no RNA was added to the mix, but preferably a control sample that was isolated from a piece of nitrocellulose that did not contain the protein-RNA complex (for example the no-antibody control). Use distinct primers (Rt1clip to Rt9clip) for the control and the different replicates or experiments. The different primers contain individual 4-nt barcode sequences that allow multiplexing of samples during high-throughput sequencing and control for cross-contamination between samples.
33. This will eliminate radioactivity from strongly labelled samples after the next step, and prevent RNA from interfering with subsequent reactions. It is possible to mix up to three samples that shall be multiplexed at this point. Alternatively, cDNA libraries of each sample can be amplified separately, and mixed after the PCR.
34. The 70 - 80 nt band is prone to producing primer artefacts in the PCR, and even if specific cDNAs are isolated, sequences are often too short to be mapped to the genome. Therefore, if binding to short RNAs (such as miRNAs) is not of interest, it is not necessary to isolate this band.
35. The test PCR in section 3.15.1 is optional. If you previously prepared libraries with the same protein and you had good radioactive RNA signal, you can estimate the number of required cycles and move directly to the preparative PCR in section 3.15.2.
36. Usually 15 - 18 cycles would be optimal. This depends on the crosslinking efficiency of the protein and can also be less or more.
37. All post-PCR work must be carried out on a specially designated bench. This cDNA must never be taken to an area where work with iCLIP RNA is done.
38. Repeat this step until samples are seen without over-amplification. The least possible PCR cycles would be preferred to reduce the PCR artefacts.

39. Although a conventional gel can be used to check the amplified library, high sensitivity systems such as TapeStation and Bioanalyzer (both Agilent Technologies) are preferred.
40. Consider that you will now be amplifying 2.5 times more concentrated cDNA, therefore one cycle less is needed than in the preliminary PCR.
41. The MinElute PCR Purification Kit will remove most of residual primers. This kit is especially helpful to concentrate low-amount libraries, which in turn allows to use less PCR cycles in the amplification step.
42. The post-PCR size selection is optional, but strongly recommended to remove residual primers and unspecific products. Removing residual primers will result in higher cluster density during high-throughput sequencing. The post-PCR size selection can also be performed using a TBE urea gel, when LabChip system is not available.
43. The kits that are used in the post-PCR processing steps for selection, visualization and quantification are optional and can be replaced by similar products. Nevertheless, high-sensitivity kits would be preferred to enable high efficiency in the high-throughput sequencing performance.

Acknowledgements

This protocol is based on previous versions that were derived with a lot of feedback and discussions in the Ule lab. We would also like to thank all iCLIP users for their feedback and comments that help to continuously improve the protocol. We acknowledge Dr. Kathi Zarnack and members of the König lab for discussions and help on the manuscript.

5. References

1. Kaper J.M. (1969) Nucleic acid-protein interactions in turnip yellow mosaic virus. *Science* 166(3902), 248–250
2. Gerstberger S., Hafner M., Tuschl T. (2014) A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15, 829–845
3. Greenberg J.R. (1979) Ultraviolet light-induced crosslinking of mRNA to proteins. *Nucleic Acids Res.* 6(2), 715–732
4. Lunde B.M., Moore C., Varani G. (2007) RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* 8, 479–490
5. Scherrer T., Mittal N., Janga S.C. et al. (2010) A screen for RNA-binding proteins in yeast indicates dual functions for many enzymes. *PLoS ONE* 5, e15499
6. Tsvetanova N.G., Klass D.M., Salzman J. et al. (2010) Proteome-wide search reveals unexpected RNA-binding proteins in *Saccharomyces cerevisiae*. *PLoS ONE* 5
7. Castello A., Fischer B., Eichelbaum K. et al. (2012) Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* 149, 1393–1406
8. Kwon S.C., Yi H., Eichelbaum K. et al. (2013) The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1122–1130
9. Baltz A.G., Munschauer M., Schwanhäusser B. et al. (2012) The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Molecular Cell* 46, 674–690
10. Cooper T.A., Wan L., Dreyfuss G. (2009) RNA and disease. *Cell* 136, 777–793
11. Chelly J., Mandel J.L. (2001) Monogenic causes of X-linked mental retardation. *Nat. Rev. Genet.* 2, 669–680
12. Mili S., Steitz J.A. (2004) Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 10, 1692–1694
13. Ule J., Jensen K., Mele A. et al. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 37, 376–386
14. Ule J., Jensen K.B., Ruggiu M. et al. (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215
15. Hafner M., Landthaler M., Burger L. et al. (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129–141
16. Licatalosi D.D., Mele A., Fak J.J. et al. (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469
17. Saulière J., Murigneux V., Wang Z. et al. (2012) CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nat. Struct. Mol. Biol.* 19, 1124–1131
18. König J., Zarnack K., Rot G. et al. (2010) iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17, 909–915
19. Sugimoto Y., König J., Hussain S. et al. (2012) Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol.* 13, R67
20. Huppertz I., Attig J., D'Ambrogio A. et al. (2014) iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* 65, 274–287
21. Jangi M., Boutz P.L., Paul P. et al. (2014) Rbfox2 controls autoregulation in RNA-binding protein networks. *Genes Dev.* 28, 637–651
22. Zarnack K., König J., Tajnik M. et al. (2013) Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 152, 453–466
23. Hussain S., Sajini A.A., Blanco S. et al. (2013) NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep* 4, 255–261
24. Broughton J.P., Pasquinelli A.E. (2013) Identifying Argonaute binding sites in *Caenorhabditis elegans* using iCLIP. *Methods* 63, 119–125
25. Bosson A.D., Zamudio J.R., Sharp P.A. (2014) Endogenous miRNA and Target Concentrations Determine Susceptibility to Potential ceRNA Competition. *Molecular Cell* 56, 347–359

4. References

1. Sutandy, F. X. R., Hildebrandt, A. & König, J. Post-Transcriptional Gene Regulation. *Methods Mol. Biol.* **1358**, 175–195 (2016).
2. Proudfoot, N. J. Ending the message: poly (A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
3. Moore, M. J. & Proudfoot, N. J. Pre-mRNA Processing Reaches Back to Transcription and Ahead to Translation. *Cell* **136**, 688–700 (2009).
4. Moore, M. J. From Birth to Death: The Complex Lives of Eukaryotic mRNAs. *Science (80-.)*. **309**, 1514–1519 (2005).
5. Kaper, J. M. Nucleic Acid-Protein Interactions in Turnip Yellow Mosaic Virus. *Science (80-.)*. **166**, 248–250 (1969).
6. Scherrer, T., Mittal, N., Janga, S. C. & Gerber, A. P. A Screen for RNA-Binding Proteins in Yeast Indicates Dual Functions for Many Enzymes. *PLoS One* **5**, (2010).
7. Tsvetanova, N. G., Klass, D. M., Salzman, J. & Brown, P. O. Proteome-wide search reveals unexpected RNA-binding proteins in *saccharomyces cerevisiae*. *PLoS One* **5**, (2010).
8. Baltz, A. G. *et al.* The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Mol. Cell* **46**, 674–690 (2012).
9. Castello, A. *et al.* Insights into RNA Biology from an Atlas of Mammalian mRNA-Binding Proteins. *Cell* **149**, 1393–1406 (2012).
10. Kwon, S. C. *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* **20**, 1122–1130 (2013).
11. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* **19**, 327–341 (2018).
12. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* **8**, 479–490 (2007).
13. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15**, 829–845 (2014).
14. Yoon, J. H. *et al.* Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat. Commun.* **4**, 1–14 (2013).
15. Dassi, E. Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. *Front. Mol. Biosci.* **4**, (2017).
16. Hudson, B. P., Martinez-Yamout, M. A., Dyson, H. J. & Wright, P. E. Recognition of the mRNA AU-rich element by the zinc finger domain of TIS11d. *Nat. Struct. Mol. Biol.* **11**, 257–264 (2004).
17. Fu, M. & Blackshear, P. J. RNA-binding proteins in immune regulation: A focus on CCCH zinc finger proteins. *Nat. Rev. Immunol.* **17**, 130–143 (2017).
18. Brown, R. S. Zinc finger proteins: Getting a grip on RNA. *Curr. Opin. Struct. Biol.* **15**, 94–98 (2005).
19. Hall, T. M. T. Multiple modes of RNA recognition by zinc finger proteins. *Curr. Opin. Struct. Biol.* **15**, 367–373 (2005).
20. Auweter, S. D., Oberstrass, F. C. & Allain, F. H. T. Sequence-specific binding of single-stranded RNA: Is there a code for recognition? *Nucleic Acids Res.* **34**, 4943–4959 (2006).
21. Beckmann, B. M., Castello, A. & Medenbach, J. The expanding universe of ribonucleoproteins: of novel RNA-binding proteins and unconventional interactions. *Pflugers Arch. Eur. J. Physiol.* **468**, 1029–1040 (2016).
22. Järvelin, A. I., Noerenberg, M., Davis, I. & Castello, A. The new (dis)order in RNA regulation. *Cell Commun. Signal.* **14**, (2016).
23. Phan, A. T. *et al.* Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat Struct Mol Biol.* **18**, 796–804 (2011).
24. Castello, A. *et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* **63**, 696–710 (2016).
25. Weber, S. C. & Brangwynne, C. P. Getting RNA and protein in phase. *Cell* **149**, 1188–1191 (2012).
26. Kato, M. *et al.* Cell-free formation of RNA granules: Low complexity sequence domains form dynamic fibers within hydrogels. *Cell* **149**, 753–767 (2012).
27. Han, T. W. *et al.* Cell-free formation of RNA granules: Bound RNAs identify features and components of cellular assemblies. *Cell* **149**, 768–779 (2012).
28. Lin, Y., Protter, D. S. W., Rosen, M. K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol. Cell* **60**, 208–219 (2015).
29. Protter, D. S. W. & Parker, R. Principles and Properties of Stress granules. *Trends Cell Biol.* **26**, 668–679 (2016).
30. Reijns, M. A. M., Alexander, R. D., Spiller, M. P. & Beggs, J. D. A role for Q / N-rich aggregation-prone regions in P-body localization. *J. Cell Sci.* **121**, 2463–2472 (2008).
31. Beckmann, B. M. *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat. Commun.* **6**, (2015).

32. Cano, F., Miranda-Saavedra, D. & Lehner, P. J. RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. *Biochem. Soc. Trans.* **38**, 1621–1626 (2010).
33. Tudek, A., Lloret-Illinares, M. & Jensen, T. H. The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Phil. Trans. R. Soc.* **373**, (2018).
34. Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **18**, 18–30 (2017).
35. Shi, Y. & Manley, J. L. The end of the message: Multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.* **29**, 889–897 (2015).
36. Goss, D. J. & Kleiman, F. E. Poly(A) Binding Proteins—Are they all created equal? *Wiley Interdiscip Rev RNA.* **4**, 167–179 (2013).
37. Darnell, J. C. & Richter, J. D. Cytoplasmic RNA-binding proteins and the control of complex brain function. *Cold Spring Harb. Perspect. Biol.* **4**, a012344 (2012).
38. Richter, J. D. & Lasko, P. Translational control in oocyte development. *Cold Spring Harb. Perspect. Biol.* **3**, a002758 (2011).
39. Kühn, U. & Wahle, E. Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* **1678**, 67–84 (2004).
40. Oszolak, F. *et al.* Resource Comprehensive Polyadenylation Site Maps in Yeast and Human Reveal Pervasive Alternative Polyadenylation. *Cell* **143**, 1018–1029 (2010).
41. Frischmeyer, P. A. *et al.* An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination. *Science (80-)*. **295**, 2258–2262 (2002).
42. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010).
43. Xie, J., Kozlov, G. & Gehring, K. The ‘tale’ of poly(A) binding protein: The MLE domain and PAM2-containing proteins. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1839**, 1062–1068 (2014).
44. Deo, R. C., Bonanno, J. B., Sonenberg, N. & Burley, S. K. Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98**, 835–845 (1999).
45. Smith, R. W. P., Blee, T. K. P. & Gray, N. K. Poly(A)-binding proteins are required for diverse biological processes in metazoans. *Biochem. Soc. Trans.* **42**, 1229–1237 (2014).
46. Baßler, J. & Hurt, E. Eukaryotic Ribosome Assembly. *Annu Rev Biochem.* **88**, 8.1–8.26 (2019).
47. Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051–1057 (2017).
48. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. *Science (80-)*. **289**, 920–930 (2000).
49. Hinnebusch, A. G. The Scanning Mechanism of Eukaryotic Translation Initiation. (2014). doi:10.1146/annurev-biochem-060713-035802
50. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* **19**, 526–541 (2018).
51. Dever, T. E., Dinman, J. D. & Green, R. Translation elongation and recoding in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **10**, (2018).
52. Hellen, C. U. T. Translation Termination and Ribosome Recycling in Eukaryotes. (2018). doi:10.1101/cshperspect.a032656
53. Dever, T. E. & Green, R. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **4**, 1–16 (2012).
54. Jackson, R. J., Hellen, C. U. T. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
55. Joazeiro, C. A. P. Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. *Annu. Rev. Cell Dev. Biol.* **2017**. **33**, (2017).
56. Hershey, J. W. B., Sonenberg, N. & Mathews, M. B. Principles of Translational Control. *Cold Spring Harb Perspect Biol.* **a032607**, (2018).
57. Kapur, M. & Ackerman, S. L. mRNA Translation Gone Awry: Translation Fidelity and Neurological Disease. *Trends Genet.* **34**, 218–231 (2018).
58. Inada, T. The Ribosome as a Platform for mRNA and Nascent Polypeptide Quality Control. *Trends Biochem. Sci.* **42**, 5–15 (2017).
59. Karousis, E. D. & Mühlemann, O. Nonsense-Mediated mRNA Decay Begins Where Translation Ends. *Cold Spring Harb. Perspect. Biol.* (2018). doi:10.1101/cshperspect.a032862
60. Shao, S. & Hegde, R. S. Target Selection during Protein Quality Control. *Trends Biochem. Sci.* **41**, 124–137 (2016).
61. Hetz, C., Chevet, E. & Oakes, S. A. Proteostasis control by the unfolded protein response. *Nat. Cell Biol.* **17**, 829–838 (2015).
62. Paknia, E., Chari, A., Stark, H. & Fischer, U. The Ribosome Cooperates with the Assembly Chaperone p1Cl_n to Initiate Formation of snRNPs. *Cell Rep.* **16**, 3103–3112 (2016).
63. Bengtson, M. H. & Joazeiro, C. A. P. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* **467**, 470–473 (2010).

64. Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nat Struct Mol Biol.* **23**, 7–15 (2016).
65. Guydosh, N. R. & Green, R. Translation of poly (A) tails leads to precise mRNA cleavage. *RNA* **23**, 749–761 (2017).
66. Shoemaker, C. J. & Green, R. Translation drives mRNA quality control. *Nat. Struct. Mol. Biol.* **19**, 594–601 (2012).
67. Kuroha, K. *et al.* Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Rep.* **11**, 956–961 (2010).
68. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006).
69. Tsuboi, T. *et al.* Dom34:Hbs1 Plays a General Role in Quality-Control Systems by Dissociation of a Stalled Ribosome at the 3' End of Aberrant mRNA. *Mol. Cell* **46**, 518–529 (2012).
70. Ito-Harashima, S., Kuroha, K., Tatematsu, T. & Inada, T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev.* **21**, 519–524 (2007).
71. Letzring, D. P., Wolf, A. S., Brule, C. E. & Grayhack, E. J. Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA* **19**, 1208–1217 (2013).
72. Juszkiwicz, S. & Hegde, R. S. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol. Cell* **65**, 743-750.e4 (2017).
73. Sundaramoorthy, E. *et al.* ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol. Cell* **65**, 751-760.e4 (2017).
74. Arthur, L. L. *et al.* Translational control by lysine-encoding A-rich sequences. *Sci. Adv.* **1**, (2015).
75. Pelechano, V., Wei, W. & Steinmetz, L. M. Extensive transcriptional heterogeneity revealed by isoform profiling. *Nature* **497**, 127–131 (2013).
76. Passos, D. O. *et al.* Analysis of Dom34 and Its Function in No-Go Decay. *Mol Biol Cell.* **20**, 3025–3032 (2009).
77. Gatfield, D. & Izaurralde, E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* **429**, 575–578 (2004).
78. Ikeuchi, K. & Inada, T. Ribosome-associated Asc1/RACK1 is required for endonucleolytic cleavage induced by stalled ribosome at the 3' end of nonstop mRNA. *Sci. Rep.* **6**, 1–10 (2016).
79. Pisareva, V. P., Skabkin, M. A., Hellen, C. U. T., Pestova, T. V & Pisarev, A. V. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO J.* **30**, 1804–1817 (2011).
80. Shoemaker, C. J. & Green, R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proc. Natl. Acad. Sci.* **108**, E1392–E1398 (2011).
81. van Hoof, A., Frischmeyer, P. A., Dietz, H. C. & Parker, R. Exosome-Mediated Recognition and Degradation of mRNAs Lacking a Termination Codon. *Science (80-.).* **295**, 1–4 (2002).
82. Schaeffer, D. & van Hoof, A. Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. *Proc. Natl. Acad. Sci.* **108**, 2366–2371 (2011).
83. Lu, J. & Deutsch, C. Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *J. Mol. Biol.* **384**, 73–86 (2008).
84. Ikeuchi, K. *et al.* Collided ribosomes form a unique structural interface to induce Hel2- driven quality control pathways. *EMBO J.* e100276 (2019). doi:10.15252/embj.2018100276
85. Winz, M.-L., Peil, L., Turowski, T. W., Rappsilber, J. & Tollervy, D. Molecular interactions between Hel2 and RNA supporting ribosome-associated quality control. *Nat. Commun.* **10**, 563 (2019).
86. Matsuo, Y. *et al.* Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat. Commun.* **8**, 159 (2017).
87. Sitron, C. S., Park, J. H. & Brandman, O. Asc1, Hel2, and Sih1 couple translation arrest to nascent chain degradation. *RNA* **23**, 798–810 (2017).
88. Brandman, O. *et al.* A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell* **151**, 1042–1054 (2012).
89. Saito, K., Horikawa, W. & Ito, K. Inhibiting K63 Polyubiquitination Abolishes No-Go Type Stalled Translation Surveillance in *Saccharomyces cerevisiae*. *PLoS Genet.* **11**, 1–18 (2015).
90. Yonashiro, R. *et al.* The Rqc2/Tae2 subunit of the ribosome-associated quality control (RQC) complex marks ribosome-stalled nascent polypeptide chains for aggregation. *Elife* **5**, (2016).
91. Simms, C. L., Yan, L. L. & Zaher, H. S. Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol. Cell* **68**, 361-373.e5 (2017).
92. Higgins, R. *et al.* The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol. Cell* **59**, 35–49 (2015).
93. Garzia, A. *et al.* The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun.* **8**, (2017).
94. Shoemaker, C. J., Eyler, D. E. & Green, R. Dom34:Hbs1 Promotes Subunit Dissociation and

- Peptidyl-tRNA Drop-Off to Initiate No-Go Decay. *Science* (80-.). **330**, 369–372 (2010).
95. Chen, L. *et al.* Structure of the Dom34-Hbs1 complex and implications for no-go decay. *Nat. Struct. Mol. Biol.* **17**, 1233–1240 (2010).
 96. Graille, M., Chaillet, M. & van Tilbeurgh, H. Structure of Yeast Dom34: A protein related to translation termination factor Erf1 and involved in No-Go decay. *J. Biol. Chem.* **283**, 7145–7154 (2008).
 97. Lee, H. H. *et al.* Structural and Functional Insights into Dom34, a Key Component of No-Go mRNA Decay. *Mol. Cell* **27**, 938–950 (2007).
 98. Kostova, K. K. *et al.* CAT-tailing as a fail-safe mechanism for efficient degradation of stalled nascent polypeptides. *Science* (80-.). **357**, 414–417 (2017).
 99. Shen, P. S. *et al.* Rqc2p and 60S ribosomal subunits mediate mRNA-independent elongation of nascent chains. *Science* (80-.). **347**, 75–78 (2015).
 100. Shao, S., Brown, A., Santhanam, B. & Hegde, R. S. Structure and assembly pathway of the ribosome quality control complex. *Mol. Cell* **57**, 433–445 (2015).
 101. Shao, S. & Hegde, R. S. Reconstitution of a Minimal Ribosome-Associated Ubiquitination Pathway with Purified Factors. *Mol. Cell* **55**, 880–890 (2014).
 102. Defenouillère, Q. *et al.* Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *PNAS* **110**, 5046–51 (2013).
 103. Verma, R. *et al.* Vms1 and ANKZF1 peptidyl-tRNA hydrolases release nascent chains from stalled ribosomes. *Nature* **557**, 446–451 (2018).
 104. Zurita Rendón, O. *et al.* Vms1p is a release factor for the ribosome-associated quality control complex. *Nat. Commun.* **9**(1), (2018).
 105. Kuroha, K., Zinoviev, A., Hellen, C. U. T. & Pestova, T. V. Release of Ubiquitinated and Non-ubiquitinated Nascent Chains from Stalled Mammalian Ribosomal Complexes by ANKZF1 and Pth1. *Mol. Cell* **72**, 1–17 (2018).
 106. Defenouillère, Q. *et al.* Rqc1 and Ltn1 Prevent C-terminal Alanine-Threonine Tail (CAT-tail) -induced Protein Aggregation by Efficient Recruitment of Cdc48 on Stalled 60S Subunits. *J. Biol. Chem.* **291**, 12245–12253 (2016).
 107. Defenouillère, Q., Namane, A., Mouaikel, J., Jacquier, A. & Fromont-Racine, M. The Ribosome-bound Quality Control complex remains associated to aberrant peptides during their proteasomal targeting and interacts with Tom1 to limit protein aggregation Quentin. *Mol Biol Cell* **28**, 1165–1176 (2017).
 108. Verma, R., Oania, R. S., Kolawa, N. J. & Deshaies, R. J. Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. *Elife* (2013). doi:10.7554/eLife.00308
 109. Dimitrova, L. N., Kuroha, K., Tatematsu, T. & Inada, T. Nascent Peptide-dependent Translation Arrest Leads to Not4p-mediated Protein Degradation by the Proteasome. *J. Biol. Chem.* **284**, 10343–10352 (2009).
 110. Halter, D., Collart, M. A. & Panasencko, O. O. The Not4 E3 ligase and CCR4 deadenylase play distinct roles in protein quality control. *PLoS One* **9**, (2014).
 111. Duttler, S., Pechmann, S. & Frydman, J. Principles of cotranslational ubiquitination and quality control at the ribosome. *Mol. Cell* **50**, 379–393 (2013).
 112. Preissler, S. *et al.* Not4-dependent translational repression is important for cellular protein homeostasis in yeast. *Embo J* **34**, 1905–1924 (2015).
 113. Shcherbik, N., Chernova, T. A., Chernoff, Y. O. & Pestov, D. G. Distinct types of translation termination generate substrates for ribosome-associated quality control. *Nucleic Acids Res.* **44**, 6840–6852 (2016).
 114. Yau, R. & Rape, M. The increasing complexity of the ubiquitin code. *Nat. Cell Biol.* **18**, 579–586 (2016).
 115. Oh, E., Akopian, D. & Rape, M. Principles of Ubiquitin- Dependent Signaling. *Annu. Rev. Cell Dev. Biol* **34**, 18.1-18.26 (2018).
 116. Berndsen, C. E. & Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol.* **21**, 301–307 (2014).
 117. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* **81**, 203–229 (2012).
 118. Rape, M. Ubiquitylation at the crossroads of development and disease. *Nat. Rev. Mol. Cell Biol.* **19**, 59–70 (2017).
 119. Finley, D., Özkaynak, E. & Varshavsky, A. The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell* **48**, 1035–1046 (1987).
 120. Zheng, N. & Shabek, N. Ubiquitin Ligases: Structure, Function, and Regulation. *Annu. Rev. Biochem.* **86**, 14.1-14.29 (2017).
 121. Budhidarmo, R., Nakatani, Y. & Day, C. L. RINGs hold the key to ubiquitin transfer. *Trends Biochem. Sci.* **37**, 58–65 (2012).
 122. Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
 123. Huibregtse, J. M., Scheffner, M., Beaudenon, S. & Howley, P. M. A family of proteins structurally and

- functionally related to the E6-AP ubiquitin-protein ligase. *Proc. Natl. Acad. Sci.* **92**, 2563–2567 (1995).
124. Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: more than just middle men. *Cell Res.* **26**, 423–440 (2016).
 125. Ikeda, F. & Dikic, I. Atypical ubiquitin chains: New molecular signals. 'Protein Modifications: Beyond the Usual Suspects' Review Series. *EMBO Rep.* **9**, 536–542 (2008).
 126. Cano, F., Rapiteanu, R., Winkler, G. S. & Lehner, P. J. A non-proteolytic role for ubiquitin in deadenylation of MHC-I mRNA by the RNA-binding E3-ligase MEX-3C. *Nat. Commun.* **6**, 8670 (2015).
 127. Gray, T. A. *et al.* The ancient source of a distinct gene family encoding proteins featuring RING and C(3)H zinc-finger motifs with abundant expression in developing brain and nervous system. *Genomics* **66**, 76–86 (2000).
 128. Böhne, A. *et al.* The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* **11**, 721 (2010).
 129. Miroci, H. *et al.* Makorin ring zinc finger protein 1 (MKRN1), a novel poly(A)-binding protein-interacting protein, stimulates translation in nerve cells. *J. Biol. Chem. Biol. Chem.* **287**, 1322–1334 (2012).
 130. Kozlov, G., Ménade, M., Rosenauer, A., Nguyen, L. & Gehring, K. Molecular Determinants of PAM2 Recognition by the MLLE Domain of Poly(A) -Binding Protein. *J. Mol. Biol.* **397**, 397–407 (2010).
 131. Lee, M.-S. *et al.* PI3K/AKT activation induces PTEN ubiquitination and destabilization accelerating tumourigenesis. *Nat. Commun.* **6**, 7769 (2015).
 132. Lee, E.-W. *et al.* Ubiquitination and degradation of the FADD adaptor protein regulate death receptor-mediated apoptosis and necroptosis. *Nat. Commun.* **3**, 978 (2012).
 133. Ko, A. *et al.* Acceleration of gastric tumorigenesis through MKRN1-mediated posttranslational regulation of p14ARF. *J. Natl. Cancer Inst.* **104**, 1660–72 (2012).
 134. Lee, E.-W. *et al.* Differential regulation of p53 and p21 by MKRN1 E3 ligase controls cell cycle arrest and apoptosis. *EMBO J.* **28**, 2100–2113 (2009).
 135. Kim, J. H. *et al.* Ubiquitin ligase MKRN1 modulates telomere length homeostasis through a proteolysis of hTERT. *Genes Dev.* **19**, 776–81 (2005).
 136. Salvatico, J., Kim, J. H., Chung, I. K. & Muller, M. T. Differentiation linked regulation of telomerase activity by Makorin-1. *Mol. Cell. Biochem.* **342**, 241–250 (2010).
 137. Omwancha, J. *et al.* Makorin RING finger protein 1 (MKRN1) has negative and positive effects on RNA polymerase II-dependent transcription. *Endocrine* **29**, 363–73 (2006).
 138. Cassar, P. A. *et al.* Integrative genomics positions MKRN1 as a novel ribonucleoprotein within the embryonic stem cell gene regulatory network. *EMBO Rep.* **16**, 1334–57 (2015).
 139. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593–597 (2013).
 140. Böhne, A. *et al.* The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* **11**, 721 (2010).
 141. Lee, S.-H. *et al.* Widespread intronic polyadenylation inactivates tumour suppressor genes in leukaemia. *Nature* **561**, 127–131 (2018).
 142. Jamar, N. H., Kritsiligkou, P. & Grant, C. M. Loss of mRNA surveillance pathways results in widespread protein aggregation. *Sci. Rep.* **8**, (2018).
 143. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
 144. Klauer, A. A. & van Hoof, A. Degradation of mRNAs that lack a stop codon: A decade of nonstop progress. *Wiley Interdiscip. Rev. RNA* **3**, 649–660 (2012).
 145. Tahmasebi, S., Khoutorsky, A., Mathews, M. B. & Sonenberg, N. Translation deregulation in human disease. *Nat. Rev. Mol. Cell Biol.* **19**, 791–807 (2018).
 146. Morita, M. *et al.* A Novel 4EHP-GIGYF2 Translational Repressor Complex Is Essential for Mammalian Development. *Mol. Cell. Biol.* **32**, 3585–3593 (2012).
 147. Chartier-Harlin, M.-C. *et al.* Translation Initiator EIF4G1 Mutations in Familial Parkinson Disease. *Am J Hum Genet.* **89**, 398–406 (2011).
 148. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
 149. Schmidt, M. & Finley, D. Regulation of proteasome activity in health and disease. *Biochim. Biophys. Acta* **1843**, 13–25 (2014).
 150. Ryu, K.-Y., Garza, J. C., Lu, X.-Y., Barsh, G. S. & Kopito, R. R. Hypothalamic neurodegeneration and adult-onset obesity in mice lacking the Ubb polyubiquitin gene. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 4016–4021 (2008).
 151. Ryu, H.-W., Park, C.-W. & Ryu, K.-Y. Disruption of polyubiquitin gene Ubb causes dysregulation of

- neural stem cell. *Sci. Rep.* **4**, (2014).
152. Atkin, G. & Paulson, H. Ubiquitin pathways in neurodegenerative disease. *Front. Mol. Neurosci.* **7**, 1–17 (2014).

II. AIMS OF THE PROJECT

The aim of this PhD project was to gain deeper insights into the biology of RNA-binding ubiquitin ligases (RBULs). RBULs have the potential to link RNA-mediated mechanisms to protein ubiquitylation. Despite this, the cellular functions, substrates and interaction partners of most RBULs remain poorly characterized.

As a first part of my PhD project, I set up new techniques to study RBULs in the lab. A powerful approach for analyzing protein functions is affinity purification (AP) combined with quantitative mass spectrometry (MS)-based proteomics. However, mapping the physiological interaction partners of RNA-binding proteins has been hampered by their intrinsic properties, in particular the existence of low-complexity regions, which are prone to engage in non-physiological interactions. Here, I used an adapted AP approach to identify the interaction partners of human RBULs harboring different RNA-binding domains. To increase the likelihood of recovering physiological interactions, I combined control and bait-expressing cells prior to lysis in order to identify only stable interactions that were originally present in the cell. To benchmark the approach in its ability to recover physiological interactions, we exploited gene function similarity between the bait proteins and interactors. With this, we aimed to demonstrate the capacity of RBULs to link posttranscriptional regulation with the ubiquitin system.

As a second part of my PhD project, I assessed the cellular and molecular function of a selected RBUL, MKRN1, in human cells. A combination of molecular and cell biological assays, together with system-wide high-throughput approaches were employed to dissect the physiological role of MKRN1's RNA-binding ability and E3 ligase activity. By performing iCLIP and RNA sequencing (RNA-Seq) experiments, I wanted to determine the RNA-binding specificity of MKRN1 and its role in post-transcriptional regulation. In order to gain insights into the molecular function of MKRN1, interactome studies were performed to identify its protein binding partners. These experiments led me to the finding that MKRN1 is involved in the ribosome-associated quality control (RQC), a quality control mechanism that leads to ribosome recycling and decay of the truncated nascent polypeptide upon the recognition of the translation of aberrant mRNAs. Based on these findings, I continued to elucidate the role of MKRN1 in RQC, for example by assessing the ubiquitylation substrates of MKRN1 with the help of ubiquitin remnant profiling.

Aims of the project

CHAPTER III

**INTERACTION PROFILING OF RNA-BINDING
UBIQUITIN LIGASES REVEALS A LINK
BETWEEN POST-TRANSCRIPTIONAL
REGULATION AND THE UBIQUITIN SYSTEM**

III. INTERACTION PROFILING OF RNA-BINDING UBIQUITIN LIGASES REVEALS A LINK BETWEEN POSTTRANSCRIPTIONAL REGULATION AND THE UBIQUITIN SYSTEM

Sci Rep. 2017 Nov 29;7(1):16582. doi: 10.1038/s41598-017-16695-6.

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

Andrea Hildebrandt¹, Gregorio Alanis-Lobato^{1,2}, Andrea Voigt¹, Kathi Zarnack³, Miguel A. Andrade-Navarro^{1,2}, Petra Beli^{1*} and Julian König^{1*}

¹Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany

²Faculty of Biology, Johannes Gutenberg University, Gresemundweg 2, 55128 Mainz, Germany

³Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, Max-von-Laue-Str. 15, 60438 Frankfurt, Germany

*Correspondence should be addressed to:

Petra Beli: p.beli@imb-mainz.de

Julian König: j.koenig@imb-mainz.de

Keywords: Mass spectrometry-based proteomics, protein-protein interactions, gene ontology similarity, RNA-binding ubiquitin ligases, PRPF19, MEX3B

Abstract

RNA-binding ubiquitin ligases (RBULs) have the potential to link RNA-mediated mechanisms to protein ubiquitylation. Despite this, the cellular functions, substrates and interaction partners of most RBULs remain poorly characterized. Affinity purification (AP) combined with quantitative mass spectrometry (MS)-based proteomics is a powerful approach for analyzing protein functions. Mapping the physiological interaction partners of RNA-binding proteins has been hampered by their intrinsic properties, in particular the existence of low-complexity regions, which are prone to engage in non-physiological interactions. Here, we used an adapted AP approach to identify the interaction partners of human RBULs harboring different RNA-binding domains. To increase the likelihood of recovering physiological interactions, we combined control and bait-expressing cells prior to lysis. In this setup, only stable interactions that were originally present in the cell will be identified. We exploit gene function similarity between the bait proteins and interactors to benchmark our approach in its ability to recover physiological interactions. We reveal that RBULs engage in stable interactions with RNA-binding proteins involved in different steps of RNA metabolism as well as with components of the ubiquitin conjugation machinery and ubiquitin-binding proteins. Our results thus demonstrate their capacity to link posttranscriptional regulation with the ubiquitin system.

Introduction

Posttranscriptional RNA processing provides a fundamental step in the regulation of gene expression. RNA-binding proteins (RBPs) play a key role in posttranscriptional regulation by determining the fate and function of their target transcripts. Recent large-scale mRNA interaction profiling studies have demonstrated that eukaryotic cells express >1,200 RBPs¹⁻³. In addition to their RNA binding capability, many of these RBPs possess catalytic activities, thus introducing an additional layer of complexity into posttranscriptional regulation¹⁻³. A particularly interesting class of proteins in this context are RNA-binding ubiquitin ligases (RBULs) that contain either a HECT, RING or ring between ring (RBR) ubiquitin ligase domain. The human genome encodes at least 26 putative RING-type RBULs which can be assigned to five different families according to their RNA-binding domains (RBDs), including CCCH zinc fingers, K-homology (KH) domains, RNA recognition motifs (RRM) and KKKTK (K-rich) regions⁴. The existence of the RNA-binding and ubiquitin ligase functions in a single protein has the potential to link posttranscriptional regulation of gene expression with the ubiquitin system. Despite

this, only a few RBULs have been characterized to date and most previous studies focused on a specific function of a single RBUL.

Mapping of protein interaction partners is frequently employed to analyze cellular functions of poorly characterized proteins. To this end, a bait protein and its interaction partners are purified by antibodies or affinity resins, and the interaction partners are identified by liquid chromatography-tandem mass spectrometry (LC-MS/MS). This approach allows to efficiently discriminate physiological interactions from background binders, when combined with quantitative MS approaches such as stable isotope labeling with amino acids in cell culture (SILAC)⁵. However, several characteristics of RBPs impede the identification of physiological interactions in proteomics approaches. RBPs often contain disordered regions, including repetitive sequences and low-complexity domains, which can promote dynamic interactions with other proteins even in cell lysates^{3,6}. Native ribonucleoprotein complexes were previously shown to extensively reassemble already during cell lysis, thereby concealing the interaction landscape of the native cell⁷. In addition, several RBPs were recently described to harbor the ability to transition into a hydrogel-like state that likely facilitates non-physiological protein associations⁸. Taken together, the above-mentioned properties of RBPs interfere with a specific enrichment of physiological interactions through affinity purification (AP) and create a demand for well-adapted sample preparation strategies.

Depending on the stability of the studied interactions, different sample preparation strategies have been developed. For instance, the conventional SILAC-AP workflow efficiently captures both stable and transient interactions by separately processing all samples until after AP (Fig. 1a). In contrast, purification after mixing (PAM)-SILAC employs combining of differentially isotope-labeled cell lysates prior to AP⁹. In this setup, transient interactors with high on/off rates quickly reach an equilibrium of light and heavy labeled forms, resulting in SILAC ratios close to background levels. Conversely, stable associations that are maintained throughout purification retain high SILAC ratios. However, neither of the two sample preparation strategies is immune to proteins that engage in non-physiological interactions already during cell lysis. Although not occurring under physiological conditions, such newly formed protein associations might display high stability, for instance if high-affinity binders had been precluded from each other *in vivo* by cellular compartmentalization. Here, we present an adapted SILAC-based AP approach (referred to adapted AP) that employs mixing of equal amounts of differentially isotope-labeled cells prior to the cell lysis (Fig. 1b). Because

light and heavy-isotope labeled forms of all proteins are mixed during lysis, newly established interactions will display low SILAC ratios irrespective of their biophysical properties. Consequently, the described approach specifically recovers stable interactions that were present within the intact cell and is complementary to other approaches, such as the conventional SILAC-AP, tandem affinity purifications (TAP) and biotin-based proximity tagging (APEX and BioID)¹⁰⁻¹². We benchmark our approach in its ability to recover physiological interactions using Gene Ontology (GO) similarity measures to score the coherence of functional annotations between the bait and putative interaction partners. Employing the described approach, we identify the interaction partners of six human RBULs (ARIH2, MEX3B, MKRN1, MKRN2, RNF17 and PRPF19). In addition to at least one RING E3 ubiquitin ligase domain, the selected RBULs harbor different RBDs, such as WD repeats (PRPF19), KH domains (MEX3B), Tudor domains (RNF17) or CCCH zinc fingers (MKRN1/2) (Fig. 2c). We further include ARIH2, which does not show an obvious RBD according to current predictions, but was previously found in association with the nuclear polyA-binding protein PABPN1¹³.

For all investigated RBULs, our analyses highlight their extensive involvement in ubiquitin-mediated functions, illustrated for instance by the interaction of five RBULs with the ubiquitin-dependent co-chaperone VCP. In addition, we find evidence that the studied RBULs are involved in different posttranscriptional pathways, including translational regulation, ribonucleoprotein assembly and splicing. The benchmarking of our approach using GO similarity testing further supports the notion that combining cells prior to lysis increases the likelihood of recovering stable physiological interactions when studying RBPs.

Results

Strategy for recovering stable interaction partners of RNA-binding ubiquitin ligases

In a conventional SILAC-based AP experiment, the bait and control pulldowns are performed separately and the enriched proteins are combined after the washing steps. This setup enables the quantitative comparison of different conditions and minimizes the intra-sample variance¹⁴. However, we reasoned that this approach might be suboptimal when analyzing RBPs, which are prone to form non-physiological interactions in the cell lysate. We therefore employed an adapted AP that disfavors dynamically exchanging or newly established protein interactions after cell lysis to analyze the stable interaction partners of RBULs. To this end, heavy isotope-labeled

cells expressing the GFP-tagged bait are combined already prior to cell lysis with light isotope-labeled control cells expressing GFP only (Fig. 1b). Interaction partners with high exchange rates will display both labels with equal probability, resulting in low, close-to-background SILAC ratios for these transient interactions. Reversely, stable physiological interaction partners that associated with the bait protein in the cell will remain constantly bound during the sample preparation, thus displaying high SILAC ratios (Fig. 1b).

In order to test the performance of our adapted AP, we analyzed its ability to identify the interaction partners of the well-characterized RBUL PRPF19. PRPF19 is part of the PRP19/CDC5L complex (also known as Nineteen complex [NTC]) and the XAB2 complex and plays a role in the catalytic activation of the spliceosome¹⁵⁻¹⁸. GFP-tagged PRPF19 was expressed in human HEK293T cells. Western blot experiments with a PRPF19-specific antibody that detects both endogenous and GFP-tagged PRPF19 suggested a 37-fold lower expression compared to the endogenous counterpart (Suppl. Fig. 3a). Apart from the previously described nuclear localization (Human Protein Atlas¹⁹, <http://www.proteinatlas.org/>), we found a fraction of GFP-PRPF19 also in the cytoplasm (Fig. 2e, Suppl. Table 6). For isotopic labeling, GFP-PRPF19-expressing cells were grown in medium containing heavy isotope-labeled arginine and lysine, while empty vector-expressing cells grown in medium containing light isotope-labeled arginine and

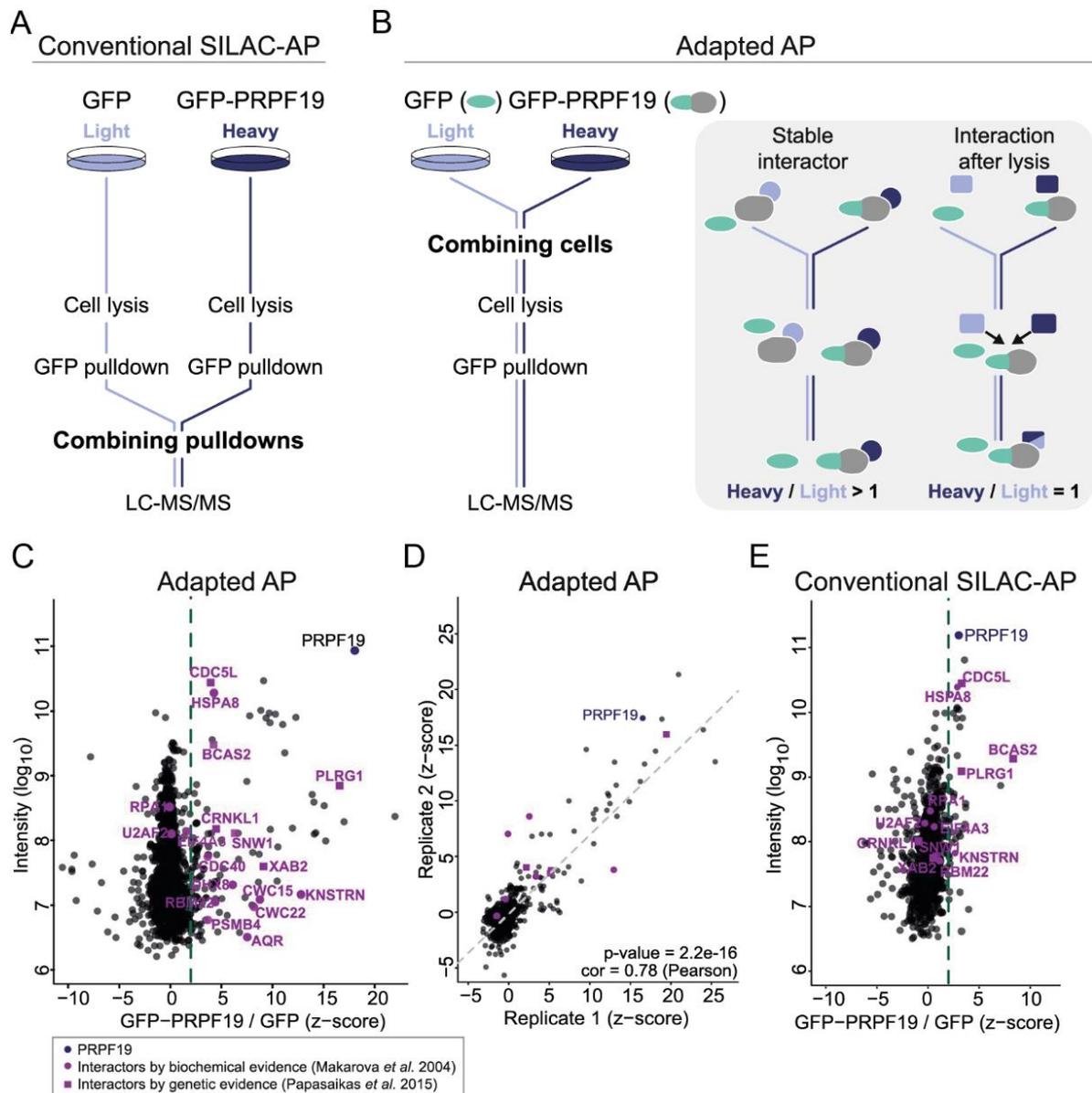


Figure III-1, The adapted SILAC-based AP approach efficiently recovers known interactors of PRPF19. **a.** In the conventional SILAC-AP, GFP alone is expressed in light-isotope labeled cells (light blue), while GFP-tagged PRPF19 is expressed in heavy-isotope labeled cells (dark blue). Upon cell lysis, GFP/GFP-tagged bait proteins are subjected to affinity purification (AP) with GFP-trap agarose beads. Enriched proteins are mixed in a 1:1 ratio and analyzed by LC-MS/MS. **b.** For the adapted AP protocol, differentially isotope-labeled cells are mixed at equal amounts and lysed as a pool. Upon AP and on-bead digest by trypsin, enriched proteins are analyzed by LC-MS/MS. In this setup, only stable interactions that were already present *in vivo* will obtain high SILAC ratio (H/L), whereas transient interactors and any associations formed after cell lysis will display equalized label occurrences. **c.** The adapted AP for GFP-PRPF19 identifies 13 previously known PRPF19 interactors. Shown are the SILAC ratios of one individual biological replicate after z-score normalization plotted against \log_{10} transformed intensities. PRPF19 is labeled in blue. Previously described PRPF19 interactors that were biochemically shown to interact¹⁵ or assigned by genetic evidence⁷² are displayed by round or square purple symbols, respectively. The applied cut-off at z-score ≥ 2 is labeled by a green dashed line. **d.** Comparison of two independent biological replicates of the adapted AP with GFP-PRPF19. Visualization and labeling as in (c). A linear regression line is indicated by a grey dashed line. **e.** Pulldowns of GFP-PRPF19 following the

conventional SILAC-AP protocol identifies one known PRPF19 interactor. SILAC ratios (after z-score normalization) are plotted against \log_{10} transformed intensities of one biological replicate. Visualization and labeling as in (c).

lysine served as control. Equal numbers of light and heavy-isotope labeled cells were combined, followed by lysis and AP using GFP-Trap agarose. Immunoprecipitated proteins were digested on-bead with trypsin, and peptides were identified by LC-MS/MS²⁰. In total, we quantified 1,061 protein groups in at least two replicate experiments (at least two peptides, including one unique peptide, and two ratio counts; Suppl. Table 3). The quantified SILAC ratios were \log_2 transformed and converted into z-scores^{21,22}. We confirmed the reproducibility of our measurements by comparing independent experimental replicates (Fig. 1d). All proteins with an average z-score ≥ 2 across five replicate experiments, meaning that they differ by at least two standard deviations from the mean of the positive values in the distribution, were considered as putative interaction partners (green dashed line, Fig. 1c). Out of 1,061 quantified proteins, 91 proteins displayed a z-score ≥ 2 in the GFP-PRPF19 pulldown compared to control, thus representing putative stable interaction partners of PRPF19 (Fig. 1c, 3, Suppl. Table 1, 2, 3). Notably, 17 out of these 91 (19%) proteins have been previously described as PRPF19 interaction partners according to the Human Integrated Protein-Protein Interaction Reference (HIPPIE) database²³ (Suppl. Table 5). Moreover, 13 of these represent well known PRPF19 interactors, including the complete NTC core module (CDC5L, PLRG1, and BCAS2) as well as XAB2 and AQR from the XAB2 complex^{15,16} (Fig. 1c). To validate the performance and specificity of our approach, we performed a conventional SILAC-AP for GFP-PRPF19, in which samples are combined after AP. In order to discriminate RNA-mediated interactions, we included an additional condition, in which the samples were subjected to optimized RNase A/RNase T1 digest after the AP (Suppl. Fig. 1, 2a). We quantified 632 protein groups, out of which 16 had a z-score ≥ 2 and were thus regarded as putative interaction partners of GFP-PRPF19 (Suppl. Table 1, 3). Among these, only one known interaction partner was recovered (compared to 13 in the pulldown applying the adapted AP) (Fig. 1e). 5 of 16 proteins were lost after RNase digestion, indicating that RNA-mediated protein interactions account for a minor fraction of detected interactions when analyzing RBPs (Suppl. Fig. 2b). In summary, we show that the adapted AP based on combining proteins already at the cell level enabled us to substantially increase the recovery rate of known interaction partners compared to the conventional SILAC-AP. These results underline the

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

increased performance and specificity of the adapted AP approach in analyzing the interaction partners of RBPs.

Functional coherence of the identified interaction profile of PRPF19

As outlined above, the adapted AP is suited to capture stable interaction partners, whereas loosely associated proteins will be categorized as background binders. We postulated that physiological interaction partners are more likely to participate in the same cellular processes as the bait protein compared to spuriously associated proteins. We therefore investigated the similarity of the associated Gene Ontology (GO) terms²⁴. To this end, we computed the Wang's Index (WI) as a measure of pair-wise similarity of Biological Process (BP) GO terms between the bait and the putative interaction partners identified by the adapted AP^{25,26}. To benchmark this approach, we extracted known high-quality protein interactions from the structurally resolved INstruct network and compared the WI distributions of connected and disconnected protein pairs^{27,28}. This analysis showed that WIs of connected proteins are significantly higher than those of disconnected pairs (P value $< 2.2 \times 10^{-16}$, Mann-Whitney U test; Fig. 2a), confirming that they offer meaningful evidence for true interactions. Based on the WI distribution for disconnected pairs, we estimated an empirical false discovery rate (FDR) and defined a cut-off at FDR $< 10\%$ (corresponding to a WI = 0.414; Fig. 2a). Under this premise, we consider a RBUL-interactor pair as 'true' (i.e. supported by GO evidence), if the associated GO-BP terms yield a WI ≥ 0.414 , enabling us to compile a reference set of 'true-positive' functionally coherent interactions. This method thus enables us to evaluate potential protein-protein interactions, identified by the adapted AP, for their probability to interact physiologically.

Based on the GO evidence, we compared the performance of the adapted AP with the conventional SILAC-AP based on Precision-Recall curves (Fig. 2b). To this end, we ordered the RBUL-interactor pairs by decreasing SILAC ratios and iteratively considered an increasing number of interaction pairs from the top of the list. For each set of interaction pairs, we computed the fraction of functionally coherent interactions (WI ≥ 0.414) with respect to the current set (precision) and the full list (recall). Putative PRPF19 interactors determined by the adapted AP showed strong signal enrichment, with increased precision associated with higher SILAC ratios. As expected, precision is sacrificed beyond a certain level of sensitivity (recall) when unspecific interactions begin to prevail. Notably, the adapted AP performs considerably better than the conventional SILAC-AP, underlining its ability to identify physiological interaction partners of RBPs. In line with the notion that some unspecific interactions could be mediated by RNA, the additional RNase digest slightly improved the performance in the GO similarity scoring.

In summary, our adapted AP in combination with computational evaluation based on GO similarities offers a useful approach to identify stable protein interactions of RBPs and can be used complementary to the conventional SILAC-AP. We benchmarked this approach on the example of PRPF19, which enabled the identification of previously known as well as novel interaction partners of this RBUL.

Defining the interaction profiles of selected RBULs

Reassured by the high efficiency in recovering known protein interaction partners of PRPF19, we applied our adapted AP and the computational evaluation steps to further RBULs. We selected MEX3B, MKRN1, MKRN2 and RNF17, which contain different RBDs and represent different RBUL families (Fig. 2c). We additionally included ARIH2, which does not harbor an RBD but was previously shown to interact with the nuclear polyA-binding protein PABPN1¹³. We first used fluorescence microscopy to confirm the subcellular localization of the tagged proteins (Fig. 2e). All GFP-tagged RBULs localized to the expected compartments as reported in the Human Protein Atlas¹⁹ and previous publications (Fig. 2e, Suppl. Table 6). For MEX3B, we additionally performed Western blots with a specific antibody that detects both endogenous and GFP-tagged MEX3B, estimating a 13-fold increased expression level of the tagged version compared to the endogenous counterpart (Suppl. Fig. 3a).

Using our adapted AP, we identified between 9 and 42 stable interaction partners which displayed a z-score ≥ 2 with the different RBULs (ARIH2, 42; MEX3B, 35; MKRN1, 9; MKRN2, 17; RNF17, 17; Fig. 2f, 3, Suppl. Table 1, 2, 3). Among these, we find a small number of likely contaminants, including HSP70 chaperone family members, such as HSPA1A, that act as disaggregases for newly translated proteins and thus represent common contaminants in AP experiments^{29,30}. In line with the notion that these interactions could be a secondary effect of ectopic RBUL expression, HSPA1A reproducibly interacted with GFP-tagged PRPF19 and MEX3B, but showed only minor interaction with endogenous PRPF19 in independent validation experiments (Fig. 4c, d, Suppl. Fig. 3b).

In order to assess the functional coherence of the identified interactomes, we employed the GO similarity scoring approach. This confirmed that the interactomes of the six RBULs are enriched for RBUL-interactor pairs with WI ≥ 0.414 , suggesting that they generally represent physiological interactors. The exception is RNF17, which is mainly annotated with superordinate GO terms such as 'cellular process' and hence not

appropriate for GO similarity scoring. The applied cut-off ($z\text{-score} \geq 2$) consistently falls within the margin of specific enrichment, supporting the choice of this threshold to detect meaningful interactions. For comparison, we subjected the five RBULs to the conventional SILAC-AP with and without RNase treatment (Suppl. Fig. 2a, b, Suppl. Table 1, 3). Similar to the initial observation with PRPF19, our adapted AP consistently outperforms the conventional SILAC-AP in the Precision-Recall evaluation for these RBULs (Fig. 2f). Accordingly, we find hardly any overlap between the interactome lists from the two approaches (Suppl. Fig. 2c). This nicely illustrates the conceptual differences of the two protocols and their target interactions, but may also result at least in parts from the inherent limitations of quantitative proteomics in terms of coverage and sensitivity³¹.

Integration of all datasets revealed that a major fraction of the interactors of each RBUL are unique and not shared with any other tested RBUL (38% [RNF17] - 88% [ARIH2]; Fig. 3), suggesting broad functional divergence. In order to get insights into their cellular functions, we used GO annotations (Biological Process and Molecular Function) to classify the interaction partners into the categories 'RNA', 'Ubiquitin' and 'Other'. Notably, the interaction partners of all studied RBULs are implicated into both, posttranscriptional processes as well as the ubiquitin system (Fig. 4a). In accordance with their ubiquitin ligase function, the interaction partners of RBULs are linked to core functionalities of the ubiquitin system. First, we find ubiquitin itself (RPS27A) as a stable interaction partner of all six RBULs, which could be indicative of auto-ubiquitylation. In addition, cross-targeting between different E3 ligases has also been described as a mode of activity regulation³². In line with this notion, we observe stable associations with other E3 ligases, including RNF20, CUL5, and MID1 (ARIH2) and HUWE1 (MEX3B & PRPF19). For PRPF19, we also identify an interaction with the ubiquitin-specific peptidases USP19 (Fig. 3). Moreover, we detect the ubiquitin-dependent co-chaperone VCP (also known as p97) in stable association with all RBULs (Fig. 3). VCP is an ATPase that structurally remodels ubiquitylated proteins before their degradation through the ubiquitin-proteasome or the autophagy-lysosome pathway³³. We verified the interaction of VCP with GFP-tagged MEX3B and PRPF19 as well as endogenous PRPF19 using pulldown experiments followed by Western blot (Fig. 4c, d, Suppl. Fig. 3b).

In addition to the ubiquitin-related functions, mapping of the RBUL interactomes reveals links to different posttranscriptional processes. For instance, several RBULs interact

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

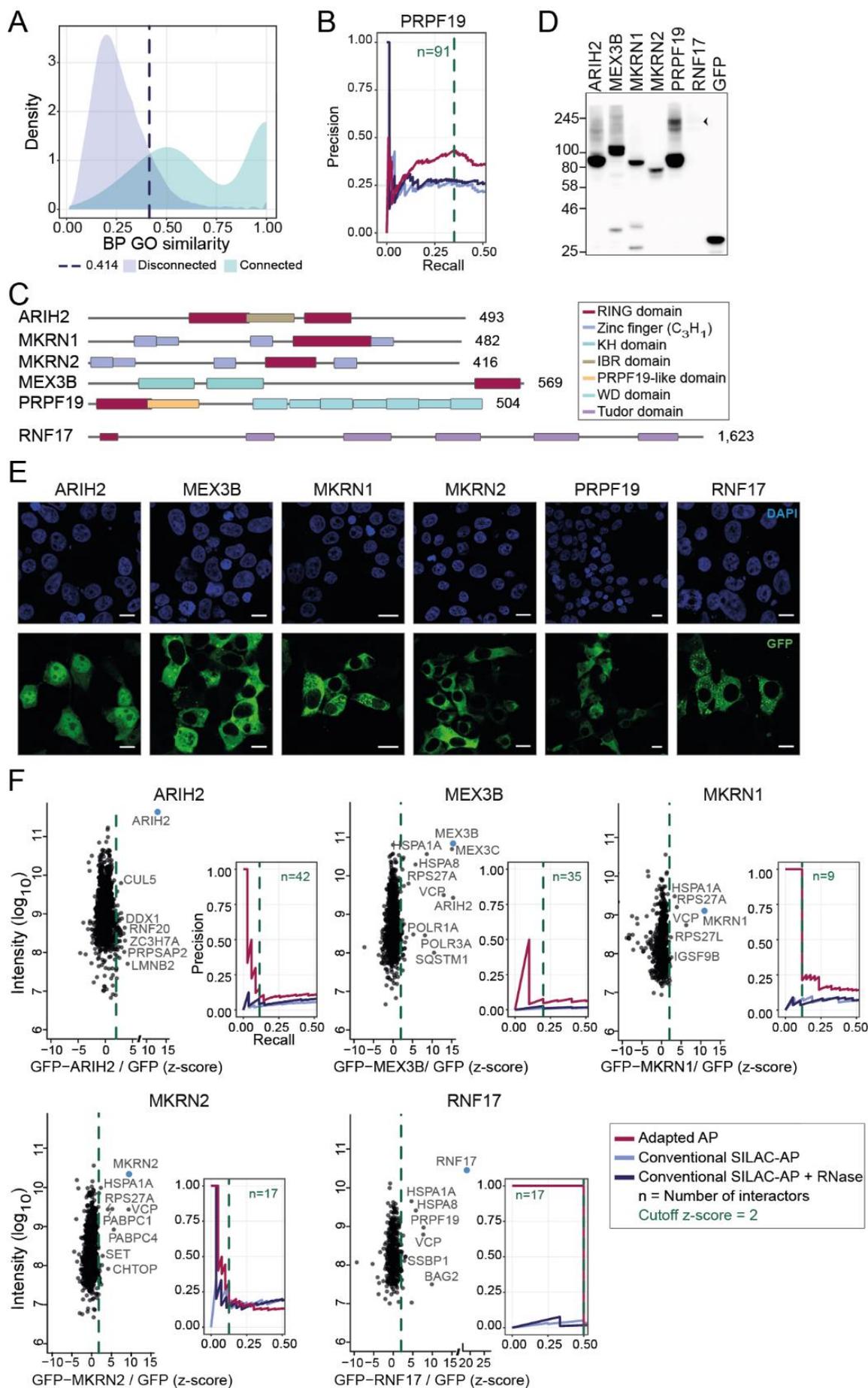


Figure III-2, Experiments following the adapted AP approach and GO similarity analyses for six RBULs. **a.** Benchmarking of the GO similarity strategy by comparing the WI distribution of interacting protein pairs from the INstruct database (turquoise) to the control set of disconnected protein pairs (lilac). Applying a 10% FDR, a 'true positive' interactor is defined by a WI of ≥ 0.414 (dashed blue line). **b.** Precision-Recall curves of the conventional SILAC-AP data without (blue) or with RNase treatment (dark blue), and the adapted AP data (red) for GFP-PRPF19. Detected proteins were ranked according to their SILAC ratios, and precision and recall of GO evidence-supported interactors were calculated based on the similarity of BP GO terms between bait and prey proteins ($WI \geq 0.414$). The dashed green line indicates applied cut-off at $z\text{-score} \geq 2$. n, number of PRPF19 interactors. Legend shown in (f). **c.** Predicted domain architecture of the six tested RBULs. Domains are indicated on the right. **d.** Expression of GFP-RBULs measured by Western blot. GFP-tagged RBULs were transiently expressed in HEK293T cells, and their expression analyzed by Western blot. A representative experiment of two biological replicates is shown. GFP-RNF17 is highlighted by an arrowhead; a longer exposure is shown in Supplementary Figure 3c. All GFP-tagged RBULs are detected at the expected size. **e.** Subcellular localization of GFP-tagged RBULs. GFP-tagged RBULs were expressed in HEK293T cells and their subcellular localization was analyzed by confocal microscopy. Scale bars indicate 10 μm . **f.** Ratio-intensity plots of adapted AP experiments and Precision-Recall curves for the RBULs ARIH2, MEX3B, MKRN1, MKRN2, and RNF17 are shown. Combined SILAC ratios from three independent biological replicates are plotted against \log_{10} transformed intensities. Bait proteins (blue) and selected preys with high SILAC ratios are indicated.

with translation initiation factors, including EIF4G1 (PRPF19 & MEX3B) and EIF1AX (ARIH2), and ribosomal proteins, such as RPLP2 (RNF17), RPS27L (MKRN1), RPL37A, and RPS12 (MKRN2). More generally, comparison of the RBUL interaction profiles with several large-scale mRNA interaction profiling screens¹⁻³ reveals that 68 out of the total of 170 interactors (40%) have been found in direct contact with poly(A)+ RNA, underlining the strong commitment of the studied RBULs to RNA-mediated processes.

Interestingly, a fraction of the identified interactors is involved in other cellular processes, such as protein transport, cell cycle or DNA damage repair ('Other', Fig. 4a). Additionally, the stable interaction of PRPF19 and other RBULs with RNA polymerase components predicts a function in transcriptional regulation (POLR2A, POLR2B [PRPF19], POLR1A, POLR1C, POLR2H [MEX3B], POLR3A [PRPF19, MEX3B & ARIH2]; Fig. 3). As PRPF19 has been previously implicated in the regulation of transcription^{34,35}, this observation underlines the capacity of our approach to identify physiologically functional networks. In line with a putative role in transcriptional regulation, we also identify multiple components of the prefoldin-like complex as stable PRPF19 interactors (PFDN1, PFDN2, VBP1/PFDN3, PFDN4 and PFDN6). Prefoldin acts as a co-chaperone in the cytoplasmic assembly of cytoskeletal and non-cytoskeletal complexes, but also modulates transcriptional activation when present in the nucleus³⁶. Moreover, we find multiple proteins that are involved in the DNA damage response to interact with PRPF19, such as DDB1, MDC1, CDC5L, SNW1 and DNAJA1. In line with this notion, PRPF19 has been recently implicated in the DNA damage response by ubiquitylation of the single-strand DNA-binding protein RPA³⁷. Finally, we

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

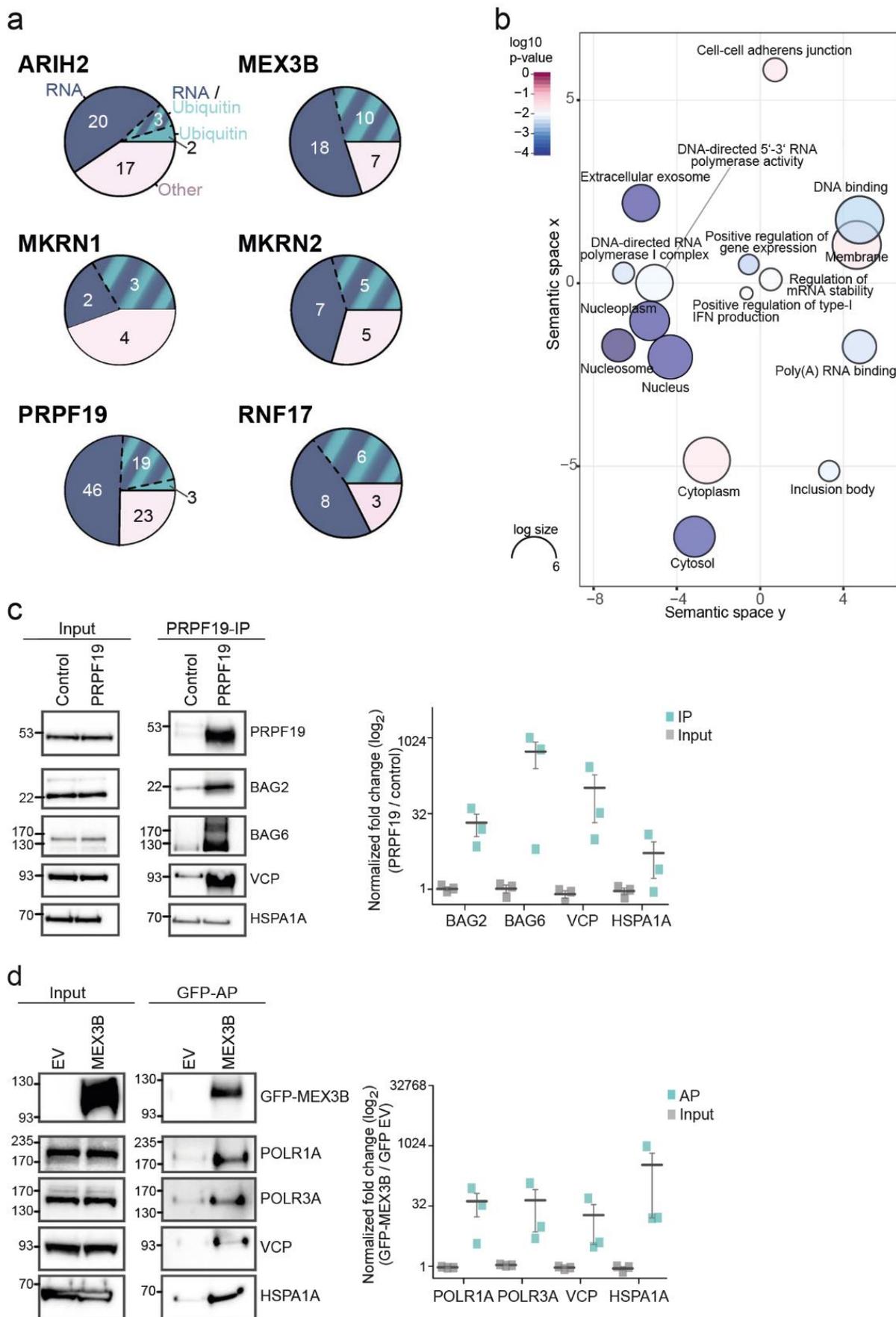


Figure III-4, The RBULs link posttranscriptional processes to the ubiquitin system. a. The BP and MF GO terms of the interactors of the six RBULs were grouped into the categories 'RNA' (blue), 'Ubiquitin' (turquoise), and 'Other' (rose). The distribution of the categories among the interactomes is shown. **b.** A summary of GO terms for MEX3B interaction partners is shown. **c & d.** Validation of RBUL interaction partners by pulldowns and Western blot. **c.** Endogenous PRPF19 was pulled down with a PRPF19-specific antibody from HEK293T cells. Experiments omitting the antibody served as control. Western blot analysis was performed with antibodies specific against BAG2, BAG6, VCP, and HSPA1A, as well as against PRPF19 itself to validate the immunoprecipitation (IP). Left: Cropped images of input and IP samples (replicate 1). Images of full membranes and different exposure times for all antibodies and replicates are presented in Supplementary Figure 4a-c. Right: Quantifications of the PRPF19-specific IPs normalized to control of three independent biological replicates are shown in a dot plot, including mean and standard error (s.e.m.). **d.** GFP (empty vector, EV) and GFP-MEX3B were expressed in HEK293T cells and pulled down with a GFP-specific antibody. Western blot analysis was performed using specific antibodies against POLR1A, POLR3A, VCP, and HSPA1A, as well as GFP. Left: Cropped images of input and AP samples (replicate 1). Images of full membranes and different exposure times for all antibodies are presented in Supplementary Figure 4d-f. Right: Quantification of the APs normalized to EV are shown as in (c).

Taking MEX3B as an example, we used the DAVID database to assess coherent functions of its interactors⁴⁰. Consistent with previous reports^{41,42}, we find that the interactors of MEX3B are involved in biological processes such as 'regulation of mRNA stability' (adjusted p -value < 0.05, Benjamini-Hochberg correction) (Fig. 3, 4b, Suppl. Table 4). Among the GO Molecular Function (MF) terms, 'poly(A) RNA binding' appears as a predominant term, reflecting the stable association of MEX3B with proteins involved in RNA metabolism, such as DHX36 and TNRC6A. Finally, a significant fraction of the identified interactors suggest a direct role of MEX3B in the regulation of transcription, including subunits of RNA polymerases (pol) I, II and III, regulatory proteins such as the transcriptional repressor ZGPAT, as well as multiple histone-associated proteins. We validated the interactions with two RNA pol I and III subunits for GFP-MEX3B (Fig. 4d). Altogether, these findings place MEX3B and the other RBULs into the context of several pathways of transcriptional and posttranscriptional gene regulation, forming a basis for future studies regarding the role of RBULs in these cellular processes. In summary, we were able to identify novel protein-protein interaction partners of six human RBULs and link them to cellular pathways by applying our adapted AP protocol followed by bioinformatics analyses.

Discussion

In this study, we analyzed the interaction profiles of six human RBULs containing different RNA-binding domains using a modified AP approach coupled to LC-MS/MS. Previous studies have suggested that RBPs are particularly prone to engage in non-physiological interactions that hamper the identification of physiological interaction partners^{7,8}. For instance, the reassembly of ribonucleoprotein complexes in cell lysates can lead to false positive results when studying the *in vivo* interaction partners of RBPs^{7,43}. Moreover, the disruption of cellular compartmentalization during cell lysis promotes artificial associations that would be impossible *in vivo* due to spatial separation⁴⁴. The propensity of RBPs for re-assortment is augmented by the prevalence of disordered regions in these proteins, which often contain repetitive sequences and low complexity domains. Such disordered regions can facilitate RNA recognition but also promote association with other molecules^{3,6}. In line with this notion, it has been found that many RBPs undergo a concentration-dependent phase transition into a hydrogel-like state^{8,45,46}. In the cellular context, this might play an important role in driving the assembly of non-membrane-bounded subcellular structures, such as RNA granules or liquid like clusters. However, this biophysical behavior is likely to interfere with reliable proteomic characterizations, as the granules reversibly aggregate and disaggregate in solution which may entrap proteins in non-physiological associations^{46,47}. Altogether, the biochemical and biophysical properties of RBPs make them challenging targets for unbiased protein-protein interaction profiling.

In order to overcome these obstacles, we modified the conventional SILAC-based AP workflow to make it more suitable for RBPs. By combining the differentially SILAC-labeled cell populations prior to cell lysis and AP, the SILAC ratios become indicative of the stability of the detected interaction. Only kinetically stable interactions will maintain high SILAC ratios, while any association that newly forms or changes in the lysate will display an equalized/background SILAC ratio. Our approach thereby advances the mixing of samples to an earlier step compared to PAM-SILAC, which is commonly employed to study stable protein interactions⁴⁸. Previous studies established that a mixing of samples at different stages of the AP protocol allows to reliably distinguish between stable and transient interactions, a strategy that was successfully applied e.g. to dissect the dynamic interactome of transcription factor complexes⁴⁹ or the human 26S proteasome⁹.

In order to facilitate quick, clean and reproducible immunoprecipitations, our adapted AP approach relies on the ectopic expression of GFP-tagged bait proteins. GFP has been evaluated as a suitable tag of choice for quantitative proteomics due to its minimal non-specific binding to cellular proteins³⁰. Moreover, it can be combined with control cells expressing unfused GFP which has been suggested as an important internal control to further eliminate non-specific contaminants⁵⁰. Our Western blot experiments show that the GFP-tagged RBULs display a certain deviation in expression levels compared to the endogenous counterpart (37-fold less and 13-fold more for PRPF19 and MEX3B, respectively; Suppl. Fig. 3a), which might partially impair protein folding and function. Accordingly, we detect recurrent interactions with several HSP70 family members, which can function as disaggregases for newly translated proteins and were previously reported as recurrent GFP-associated contaminants^{29,30}. Nevertheless, we reliably recover a large number of known physiological interaction partners, indicating that the overall functionality of the GFP-tagged RBULs is preserved. Moreover, the GFP-tagged RBULs mostly localize to the correct cellular compartment, supporting the notion that the majority of the tagged proteins is properly folded and delivered. It is important to note that our protocol is also compatible with standard antibodies against endogenous proteins which usually support an efficient immunoprecipitation on a similar timescale. This would allow to circumvent the ectopic expression of tagged bait proteins in future applications.

In order to validate the performance of our approach, we used the well-characterized RBUL PRPF19 as an example. PRPF19 has been described to engage in at least three different sub-complexes that confer distinct cellular functions in human cells⁵¹. Using the adapted AP, we recover the interactions of PRPF19 with the NTC core complex as well as components of the XAB2 complex, a second PRPF19 sub-complex involved in genome maintenance³⁴. Moreover, we identify multiple components of the ubiquitin-proteasome system such as VCP, HSP90AB1, UBL4A, GET4 as well as several proteasome subunits (PSMC1, PSMC2, PSMC5, PSMC6, PSMD2). Although we do not detect PSMB4, a subunit of the catalytic core of the proteasome that was previously reported to interact with PRPF19, the identified interaction partners strongly support the idea of a proteasome-dependent function of PRPF19-mediated ubiquitylation. The interactions with proteasome components were previously interpreted as a potential link of PRPF19 to protein degradation via substrate delivery to the proteasome⁵¹. Finally, the observation that PRPF19 interacts with the BAT3 complex, which we independently validated for both the GFP-tagged as well as the endogenous protein, suggests a role in

chaperoning and targeting proteins within the cytosol³⁸. In summary, the recovery of known PRPF19 interactors supports the validity of our approach to identify physiological protein interactions from the native environment of the cell.

The identified protein-protein interactions could be regulators as well as ubiquitylation substrates of the RBUL. For instance, PRPF19 was previously shown to transfer a non-proteolytic K63-linked ubiquitin chain to PRPF3 within the U4 snRNP, thereby stabilizing the U4/U6.U5 tri-snRNP during spliceosome remodeling^{51,52}. Subsequent de-ubiquitylation is a prerequisite for spliceosome activation. In the adapted AP, we do not identify PRPF3; however, PRPF8 as well as several other components of the tri-snRNP are found (such as SNRPE, SNRPD3, SNRNP40 and SNRNP200), which might represent regulators or substrates of PRPF19. To distinguish ubiquitylation substrates from regulators of protein function, further experiments including ubiquitin remnant profiling in ubiquitin ligase knockdown cells or pulldowns using catalytically dead mutants could be performed^{53,54}.

Using the adapted AP, we characterized the core interactomes of five additional RBULs, including ARIH2/TRIAD1, MEX3B, MKRN1, MKRN2 and RNF17. In total, we detected 170 interactors, which will be a valuable resource for future studies on RBUL functions in human cells. In line with the composite domain architecture, our results suggest that all tested RBULs interact both with components of the ubiquitin system as well as with RBPs involved in different posttranscriptional processes (116 out of 170 interactors belonging to either category; 68%). For example, the hnRNP family of RNA-binding proteins has been reported to interact with ubiquitin ligases such as TRIM28, PRPF19, and CUL1⁵⁵. In accordance, we find PRPF19 in stable association with HNRNPC.

There are multiple known examples of regulatory associations between ubiquitylation and RNA. On the one hand, ubiquitylation can control the activity of large RNA processing machineries, as mentioned above for the modification of PRPF3 by PRPF19 during spliceosome remodeling. Similarly, non-proteolytic ubiquitylation of CNOT7 by the RBUL MEX3C activates de-adenylation by the CCR4-NOT machinery, which is the rate-limiting step in eukaryotic mRNA decay. Notably, the ubiquitin modification has no effect on basal de-adenylation activity of CNOT7 but is important for the decay of specific mRNAs that are recognized by MEX3C on the level of RNA binding⁵⁵. The RNA-binding/ubiquitylation-mediated CCR4-NOT activation thus introduces an additional layer of specificity for distinct RNAs that might be relevant under certain cellular conditions such as during cell differentiation or disease⁵⁶. On the other hand, a

reverse relationship has also been observed in which the RNA functions as a regulatory molecule to control activities within the ubiquitin system. For instance, RNA binding of the Roquin paralog RC3H2 influences its E3 ubiquitin ligase function *in vitro*⁵⁷. Moreover, the long non-coding RNA *HOTAIR* was found to act as a molecular assembly platform for protein ubiquitylation⁵⁸. This is achieved through adjacent RNA binding of the RBULs DZIP3 and MEX3B together with their respective client proteins Ataxin-1 and Snuportin-1, which facilitates their ubiquitylation and accelerates their degradation, e.g. during cellular senescence. It is conceivable that similar mechanisms are at work to control transcription and translation.

Notably, we find that RBULs form stable interactions with RNA polymerase subunits as well as with components of the translation machinery. In line with the notion that the RBULs could be involved in translational regulation, several studies demonstrated a role for ubiquitin as a modulator of translation. For instance, K63-linked poly-ubiquitylation of ribosomal proteins and translation elongation factors was recently shown to promote translation during oxidative stress in baker's yeast⁵⁹. Similarly, 40S ribosomal proteins in human cells were found to be ubiquitylated due to the activation of the unfolded protein response and inhibition of translation⁶⁰. Moreover, the E3 ubiquitin ligase CUL3 has been implicated in the ubiquitin-dependent formation of a ribosome modification platform that alters the translation of specific mRNAs⁶¹. Future studies will be needed to understand the molecular mechanisms by which RBULs link ubiquitylation to posttranscriptional processes. Our results provide evidence for an extensive crosstalk between these two modes of gene regulation.

In summary, we conclude that the adapted AP approach is capable of detecting so far unknown protein-protein interactions of RBPs and enables embedding those protein-protein interactions into a functional context. It proves particularly valuable for studying RBPs that are prone to engage into non-physiological interactions during cell lysis. Our approach thereby adds a useful tool to complement existing protocols.

Material and methods

Cell culture

HEK293T cells were purchased from DSMZ (Catalog no. ACC 635) and maintained in DMEM (Life Technologies, 21969035) complemented with 1% penicillin/streptomycin (Life Technologies, 15140-122), 1% L-glutamine (Life Technologies, 25030-029) and 10% fetal bovine serum (Life Technologies, A15-101). All cells were cultured at 37°C in a humidified incubator containing 5% CO₂. For SILAC labeling, cells were cultured in media containing either L-arginine and L-lysine (light SILAC label), L-arginine (¹³C₆) and L-lysine (²H₄) (medium SILAC label), or L-arginine (¹³C₆-¹⁵N₄) and L-lysine (¹³C₆-¹⁵N₂) (heavy SILAC label) (Cambridge Isotope Laboratories).

Vectors/plasmids

The following entry vectors, suitable for Gateway Cloning, were obtained from the IMB Core Facility ORFeome Collection⁶²: pENTR221-ARIH2, pENTR201-MEX3B, pENTR221-MKRN1, pENTR221-MKRN2, pENTR221-PRPF19 and pENTR223.1-RNF17. Coding sequences of six RNA-binding ubiquitin ligases were cloned into the mammalian expression vector pMX-DEST53-IP-GFP by LR Gateway cloning according to manufacturer's recommendations (Gateway® LR Clonase® II Enzyme mix; Life Technologies, 11791-100), then ectopically expressed in HEK293T cells using Polyethylenimine MAX 4000 (Polysciences, 24885-2).

Immunoprecipitations

For GFP-APs, HEK293T cells transiently expressing GFP empty vector or a GFP-tagged RBUL were used. For endogenous APs, untransfected HEK293T cells were used. The cells were washed with ice-cold phosphate-buffered saline (PBS) and lysed in modified RIPA (mRIPA) buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% sodium deoxycholate). The mRIPA buffer was supplemented with protease inhibitors (protease inhibitor cocktail, Sigma). Lysates were cleared at 16,000 × g for 15 min. Protein concentrations of the cleared lysates were estimated by BCA Protein Assay (Thermo Scientific). GFP-trap agarose beads (Chromotek) were washed in mRIPA buffer before they were incubated with the cleared lysate for 1 h at 4°C. After three washes with NET buffer and three washes with ultrapure water, the

beads were resuspended in LDS sample buffer (Life Technologies) and boiled at 70°C for 10 min.

Western blotting

Denatured proteins were separated by SDS-PAGE on a NuPAGE 4-12% Bis-Tris protein gel (Life Technologies) and transferred to a 0.45 µm nitrocellulose membrane (VWR). For detection, HRP-conjugated secondary antibodies and WesternBright Chemiluminescent Substrate (Biozym Scientific) or SuperSignal West Pico Chemiluminescent Substrate (Life Technologies) were used. Western blots were quantified by determining the background-subtracted densities of the protein of interest using Image J⁶³. The signal from the IP (against endogenous PRPF19) or AP (against GFP-tagged PRPF19 and MEX3B) was normalized to the respective control samples omitting the antibody or expressing the empty vector, respectively.

Antibodies

The following antibodies were used: anti-GFP (B-2 clone; Santa Cruz; sc-9996), anti-HSP70 (Enzo Life Sciences; N15F2-5), anti-BAG6 (Cell Signaling Technology, 8523), anti-VCP (Cell Signaling Technology; 2649), anti-BAG2 (Sigma Alrich; HPA018862), anti-PRPF19 (Abcam; ab27692), anti-POLR1A (Santa Cruz; sc-48385), anti-POLR3A (D5Y2D; Cell Signaling Technology, 12825).

Immunofluorescence microscopy

HEK293T cells were seeded on microscopy cover slips and transfected with GFP-tagged RBULs using Polyethylenimine 'Max' (Mw 4,000) (Polysciences Inc). Cells were washed with ice-cold PBS and fixed with 4% paraformaldehyde (Affymetrix) for 20 min. Cells were stained with DAPI (Sigma) and rinsed in PBS, Wash Buffer (10 mM Tris-HCl, pH 7.5) and water. Cover slips were mounted with ProLong Diamond Antifade Mountant (Life Technologies). Images were acquired with the TCS SP5 Confocal Microscope 546 (Leica) using a 63x oil objective, and four pictures were taken per frame for each RBUL. Image analysis was performed in Fiji⁶⁴.

Sample preparation for the conventional SILAC workflow

HEK293T cells transiently expressing GFP were cultured in light SILAC medium, while cells expressing GFP-tagged RBULs were cultured in medium and heavy SILAC medium. In the conventional SILAC-AP experiments, enriched protein fractions extracted from GFP-RBUL expressing heavy isotope labeled cells were used for RNase digestion. The cells were washed with ice-cold PBS and lysed in mRIPA buffer. The mRIPA buffer was supplemented with protease inhibitors, 10 mM N-ethylmaleimide (NEM), 1 mM sodium orthovanadate, 5 mM β -glycerophosphate, and 5 mM sodium fluoride. Lysates were cleared at 16,000 \times g for 15 min. The Pierce BCA Protein Assay Kit (Thermo Fisher) was used to estimate protein concentrations of the cleared lysates. GFP-trap agarose beads were washed in mRIPA buffer before they were incubated with the cleared lysate for 1 h at 4°C. After two washes in NET buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.1% Triton X-100), the heavy SILAC labeled lysate was incubated with 0.5 U/ μ l RNase A (Qiagen) and 20 U/ μ l RNase T1 (Thermo Fisher Scientific) for 30 min at 4°C, while the light and medium SILAC labeled protein precipitates were kept on ice. All three differentially SILAC-labeled pulldown samples were washed twice with NET buffer and twice with ultrapure water in parallel, then combined, and washed once more with ultrapure water. The beads were resuspended in digestion buffer and incubated with 10 mM dithiothreitol (DTT; Sigma, D5545) for 20 min at RT and with 50 mM 2-Chloroacetamide (CAA; Sigma, C0267) for 20 min at RT in the dark²⁰.

Sample preparation for the adapted SILAC-based AP protocol

HEK293T cells transiently expressing GFP were cultured in light SILAC medium, while cells expressing GFP-tagged RBULs were cultured in heavy SILAC medium. Cells were washed with ice-cold phosphate-buffered saline. Light and heavy SILAC labeled cells were mixed in equal amounts. Then, the combined cell fraction was lysed in mRIPA (as described above). Lysates were cleared at 16,000 \times g for 15 min. Pierce BCA Protein Assay Kit was used to estimate protein concentrations of the cleared lysates. GFP-trap agarose beads were washed in mRIPA buffer before they were incubated with the cleared lysate for 1 h at 4°C. After three washes with NET buffer and three washes with ultrapure water, the beads were resuspended in digestion buffer and incubated with 10 mM DTT for 20 min at RT and with 50 mM CAA for 20 min at RT in the dark²⁰.

MS sample preparation

The enriched proteins were digested on-bead with trypsin for 2 h at RT. Supernatants were transferred to into new microcentrifuge tubes, while the beads were re-incubated with digestion buffer for 5 min at RT. Afterwards, supernatants were combined with the supernatants collected before and trypsin was added to continue protein digestion overnight at RT²⁰. For two PRPF19 replicates (replicates 1 and 2 of the conventional SILAC-AP and the adapted AP), proteins were resolved by SDS-PAGE on a NuPAGE 4-12% Bis-Tris protein gel (Thermo Fisher Scientific) and submitted to in-gel digest using trypsin. Subsequently, peptides were extracted from the gel. To concentrate, clear and acidify the peptides, they were bound to C18 StageTips as described previously^{65,66}.

MS analysis

Peptide fractions were analyzed on a quadrupole Orbitrap mass spectrometer (Thermo Q Exactive Plus, Thermo Scientific) coupled to an UHPLC system (EASY-nLC 1000, Thermo Scientific)⁶⁷. Peptide samples were separated on a C18 reversed phase column (length: 20 cm, inner diameter: 75 μ m, bead size: 1.9 μ m) and eluted in a linear gradient from 8 to 40% acetonitrile containing 0.1% formic acid in 105 min. The mass spectrometer was operated in data-dependent mode, automatically switching between MS and MS² acquisition. In the Orbitrap, the full scan MS spectra (m/z 300 - 1650) was acquired. Sequential isolation and fragmentation of the ten most abundant ions was performed by higher-energy collisional dissociation (HCD)⁶⁸. Peptides with unassigned charge states, as well as with charge states less than +2 were excluded from fragmentation. The Orbitrap mass analyzer was used for acquisition of fragment spectra.

Peptide identification and quantification

Using the MaxQuant software (version 1.5.28), raw data files for each RBUL (conventional and adapted AP) were analyzed and peptides were identified²⁴. Parent ion and MS² spectra were compared to a database containing 92,578 human protein sequences obtained from UniProtKB, released in May 2016, coupled to the Andromeda search engine⁶⁹. Cysteine carbamidomethylation, oxidation, and NEM were set as fixed modifications. The mass tolerance for the spectra search was set to be lower than 6 ppm in MS and 20 ppm in HCD MS² mode. Spectra were also searched with strict

trypsin specificity and allowing up to three miscleavages. Site localization probabilities were determined by MaxQuant using the PTM scoring algorithm as described previously^{21,69}. Filtering of the dataset was based on the posterior error probability to arrive at a false discovery rate below 1% estimated using a target-decoy approach⁷⁰. Proteins that were categorized as "only identified by site", known contaminants and reverse hits were removed. Only proteins identified with at least two peptides (including at least one unique peptide) and a SILAC ratio count at least two were used for analysis. The SILAC ratios were \log_2 transformed and converted into an asymmetric z-score based on the mean and interquartile range of the distribution as described previously²¹. In order to extract interactors for the six RBULs, we \log_2 transformed the SILAC ratios. For proteins that were detected in at least two replicate experiments for a given RBUL and AP protocol, we applied a cut-off at mean z-score ≥ 2 to identify putative interaction partners^{21,22}. Interactions from the conventional SILAC-AP experiments were additionally classified as RNA-dependent or RNA-independent based on z-score ≤ -1 or > -1 , respectively, when comparing GFP-RBUL + RNase vs. GFP-RBUL SILAC ratios determined by conventional SILAC-AP. Ratio-Intensity plots were created in R (version 3.2.3).

GO similarity approach and GO enrichment analysis

The guilt-by-association principle states that proteins that interact with each other usually participate in the same biological process or pathway²⁴. Under this premise, we considered the interactions between RBULs and putative protein partners as biologically meaningful if the Gene Ontology (GO) similarity between the putative interactors was high. We used the R package GOSemSim²⁵ to compute GO similarities between RBULs and their interaction partners based on the Wang's Index (WI) comparing their Biological Process (BP) GO terms (state June 2017). WI was developed specifically for GO and emits similarities that are consistent with manually curated GO-gene associations²⁶. To benchmark our approach, we calculated the WI distribution of interacting proteins in INstruct, a high-quality database of structurally resolved protein interactions, containing 3,354 proteins (N) with 6,093 edges (L, interactions) between them^{27,28}. As a negative control, we used all disconnected protein pairs that are not connected by edges within the INstruct network ($N*(N-1)/2-L = 5,616,888$ non-edges). As expected, the BP GO similarity of connected protein pairs in the INstruct network is significantly higher than that of disconnected pairs (p-value $< 2.2 \times 10^{-16}$, Mann-Whitney

U test). Based on the control distribution, we estimated an empirical false discovery rate (FDR) and a cut-off at $FDR < 10\%$, corresponding to $WI = 0.414$ (Fig. 2a).

In order to construct the Precision-Recall curves, we sorted the list of RBUL-interactor pairs in decreasing order by their normalized SILAC ratios from a given AP approach, and considered an interaction as ‘*true*’ i.e. functionally coherent if its WI was at or above 0.414. We then progressively went over the entire range of ratios, assessing an increasing number of protein interaction pairs (using stepwise increase of one interaction pair in each iteration). In each iteration, the fraction of *true* protein pairs from the total determines the ‘Precision’, while the fraction of *true* protein pairs at the current step from the total of *true* interactions in the list determines the ‘Recall’. After the evaluation of the full range of SILAC ratios, all Precision-Recall value pairs were used to construct a curve that measures the ability of the AP approach to preferentially assign high SILAC ratios to functionally coherent interactions. If no GO term was available for a prey, the WI was set to -1.

All computational analyses were executed on a Lenovo ThinkPad 64-bit with 7.7 GB of RAM and an Intel Core i7-4600U CPU @ 2.10 GHz × 4, running Ubuntu 16.04 LTS.

Functional annotation of RBUL interactors

To assess the involvement of interacting proteins in posttranscriptional (‘RNA’) or in ubiquitin-mediated processes (‘Ubiquitin’, Fig. 4a) we used a manually curated list of GO terms for ‘Biological Process’, ‘Molecular Function’ and ‘Cellular Component’. GO enrichment analyses were performed using the DAVID (Database for Annotation, Visualization and Integrated Discovery) tool for all three GO domains⁴⁰. Enriched GO terms (adjusted p -value < 0.05 , Benjamini-Hochberg correction) were visualized using REVIGO⁷¹ (Reduce & Visualize Gene Ontology), allowing medium GO term similarity (Fig. 4b). All identified RBUL-interactor pairs with a z -score ≥ 2 from the adapted AP (170 in total) were queried against the Human Integrated Protein-Protein Interaction Reference (HIPPIE) database²³. Out of 170 interactions in total, only 19 were reported to date. HIPPIE confidence scores, number of interactions reported for each of the six RBULs in HIPPIE and the kind of experiment used to measure in combination with PMIDs and reporting databases were extracted from HIPPIE.

Data availability

All datasets generated in the current study are available from the corresponding author on request.

Author contributions

A.H. performed all mass spectrometry experiments. G.A. and M.A.A. developed a method for evaluation of the adapted AP. G.A. performed computational functional analyses. J.K. and P.B. designed and supervised the experiments. P.B., J.K., G.A., and A.H. analyzed the data. A.V. and A.H. performed the Western Blot validations. K.Z. contributed ideas; and K.Z. and A.H. wrote the manuscript. All authors read and commented on the manuscript.

The authors have declared no conflict of interest nor competing financial interests.

Acknowledgments

We thank Anja Freiwald for the assistance with mass spectrometry. The authors thank the IMB Microscopy Core Facility for technical support. K.Z. was supported by the LOEWE program Ubiquitin Networks (Ub-Net) of the State of Hesse (Germany) and the SFB 902 of the German Research Foundation. P.B. was supported by the Emmy Noether Program (BE 5342/1-1), the SFB 1177 of the German Research Foundation and the Marie Curie Career Integration Grant from the European Commission (grant agreement number: 630763). J.K. was supported by the SPP 1935 of the German Research Foundation (KO 4566/3-1).

References

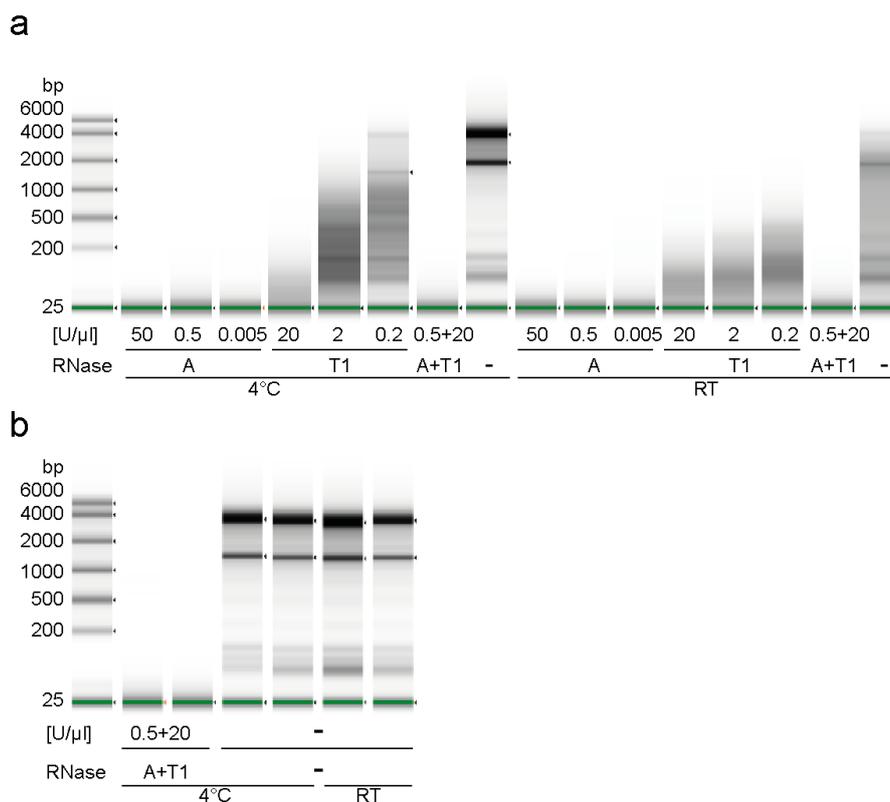
1. Baltz, A.G. et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* 46, 674-690 (2012).
2. Beckmann, B.M. et al. The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun* 6, 10127 (2015).
3. Castello, A. et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393-1406 (2012).
4. Cano, F., Miranda-Saavedra, D. & Lehner, P.J. RNA-binding E3 ubiquitin ligases: novel players in nucleic acid regulation. *Biochem Soc Trans* 38, 1621-1626 (2010).
5. Ong, S.E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1, 376-386 (2002).
6. Lunde, B.M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* 8, 479-490 (2007).
7. Mili, S. & Steitz, J.A. Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 10, 1692-1694 (2004).

8. Weber, S.C. & Brangwynne, C.P. Getting RNA and protein in phase. *Cell* 149, 1188-1191 (2012).
9. Wang, X. & Huang, L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol Cell Proteomics* 7, 46-57 (2008).
10. Puig, O. et al. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24, 218-229 (2001).
11. Roux, K.J., Kim, D.I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196, 801-810 (2012).
12. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat Protoc* 11, 456-475 (2016).
13. Raz, V. et al. A novel feed-forward loop between ARIH2 E3-ligase and PABPN1 regulates aging-associated muscle degeneration. *Am J Pathol* 184, 1119-1131 (2014).
14. Ong, S.E., Foster, L.J. & Mann, M. Mass spectrometric-based approaches in quantitative proteomics. *Methods* 29, 124-130 (2003).
15. Makarova, O.V. et al. A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J* 23, 2381-2391 (2004).
16. Grote, M. et al. Molecular architecture of the human Prp19/CDC5L complex. *Mol Cell Biol* 30, 2105-2119 (2010).
17. Chan, S.P., Kao, D.I., Tsai, W.Y. & Cheng, S.C. The Prp19p-associated complex in spliceosome activation. *Science* 302, 279-282 (2003).
18. Ohi, M.D. & Gould, K.L. Characterization of interactions among the Cef1p-Prp19p-associated splicing complex. *RNA* 8, 798-815 (2002).
19. Thul, P.J. et al. A subcellular map of the human proteome. *Science* 356 (2017).
20. Baymaz, H.I., Spruijt, C.G. & Vermeulen, M. Identifying nuclear protein-protein interactions using GFP affinity purification and SILAC-based quantitative mass spectrometry. *Methods Mol Biol* 1188, 207-226 (2014).
21. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367-1372 (2008).
22. Brannan, K.W. et al. SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell* 64, 282-293 (2016).
23. Alanis-Lobato, G., Andrade-Navarro, M.A. & Schaefer, M.H. HIPPIE v2.0: enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Res* 45, D408-D414 (2017).
24. Oliver, S. Guilt-by-association goes global. *Nature* 403, 601-603 (2000).
25. Yu, G. et al. GOSemSim: an R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* 26, 976-978 (2010).
26. Wang, J.Z., Du, Z., Payattakool, R., Yu, P.S. & Chen, C.F. A new method to measure the semantic similarity of GO terms. *Bioinformatics* 23, 1274-1281 (2007).
27. Wang, X. et al. Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat Biotechnol* 30, 159-164 (2012).
28. Meyer, M.J., Das, J., Wang, X. & Yu, H. INstruct: a database of high-quality 3D structurally resolved protein interactome networks. *Bioinformatics* 29, 1577-1579 (2013).
29. Mellacheruvu, D. et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat Methods* 10, 730-736 (2013).
30. Trinkle-Mulcahy, L. et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol* 183, 223-239 (2008).
31. Röst, H.L., Malmström, L. & Aebersold, R. Reproducible quantitative proteotype data matrices for systems biology. *Mol Biol Cell* 26, 3926-3931 (2015).
32. de Bie, P. & Ciechanover, A. Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ* 18, 1393-1402 (2011).
33. Meyer, H., Bug, M. & Bremer, S. Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat Cell Biol* 14, 117-123 (2012).
34. Kuraoka, I. et al. Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J Biol Chem* 283, 940-950 (2008).
35. David, C.J., Boyne, A.R., Millhouse, S.R. & Manley, J.L. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev* 25, 972-983 (2011).
36. Millan-Zambrano, G. & Chavez, S. Nuclear functions of prefoldin. *Open Biol* 4 (2014).
37. Marechal, A. et al. PRP19 transforms into a sensor of RPA-ssDNA after DNA damage and drives ATR activation via a ubiquitin-mediated circuitry. *Mol Cell* 53, 235-246 (2014).
38. Mariappan, M. et al. A ribosome-associating factor chaperones tail-anchored membrane proteins. *Nature* 466, 1120-1124 (2010).

39. Dai, Q. et al. Regulation of the cytoplasmic quality control protein degradation pathway by BAG2. *J Biol Chem* 280, 38673-38681 (2005).
40. Jiao, X. et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* 28, 1805-1806 (2012).
41. Takada, H. et al. The RNA-binding protein Mex3b has a fine-tuning system for mRNA regulation in early *Xenopus* development. *Development* 136, 2413-2422 (2009).
42. Pereira, B., Le Borgne, M., Chartier, N.T., Billaud, M. & Almeida, R. MEX-3 proteins: recent insights on novel post-transcriptional regulators. *Trends Biochem Sci* 38, 477-479 (2013).
43. Stubbs, S.H. & Conrad, N.K. Analysis of RNA-protein interactions by cell mixing. *Methods Enzymol* 539, 67-80 (2014).
44. Kittur, N., Darzacq, X., Roy, S., Singer, R.H. & Meier, U.T. Dynamic association and localization of human H/ACA RNP proteins. *RNA* 12, 2057-2062 (2006).
45. Kato, M. et al. Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels. *Cell* 149, 753-767 (2012).
46. Lin, Y., Protter, D.S., Rosen, M.K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell* 60, 208-219 (2015).
47. Kwon, I. et al. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* 345, 1139-1145 (2014).
48. Hein, M.Y. et al. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* 163, 712-723 (2015).
49. Mousson, F., Kolkman, A., Pijnappel, W.W., Timmers, H.T. & Heck, A.J. Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol Cell Proteomics* 7, 845-852 (2008).
50. Trinkle-Mulcahy, L. et al. Repo-Man recruits PP1 gamma to chromatin and is essential for cell viability. *J Cell Biol* 172, 679-692 (2006).
51. Chanarat, S. & Strasser, K. Splicing and beyond: the many faces of the Prp19 complex. *Biochim Biophys Acta* 1833, 2126-2134 (2013).
52. Song, E.J. et al. The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* 24, 1434-1447 (2010).
53. Wagner, S.A. et al. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* 10, M111 013284 (2011).
54. Heidelberg, J.B., Wagner, S.A. & Beli, P. Mass Spectrometry-Based Proteomics for Investigating DNA Damage-Associated Protein Ubiquitylation. *Front Genet* 7, 109 (2016).
55. Cano, F., Rapiteanu, R., Sebastiaan Winkler, G. & Lehner, P.J. A non-proteolytic role for ubiquitin in deadenylation of MHC-I mRNA by the RNA-binding E3-ligase MEX-3C. *Nat Commun* 6, 8670 (2015).
56. Wahle, E. & Winkler, G.S. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. *Biochim Biophys Acta* 1829, 561-570 (2013).
57. Zhang, Q. et al. New Insights into the RNA-Binding and E3 Ubiquitin Ligase Activities of Roquins. *Sci Rep* 5, 15660 (2015).
58. Yoon, J.H. et al. Scaffold function of long non-coding RNA HOTAIR in protein ubiquitination. *Nat Commun* 4, 2939 (2013).
59. Silva, G.M., Finley, D. & Vogel, C. K63 polyubiquitination is a new modulator of the oxidative stress response. *Nat Struct Mol Biol* 22, 116-123 (2015).
60. Higgins, R. et al. The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol Cell* 59, 35-49 (2015).
61. Werner, A. et al. Cell-fate determination by ubiquitin-dependent regulation of translation. *Nature* 525, 523-527 (2015).
62. Collaboration, O. The ORFeome Collaboration: a genome-scale human ORF-clone resource. *Nat Methods* 13, 191-192 (2016).
63. Schindelin, J., Rueden, C.T., Hiner, M.C. & Eliceiri, K.W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 82, 518-529 (2015).
64. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9, 676-682 (2012).
65. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2, 1896-1906 (2007).
66. Cox, J. et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4, 698-705 (2009).
67. Michalski, A. et al. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* 10, M111 011015 (2011).
68. Olsen, J.V. et al. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4, 709-712 (2007).

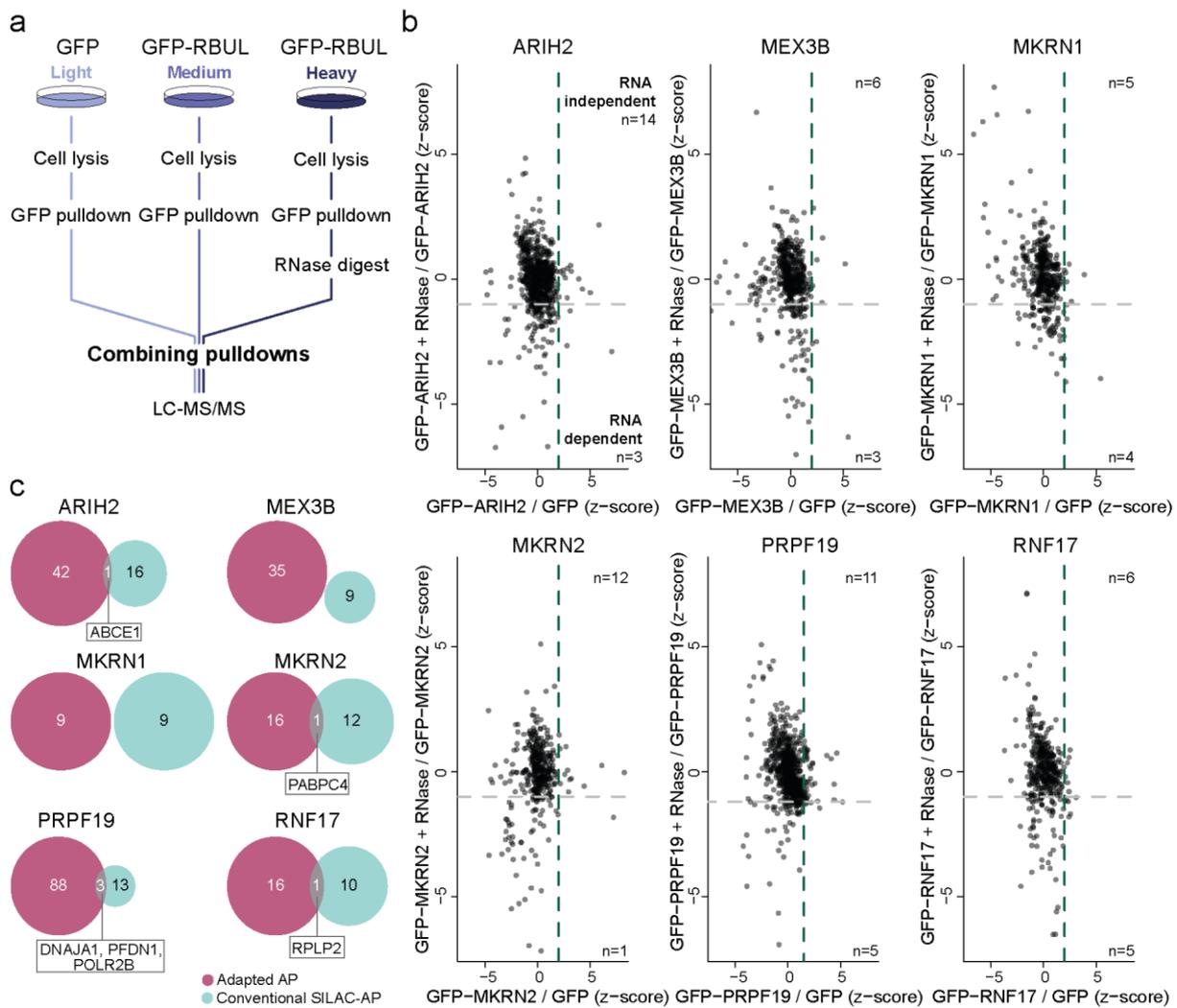
69. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10, 1794-1805 (2011).
70. Elias, J.E. & Gygi, S.P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4, 207-214 (2007).
71. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6, e21800 (2011).
72. Papasaikas, P., Tejedor, J.R., Vigevani, L. & Valcarcel, J. Functional splicing network reveals extensive regulatory potential of the core spliceosomal machinery. *Mol Cell* 57, 7-22 (2015).

Supplementary Figures



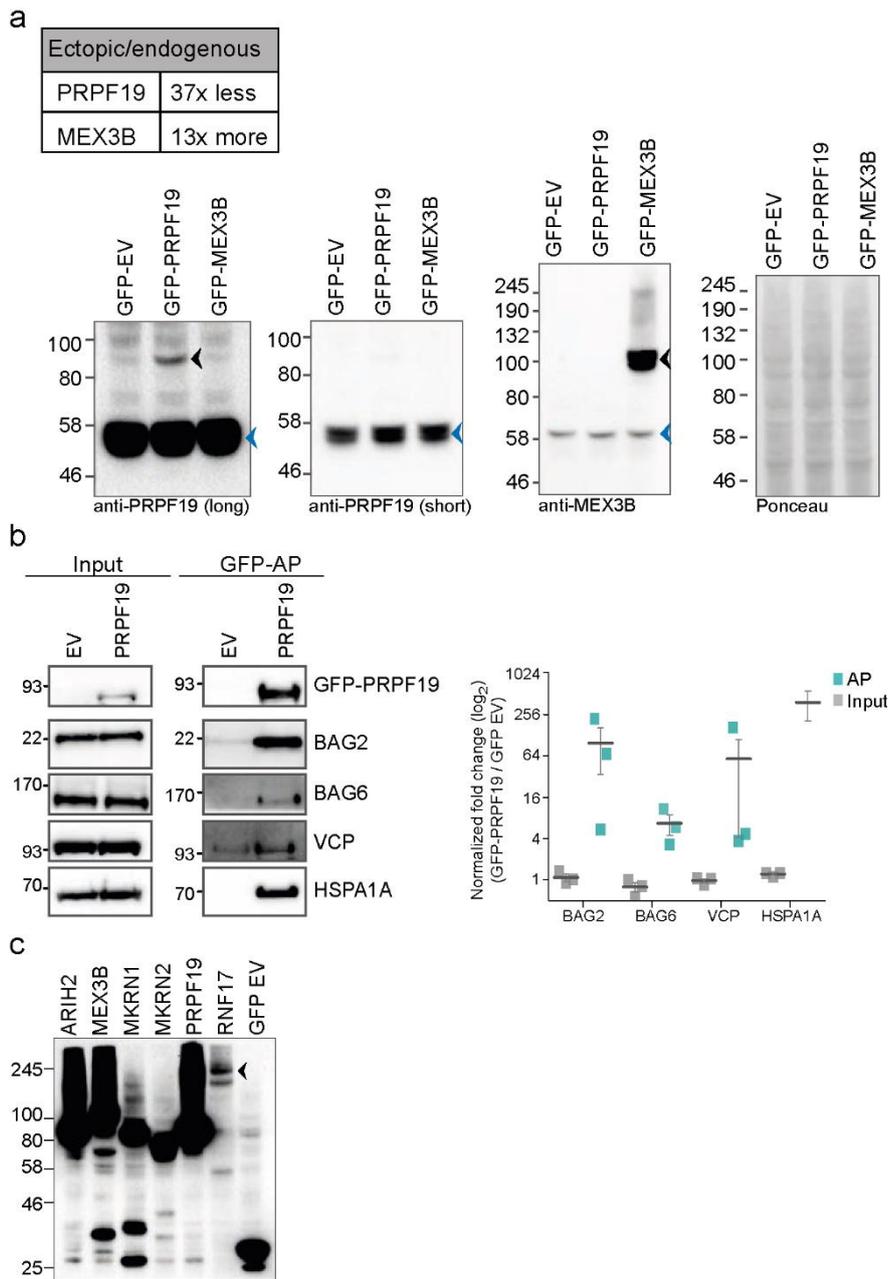
Supplementary Figure III-1, RNase optimization for the conventional SILAC-AP workflow. Cells were lysed and incubated with different concentrations of RNase A and/or RNase T1 for 30 min at 4°C or room temperature (RT). RNA was isolated using Trizol and analyzed by capillary gel electrophoresis. **a.** The level of RNA degradation using different RNase concentrations and combinations at 4°C or RT was tested. **b.** RNA treated with conditions chosen from A (0.5 U/μl RNase A + 2 U/μl RNase T1, 4°C, 30 min) was compared to undigested RNA levels kept at 4°C or RT for 30 min in duplicates by capillary gel electrophoresis.

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system



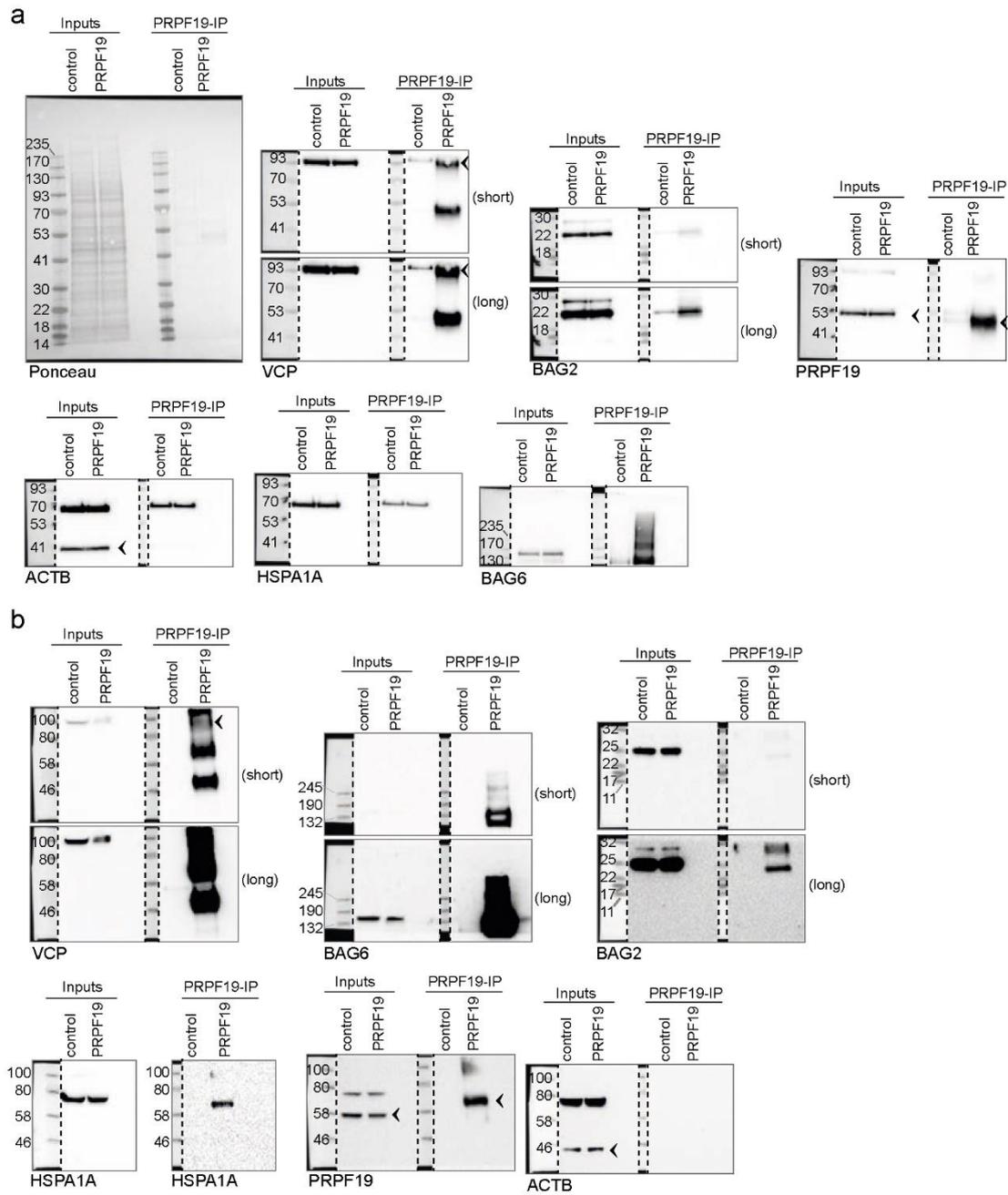
Supplementary Figure III-2, Conventional SILAC-AP in combination with RNase digest enables differentiation between RNA-dependent and RNA-independent protein interactions of the six studied RBULs. a. In the conventional SILAC-AP in combination with RNase digest, GFP alone was expressed in light-isotope labeled cells (light blue), while a GFP-tagged RBUL was expressed in medium-isotope (medium blue) and heavy-isotope (dark blue) labeled cells. Lysates were subjected to GFP-specific AP. After washing, heavy-isotope labeled enriched proteins bound to the GFP-trap agarose beads were incubated with RNase A and RNase T1 for 30 min at 4°C. Light- and medium-isotope labeled proteins were kept on ice without RNase digest. After washing, enriched proteins were mixed in a 1:1 ratio, digested on-bead with trypsin, and analyzed by LC-MS/MS. **b.** Conventional SILAC-APs are shown in dependence on RNase digest. GFP-RBUL vs. GFP SILAC ratios after z-score normalization are plotted against z-score normalized SILAC ratios of GFP-RBUL with RNase digest vs. GFP-RBUL. The dashed lines indicate the applied cut-offs at z-score ≥ 2 for GFP-RBUL over GFP (green; x-axis) and at z-score ≤ -1 for GFP-RBUL + RNase over GFP-RBUL (grey; y-axis). n, number of interactors in the two categories. **c.** Comparisons of interaction partners for the adapted AP and the conventional SILAC-AP are shown. Proteins are only considered for this analysis if they are detected in at least two replicates and pass the z-score ≥ 2 cut-off for both AP approaches. Proteins detected in both AP approaches are named below the Venn diagrams.

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

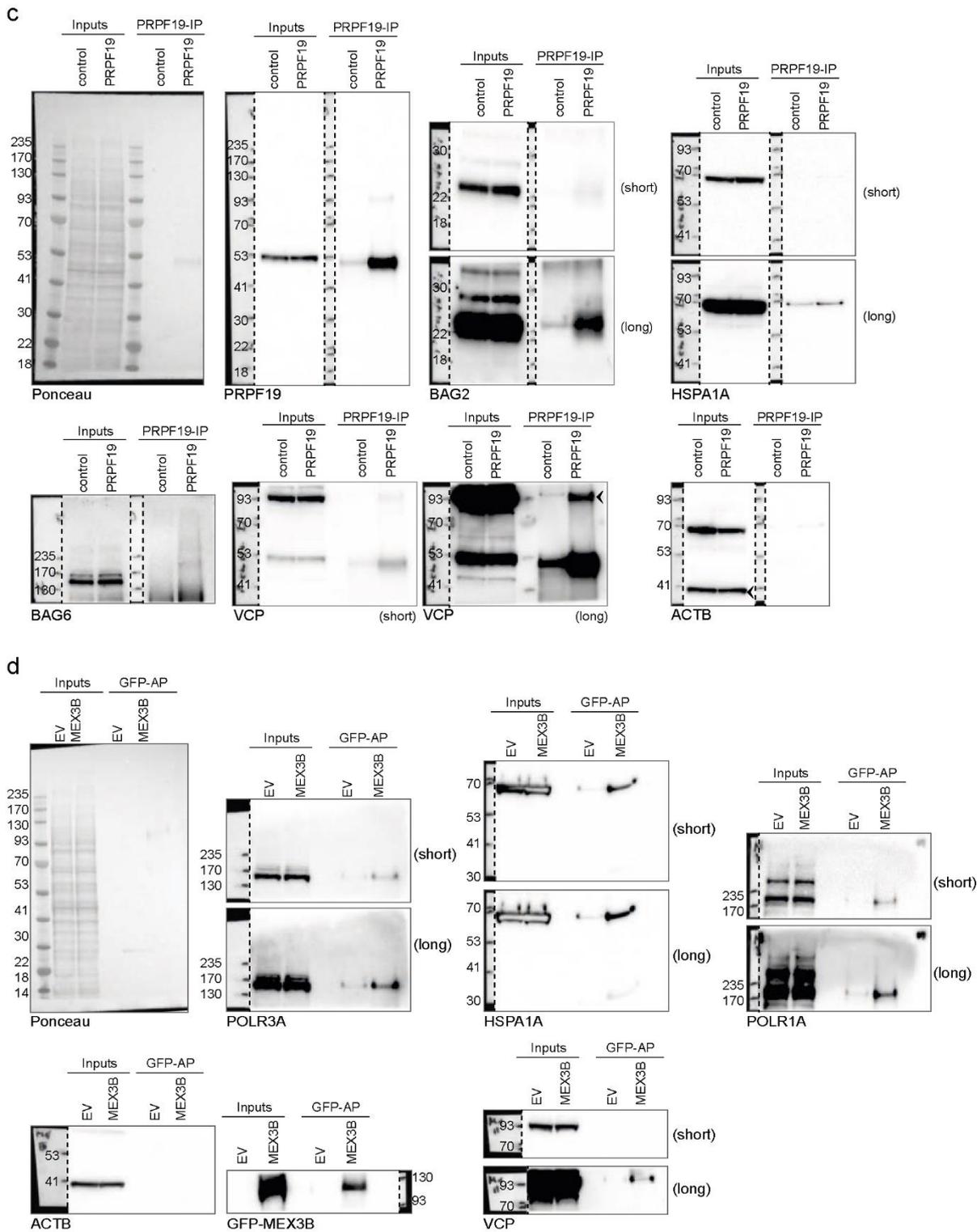


Supplementary Figure III-3, Validation of RBUL interaction partners by pulldowns and Western blot (corresponding to Figure 4c, d and 2d). **a**. Comparison of the expression levels of GFP-tagged and endogenous PRPF19 and MEX3B. The ratios of GFP-tagged vs. endogenous PRPF19 and MEX3B are given above. Western blot analyses show the expression of GFP-PRPF19 (black arrowhead) and endogenous PRPF19 (blue arrowhead) in short and long exposure. Black and blue arrowheads indicate extrinsic GFP-PRPF19 and endogenous PRPF19, respectively. GFP-MEX3B (black arrowhead) is compared to endogenous MEX3B (blue arrowhead) by Western blot. The empty vector (EV) expressing only GFP was transfected as a control. Ponceau staining is shown as a control. The experiment was performed in two replicates (see Supplementary Figure 4j). **b**. GFP and GFP-PRPF19 were expressed in HEK293T cells and pulled down with a GFP-specific antibody. Western blot analysis was performed with antibodies against BAG2, BAG5, VCP, and HSPA1A, as well as GFP to validate the AP. Left: Cropped images of input and AP samples (replicate 1). After the GFP-specific Western blot, the membrane was cut. All membrane pieces are depicted in full size for all replicates in Supplementary Figure 4g-i. Right: Quantifications of the APs normalized to EV for three independent biological replicates are shown in a dot plot, including mean and standard error (s.e.m.). **c**. Expression of GFP-RBULs and GFP (EV) measured by Western blot. A long exposure of the Western blot from Figure 2d is shown. GFP-RNF17 is indicated by an arrowhead.

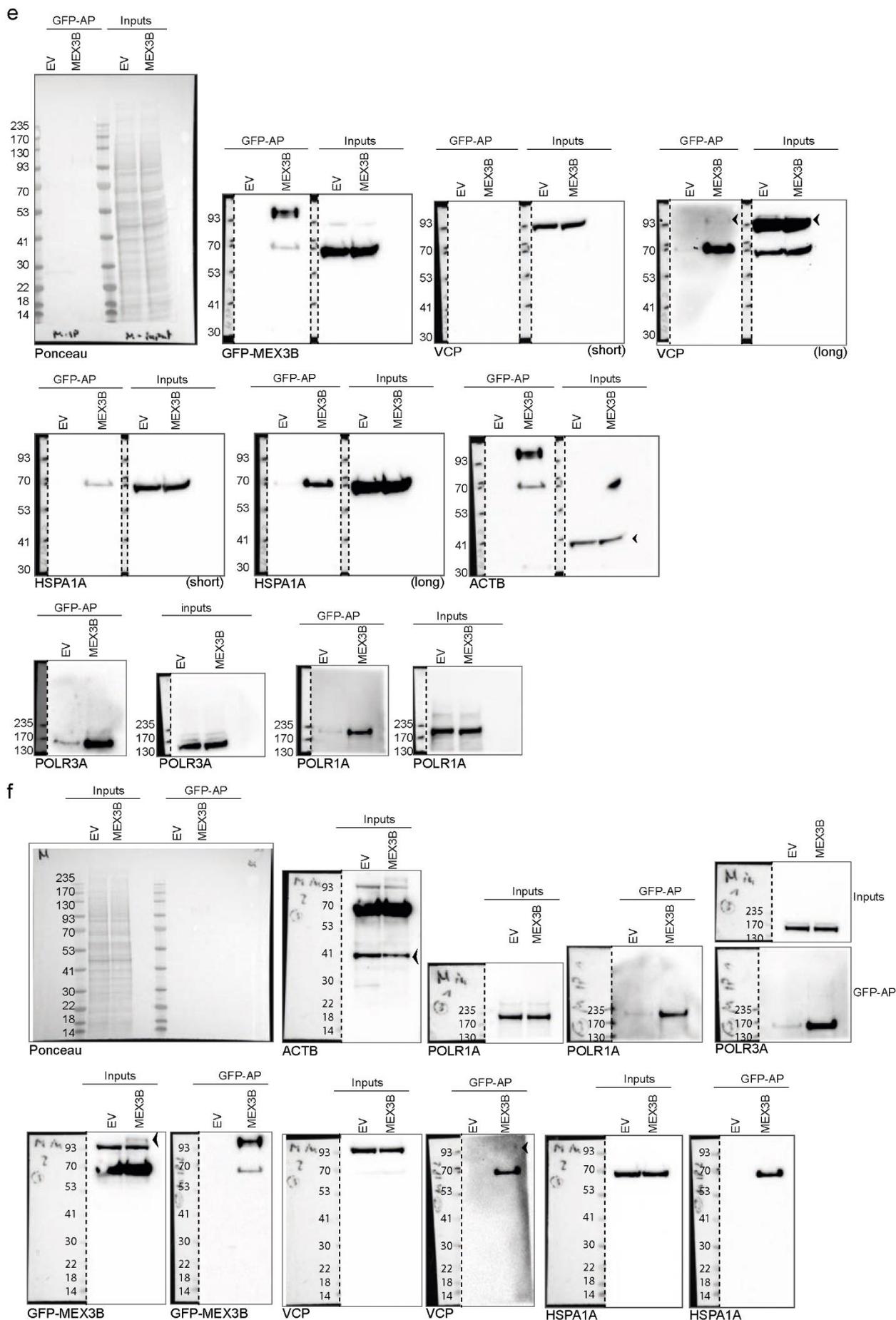
Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system



Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

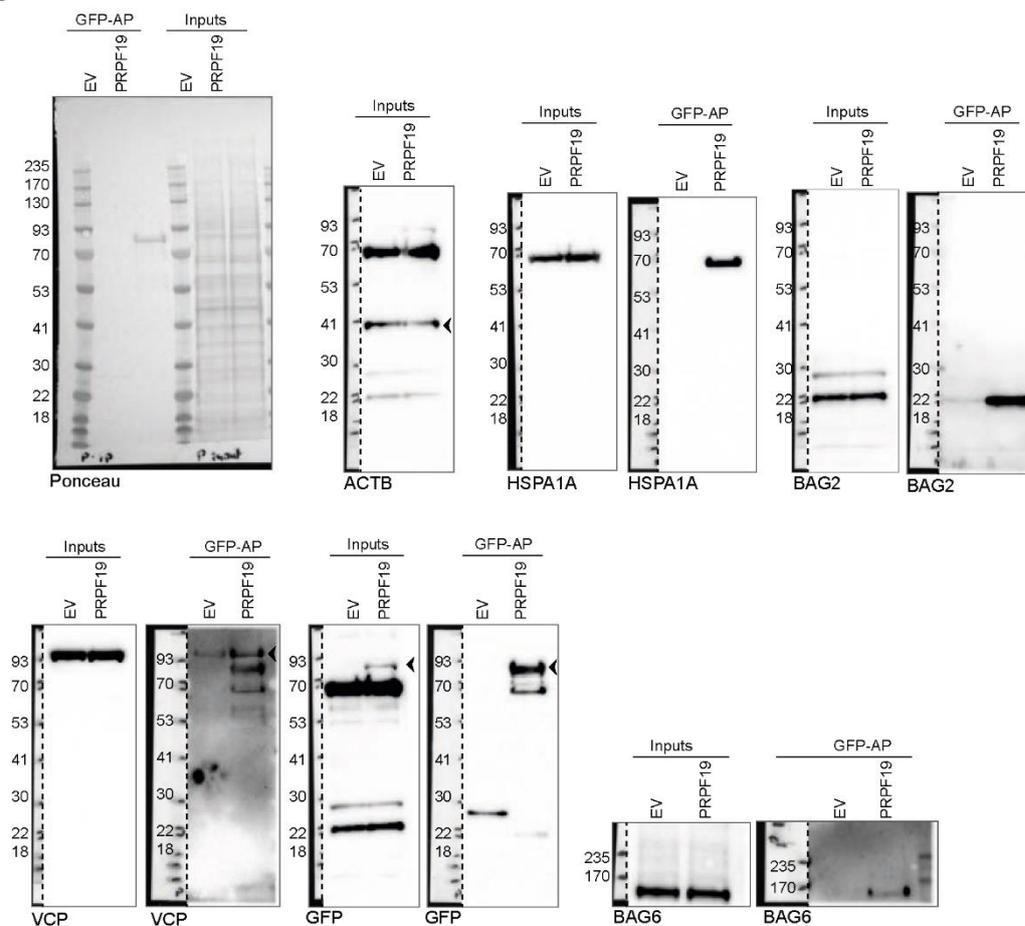


Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

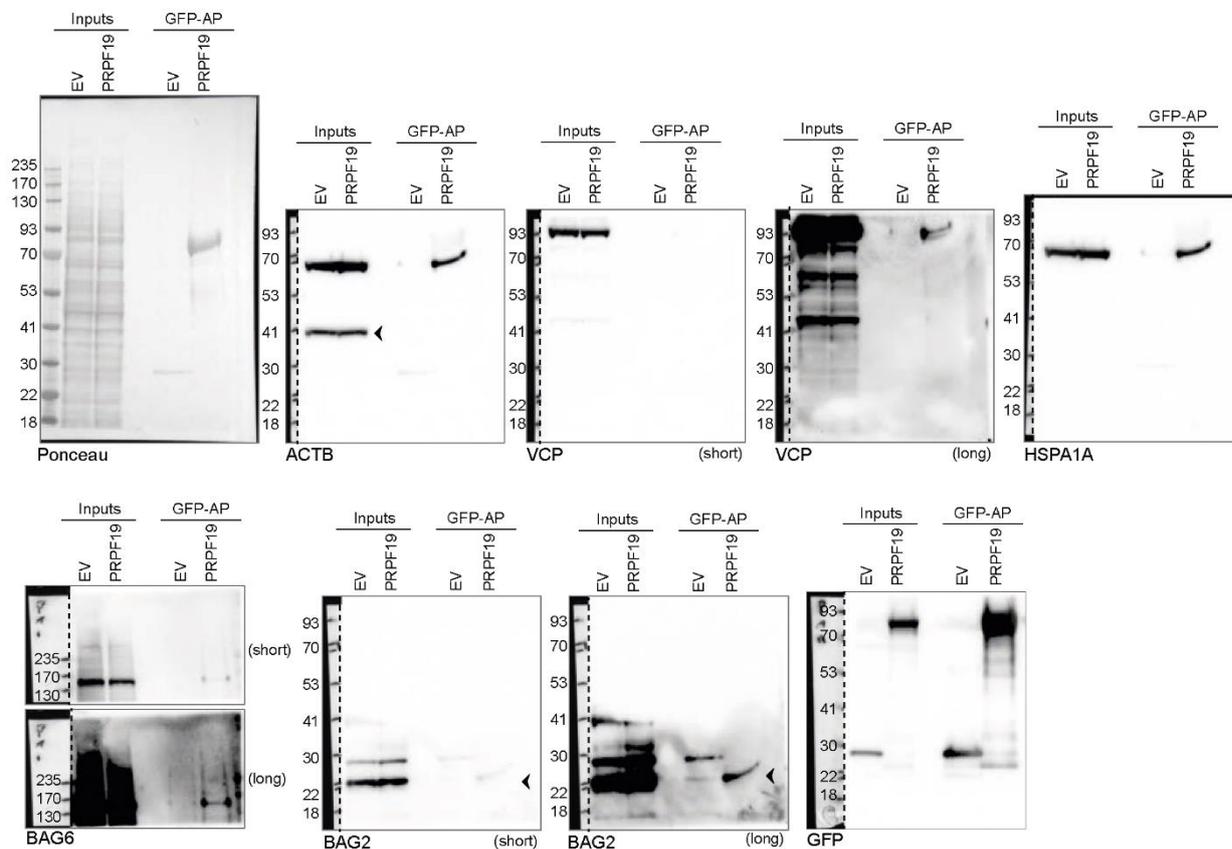


Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

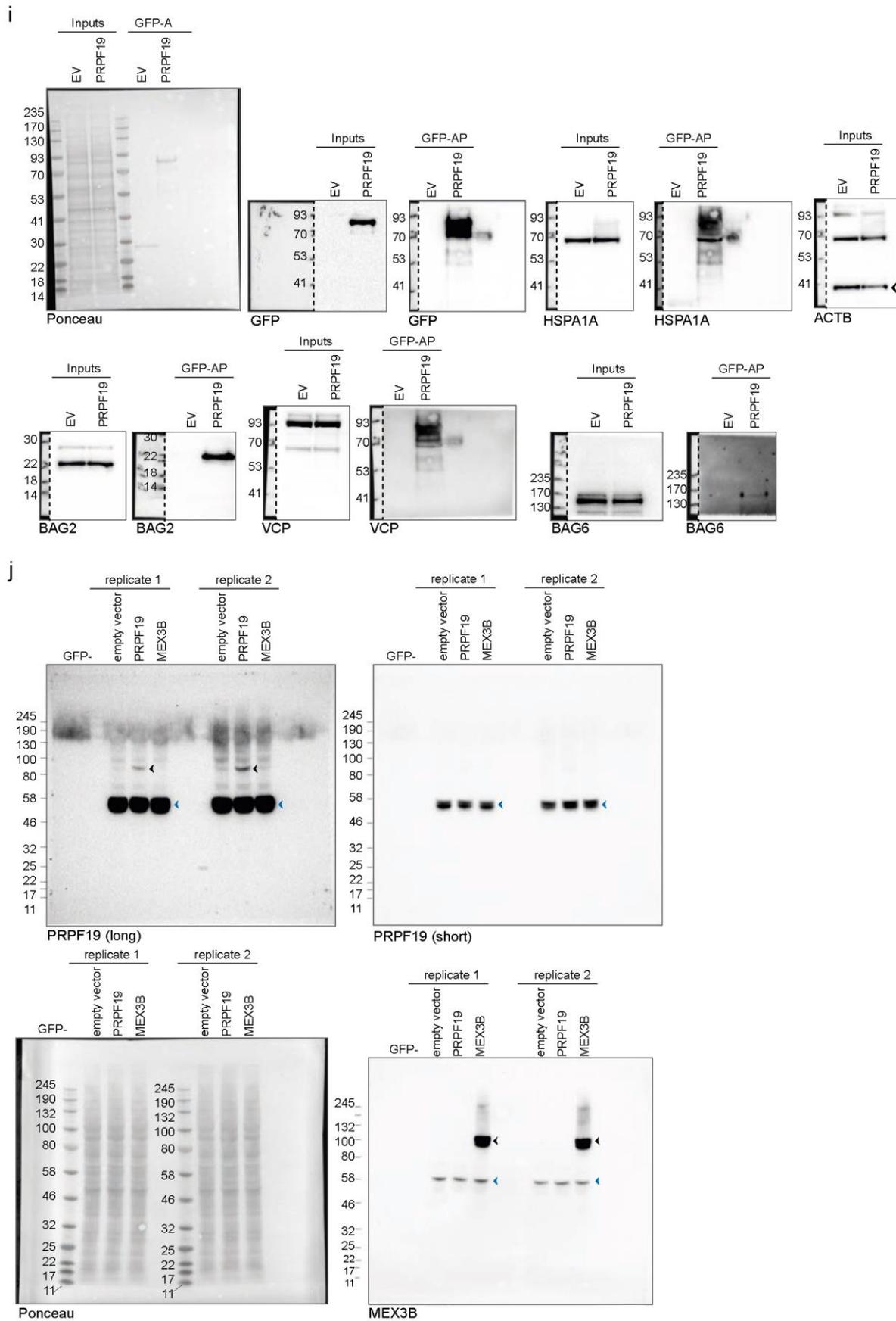
g



h



Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system



Supplementary Figure III-4, Images of full membranes and different exposure times for Western blot analyses in Figure 4c, d, Supplementary Figure 3a and b. a – c. Immunoprecipitation of endogenous PRPF19 with a PRPF19-specific antibody; extension of Figure 4c. Western blot analysis was performed with antibodies specific against BAG2, BAG6, VCP, and HSPA1A, as well as against PRPF19 itself to validate the immunoprecipitation (IP). Images of full membranes and different exposure times for all antibodies and replicates are shown. **d – f.** Affinity purification of GFP (empty vector, EV) and GFP-MEX3B with a GFP-specific antibody; extension of Figure 4d. Western blot analysis was performed with antibodies specific against POLR1A, POLR3A, VCP, and HSPA1A, as well as GFP to validate the AP. Images of full membranes and different exposure times for all antibodies and replicates are shown. **g – i.** Affinity purification of GFP (empty vector, EV) and GFP-PRPF19 with a GFP-specific antibody; extension of Supplementary Figure 3b. Western blot analysis was performed with antibodies specific against BAG2, BAG6, VCP, and HSPA1A, as well as GFP to validate the AP. Images of full membranes and different exposure times for all antibodies and replicates are shown. **j.** Images of full membranes of the comparison of the expression levels of GFP-tagged and endogenous PRPF19 and MEX3B are shown; extension of Supplementary Figure 3c. Black arrowheads indicate the GFP-tagged RBUL, while blue arrowheads indicate the endogenous RBUL (PRPF19 or MEX3B). The empty vector (EV) expressing only GFP was transfected as a control. Ponceau staining is shown as a control. The experiment was performed in two replicates.

Supplementary Tables

Supplementary Table III-1, MaxQuant analysis of MS data from the conventional SILAC-AP and the adapted AP for the six RBULS. MS raw data files for the conventional SILAC-AP and the adapted AP experiments for ARIH2, MEX3B, MKRN1, MKRN2, PRPF19, and RNF17 were analyzed with MaxQuant as described in the methods section. Single tabs contain the data of the conventional SILAC-AP or the adapted AP for the respective RBUL. Protein and gene names are highlighted in light blue. Columns containing the z-score of the log-transformed GFP-RBUL vs. GFP SILAC ratios ("z-score Ratio M/L"), the z-score of the log-transformed GFP-RBUL + RNase vs. GFP SILAC ratios ("z-score Ratio H/L"), and the z-score of the log-transformed GFP-RBUL + RNase vs. GFP-RBUL SILAC ratios ("z-score Ratio H/M") from the conventional SILAC-AP experiments ("RBUL_conventional_SILAC_AP") as well as the z-scores of the log-transformed GFP-RBUL + RNase vs. GFP SILAC ratios ("z-score Ratio H/L") from the adapted AP experiments ("RBUL_adapted_AP") are highlighted in green. *Available as Excel file with publication*

Supplementary Table III-2, Summary of adapted AP interactome analyses for the six RBULS. Gene names, protein names, annotated Biological Process (BP) GO terms, SILAC H/L ratios (z-score), Wang's Index (WI) comparing all BP GO terms of prey and bait, and interactions with WI ≥ 0.414 for each RBUL (ARIH2, MEX3B, MKRN1, MKRN2, PRPF19, and RNF17) are given. It is indicated if the respective interaction was considered as "TRUE" or "FALSE" considering the WI. In case of unavailability of a GO term for a prey, the WI was set to -1. Interaction partners with a z-score ≥ 2 in the adapted AP are highlighted in blue. The respective RBUL is highlighted in green. *Available as Excel file with publication*

Supplementary Table III-3, Comparison of numbers of interaction partners in the conventional SILAC-AP to the adapted AP experiments. Numbers of total quantified protein groups as well as of interaction partners with a z-score ≥ 2 and detected at least in two replicates are given for the conventional SILAC-AP and the adapted AP with the six RBULs (ARIH2, MEX3B, MKRN1, MKRN2, PRPF19 and RNF17). For the conventional SILAC-AP, the number of proteins is separated into RNase-dependent and RNase-independent interactions.

	Conventional SILAC-AP			Adapted AP	
	Total # of quantified proteins	# of proteins with z-score ≥ 2 RNase-independent	# of proteins lost upon RNase treatment	Total # of quantified proteins	# of proteins with z-score ≥ 2
ARIH2	653	14	3	963	42
MEX3B	470	6	3	795	35
MKRN1	408	5	4	572	9
MKRN2	387	12	1	717	17
PRPF19	632	11	5	1,061	91
RNF17	395	6	5	377	17

Supplementary Table III-4, DAVID GO ontology enrichment for GFP-MEX3B interactors for all three GO domains. GO enrichment analyses for GFP-MEX3B interactors were performed using DAVID for the GO terms from the domains Biological Process (BP), Cellular Component (CC), and Molecular Function (MF). Enriched GO terms (adjusted p -value < 0.05 , Benjamini-Hochberg correction) are highlighted in green.

Supplementary Table III-5, DAVID GO ontology enrichment for GFP-MEX3B interactors for all three GO domains.

GO	ID	Term	#	%	Interactors	Total	Adjusted p-value
BP	GO:0045815	Positive regulation of gene expression, epigenetic	5	13.9	POLR2H, HIST1H4A, POLR1A, H2AFY, POLR1C	62	2,75E-03
BP	GO:0043488	Regulation of mRNA stability	5	13.9	EIF4G1, PSMC5, HSPA1A, RPS27A, HSPA8	103	1.02E-02
BP	GO:0006342	Chromatin silencing	4	11.1	HIST1H2AC, H2AFV, H2AFY, HIST1H2AJ	45	1.25E-02
BP	GO:0032481	Positive regulation of type I interferon production	4	11.1	POLR2H, POLR3A, POLR1C, DHX36	51	1.36E-02
CC	GO:0000786	Nucleosome	6	16.7	HIST1H2AC, HIST1H2BC, HIST1H4A, H2AFV, H2AFY, HIST1H2AJ	94	7.70E-05
CC	GO:0070062	Extracellular exosome	18	50.0	HIST1H2AC, HIST1H2BC, PFKL, GART, PSMC5, HIST1H4A, VCP, HUWE1, TXNDC17, H2AFV, CSE1L, SQSTM1, RPL34, H2AFY, DHX36, HIST1H2AJ,	2811	1.19E-04

Supplementary Table III-5, DAVID GO ontology enrichment for GFP-MEX3B interactors for all three GO domains.

GO	ID	Term	#	%	Interactors	Total	Adjusted p-value
					HSPA8, RPS27A		
CC	GO:0005634	Nucleus	24	66.7	POLR2H, HIST1H2AC, HIST1H2BC, MEX3B, AKAP8L, POLR1A, EIF4G1, RECQL, ZGPAT, ARIH2, PSMC5, HIST1H4A, CSE1L, H2AFV, VCP, HUWE1, DPM1, MEX3C, H2AFY, DHX36, RANBP1, HIST1H2AJ, RPS27A, HSPA8	5415	1.28E-04
CC	GO:0005829	Cytosol	19	52.8	POLR2H, PFKL, POLR3A, HSPA1A, POLR1C, GART, EIF4G1, PSMC5, VCP, HUWE1, TXNDC17, CSE1L, SEH1L, SQSTM1, RPL34, DHX36, TNRC6A, RPS27A, HSPA8	3315	1.60E-04
CC	GO:0005654	Nucleoplasm	17	47.2	POLR2H, HIST1H2BC, AKAP8L, POLR1A, POLR3A, HSPA1A, POLR1C, RECQL, PSMC5, VCP, HUWE1, CSE1L, HIST1H4A, SQSTM1, TNRC6A, HSPA8, RPS27A	2784	2.24E-04
CC	GO:0005736	DNA-directed RNA	3	08.3	POLR2H, POLR1A, POLR1C	13	4.11E-03
CC	GO:0005666	DNA-directed RNA polymerase III complex	3	08.3	POLR2H, POLR3A, POLR1C	19	6.70E-03
CC	GO:0016234	Inclusion body	3	08.3	PSMC5, SQSTM1, HSPA1A	18	6.86E-03
CC	GO:0005737	Cytoplasm	19	52.8	HIST1H2BC, PFKL, AKAP8L, POLR1A, HSPA1A, GART, EIF4G1, RECQL, ARIH2, PSMC5, VCP, HUWE1, CSE1L, SQSTM1, RPL34, MEX3C, DHX36, RANBP1, RPS27A	5222	2.57E-02
CC	GO:0005913	Cell-cell adherens junction	5	13.9	EIF4G1, RPL34, RANBP1, HSPA1A, HSPA8	323	2,77E-02
CC	GO:0000790	Nuclear chromatin	4	11.1	HIST1H2AC, H2AFV, H2AFY, HIST1H2AJ	193	4.27E-02
CC	GO:0016020	Membrane	11	30.6	EIF4G1, RECQL, PSMC5, HIST1H4A, CSE1L, PFKL, HUWE1, DPM1, POLR3A, RPS27A, HSPA8	2200	4.62E-02
MF	GO:0003677	DNA binding	13	36.1	POLR2H, HIST1H2AC, HIST1H2BC, POLR1A, AKAP8L, POLR3A, POLR1C, RECQL, HUWE1, H2AFV, HIST1H4A, H2AFY, HIST1H2AJ	1674	3.24E-03
MF	GO:0044822	poly(A) RNA binding	11	30.6	EIF4G1, HIST1H4A, HUWE1, VCP, MEX3B, MEX3C, AKAP8L, DHX36, TNRC6A, RPS27A, HSPA8	1129	3.98E-03
MF	GO:000	DNA-	4	11.1	POLR2H, POLR1A, POLR3A,	38	7.70E-03

Supplementary Table III-5, DAVID GO ontology enrichment for GFP-MEX3B interactors for all three GO domains.

GO	ID	Term	#	%	Interactors	Total	Adjusted p-value
	3899	directed RNA polymerase activity			POLR1C		
MF	GO:0001054	RNA polymerase I activity	3	08.3	POLR2H, POLR1A, POLR1C	12	7.99E-03
MF	GO:0001056	RNA polymerase III activity	3	08.3	POLR2H, POLR3A, POLR1C	18	1.47E-02

Supplementary Table III-6, Known and novel protein-protein interactions identified from the adapted APs of the six RBULs. A comparison of identified interaction partners from the adapted AP to known interactions from the literature is shown. Out of 170 interactions in total, only 19 were reported in literature to date. HIPPIE confidence scores, numbers of interactions reported for each of the six RBULs in HIPPIE and the kind of experiment used to measure in combination with PubMed identifiers (PMIDs) and reporting databases are given.

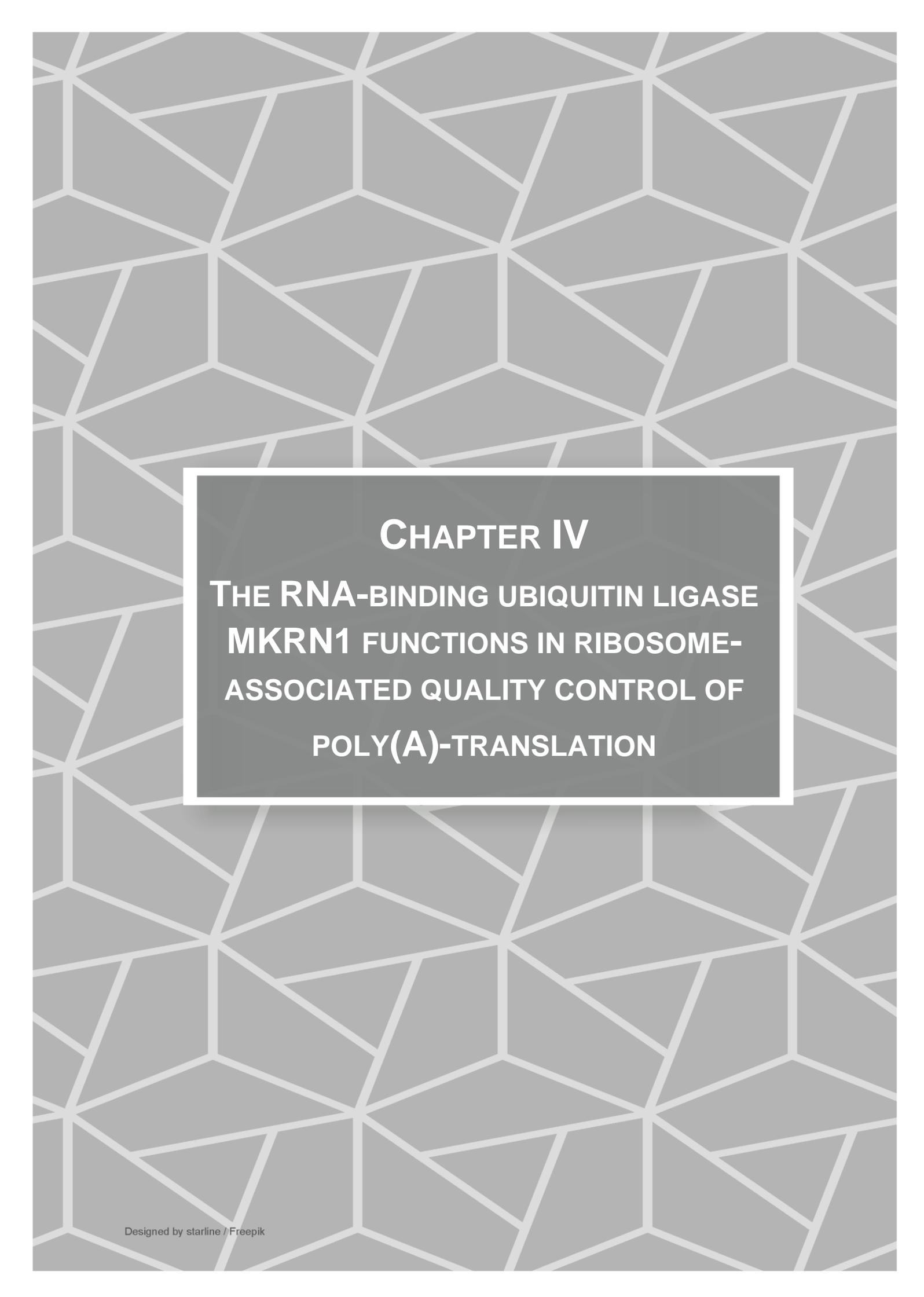
Bait	HIPPIE interactions	Interactor	Score	Sources
ARIH2	4	CUL5	0.78	enzymatic study, affinity enzymatic study, affinity chromatography technology, pull down, anti tag coimmunoprecipitation, anti bait coimmunoprecipitation (PMID: 24076655; sources: BioGRID, IntAct) enzymatic study, affinity chromatography technology, pull down, anti tag coimmunoprecipitation, anti bait coimmunoprecipitation (PMID: 24076655; sources: BioGRID, IntAct)
		TCEB1	0.72	affinity chromatography technology, anti tag coimmunoprecipitation (PMIDs: 24076655; sources: BioGRID, IntAct)
		HK2	0.63	affinity chromatography technology (PMID: 22863883; source: BioGRID)
		VCP	0.63	affinity chromatography technology (PMID: 23383273; source: BioGRID)
MEX3B	Not present	ZC3H7A, RNF20, PRPSAP2, PPM1G, DDX1, GSTO1, PAPSS1, CHERP, Nbla03646, MID1, GNE, POLR3A, APPL1, GTF2H2, PRKCI, PYGB, DOCK7, ASNS, ALDH2, TUBB4A, CAPN1, CAPN2, TRMT1, RPS27A, PGM1, RFC5, ABCE1, SRP54, ZGPAT, EIF1AX, ACTR2, BCKDK, GANAB, RPL34, UROD, ASCC3, DHX15, MEX3B		
		RPS27A, MEX3C, SQSTM1, HSPA1A, HSPA8, POLR1A, POLR2H, PDE12, PFKL, H2AFV, POLR1C, HIST1H4A, HIST1H2BC, HUWE1, EIF4G1, HIST1H2AJ, DHX36, CSE1L, H2AFY, DPM1, AKAP8L, HIST1H2AC, PSMC5, TXNDC17, RECQL, METTL13, GART, RANBP1, SEH1L, TNRC6A, RPL34, ZGPAT, POLR3A, VCP		
MKRN1	Not present	VCP, RPS27A, HSPA1A, IGSF9B, RPS27L, EXD2, ZNF512, C3orf17, UBAP2L, VCP, RPS27A		
MKRN2	Not present	RPS27A, VCP, RPL34, IGSF9B, PABPC4, PABPC1, CHTOP, SET, GOLGA4, ST13, RPS12, RPL37A, YWHAZ, ATP2B1, MSH3, HSPA1A, HSPA8		
PRPF19	17	CDC5L	0.9	affinity Capture-MS, copurification, pull down, affinity chromatography technology, anti bait

Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system

Bait	HIPPIE interactions	Interactor	Score	Sources
				coimmunoprecipitation, far western blotting (PMIDs: 11101529, 20467437, 19633697, 20595234, 20176811, 17276391, 22939629, 26344197; sources: BioGRID, MINT, IntAct)
		BCAS2	0.89	two-hybrid, anti tag coimmunoprecipitation, affinity chromatography technology, far western blotting, pull down (PMIDs: 22365833, 20176811, 21536736, 22939629, 26344197; sources: MINT, BioGRID, IntAct)
		PLRG1	0.87	affinity Capture-MS, anti bait coimmunoprecipitation, affinity chromatography technology (PMIDs: 17353931, 20176811, 17276391, 26344197; sources: BioGRID, IntAct, I2D)
		PRPF8	0.87	two-hybrid, anti tag coimmunoprecipitation, affinity chromatography technology (PMIDs: 22365833, 22939629, 26344197; sources: MINT, IntAct, BioGRID)
		CDC40	0.76	two-hybrid, anti tag coimmunoprecipitation, affinity chromatography technology (PMID: 22365833; sources: MINT, IntAct, BioGRID)
		SNW1	0.73	anti bait coimmunoprecipitation (PMIDs: :20467437, 26344197; sources: MINT, IntAct, BioGRID)
		HSP90AA1	0.63	affinity chromatography technology (PMID: 22939624; source: BioGRID)
		HUWE1	0.63	affinity chromatography technology (PMID: 25147182; source: BioGRID)
		VCP	0.63	affinity chromatography technology (PMID: 23443559; source: BioGRID)
		SF3B3	0.59	PMIDs: 22939629, 26344197; source: BioGRID
		SNRNP200	0.59	PMIDs: 22939629, 26344197; source: BioGRID
		XAB2	0.59	PMIDs: 22939629, 26344197; source: BioGRID
		AQR	0.55	cosedimentation through density gradient (PMID: 24304693; source: IntAct)
		DHX8	0.49	PMIDs: 26344197; source: BioGRID
		EFTUD2	0.49	PMID: 22939629; source: BioGRID
		HSP90AB1	0.49	PMID: 22939624; source: IntAct
		RNPS1	0.49	PMID: 22939629; source: BioGRID
	Not present	HIST1H4A, HIST1H2BC, EIF4G1, HIST1H2AJ, H2AFY, HIST1H2AC, IGSF9B, CHTOP, PFDN4, UBL4A, VBP1, PFDN1, BAT3, PFDN2, CCT2, ASNA1, PFDN6, BAG2, CCT8, CCT5, CCT6A, GET4, CCT4, CCT7, CCT3, TCP1, HSPA6, TAB1, SPAG9, RHOT2, PSMD2, SUGT1, RPAP3, NACA, PSMC6, RBM22, CRNKL1, EPB41L2, DDB1, DNAJB6, USP19, RBM8A, FBL, POLR2B, DNAJC7, DNAJA1, PSMC4, SNRNP40, HNRNPC, TP53BP2, ZNF326, SNRPD3, TBC1D4, IRS4, TCOF1, TRA2A, PNN, DNAJA2, PSMC1, PSMC2, STIP1, MDC1, H1F0, GNB2L1, FTSJ3, RSL1D1, RBMX, SF3B2, SNRPE, SNRPN, RNF17, HSPA8, HSPA1A, RPS27A, POLR3A		
RNF17	Not present	VCP, RPS27A, HSPA1A, CCT2, BAG2, CCT8, CCT6A, CCT4, HSPA6, HSPA8, GOLGA4, ACIN1, SSBP1, PGAM5, RAB1B, RPLP2		

Supplementary Table III-7, A comparison of the cellular localization of the RBULs. The cellular localization according to the Human Protein Atlas (“Cellular localization - human protein atlas”), as well as the observed cellular localization in our microscopy experiments (“Cellular localization - our data”) are given for each RBUL. References for further localization studies are mentioned in the “Additional References” column.

RBUL	Cellular localization - Human Protein Atlas	Cellular localization - our data	Additional references
ARIH2	Nucleoplasm	Nucleoplasm and cytosol	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423, 19340006
MEX3B	Nucleoplasm and cytosol (shuttling); cytosol and cytosolic granules	Nucleoplasm periphery, cytosol, and cytosolic granules	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423, 17267406, 18779327
MKRN1	Nucleus, nuclear membrane, and cytosol	Nucleus, cytosol, and cytosolic granules	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423, 22128154
MKRN2	Nucleoplasm and cytosol	Cytosol	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423
PRPF19	Nuclear speckles	Nucleoplasm and cytosol	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423, 20176811
RNF17	No data	Cytosol and cytosolic granules	PMIDs: 25613900, 21139605, 18669619, 16127175, 11435423



CHAPTER IV
THE RNA-BINDING UBIQUITIN LIGASE
MKRN1 FUNCTIONS IN RIBOSOME-
ASSOCIATED QUALITY CONTROL OF
POLY(A)-TRANSLATION

IV. THE RNA-BINDING UBIQUITIN LIGASE MKRN1 FUNCTIONS IN RIBOSOME-ASSOCIATED QUALITY CONTROL OF POLY(A)-TRANSLATION

bioRxiv 516005; doi: <https://doi.org/10.1101/516005>

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

Andrea Hildebrandt¹, Mirko Brüggemann², Susan Boerner², Cornelia Rücklé², Jan Bernhard Heidelberger¹, Annabelle Dold¹, Anke Busch¹, Heike Hänel¹, Andrea Voigt¹, Stefanie Ebersberger¹, Ingo Ebersberger^{3,4}, Jean-Yves Roignant¹, Kathi Zarnack^{2§}, Julian König^{1§}, Petra Beli^{1§}

¹ Institute of Molecular Biology (IMB), Ackermannweg 4, 55128 Mainz, Germany

² Buchmann Institute of Molecular Life Sciences (BMLS), Goethe University, Max-von-Laue-Str. 15, 60438 Frankfurt am Main, Germany

³ Department for Applied Bioinformatics, Institute of Cell Biology and Neuroscience, Goethe University, Max-von-Laue-Str. 13, 60438 Frankfurt am Main, Germany

⁴ Senckenberg Biodiversity and Climate Research Centre (BiK-F), Georg-Voigt-Straße 14-16, 60325 Frankfurt am Main, Germany

§ Corresponding authors:

KZ (kathi.zarnack@bmls.de), JK (j.koenig@imb-mainz.de) and PB (p.beli@imb-mainz.de)

Keywords: MKRN1, ubiquitylation, RNA binding, ribosome-associated quality control, RQC, poly(A), iCLIP, ubiquitin remnant profiling, translation

Abstract

Cells have evolved quality control mechanisms to ensure protein homeostasis by detecting and degrading aberrant mRNAs and proteins. A common source of aberrant mRNAs is premature polyadenylation, which can result in non-functional protein products. Translating ribosomes that encounter poly(A) sequences are terminally stalled, followed by ribosome recycling and decay of the truncated nascent polypeptide via the ribosome-associated quality control (RQC). Here, we demonstrate that the conserved RNA-binding E3 ubiquitin ligase Makorin Ring Finger Protein 1 (MKRN1) promotes ribosome stalling at poly(A) sequences during RQC. We show that MKRN1 interacts with the cytoplasmic poly(A)-binding protein (PABP) and is positioned upstream of poly(A) tails in mRNAs. Ubiquitin remnant profiling uncovers PABP and ribosomal protein RPS10, as well as additional translational regulators as main ubiquitylation substrates of MKRN1. We propose that MKRN1 serves as a first line of poly(A) recognition at the mRNA level to prevent production of erroneous proteins, thus maintaining proteome integrity.

Abbreviations: Adenosine (A), Makorin Ring Finger Protein 1 (MKRN1), ribosome-associated quality control (RQC), poly(A)-binding protein (PABP), mass spectrometry (MS), affinity purification (AP), stable isotope labelling with amino acids in cell culture (SILAC), false discovery rate (FDR), Gene Ontology (GO), Biological Process (BP), Molecular Function (MF), mRNA ribonucleoprotein particle (mRNP), PABP-interacting motif (PAM2), MKRN1 variant with point mutations in the PAM2 motif (GFP-MKRN1^{PAM2mut}), MKRN1 variant with point mutation in RING domain (GFP-MKRN1^{RINGmut}), individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP), signal-over-background (SOB), nucleotides (nt), 4-thiouridine (4SU), Pearson correlation coefficients (r), A-rich stretches (A-stretches), lysine (K), knock down (KD), RNA recognition motif (RRM), Polyethylenimine (PEI), modified RIPA (mRIPA), N-ethylmaleimide (NEM), dithiothreitol (DTT), chloroacetamide (CAA), higher-energy collisional dissociation (HCD), glycine-glycine (GlyGly), strong-cation exchange chromatography (SCX)

Introduction

During gene expression, quality control pathways monitor each step to detect aberrant mRNAs and proteins. These mechanisms ensure protein homeostasis and are essential to prevent neurodegenerative diseases (Chu et al. 2009). A common source of aberrant mRNAs is premature polyadenylation, often in combination with mis-splicing, which results in truncated non-functional protein products (Kaida et al. 2010). Therefore, mechanisms are in place that recognise such homopolymeric adenosine (poly(A)) sequences and abrogate their translation (Bengtson and Joazeiro 2010).

In eukaryotes, ribosomes that terminally stall for diverse reasons during translation are detected by the ribosome-associated quality control (RQC) (reviewed in Brandman and Hegde 2016; Joazeiro 2017). Upon splitting of the 60S and 40S ribosomal subunits, the RQC complex assembles on the stalled 60S subunit to initiate the release and rapid degradation of the truncated tRNA-bound polypeptide. The E3 ubiquitin ligase Listerin (LTN1) modifies the truncated polypeptide with K48-linked ubiquitin chains to target it for degradation in a p97-dependent manner through the proteasome (Bengtson and Joazeiro 2010; Brandman et al. 2012; Verma et al. 2013). Whereas peptide release and ribosome recycling by the RQC complex are relatively well understood, less is known about the mechanisms that promote poly(A) recognition and initial ribosome stalling.

Several recent studies demonstrated a role for the RNA-binding E3 ubiquitin ligase ZNF598 in initiating RQC for prematurely polyadenylated mRNAs (Garzia et al. 2017; Juszkiwicz e and Hegde 2017; Sundaramoorthy et al. 2017). It was suggested that ZNF598 senses the translation of poly(A) segments through binding the cognate lysine tRNAs (Garzia et al. 2017). In addition, ZNF598 recognises the collided di-ribosome structure that arises when a trailing ribosome encounters a slower leading ribosome (Juszkiwicz et al. 2018). This is followed by site-specific, regulatory ubiquitylation of the 40S ribosomal proteins RPS10 and RPS20 by ZNF598. In addition to ZNF598, the 40S ribosomal subunit-associated protein RACK1 was shown to regulate ubiquitylation of RPS2 and RPS3 upstream of ribosomal rescue (Sundaramoorthy et al. 2017).

Makorin Ring Finger Protein 1 (MKRN1) belongs to a family of evolutionary conserved RNA-binding E3 ubiquitin ligases. Up to four paralogs exist in vertebrates (MKRN1-4), which combine a RING domain with one or more CCCH zinc finger domains (Gray et al. 2000; Böhne et al. 2010) (**Supplemental Fig. 1A**). MKRN1 has been implicated in the regulation of telomere length, RNA polymerase II transcription and the turnover of tumour suppressor protein p53 and cell cycle regulator p21 (Kim et al. 2005; Omwancha

et al. 2006; Lee et al. 2009; Salvatico et al. 2010), but its RNA-related functions remain poorly understood. A study in mouse embryonic stem cells (mESC) reported its interaction with hundreds of mRNAs as well as multiple RNA-binding proteins (RBPs), including the cytoplasmic poly(A)-binding protein (PABP) PABPC1, IGF2BP1 and ELAVL1 (Cassar et al. 2015). The interaction with PABP was further corroborated in human HEK293 cells (Miroci et al. 2012). The same study demonstrated that a shortened isoform of MKRN1 controls local translation via its PABP-interacting motif 2 (PAM2 motif) in rat neurons (Miroci et al. 2012). In line with a role in translation, MKRN1 was found in association with ribosomes, from which it could be released together with PABP and other proteins by RNase digestion (Simsek et al. 2017). Nevertheless, the RNA binding specificity and functional role of MKRN1 in human cells remained largely elusive.

Here, we introduce MKRN1 as a novel factor in RQC. We propose that MKRN1 is recruited to A-rich sequences in mRNAs in a PABP-dependent manner, where it acts as a first line of defence against poly(A) translation. MKRN1 depletion abrogates ribosome stalling in reporter assays, accompanied by reduced ubiquitylation of RQC-related proteins. We therefore hypothesise that MKRN1 allows recognition of poly(A) sequences prior to their translation.

Results

MKRN1 interacts with PABPC1 and other RBPs

In order to learn about potential functions, we first characterised the protein interaction profile of MKRN1 in HEK293T cells. To this end, we used affinity purification (AP) coupled to stable isotope labelling with amino acids in cell culture (SILAC)-based quantitative mass spectrometry (MS) using GFP-MKRN1^{wt} or GFP as a bait. We identified 53 proteins that were significantly enriched in GFP-MKRN1^{wt} compared to the control APs (false discovery rate [FDR] < 5%, combined ratios of three independent experiments). In line with previous reports (Miroci et al. 2012; Cassar et al. 2015; Hildebrandt et al. 2017), we found the cytoplasmic poly(A)-binding proteins (PABP) PABPC1 and PABPC4 among the highly enriched MKRN1 interactors (z-score > 4, corrected *P* values = 7.18e-10 and 6.16e-16, respectively) (**Fig. 1A, Supplemental Fig. S2A, and Supplemental Table S1**). Moreover, we detected 14 ribosomal proteins as well as four proteins that were previously shown to co-purify with ribosomes (Simsek et al. 2017), including IGF2BP1, LARP1, UPF1, and ELAVL1 (**Fig. 1A**). Consistently,

“translation” was among the significantly enriched Gene Ontology (GO) terms for the MKRN1 interaction partners (Biological Process [BP], **Supplemental Fig. S2B**). Almost all interactors were previously found in association with polyadenylated transcripts (50 out of 53 proteins have been annotated with the GO term “poly(A) RNA binding”, Molecular Function [MF], **Supplemental Fig. S2B**). We confirmed the MS results in reciprocal AP experiments with GFP-tagged PABPC1, ELAVL1, and IGF2BP1 as baits followed by Western blot for endogenous MKRN1 (**Supplemental Fig. S2C**). All detected interactions persisted in the presence of RNases (RNase A and T1), demonstrating that MKRN1 interacts with these proteins in an RNA-independent manner (**Supplemental Fig. S2C**). Together, these observations suggest that MKRN1 is part of a larger mRNA ribonucleoprotein particle (mRNP) together with PABP and other RBPs. This is further supported by a parallel study on the Mkrn1 ortholog in *Drosophila melanogaster*, which consistently identified pAbp, Larp, Upf1 and Imp (IGF2BP in mammals) as interaction partners (Dold et al, parallel submission; preprint available at bioRxiv, doi: 10.1101/501643).

Many proteins interact with PABP via a PABP-interacting motif (PAM2) motif, which specifically binds to the MLLE domain present almost exclusively in PABP (Deo et al. 2001; Kotzlov et al. 2010). Accordingly, a previous study demonstrated that MKRN1 associates with PABP via a PAM2 motif at amino acid positions 161-193 (Miroci et al. 2012). In support of a putative functional relevance, a phylogenetic analysis illustrated that the presence and positioning of the PAM2 motif are preserved in MKRN1 orthologs across metazoans (**Supplemental Fig. S1A,B**). AP of a MKRN1 variant with point mutations in the PAM2 motif (GFP-MKRN1^{PAM2mut}) (Pohlmann et al. 2015) no longer recovered PABPC1 and PABPC4 (Fig. 1B-D and Supplemental Table S1). For comparison, we also tested a previously described point mutation in the RING domain that abolishes the E3 ubiquitin ligase function (ligase-dead, GFP-MKRN1RINGmut) (Kim et al 2005). This mutation did not impair the interaction of MKRN1 with PABPC1, but led to a slight increase, possibly due to stabilisation of MKRN1 and/or PABPC1 (Fig. 1C, Supplemental Fig. S7, and Supplemental Table S1). Surprisingly, MKRN1PAM2mut lost interaction not only with PABPC1 and PABPC4, but also with several other identified proteins (Fig. 1D), suggesting that MKRN1PAM2mut no longer resided within the mRNPs. These results confirmed that MKRN1 interacts with PABP proteins, and suggested that this association is required for mRNP formation.

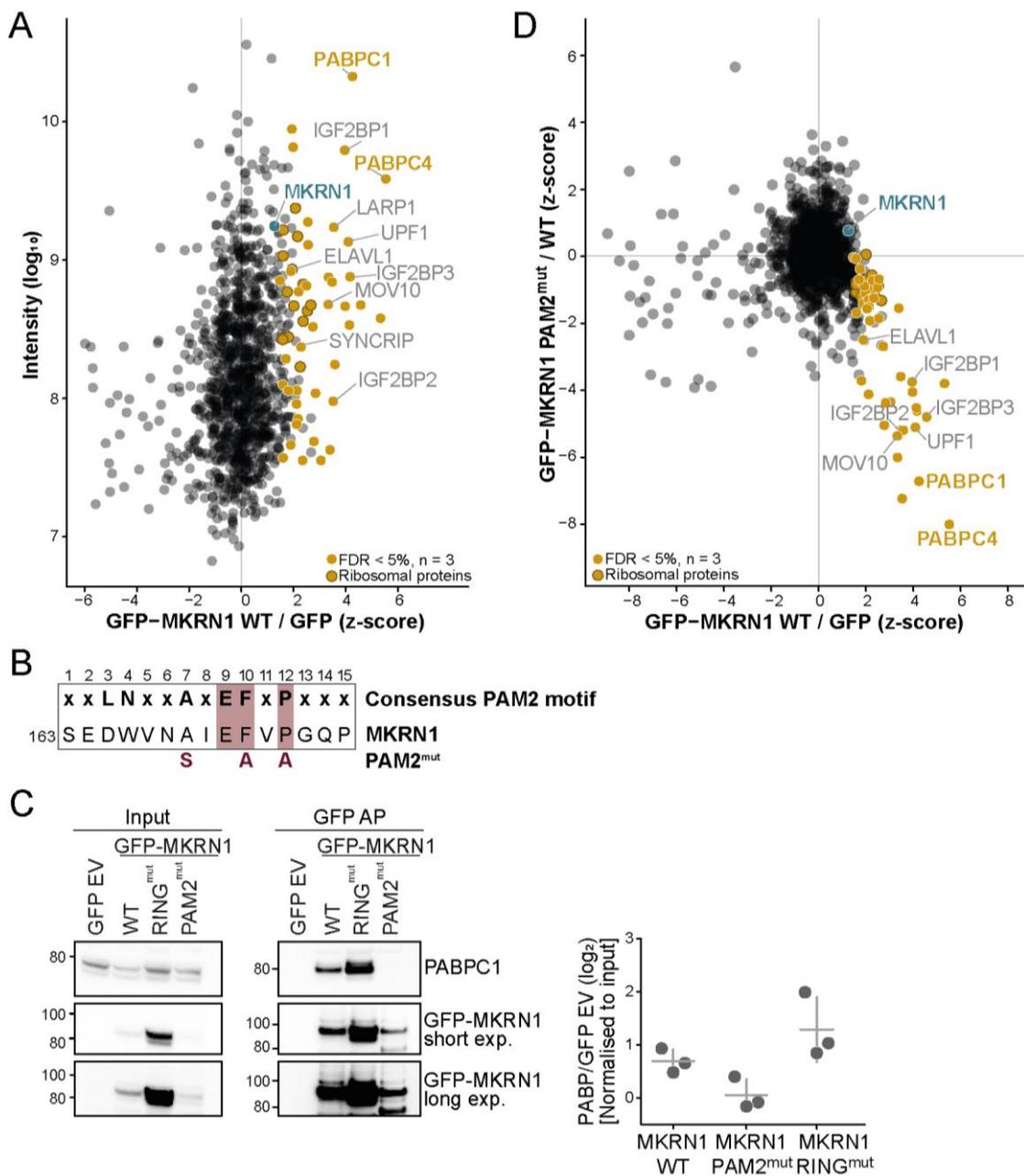


Figure IV-1, MKRN1 interacts with PABP and other regulators of translation and RNA stability. (A) Protein interactome of GFP-MKRN1wt in HEK293T cells analysed by quantitative MS-based proteomics. Combined SILAC ratios (n = 3 replicates) after z-score normalisation are plotted against log₁₀-transformed intensities. 1,100 protein groups were quantified in at least two out of three replicate experiments. MKRN1 and significant interactors are highlighted (FDR < 5%). (B) A PAM2 motif similar to the previously reported consensus (shown on top; Supplemental Fig. S1B) (Albrecht and Lengauer 2004) is present in MKRN1 (first amino acid position indicated on the left). Introduced mutations in MKRN1PAM2mut are indicated in red below. Relevant positions are highlighted (Supplemental Fig. S1B). (C) Endogenous PABP interacts with MKRN1wt and MKRN1RINGmut, but not with MKRN1PAM2mut. Western blots for endogenous PABPC1 and GFP (two exposure times, exp.) after AP of GFP-MKRN1 (wt and mutants). Ratios of PABP signal (normalised to input) in GFP-MKRN1 APs over control (GFP empty vector, EV) are shown on the right. Replicates 2, 3, and uncropped gel images are shown in Supplemental Fig. S9A-C. (D) Quantitative comparison of the interactomes of GFP-MKRN1wt and GFP-MKRN1PAM2mut shows that PABP and several other interactors are lost upon PAM2 mutation. Combined ratios of three replicates are shown in a scatter plot. Only proteins detected in at least two out of three replicates are shown. MKRN1wt significant interactors (from A) are highlighted as in (A) (FDR < 5% in MKRN1wt).

MKRN1 binds to poly(A) tails and at internal A-stretches

In order to characterise the RNA-binding behaviour of human MKRN1 *in vivo*, we performed individual-nucleotide resolution UV crosslinking and immunoprecipitation (iCLIP) (König et al. 2010) in combination with 4-thiouridine (4SU) labelling to enhance UV crosslinking (Hafner et al. 2010). In three replicate experiments with GFP-tagged MKRN1 (GFP-MKRN1^{wt}) expressed in HEK293T cells, we identified more than 4,6 million unique crosslink events, cumulating into 7,331 MKRN1 binding sites (see Materials and methods; **Supplemental Table S2**). These were further ranked according to the strength of MKRN1 binding, which was estimated from the enrichment of crosslink events within a binding site relative to its local surrounding, which served as a proxy for transcript abundance (“signal-over-background”, SOB; see Materials and Methods) (Sutandy et al. 2018). SOB values were highly reproducible between replicates (Pearson correlation coefficients $r > 0.72$, **Supplemental Fig. S3**).

Across the transcriptome, MKRN1 almost exclusively bound to protein-coding mRNAs with a strong tendency to locate in 3' UTRs (**Fig. 2A,E**). Binding sites generally harboured uridine-rich tetramers (**Supplemental Fig. S4A**), likely reflecting 4SU-based UV crosslinking (Hafner et al. 2010). Strikingly, the top 20% MKRN1 binding sites were massively enriched in AAAA tetramers (A, adenosine) within 5-50 nucleotides (nt) downstream of the binding sites (**Fig. 2B** and **Supplemental Fig. S4A**). These were situated within A-rich stretches (A-stretches), which ranged from 8-30 nt in length (**Supplemental Fig. S4B**; see Materials and methods). Within 3' UTRs, 30% (1,848 out of 6,165) of MKRN1 binding sites resided immediately upstream of an A-stretch (**Fig. 2C,E**) and longer A-stretches associated with stronger MKRN1 binding (**Supplemental Fig. S4C,D**). Intriguingly, we detected a requirement for a continuous run of at least 8 A's to confer strong MKRN1 binding (**Fig. 2D**), which precisely matched the RNA footprint of PABP (Webster et al. 2018). Since PABP was previously reported to also bind within 3' UTRs (Bag 2001; Lyabin et al. 2011; Kini et al. 2016), these observations indicated that MKRN1 binds together with PABP to mRNAs.

Prompted by this notion, we analysed the unusually high fraction of unmapped iCLIP reads in the MKRN1 dataset (**Supplemental Table S2**). In accordance with binding of MKRN1 immediately upstream of poly(A) tails, more than 13% of the unmapped reads displayed an increased A-content (**Fig. 3B**), compared to only 2% for an unrelated control RBP (Braun et al. 2018). In addition, the mapped GFP-MKRN1^{wt} crosslink

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

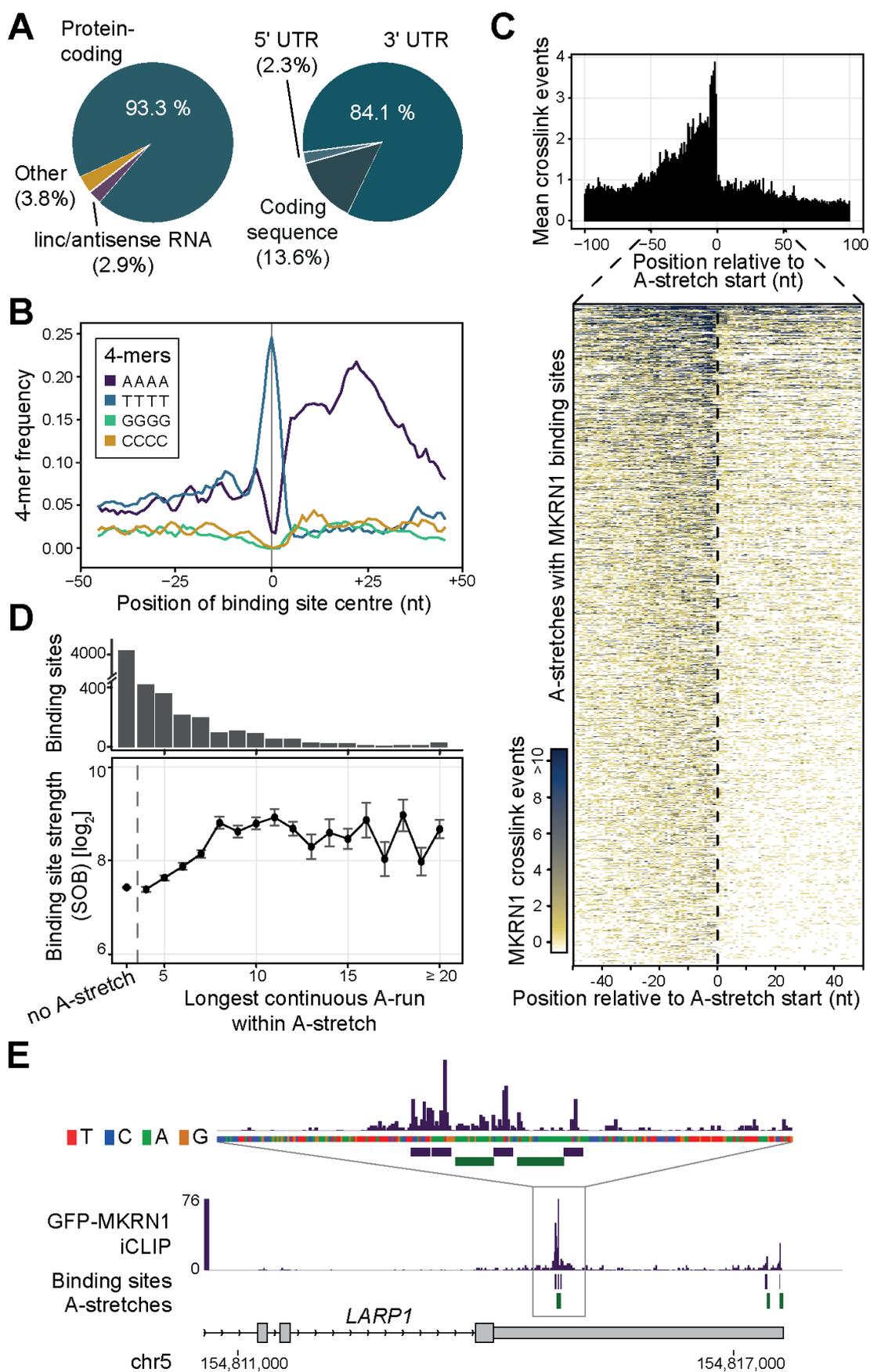


Figure IV-2, MKRN1 binds upstream of A-stretches in 3' UTRs. (A) MKRN1 predominantly binds in the 3' UTR of protein-coding genes. Piecharts summarising the distribution of MKRN1 binding sites to different RNA biotypes (7,331 binding sites, left) and different regions within protein-coding transcripts (6,913 binding sites, right). (B) MKRN1 binding sites display a downstream enrichment of AAAA homopolymers. Frequency per nucleotide (nt) for four homopolymeric 4-mers in a 101-nt window around the midpoints of the top 20% MKRN1 binding sites (according to signal-over-background; see Material and methods). (C) MKRN1 crosslink events accumulate upstream of A-stretches. Metaprofile (top) shows the mean crosslink events per nt in a 201-nt window around the start position of 1,412 MKRN1-associated A-stretches in 3' UTRs. Heatmap visualisation (bottom) displays crosslink events per nt (see colour scale) in a 101-nt window around the MKRN1-associated A-stretches. (D) MKRN1 binding site strength (signal-over-background, SOB) increases with length of longest continuous run of A's (LCA) within the A-stretch. Mean and standard deviation of MKRN1 binding sites associated with A-stretches harbouring LCAs of increasing length (x-axis). MKRN1 binding sites without associated A-stretches are shown for comparison on the left. Number of binding sites in each category indicated as barchart above. (E) MKRN1 binds upstream of A-stretches in the 3' UTR of the LARP1 gene. Genome browser view of GFP-MKRN1 iCLIP data showing crosslink events per nt (merged replicates, turquoise) together with binding sites (lilac) and associated A-stretches (dark green).

events were enriched upstream of annotated polyadenylation sites, as exemplified in the *SRSF4* gene (**Fig. 2E** and **Fig. 3A,C**). Together, these results support the notion that MKRN1 binds upstream of poly(A) tails, possibly in conjunction with PABP. In order to test whether PABP is required for MKRN1 binding, we performed UV crosslinking experiments with GFP-MKRN1^{PAM2mut}, which no longer interacts with PABP (**Fig. 1C,D**). Strikingly, RNA binding of this mutant was globally reduced compared to GFP-MKRN1^{wt} (**Fig. 3D** and **Supplemental Fig. S5**), indicating that PABP might recruit MKRN1 to RNA. In summary, these results strongly imply that MKRN1 binds upstream of poly(A) tails, which could be implemented via its interaction with PABP. In a concordant scenario, it was found that *Drosophila* Mkrn1 bound before an extended A-stretch in the 3' UTR of *oskar* mRNA, and that this binding was significantly reduced upon depletion of pAbp (Dold et al., parallel submission; preprint available at bioRxiv, doi: 10.1101/501643).

MKRN1 promotes ribosome stalling at poly(A) sequences

As outlined above, our iCLIP data evidenced that MKRN1 marks the beginning of poly(A) tails. Hence, it is conceivable that MKRN1 will also bind upstream of premature polyadenylation events within open reading frames. Based on MKRN1's binding pattern, its interaction partners and its previously reported association with ribosomes (Simsek et al. 2017), we hypothesised that MKRN1 may be involved in the clearance of such transcripts by ribosome-associated quality control (RQC). In this process, ribosomes that translate into a poly(A) sequence, for instance upon stop codon readthrough and premature polyadenylation, are stalled and eventually recycled (Brandman and Hegde 2016; Joazeiro 2017). To test this hypothesis, we employed a recently introduced flow

cytometry-based assay that monitors ribosome stalling in a dual fluorescence reporter (Juszkiewicz and Hegde 2017) (**Fig. 4A**).

As reported previously, inserting a K(AAA)₂₀ linker (encoding for 20 lysine residues) into the reporter resulted in predominant ribosome stalling compared to the starting vector (K₀, **Fig. 4B** and **Supplemental Fig. 6A**). Importantly, *MKRN1* depletion with two independent siRNA sequences led to a reproducible recovery of RFP expression downstream of K(AAA)₂₀, demonstrating that many ribosomes failed to stall at K(AAA)₂₀ (*MKRN1* KD1 and KD2; **Fig. 4C** and **Supplemental Fig. S6A,B**). *MKRN1* KD2 seemed slightly more effective, possibly because this siRNA simultaneously decreased the transcript levels of the close paralogue *MKRN2* (**Supplemental Fig. S6C**). Notably, *MKRN1* KD2 impaired ribosome stalling to a similar level as KD of *ZNF598*, the E3 ubiquitin ligase that was recently reported to function in RQC (Garzia et al. 2017; Juszkiewicz and Hegde 2017; Sundaramoorthy et al. 2017).

Moreover, simultaneous depletion of *MKRN1* and *ZNF598* was not additive, indicating that both proteins are necessary for function (**Fig. 4C** and **Supplemental Fig. S6A**). In addition, we noted a certain level of cross-regulation, such that *ZNF598* expression was decreased in *MKRN1* KD1 (but not in *MKRN1* KD2), whereas *ZNF598* overexpression reduced *MKRN1* expression (**Supplemental Fig. S6D,E**). Taken together, we conclude that MKRN1 contributes to efficient ribosome stalling in RQC.

MKRN1 mediates the ubiquitylation of ribosome-associated proteins

RQC builds on a series of ubiquitylation events by multiple E3 ubiquitin ligases, including Listerin and *ZNF598* (Brandman and Hegde 2016). In order to identify putative ubiquitylation substrates of MKRN1, we first determined the protein interactome of the ligase-deficient mutant GFP-MKRN1^{RINGmut}. In three replicate experiments, we quantified 1,097 protein groups present in at least two out of three replicates (**Supplemental Table S1**), revealing 137 proteins that were significantly enriched compared to GFP-MKRN1^{wt} (**Supplemental Fig. S7**). Intriguingly, these included RPS10, a ribosomal protein that was previously reported to be modified by *ZNF598* during RQC (Garzia et al. 2017; Juszkiewicz and Hegde 2017; Sundaramoorthy et al. 2017).

In order to directly test for ubiquitylation of putative substrates of MKRN1, we performed ubiquitin remnant profiling to compare the relative abundance of di-glycine-modified lysines in wild type and *MKRN1* KD cells. We quantified 2,324 ubiquitylation sites

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

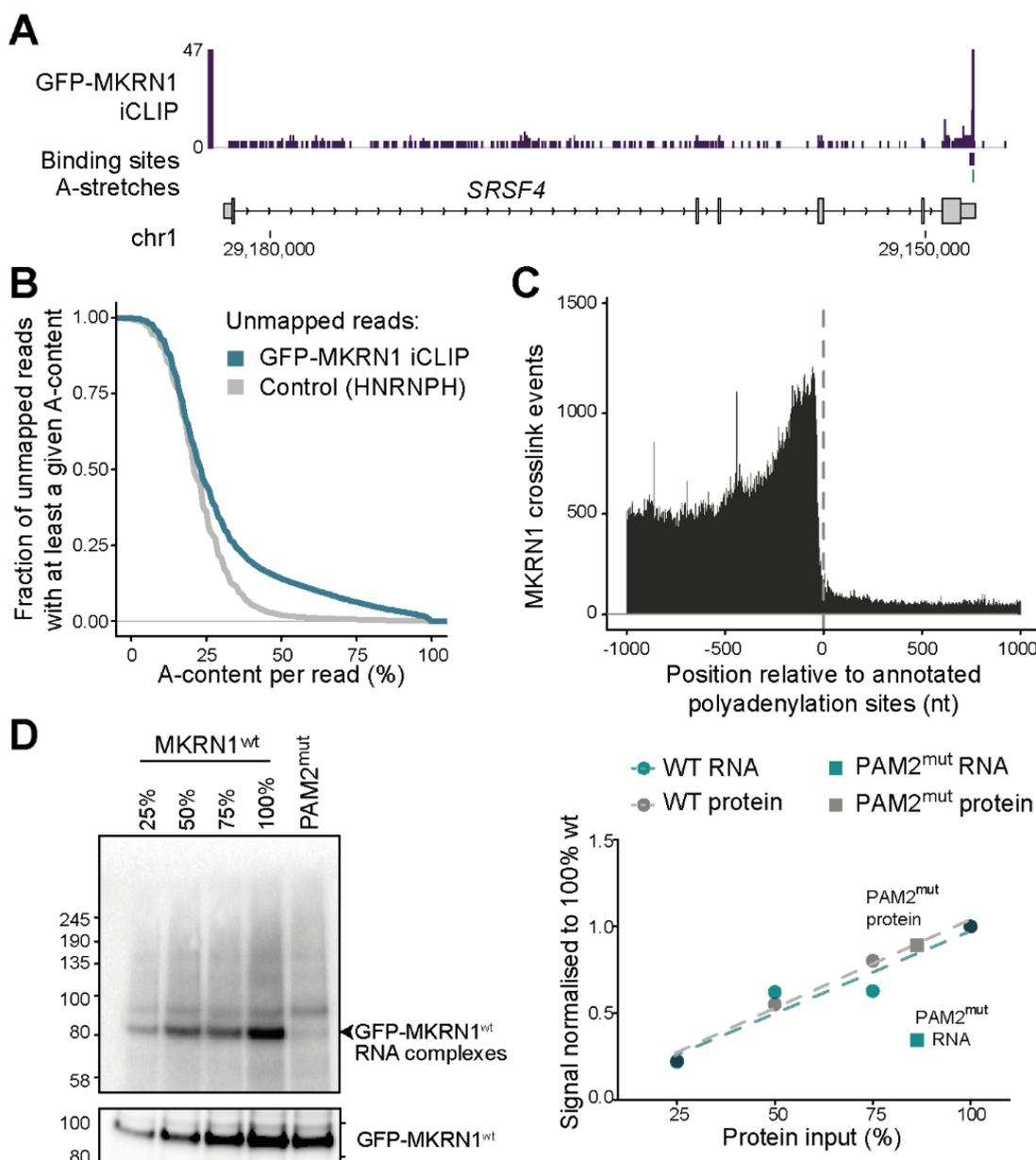


Figure IV-3, MKRN1 binds at poly(A) tails. (A) MKRN1 binds at the annotated transcript 3' end in the SRSF4 gene. Genome browser view as in Fig. 2E. (B) Unmapped MKRN1 iCLIP reads display increased A-content (more than half of all nucleotides in the read), evidencing poly(A) tail binding. Cumulative fraction of iCLIP reads (y-axis, merged replicates) that could not be mapped to the human genome (see Materials and methods) and show at least a given A-content (x-axis). iCLIP data for the unrelated RBP HNRNPH (Braun et al. 2018) are shown for comparison. (C) MKRN1 crosslink events increase towards 3'UTR ends. Metaprofile shows the sum of crosslink events per nt in a 2001-nt window around annotated transcript 3' ends of transcripts with >1 kb 3' UTRs (n = 63,030). (D) Overall RNA binding of MKRN1 is strongly reduced when abrogating PABP interaction. Autoradiograph (left) of UV crosslinking experiments (replicate 1, with 4SU and UV crosslinking at 365 nm; replicates 2 and 3 in Supplemental Fig. S5) comparing GFP-MKRN1PAM2mut with GFP-MKRN1wt at different dilution steps for calibration. Quantification of radioactive signal of protein-RNA complexes and corresponding Western blots shown on the right. Uncropped gel images are shown in Supplemental Fig. S10.

(in 1,264 proteins) that were detected in all four replicate experiments (**Supplemental Table S3**). Notably, *MKRN1* depletion led to a significantly decreased abundance of 29 ubiquitylation sites on 21 proteins (FDR < 10%, **Fig. 5A**). The majority of the ubiquitylation targets assembled into a coherent cluster of translational regulators based

on previously reported protein-protein interactions and functional annotations (**Fig. 5B,C** and **Supplemental Fig. 8A**). Among these proteins, we had already detected PABPC1/4, IGF2BP1, ELAVL1, MOV10, LARP1, and RPS10 as significant interactors of GFP-MKRN1^{wt} and/or GFP-MKRN1^{RINGmut} (**Fig. 1A, Fig. 5F** and **Supplemental Fig. S7**). Importantly, we detected a significant decrease in ubiquitylation at lysine 107 of RPS10 (K107; **Fig. 5D**). In order to distinguish differential ubiquitylation from protein level changes, we also measured the total protein levels in *MKRN1* KD cells and did not observe changes in RPS10, PABPC1/4, IGF2BP1/2/3, ELAVL1, and MOV10 protein levels (**Supplemental Fig. 8B** and **Supplemental Table S4**). Taken together, we conclude that MKRN1 mediates ubiquitylation of the ribosomal protein RPS10 and several translational regulators during ribosome-associated quality control.

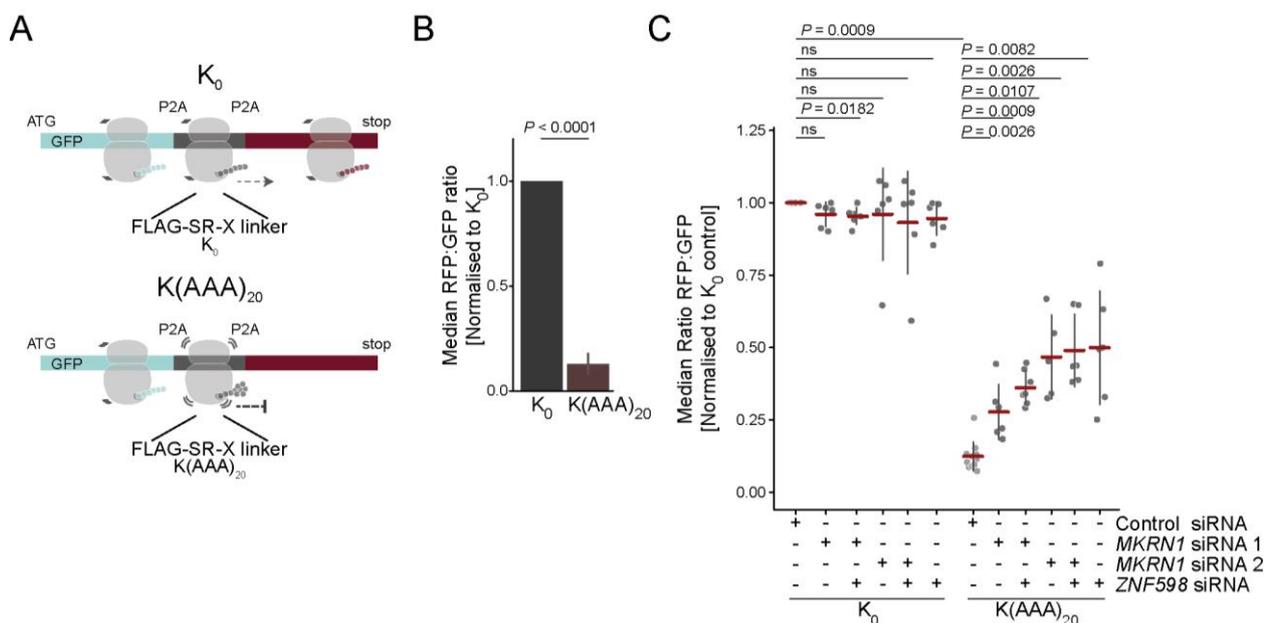


Figure IV-4, MKRN1 stalls ribosomes at poly(A) sequences. (A) The dual fluorescence reporter harbours an N-terminal GFP, followed by a FLAG-SR-X linker and a C-terminal RFP, which are separated by P2A sites to ensure translation into three separate proteins (Juszkiewicz and Hegde 2017). The resulting GFP:RFP ratio was determined using flow cytometry. The inserted fragment K(AAA)₂₀ encodes 20 lysines by repeating the codon AAA. The starting vector without insert (K₀) served as control. Schematic ribosomes illustrate translation of the respective reporter segments. (B) Ribosomes are efficiently stalled at K(AAA)₂₀ in HEK293T cells. Median RFP:GFP ratios, normalised to K₀, are shown. Error bars represent standard deviation of the mean (s.d.m., n = 6 replicates). P value indicated above (paired two-tailed t-test). (C) Ribosomes fail to stall in the absence of MKRN1. HEK293T cells were transfected with control siRNA or siRNAs targeting MKRN1 (KD1 and KD2) or ZNF598 for 24 h, followed by transfection of the reporter plasmids for 48 h. Western blots for KDs are shown in Supplemental Fig. S6B. RFP and GFP signals were analysed by flow cytometry. Median RFP:GFP ratios, normalised to K₀ in control, are shown. Error bars represent s.d.m.; P values indicated above (paired two-tailed t-test, Benjamini-Hochberg correction, n ≥ 6 replicates; ns, not significant).

Discussion

Ribosome-associated quality control is essential to recognise and clear terminally stalled ribosomes. Here, we uncover MKRN1 as a novel factor in RQC. Our data indicate that MKRN1 is positioned upstream of poly(A) sequences through direct interaction with PABP, thereby marking the beginning of poly(A) tails. We propose that in case of premature polyadenylation, MKRN1 stalls the translating ribosome and initiates RQC by ubiquitylating ribosomal protein RPS10, PABP and other translational regulators (**Fig. 6**).

PABP recruits MKRN1 upstream of A-stretches and poly(A) tails

Central to our model is the specific RNA-binding behaviour of MKRN1, which is recruited to mRNA by PABP to mark the beginning of poly(A) tails. This builds on the following observations: (i) We and others show that MKRN1 and PABP interact via the PAM2 motif (Miroci et al. 2012). (ii) MKRN1 binding to RNA is strongly reduced when interaction with PABP is abolished. (iii) The association of strong MKRN1 binding with continuous A-runs of ≥ 8 A's mirrors the footprint of one RNA recognition motif (RRM) domain of PABP, indicating that the binding of one RRM to poly(A) is sufficient for MKRN1 recruitment (Webster et al. 2018). On such short A-stretches, MKRN1 might stabilise PABP binding, while on longer A-stretches, PABP might be the major driving force to recruit MKRN1. This interaction might also anchor the first PABP at the beginning of the poly(A) tail. One possible function could be the stabilisation of PABP on short poly(A) tails to promote efficient translation (Lima et al. 2017). In yeast, where a MKRN1 ortholog is missing (see below), this anchoring is thought to be achieved by Pab1p itself via its fourth RRM domain (Webster et al. 2018). Of note, a parallel study with the Mkrn1 ortholog from *D. melanogaster* demonstrates binding of a Mkrn1/pAbp complex at an A-stretch in the 3' UTR of *oskar* mRNA, which is involved in translational control and required for oogenesis (Dold et al., parallel submission; preprint available at bioRxiv, doi: 10.1101/501643).

MKRN1 ubiquitylates RPS10 and translational regulators to stall ribosomes

Our data suggest that ribosomes encountering the MKRN1-PABP complex are stalled, possibly via ubiquitylation of RPS10 and other MKRN1 interactors. Concordantly, ZNF598, a factor that was recently shown to function in RQC, was also found to mediate ubiquitylation of RPS10 (Juszkiewicz et al. 2018). In conjunction with its unique RNA-binding behaviour, we therefore hypothesise that MKRN1 acts as a first line of

Figure IV-5, MKRN1 ubiquitylates ribosomal protein RPS10 and translational regulators. (A) Ubiquitin remnant profiling to compare the relative abundance of ubiquitylation sites in MKRN1 KD2 and control HEK293T cells. Ubiquitin remnant peptides were enriched and analysed by quantitative mass spectrometry, quantifying a total of 15,528 ubiquitylation sites. 29 putative MKRN1 target sites with significantly decreased ubiquitylation upon MKRN1 KD2 (FDR < 10%, n = 4 replicates) are highlighted and labelled with the respective protein name. Note that many proteins contain several differentially regulated ubiquitylation sites. (B) Protein interaction network of 21 proteins with putative MKRN1 ubiquitylation target sites (significantly reduced, shown in (A)). The functional interactions were obtained from the STRING and BioGrid databases and our study. Visualisation by Cytoscape. (C) Ubiquitin remnant profiling results for significantly regulated ubiquitylation sites (FDR < 10%) in proteins from network in (B). Mean and standard deviation of the mean (s.d.m., error bars) are given together with all data points. (D) Ubiquitin remnant profiling results for seven quantified ubiquitylation sites in RPS10 and RPS20. Significant changes are shown in black (FDR < 10%) and non-significant changes in grey. Representation as in (C). (E) Comparison of ubiquitylation sites in selected target proteins that are modified by ZNF598 and MKRN1 during RQC. (F) Comparison of enriched proteins from the interactomes for GFP-MKRN1^{wt} (over GFP, see Fig. 1A) and GFP-MKRN1^{RINGmut} (over GFP-MKRN1^{wt}, see Supplemental Fig. S7B) with the proteins containing MKRN1 ubiquitylation targets sites (UB, see (A)). Protein names of overlapping targets are given.

defence against poly(A) translation. We propose that MKRN1 is recruited by PABP to the beginning of poly(A) tails, including premature polyadenylation events within open reading frames, where it represents a physical "roadblock" to the translating ribosome. Upon contact with the translating ribosome, MKRN1 ubiquitylates K107 on RPS10, thereby stalling the ribosome before it translates the poly(A) tail. Subsequently, the trailing ribosomes collide with the initially stalled ribosome. ZNF598 recognises the collision interface and ubiquitylates the collided ribosomes (Simms et al. 2017; Juskiewicz et al. 2018). In summary, we suggest that a sequence of MKRN1-mediated and ZNF598-mediated ubiquitylation events on ribosomal proteins and possibly other factors, including PABPC1, triggers ribosome-associated quality control.

Differences between human and yeast RQC explain the requirement for MKRN1

Many known components of the RQC machinery, such as Listerin (Ltn1p in yeast) and ZNF598 (Hel2p in yeast), are identical from yeast to human, however the molecular signals that are recognised differ partially. In yeast, RQC can be triggered by an excess of positively charged amino acids (lysine and arginine), which are sensed while they pass through the ribosomal exit tunnel (Lu and Deutsch 2008; Letzring et al. 2013). In contrast, in human, sensing the aberrant mRNAs does not occur via the encoded amino acids but at the level of the mRNA sequence and corresponding tRNAs, such that only poly(A) effectively results in ribosome stalling (Arthur et al. 2015; Garzia et al. 2017; Juskiewicz and Hegde 2017). We propose that MKRN1 acts as direct reader of poly(A) sequences based on its interaction with PABP. Consistent with this conceptual difference, there is no functionally equivalent ortholog of MKRN1 in yeast (Yth1p and

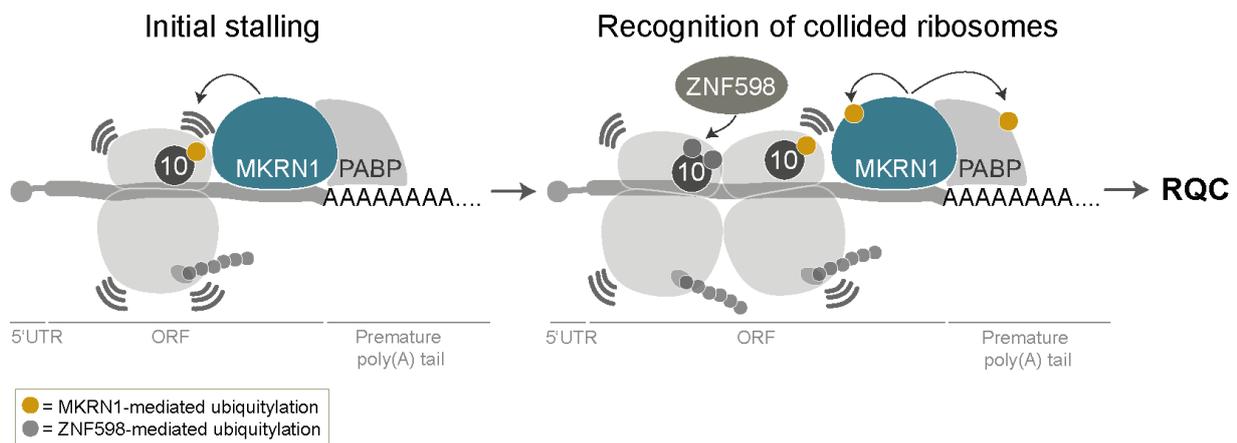


Figure IV-6, MKRN1 is a sensor for poly(A) sequences that stalls ribosomes to initiate ribosome-associated quality control. Proposed model of MKRN1 function: MKRN1 is positioned upstream of (premature) poly(A) tails via interaction with PABP. Ribosomes translating the open reading frame run into MKRN1 that acts as a roadblock to prohibit poly(A) translation. Upon contact with the translating ribosome, MKRN1 ubiquitylates the 40S ribosomal protein RPS10. This stalls the ribosome, causing the trailing ribosomes to collide. ZNF598 recognises the collided ribosomes and ubiquitylates ribosomal proteins to promote RQC.

Lee1p are similar, but lack RING domain and PAM2 motif; **Supplemental Fig. 1C**). Why yeast and human employ partially different mechanisms to detect poly(A) translation is currently unclear, but it has been suggested that spurious translation of poly-lysine stretches from long human poly(A) tails might target the aberrant proteins to the nucleus (Juzskiewicz and Hegde 2017). Loss of mRNA surveillance and RQC deficiency can lead to protein aggregation and culminate in proteotoxic stress, which in turn is lined to neurological disorders such as amyotrophic lateral sclerosis (Choe et al. 2016; Jamar et al. 2018). Hence, recognition of poly(A) sequences prior to their translation might be particularly beneficial in humans.

Materials and methods

Cell culture

HEK293T cells were obtained from DSMZ and cultured in DMEM (Life Technologies) with 10% fetal bovine serum (Life Technologies), 1% penicillin/streptomycin (Life Technologies), and 1% L-glutamine (Life Technologies). All cells were maintained at 37°C in a humidified incubator containing 5% CO₂ and routinely tested for mycoplasma infection. For SILAC labelling, cells were maintained in media containing either L-arginine and L-lysine (light SILAC label), L-arginine (¹³C₆) and L-lysine (²H₄) (medium SILAC label), or L-arginine (¹³C₆-¹⁵N₄) and L-lysine (¹³C₆-¹⁵N₂) (heavy SILAC label) (Cambridge Isotope Laboratories).

Vectors

The following vectors, suitable for Gateway Cloning, were obtained either from the IMB Core Facility ORFeome Collection (Collaboration 2016) or from the Harvard PlasmID Repository (<https://plasmid.med.harvard.edu/PLASMID/>): pENTR221-MKRN1, pENTR221-PABPC1, pENTR223.1-IGF2BP1, pENTR221-ELAVL1, pCMV-SPORT-ZNF598. Coding sequences from the entry vectors were cloned into the mammalian expression vectors pMX-DEST53-IP-GFP by LR Gateway cloning according to the manufacturer's recommendations (Gateway LR Clonase II Enzyme mix; Life Technologies). Dual fluorescence reporter plasmids (pmGFP-P2A-K₀-P2A-RFP, pmGFP-P2A-(K^{AAA})₁₂-P2A-RFP, pmGFP-P2A-(K^{AAA})₂₀-P2A-RFP, and pmGFP-P2A-(R^{CGA})₁₀-P2A-RFP) were generously provided by Ramanujan S. Hegde (MRC Laboratory of Molecular Biology, Cambridge, UK) (Juszkiewicz and Hegde 2017).

Cloning

All MKRN1 mutant plasmids were generated with the Q5 Site-Directed Mutagenesis Kit (NEB) according to the manufacturer's recommendations. In order to disrupt MKRN1's interaction with PABP (MKRN1^{PAM2mut}), three point mutations were introduced into the PAM2 motif (A169S, F172A, P174A; **Fig. 1B**) as previously described (Pohlmann et al. 2015). In MKRN1^{RINGmut}, a previously described mutation in the RING domain (H307E) was introduced to abolish E3 ubiquitin ligase function (Kim et al. 2005). All primers used for introducing mutations into MKRN1 are listed in **Supplemental Table S5**.

Transfections

Overexpression of vectors was performed using Polyethylenimine MAX 4000 (Polysciences, 24885-2) with a DNA:PEI ratio of 1:10. Knockdowns were performed with siRNAs (**Supplemental Table S6**) using Lipofectamine RNAiMAX (Life Technologies) according to the manufacturer's recommendations.

Affinity purification (AP) for Western blot analyses

GFP-based affinity purifications (APs) were performed as described before (Hildebrandt et al. 2017). In brief, HEK293T cells transiently expressing GFP (empty vector) or a GFP-tagged target protein were used. The cells were lysed in modified RIPA (mRIPA) buffer supplemented with protease inhibitors (protease inhibitor cocktail, Sigma), 1 mM

sodium orthovanadate, 5 mM β -glycerophosphate, 5 mM sodium fluoride, and 10 mM N-ethylmaleimide (NEM) (all from Sigma). Protein concentrations were determined using the Pierce BCA Protein Assay Kit (Thermo Fisher). GFP-trap agarose beads (Chromotek) were incubated with the cleared lysate for 1 h at 4°C. After five washes with mRIPA buffer, the beads were resuspended in LDS sample buffer (Life Technologies) and heated to 70°C for 10 min. For RNase digests, the enriched proteins were incubated with 0.5 U/ μ l RNase A (Qiagen) and 20 U/ μ l RNase T1 (Thermo Fisher Scientific) for 30 min at 4°C after the first two washes in mRIPA buffer.

Sample preparation for the protein interactome analysis

GFP-based APs were performed as described before (Hildebrandt et al. 2017). In brief, HEK293T cells transiently expressing GFP (empty vector) were cultured in light SILAC medium, while cells expressing N-terminally GFP-tagged MKRN1 wt or mutants were cultured in medium or heavy SILAC medium. The cells were lysed as described above. After washing in mRIPA buffer, GFP-trap agarose beads were incubated with the cleared lysate for 1 h at 4°C. All AP samples were washed four times with mRIPA buffer, combined and washed again in mRIPA buffer. The beads were heated in LDS sample buffer, supplemented with 1 mM dithiothreitol (DTT; Sigma, D5545) for 10 min at 70°C and alkylated using 5.5 mM 2-chloroacetamide (CAA; Sigma, C0267) for 30 min at RT in the dark (Nielsen et al. 2008).

Sample preparation for the proteome analysis

MKRN1 KD using siRNA2 was performed in heavy labelled SILAC cells and control KD was performed in light labelled SILAC cells in two replicates. For the third replicate, a label swop was performed, knocking down *MKRN1* (siRNA2) in light labelled SILAC cells and control in heavy labelled SILAC cells. For proteome analysis, cells were lysed as described above. Subsequently, 25 μ g protein from each SILAC condition (50 μ g in total) were pooled and processed as described below.

Sample preparation for mass spectrometry

The enriched proteins were resolved by SDS-PAGE on a NuPAGE 4-12% Bis-Tris protein gel (Thermo Fisher Scientific) and stained using the Colloidal Blue Staining Kit (Life Technologies). Proteins were in-gel digested using trypsin, before peptides were

extracted from the gel. To concentrate, clear and acidify the peptides, they were bound to C18 StageTips as described previously (Rappsilber et al. 2007).

Mass spectrometry data acquisition

Peptide fractions were analysed on a quadrupole Orbitrap mass spectrometer (Thermo Q Exactive Plus, Thermo Scientific) coupled to an uHPLC system (EASY-nLC 1000, Thermo Scientific) (Michalski et al. 2011). Peptide samples were separated on a C18 reversed phase column (length: 20 cm, inner diameter: 75 μ m, bead size: 1.9 μ m) and eluted in a linear gradient from 8 to 40% acetonitrile containing 0.1% formic acid in 105 min for the interactome analyses, in 175 min for the proteome analyses, or in 125 min for the ubiquitylome analyses. The mass spectrometer was operated in data-dependent positive mode, automatically switching between MS and MS² acquisition. The full scan MS spectra (m/z 300–1650) were acquired in the Orbitrap. Sequential isolation and fragmentation of the ten most abundant ions was performed by higher-energy collisional dissociation (HCD) (Olsen et al. 2007). Peptides with unassigned charge states, as well as with charge states less than +2 were excluded from fragmentation. The Orbitrap mass analyser was used for acquisition of fragment spectra.

Peptide identification and quantification

Raw data files were analysed and peptides were identified using the MaxQuant software (version 1.5.28) (Cox et al. 2009). Parent ion and MS² spectra were compared to a database containing 92,578 human protein sequences obtained from UniProtKB (release June 2018), coupled to the Andromeda search engine (Cox et al. 2011). Cysteine carbamidomethylation was set as a fixed modification. N-terminal acetylation, oxidation, and N-ethylmaleimide (NEM) were set as variable modifications. For ubiquitylome data analysis, glycine-glycine (GlyGly) modification of lysine was additionally set as a variable modification. The mass tolerance for the spectra search was set to be lower than 6 ppm in MS and 20 ppm in HCD MS² mode. Spectra were searched with strict trypsin specificity and allowing for up to three mis-cleavages. Site localisation probabilities were determined by MaxQuant using the PTM scoring algorithm as described previously (Elias and Gygi 2007; Cox and Mann 2008). Filtering of the dataset was based on the posterior error probability to arrive at a false discovery rate (FDR) < 1% estimated using a target-decoy approach. Proteins that were categorised as “only identified by site”, potential contaminants and reverse hits were

removed. Only proteins identified with at least two peptides (including at least one unique peptide) and a SILAC ratio count of at least two were used for analysis. For AP experiments, proteins that were quantified in at least two out of three experiments were kept for further analysis. In total, we quantified 1,106 and 1,097 protein groups in the AP experiments with GFP-MKRN1^{wt} (**Fig. 1A**), GFP-MKRN1^{PAM2mut} (**Fig. 1D**) and GFP-MKRN1^{RINGmut} (**Supplemental Fig. S7**), respectively (**Supplemental Table S1**). The SILAC ratios were log₂ transformed and converted into an asymmetric z-score based on the mean and interquartile range of the distribution as described previously (Cox and Mann 2008). For statistical analysis, a moderated t-test from the limma algorithm was used (Ritchie et al. 2015). Enriched proteins with an FDR < 5% were determined to be significantly enriched interactors (for GFP-MKRN1^{wt}). For proteins enriched in GFP-MKRN1^{RINGmut} over GFP-MKRN1^{wt}, proteins with an FDR < 5% and a GFP-MKRN1^{wt}/GFP z-score > 1 were selected. In the proteome experiment, we quantified 6,439 protein groups, present in all three replicates. Ratio-ratio and ratio-intensity plots were created in R (version 3.4.3) using RStudio (<http://www.rstudio.com/>).

Functional annotation of MKRN1 interactors and MKRN1-ubiquitylation targets.

In order to assess the functions of MKRN1-interacting proteins and proteins with MKRN1-dependent ubiquitylation sites, we performed gene ontology (GO) enrichment analyses using the Database for Annotation, Visualization and Integrated Discovery (DAVID 6.7) for three GO domains (Jiao et al. 2012). Enriched GO terms (modified Fisher exact test, adjusted *P* value < 0.05, Benjamini-Hochberg correction; **Supplemental Fig. S2B, S8A**) were visualised using REVIGO (Reduce & Visualize Gene Ontology) allowing medium GO term similarity (Supek et al. 2011).

Western blot

Denatured proteins were separated by SDS-PAGE on a NuPAGE 4-12% Bis-Tris protein gel (Life Technologies) and transferred to a 0.45 µm nitrocellulose membrane (VWR). For detection, either fluorophore-coupled secondary antibodies or HRP-conjugated secondary antibodies and WesternBright Chemiluminescent Substrate (Biozym Scientific) or SuperSignal West Pico Chemiluminescent Substrate (Life Technologies) were used. Western blots were quantified by determining the background-subtracted densities of the protein of interest using ImageJ (Schindelin et al. 2015). The signal from the AP (against GFP-tagged protein of interest) was

normalised to the respective control samples expressing the empty vector or to the input.

Antibodies

The following antibodies were used: anti-GFP (B-2 clone; Santa Cruz; sc-9996), anti-MKRN1 (Bethyl Laboratories, A300-990A), anti-PABPC1/3 (Cell Signaling, 4992), anti-Znf598 (N1N3; GeneTex; GTX119245), anti- α Tubulin (Sigma Aldrich, T-5168), anti-Rabbit IgG (Cell Signaling; 7074), anti-Mouse IgG (Cell Signaling; 7076), IRDye 680RD Goat anti-Mouse IgG (P/N 925-68070), and IRDye 800CW Goat anti-Rabbit IgG (P/N 925-32211) (both LI-COR Biosciences GmbH).

RNA isolation, cDNA synthesis and qPCR

Cells were washed twice in ice-cold PBS and harvested. RNA was isolated using the RNeasy Plus Mini Kit (Qiagen) according to the manufacturer's recommendations. 500 ng total RNA was transcribed into cDNA using random hexamer primers (Thermo Scientific) and the RevertAid Reverse Transcriptase (Thermo Scientific) according to the manufacturer's recommendations. qPCR was performed using the Luminaris HiGreen qPCR Master Mix, low ROX (Thermo Scientific) according to the manufacturer's recommendations with 10 μ M forward and reverse primers (**Supplemental Table S5**).

iCLIP experiments and data processing

iCLIP libraries were prepared as described previously (Huppertz et al. 2014; Sutandy et al. 2016). HEK293T cells ectopically expressing either GFP alone (empty vector) or N-terminally GFP-tagged MKRN1 wild type (GFP-MKRN1^{wt}), GFP-MKRN1^{PAM2mut}, or GFP-MKRN1^{RINGmut} were used. For crosslinking, confluent cells were irradiated once with 150 mJ/cm² at 254 nm in a Stratalinker 2400 or treated with 4-thiouridine (100 μ M for 16 h) and irradiated with 3x 300 mJ/cm² in a Stratalinker 2400 with 365 nm bulbs. For IP, 10.5 μ g anti-GFP antibody (goat, Protein Unit, MPI-CBG, Dresden) were used per sample. The libraries were sequenced as 50-nt single-end reads on an Illumina MiSeq platform (**Supplemental Table S2**).

Basic sequencing quality checks were applied to all reads using FastQC (version 0.11.5) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Afterwards, reads

were filtered based on sequencing qualities (Phred score) of the barcode region. Only reads with at most one position with a sequencing quality < 20 in the experimental barcode (positions 4 to 7) and without any position with a sequencing quality < 17 in the random barcode (positions 1-3 and 8-9) were kept for further analysis. Remaining reads were de-multiplexed based on the experimental barcode on positions 4 to 7 using Flexbar (version 3.0.0) (Dodt et al. 2012) without allowing mismatches.

All following steps of the analysis were performed on all individual samples after de-multiplexing. Remaining adapter sequences were trimmed from the right end of the reads using Flexbar (version 3.0.0) allowing up to one mismatch in 10 nt, requiring a minimal overlap of 1 nt of read and adapter. After trimming off the adapter, the barcode is trimmed off of the left end of the reads (first 9 nt) and added to the header of the read, such that the information is kept available for downstream analysis. Reads shorter than 15 nt were removed from further analysis.

Trimmed and filtered reads were mapped to the human genome (assembly version GRCh38) and its annotation based on GENCODE release 25 (Harrow et al. 2012) using STAR (version 2.5.4b) (Dobin et al. 2013). When running STAR, up to two mismatches were allowed, soft-clipping was prohibited and only uniquely mapping reads were kept for further analysis.

Following mapping, duplicate reads were marked using the dedup function of bamUtil (version 1.0.13), which defines duplicates as reads whose 5' ends map to the same position in the genome (<https://github.com/statgen/bamUtil>). Subsequently, marked duplicates with identical random barcodes were removed since they are considered technical duplicates, while biological duplicates showing unequal random barcodes were kept.

Resulting bam files were sorted and indexed using SAMtools (version 1.5) (Li et al. 2009). Based on the bam files, bedgraph files were created using bamToBed of the BEDTools suite (version 2.25.0) (Quinlan and Hall 2010), considering only the position upstream of the 5' mapping position of the read, since this nucleotide is considered as the crosslinked nucleotide. bedgraph files were then transformed to bigWig file format using bedGraphToBigWig of the UCSC tool suite (Kent et al. 2010).

Identification and characterisation of MKRN1 binding sites

Peak calling was performed on merged iCLIP coverage tracks (crosslink events per nucleotide) from the three replicates based on GENCODE annotation (release 27, GRCh38) using ASPeak (version 2.0; default setting plus `-nornaseq` to estimate parameters p and r for the negative binomial distributions in a 500-nt window around each peak) (Kucukural et al. 2013). The initially predicted peaks were resized to uniform 9-nt windows around their weighted centred as defined by ASPeak. To avoid artefacts, we removed sparsely covered peaks that harbour crosslink events on less than three nucleotides within the 9-nt region window. We iteratively merged all remaining windows if overlapping by at least 1 nt, by defining the position with the cumulative half maximum count of crosslink events as new window centre. We further kept only reproducible windows with at least three crosslink events from any two replicates. Finally, we excluded all windows overlapping with none or multiple protein-coding genes (GENCODE annotations support level ≥ 2 and transcript support level ≥ 3), and assign each binding site to a distinct genomic region (3' UTR, 5' UTR, CDS, intron). Consistent with the mostly cytoplasmic localisation of MKRN1 (Miroci et al. 2012; Cassar et al. 2015; Hildebrandt et al. 2017), less than 6% of the binding sites were predicted within introns, which were excluded from further analysis. This procedure yielded a total of 7,331 MKRN1 binding sites in 2,163 genes.

In order to estimate binding site strength and to facilitate comparisons between binding sites (**Fig. 2B,D** and **Supplemental Fig. 3D-F, 4A,C,D**), we corrected for transcript abundance by representing the crosslink events within a binding site as a 'signal-over-background' ratio (SOB). The respective background was calculated as the sum of crosslink events outside of binding sites (plus 5 nt to either side) by the merged length of all exons. 3' UTR lengths were restricted to 10 nt past the last MKRN1 binding site or 500 nt if no binding site was present. SOB calculations were performed separately for each replicate and then averaged. No SOB value was assigned for genes with a background of < 10 crosslink events, resulting in SOB values for 97% of all binding sites.

In order to assess the local RNA sequence context of MKRN1 binding sites (**Fig. 2B** and **Supplemental Fig. S4A**), enriched 4-mers were counted inside the 9-nt binding sites as well as within 40-nt before and after. To estimate an empirical background distribution, 1,000 9-nt windows were randomly picked in 3' UTRs and 4-mer frequencies were counted in the same windows. This process was repeated 100 times,

and the resulting mean and standard deviation were used to calculate the z-score for each 4-mer.

In order to define the A-rich regions downstream of MKRN1 binding sites in 3' UTRs (A-rich stretches), we used a maximisation approach in a 55-nt search space starting from the binding site centre. Within this space, we calculated the percentage of A nucleotides (A-content) for windows of increasing size (8-30 nt) and selected the stretch with highest value for each window size. In case of ties, the window closer to the binding site was preferred, resulting in a set of 23 candidate A-stretches with the maximal A-content for each length. Next, we computed the longest continuous A run (LCA) and a weighted A-content (multiplying the A-content with the number of A nucleotides) for each candidate A-stretch. Candidate A-stretches with an A-content < 70%, a weighted A-content < 11 and an LCA < 4 were excluded. The final A-stretch for each binding site was then selected in a hierarchical manner, preferring LCA over weighted A-content. Lastly, overlapping A-stretches of neighbouring binding sites were merged by selecting the highest scoring A-stretch, based on LCA and weighted A-content. In total, this procedure identified 1,412 non-overlapping A-stretches, associated with 1,848 binding sites.

In order to estimate the extent of MKRN1 binding to poly(A) tails (**Fig. 3B**), we evaluated the percentage of adenosine within the iCLIP reads that could not be mapped to the human genome without soft-clipping (see above). iCLIP data for heterogeneous nuclear ribonucleoprotein H (HNRNPH) served as control (Braun et al. 2018). Annotated transcript 3' ends (i.e. polyadenylation sites) were taken from GENCODE (all annotated protein-coding transcripts with support level ≤ 2 and transcript support level ≤ 3 ; release 28, GRCh38.p12; www.gencodegenes.org). For **Fig. 3C**, all crosslink events within a 2-kb window around the polyadenylation sites for 3' UTR longer than 1 kb were counted.

Evolutionary characterisation of Makorin protein family

Four different ortholog searches were performed using HaMStR-OneSeq (Ebersberger et al. 2014) against the Quest for Orthologs Consortium protein set, containing 78 species (release 2017_04) (Sonnhammer et al. 2014). For each run, a different seed protein was chosen: human MKRN1-3 (UniProt identifiers Q9UHC7, Q9H000 and Q13064) and MKRN4 from zebrafish (A9C4A6). In order to identify proteins with a similar domain architecture, we calculated a unidirectional feature architecture similarity

(FAS) score which compares the domain architecture of the seed protein and the predicted ortholog (Koestler 2010). Predicted orthologues with FAS < 0.7 were removed after initial assessment. Finally, all vertebrate species and selected invertebrate species were used for reconstruction of a maximum likelihood (ML) tree. For this, protein sequences were aligned using MAFFT v7.294b L-INS-I (Kato and Standley 2013), and ML trees with 100 bootstrap replicates were calculated using RAxML version 8.1.9 (Stamatakis 2014). Settings for a rapid bootstrap analysis and searching for the best scoring ML tree in one program run (-f a) and an automatic selection of the best fitting amino acid substitution model (-m PROTGAMMAAUTO) were chosen. Reconstructed trees were visualised using FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

The phylogenetic tree and FASTA sequences from the ortholog dataset were loaded into DoMosaics (<http://www.domosaics.net>) and Pfam domains were annotated with HMMER (<http://hmmer.org/>, default parameters). Since the PAM2 motif in all Makorin proteins differs from the described consensus motif (Albrecht and Lengauer 2004), a custom Hidden Markov Model was trained on PAM2 motifs from selected Makorin orthologs and used for a HMMER scan of the orthologs (no E-value cutoff). The same procedure was repeated for the PAM2-like motif (PAM2L) (Pohlmann et al. 2015).

Dual fluorescence translation stall assay via flow cytometry

Knockdowns were performed for 24 h, before the dual fluorescence reporter plasmids were ectopically expressed for 48 h. Cells were washed in PBS and trypsinised. After sedimentation, cells were resuspended in DPBS supplemented with 2 mM EDTA. Cellular GFP and RFP fluorescence was measured using flow cytometry on a LSRFortessa SORP (BD Biosciences). Data analysis was done using FlowJo (v10) (FlowJo, LLC). For statistical testing, paired two-tailed Student's t-tests with Benjamini-Hochberg correction were performed on $n \geq 6$ replicates.

Ubiquitin remnant profiling

Di-glycine remnant profiling was performed as described before (Wagner et al. 2011;L Heidelberg et al. 2018). In four different experiments, isotope labels were assigned as follows: experiment 1, *MKRN1* KD1 (siRNA1), *MKRN1* KD2 (siRNA2) and control siRNA with light, medium and heavy SILAC labels, respectively; experiment 2, *MKRN1* KD2 (siRNA2) and control siRNA with heavy and light SILAC labels, respectively; experiment 3, *MKRN1* KD2 (siRNA2) and control siRNA with heavy and light SILAC labels,

respectively; experiment 3, *MKRN1* KD2 (siRNA2) and control siRNA with light and heavy SILAC labels, respectively. Cells were treated with the proteasome inhibitors bortezomib (1 μ M, 8h, replicate 1; Santa Cruz Biotechnology) or MG132 (10 μ M, 2 h, replicates 2, 3, 4; Sigma). Proteins were precipitated in acetone. Proteins were digested with endoproteinase Lys-C (Wako Chemicals) and sequencing-grade modified trypsin (Sigma). To purify the peptides, reversed-phase Sep-Pak C18 cartridges (Waters) were used. Modified peptides were enriched using di-glycine-lysine antibody resin (Cell Signaling Technology). The enriched peptides were eluted with 0.15% trifluoroacetic acid in water, then fractionated using micro-column-based strong-cation exchange chromatography (SCX) (Weinert et al. 2013) before being desalted on reversed-phase C18 StageTips (Rappsilber et al. 2007). Samples were analysed by quantitative mass spectrometry and MaxQuant as described above. To identify significantly regulated ubiquitylation sites, the limma algorithm was applied (Ritchie et al. 2015). A *P* value < 0.1 after multiple testing correction was used as a cut-off to determine up- and downregulated ubiquitylation sites. Volcano and dot plots were created in R (version 3.4.3).

Functional interaction network of MKRN1 ubiquitylation target proteins

The functional protein interaction network analysis was performed by integrating interaction data from the STRING database (score > 0.4), the BioGrid database and our own findings (Franceschini et al. 2013; Chatr-Aryamontri et al. 2017). Cytoscape (version 3.6.1) was used to visualise the protein interaction network (Saito et. al 2012).

Availability of data and materials

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository with the dataset identifier PXD011772.

Raw and processed iCLIP data are available at GEO under the accession number GSE122869.

Acknowledgements

We would like to thank all members of the Zarnack, König and Beli labs, as well as René Ketting, Nadine Wittkopp, and Miguel Almeida for fruitful discussions. The authors

gratefully thank the Ramanujan S. Hegde for providing the dual fluorescence reporter plasmids. We thank Anja Freiwald for assistance with mass spectrometry analysis and Dr. Stefan Simm for assistance with evaluating PAM2 mutations. The support of the IMB Core Facilities Bioinformatics, Flow Cytometry, Genomics, and the use of its Illumina MiSeq, as well as the DFG-funded Mass Spectrometer Q Exactive Plus (INST 247/766-1 FUGG) are gratefully acknowledged. K.Z. was supported by the LOEWE program Ubiquitin Networks (Ub-Net) of the State of Hesse (Germany) and the SFB 902 of the German Research Foundation. P.B. is supported by the Emmy Noether Program (BE 5342/1-1), the SFB 1177 of the German Research Foundation and the Marie Curie Career Integration Grant from the European Commission (grant agreement number: 630763). The project was funded by the German Research Foundation (DFG) as part of SPP1925 to J.-Y.R. (RO 4681/4-1) and J.K. (KO 4566/3-1). Animal shapes in Supplemental Fig. S1A were obtained from PhyloPic and are used under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported license.

Author contributions

A.H. performed iCLIP experiments, flow cytometry measurements of dual fluorescence reporters and most proteomics experiments. M.B. performed most bioinformatics analyses of MKRN1 iCLIP data. C.R. analysed MKRN1 binding at polyadenylation sites and poly(A) tails. A.B. and S.B. performed initial iCLIP data processing and analysis. A.H. and A.B. analysed the proteomics data. J.B.H. and A.V. contributed to replicate ubiquitin remnant profiling experiments and AP-Western blot experiments, respectively. H.H. performed replicate iCLIP and replicate AP-Western blot experiments. C.R. and I.E. contributed evolutionary characterisation of Makorin proteins. A.D. and J.-Y.R. performed complementary studies in *D. melanogaster*. S.E. and K.Z. supervised the bioinformatics analyses. J.K. and P.B. conceived the project with K.Z. and supervised the experimental work. A.H., J.K., K.Z. and P.B. wrote the manuscript with help and comments from all co-authors.

Competing interests

The authors declare that they have no competing interests.

References

- Albrecht M, Lengauer T. 2004. Survey on the PABC recognition motif PAM2. *Biochem Biophys Res Commun* 316: 129-138.
- Arthur L, Pavlovic-Djuranovic S, Smith-Koutmou K, Green R, Szczesny P, Djuranovic S. 2015. Translational control by lysine-encoding A-rich sequences. *Sci Adv* 1.
- Bag J. 2001. Feedback inhibition of poly(A)-binding protein mRNA translation. A possible mechanism of translation arrest by stalled 40 S ribosomal subunits. *J Biol Chem* 276: 47352-47360.
- Bengtson MH, Joazeiro CA. 2010. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467: 470-473.
- Böhne A, Darras A, D'Cotta H, Baroiller JF, Galiana-Arnoux D, Volff JN. 2010. The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* 11: 721.
- Brandman O, Hegde RS. 2016. Ribosome-associated protein quality control. *Nat Struct Mol Biol* 23: 7-15.
- Brandman O, Stewart-Ornstein J, Wong D, Larson A, Williams CC, Li GW, Zhou S, King D, Shen PS, Weibezahn J et al. 2012. A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* 151: 1042-1054.
- Braun S, Enculescu M, Setty ST, Cortés-López M, de Almeida BP, Sutandy FXR, Schulz L, Busch A, Seiler M, Ebersberger S et al. 2018. Decoding a cancer-relevant splicing decision in the RON proto-oncogene using high-throughput mutagenesis. *Nat Commun* 9: 3315.
- Cassar PA, Carpenedo RL, Samavarchi-Tehrani P, Olsen JB, Park CJ, Chang WY, Chen Z, Choey C, Delaney S, Guo H et al. 2015. Integrative genomics positions MKRN1 as a novel ribonucleoprotein within the embryonic stem cell gene regulatory network. *EMBO Rep* 16: 1334-1357.
- Chatr-Aryamontri A, Oughtred R, Boucher L, Rust J, Chang C, Kolas NK, O'Donnell L, Oster S, Theesfeld C, Sellam A et al. 2017. The BioGRID interaction database: 2017 update. *Nucleic Acids Res* 45: D369-D379.
- Choe YJ, Park SH, Hassemer T, Korner R, Vincenz-Donnelly L, Hayer-Hartl M, Hartl FU. 2016. Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature* 531: 191-195.
- Chu J, Hong NA, Masuda CA, Jenkins BV, Nelms KA, Goodnow CC, Glynne RJ, Wu H, Masliah E, Joazeiro CA et al. 2009. A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. *Proc Natl Acad Sci U S A* 106: 2097-2103.
- Collaboration O. 2016. The ORFeome Collaboration: a genome-scale human ORF-clone resource. *Nat Methods* 13: 191-192.
- Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26: 1367-1372.
- Cox J, Matic I, Hilger M, Nagaraj N, Selbach M, Olsen JV, Mann M. 2009. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat Protoc* 4: 698-705.
- Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. 2011. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 10: 1794-1805.
- Deo RC, Sonenberg N, Burley SK. 2001. X-ray structure of the human hyperplastic discs protein: an ortholog of the C-terminal domain of poly(A)-binding protein. *Proc Natl Acad Sci U S A* 98: 4414-4419.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15-21.
- Dodt M, Roehr JT, Ahmed R, Dieterich C. 2012. FLEXBAR-Flexible Barcode and Adapter Processing for Next-Generation Sequencing Platforms. *Biology (Basel)* 1: 895-905.
- Ebersberger I, Simm S, Leisegang MS, Schmitzberger P, Mirus O, von Haeseler A, Bohnsack MT, Schleiff E. 2014. The evolution of the ribosome biogenesis pathway from a yeast perspective. *Nucleic Acids Res* 42: 1509-1523.
- Elias JE, Gygi SP. 2007. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 4: 207-214.
- Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C et al. 2013. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 41: D808-815.
- Garzia A, Jafarnejad SM, Meyer C, Chapat C, Gogakos T, Morozov P, Amiri M, Shapiro M, Molina H, Tuschl T et al. 2017. The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun* 8: 16056.
- Gray TA, Hernandez L, Carey AH, Schaldach MA, Smithwick MJ, Rus K, Marshall Graves JA, Stewart CL, Nicholls RD. 2000. The ancient source of a distinct gene family encoding proteins featuring RING and C(3)H zinc-finger motifs with abundant expression in developing brain and nervous system. *Genomics* 66: 76-86.

- Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M et al. 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141: 129-141.
- Harrow J, Frankish A, Gonzalez JM, Tapanari E, Diekhans M, Kokocinski F, Aken BL, Barrell D, Zadissa A, Searle S et al. 2012. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* 22: 1760-1774.
- Heidelberger JB, Voigt A, Borisova ME, Petrosino G, Ruf S, Wagner SA, Beli P. 2018. Proteomic profiling of VCP substrates links VCP to K6-linked ubiquitylation and c-Myc function. *EMBO Rep* 19.
- Hildebrandt A, Alanis-Lobato G, Voigt A, Zarnack K, Andrade-Navarro MA, Beli P, König J. 2017. Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system. *Sci Rep* 7: 16582.
- Huppertz I, Attig J, D'Ambrogio A, Easton LE, Sibley CR, Sugimoto Y, Tajnik M, König J, Ule J. 2014. iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* 65: 274-287.
- Jamar NH, Kritsiligkou P, Grant CM. 2018. Loss of mRNA surveillance pathways results in widespread protein aggregation. *Sci Rep* 8: 3894.
- Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, Lempicki RA. 2012. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* 28: 1805-1806.
- Joazeiro CAP. 2017. Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. *Annu Rev Cell Dev Biol* 33: 343-368.
- Juszkiewicz S, Chandrasekaran V, Lin Z, Kraatz S, Ramakrishnan V, Hegde RS. 2018. ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol Cell* 72: 469-481 e467.
- Juszkiewicz S, Hegde RS. 2017. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell* 65: 743-750 e744.
- Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. 2010. U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468: 664-668.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30: 772-780.
- Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. 2010. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* 26: 2204-2207.
- Kim JH, Park KW, Lee EW, Jang WS, Seo J, Shin S, Hwang KA, Song J. 2014. Suppression of PPARgamma through MKRN1-mediated ubiquitination and degradation prevents adipocyte differentiation. *Cell Death Differ* 21: 594-603.
- Kim JH, Park SM, Kang MR, Oh SY, Lee TH, Muller MT, Chung IK. 2005. Ubiquitin ligase MKRN1 modulates telomere length homeostasis through a proteolysis of hTERT. *Genes Dev* 19: 776-781.
- Kini HK, Silverman IM, Ji X, Gregory BD, Liebhaber SA. 2016. Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. *RNA* 22: 61-74.
- Ko A, Shin JY, Seo J, Lee KD, Lee EW, Lee MS, Lee HW, Choi IJ, Jeong JS, Chun KH et al. 2012. Acceleration of gastric tumorigenesis through MKRN1-mediated posttranslational regulation of p14ARF. *J Natl Cancer Inst* 104: 1660-1672.
- Koestler T, von Haeseler A, Ebersberger I. 2010. FACT: functional annotation transfer between proteins with similar feature architectures. *BMC Bioinformatics* 11: 417.
- König J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J. 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* 17: 909-915.
- Kozlov G, De Crescenzo G, Lim NS, Siddiqui N, Fantus D, Kahvejian A, Trempe JF, Elias D, Ekiel I, Sonenberg N et al. 2004. Structural basis of ligand recognition by PABC, a highly specific peptide-binding domain found in poly(A)-binding protein and a HECT ubiquitin ligase. *EMBO J* 23: 272-281.
- Kozlov G, Menade M, Rosenauer A, Nguyen L, Gehring K. 2010. Molecular determinants of PAM2 recognition by the MLL domain of poly(A)-binding protein. *J Mol Biol* 397: 397-407.
- Kozlov G, Trempe JF, Khaleghpour K, Kahvejian A, Ekiel I, Gehring K. 2001. Structure and function of the C-terminal PABC domain of human poly(A)-binding protein. *Proc Natl Acad Sci U S A* 98: 4409-4413.
- Kucukural A, Ozadam H, Singh G, Moore MJ, Cenik C. 2013. ASPeak: an abundance sensitive peak detection algorithm for RIP-Seq. *Bioinformatics* 29: 2485-2486.
- Lee EW, Lee MS, Camus S, Ghim J, Yang MR, Oh W, Ha NC, Lane DP, Song J. 2009. Differential regulation of p53 and p21 by MKRN1 E3 ligase controls cell cycle arrest and apoptosis. *EMBO J* 28: 2100-2113.
- Letzring DP, Wolf AS, Brule CE, Grayhack EJ. 2013. Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA* 19: 1208-1217.

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078-2079.
- Lima SA, Chipman LB, Nicholson AL, Chen YH, Yee BA, Yeo GW, Collier J, Pasquinelli AE. 2017. Short poly(A) tails are a conserved feature of highly expressed genes. *Nat Struct Mol Biol* 24: 1057-1063.
- Lu J, Deutsch C. 2008. Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J Mol Biol* 384: 73-86.
- Lyabin DN, Eliseeva IA, Skabkina OV, Ovchinnikov LP. 2011. Interplay between Y-box-binding protein 1 (YB-1) and poly(A) binding protein (PABP) in specific regulation of YB-1 mRNA translation. *RNA Biol* 8: 883-892.
- Michalski A, Damoc E, Hauschild JP, Lange O, Wieghaus A, Makarov A, Nagaraj N, Cox J, Mann M, Horning S. 2011. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol Cell Proteomics* 10: M111 011015.
- Miroci H, Schob C, Kindler S, Olschlager-Schutt J, Fehr S, Jungenitz T, Schwarzacher SW, Bagni C, Mohr E. 2012. Makorin ring zinc finger protein 1 (MKRN1), a novel poly(A)-binding protein-interacting protein, stimulates translation in nerve cells. *J Biol Chem* 287: 1322-1334.
- Nielsen ML, Vermeulen M, Bonaldi T, Cox J, Moroder L, Mann M. 2008. Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat Methods* 5: 459-460.
- Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M. 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4: 709-712.
- Omwancha J, Zhou XF, Chen SY, Baslan T, Fisher CJ, Zheng Z, Cai C, Shemshedini L. 2006. Makorin RING finger protein 1 (MKRN1) has negative and positive effects on RNA polymerase II-dependent transcription. *Endocrine* 29: 363-373.
- Pohlmann T, Baumann S, Haag C, Albrecht M, Feldbrügge M. 2015. A FYVE zinc finger domain protein specifically links mRNA transport to endosome trafficking. *Elife* 4.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-842.
- Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2: 1896-1906.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47.
- Saito R, Smoot ME, Ono K, Ruscheinski J, Wang PL, Lotia S, Pico AR, Bader GD, Ideker T. 2012. A travel guide to Cytoscape plugins. *Nat Methods* 9: 1069-1076.
- Salvatico J, Kim JH, Chung IK, Muller MT. 2010. Differentiation linked regulation of telomerase activity by Makorin-1. *Mol Cell Biochem* 342: 241-250.
- Schindelin J, Rueden CT, Hiner MC, Eliceiri KW. 2015. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol Reprod Dev* 82: 518-529.
- Simms CL, Yan LL, Zaher HS. 2017. Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol Cell* 68: 361-373 e365.
- Simsek D, Tiu GC, Flynn RA, Byeon GW, Leppek K, Xu AF, Chang HY, Barna M. 2017. The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* 169: 1051-1065 e1018.
- Sonnhammer EL, Gabaldon T, Sousa da Silva AW, Martin M, Robinson-Rechavi M, Boeckmann B, Thomas PD, Dessimoz C, Quest for Orthologs c. 2014. Big data and other challenges in the quest for orthologs. *Bioinformatics* 30: 2993-2998.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312-1313.
- Sundaramoorthy E, Leonard M, Mak R, Liao J, Fulzele A, Bennett EJ. 2017. ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol Cell* 65: 751-760 e754.
- Supek F, Bosnjak M, Skunca N, Smuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* 6: e21800.
- Sutandy FXR, Ebersberger S, Huang L, Busch A, Bach M, Kang HS, Fallmann J, Maticzka D, Backofen R, Stadler PF et al. 2018. In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. *Genome Res* 28: 699-713.
- Sutandy FXR, Hildebrandt A, König J. 2016. Profiling the Binding Sites of RNA-Binding Proteins with Nucleotide Resolution Using iCLIP. *Methods Mol Biol* 1358: 175-195.
- Verma R, Oania RS, Kolawa NJ, Deshaies RJ. 2013. Cdc48/p97 promotes degradation of aberrant nascent polypeptides bound to the ribosome. *Elife* 2: e00308.
- Wagner SA, Beli P, Weinert BT, Nielsen ML, Cox J, Mann M, Choudhary C. 2011. A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. *Mol Cell Proteomics* 10: M111 013284.

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

- Webster MW, Chen YH, Stowell JAW, Alhusaini N, Sweet T, Graveley BR, Collier J, Passmore LA. 2018. mRNA Deadenylation Is Coupled to Translation Rates by the Differential Activities of Ccr4-Not Nucleases. *Mol Cell* 70: 1089-1100 e1088.
- Weinert BT, Scholz C, Wagner SA, Iesmantavicius V, Su D, Daniel JA, Choudhary C. 2013. Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. *Cell Rep* 4: 842-851.

Supplementary Tables and Figures

Supplementary Tables

Supplementary Table IV-1 is provided as an Excel file with the publication

Supplementary Table IV-2, Summary of MKRN1 iCLIP experiments. iCLIP experiments with GFP-MKRN1 were performed in three independent replicates.

	Sequenced reads	Uniquely mapped reads	Crosslink events	Binding sites
Replicate 1	3,418,021	1,561,445	957,097	-
Replicate 2	6,527,256	3,149,583	1,972,821	-
Replicate 3	4,660,274	2,515,161	2,293,633	-
Total	14,605,551	7,226,189	5,223,551	7,331

Supplementary Table IV-3 is provided as an Excel file with the publication

Supplementary Table IV-4 is provided as an Excel file with the publication

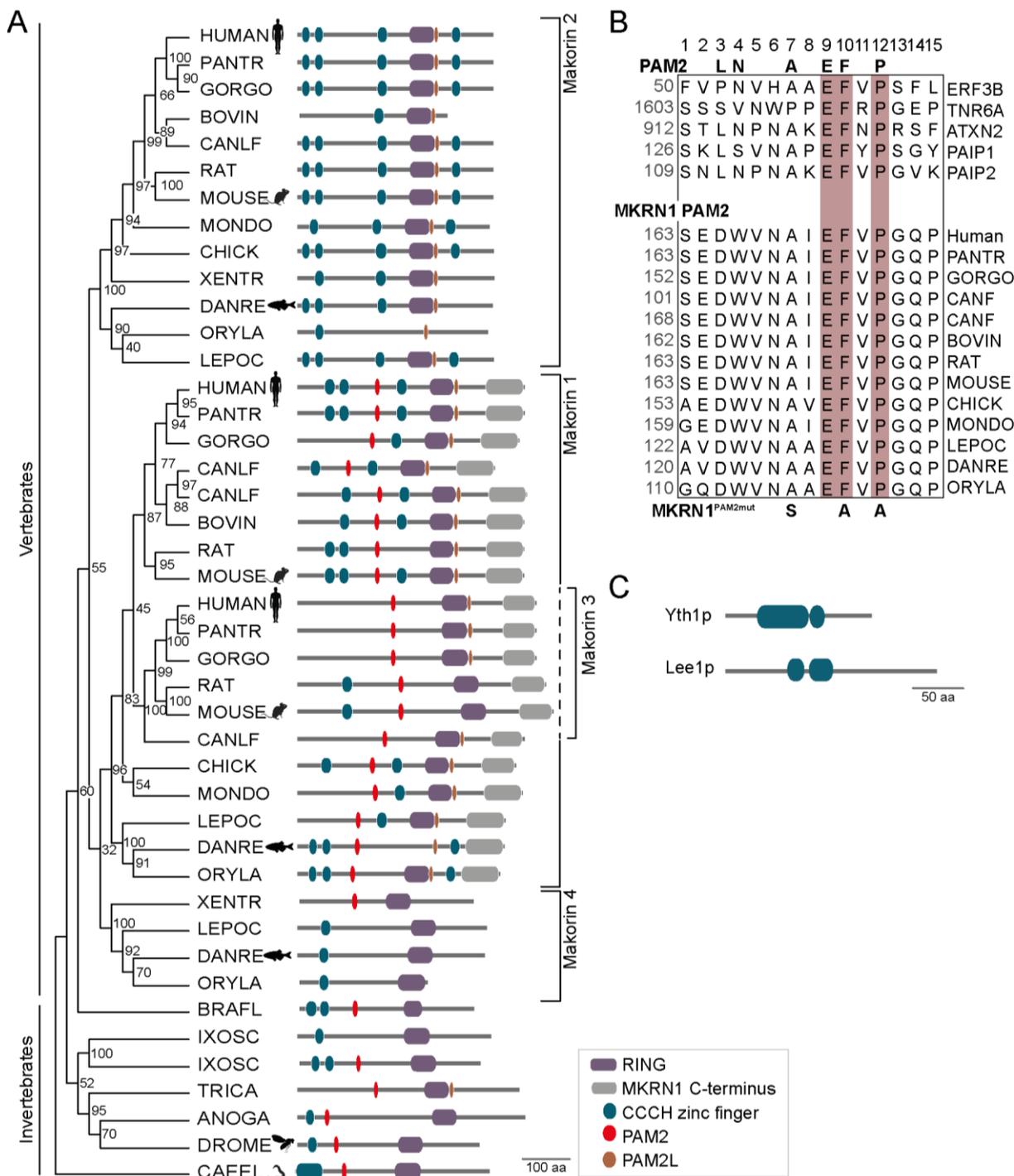
Supplementary Table IV-5, Primers used in this study.

Name	Sequence 5' – 3'	Comment
PAMmut	GCCGTTGCCGGGCAACCCTACTGTGGC	MKRN1 ^{PAM2mut}
	CTCAATAGAATTCACCCAGTCCTCTGAACC	mutant
H307E	CAACTGCAACGAAACCTACTGTCTCAAG	MKRN1 ^{RINGmut}
	GAGAGGATCCCGAAGCGG	mutant
MKRN1 qPCR	CGATACGGGGAGAACTGTGT	<i>MKRN1</i> qPCR
	CCTTCTCATGGGCCTCAAT	primer
MKRN2 qPCR	ACTCACATGAACCCGGAAAG	<i>MKRN2</i> qPCR
	AGCTGCCTGGATTACTCACC	primer
ZNF598 qPCR	AACCTCGACAAATGGTCCTG	<i>ZNF598</i> qPCR
	GTCTTCGTCCTTGAGCTTCG	primer
βActin qPCR	TCCTCCCTGGAGAAGAGCTAC	<i>β-actin</i> qPCR
	TGGAGTTGAAGGTAGTTCGTG	primer

Supplementary Table IV-6, siRNAs used in this study.

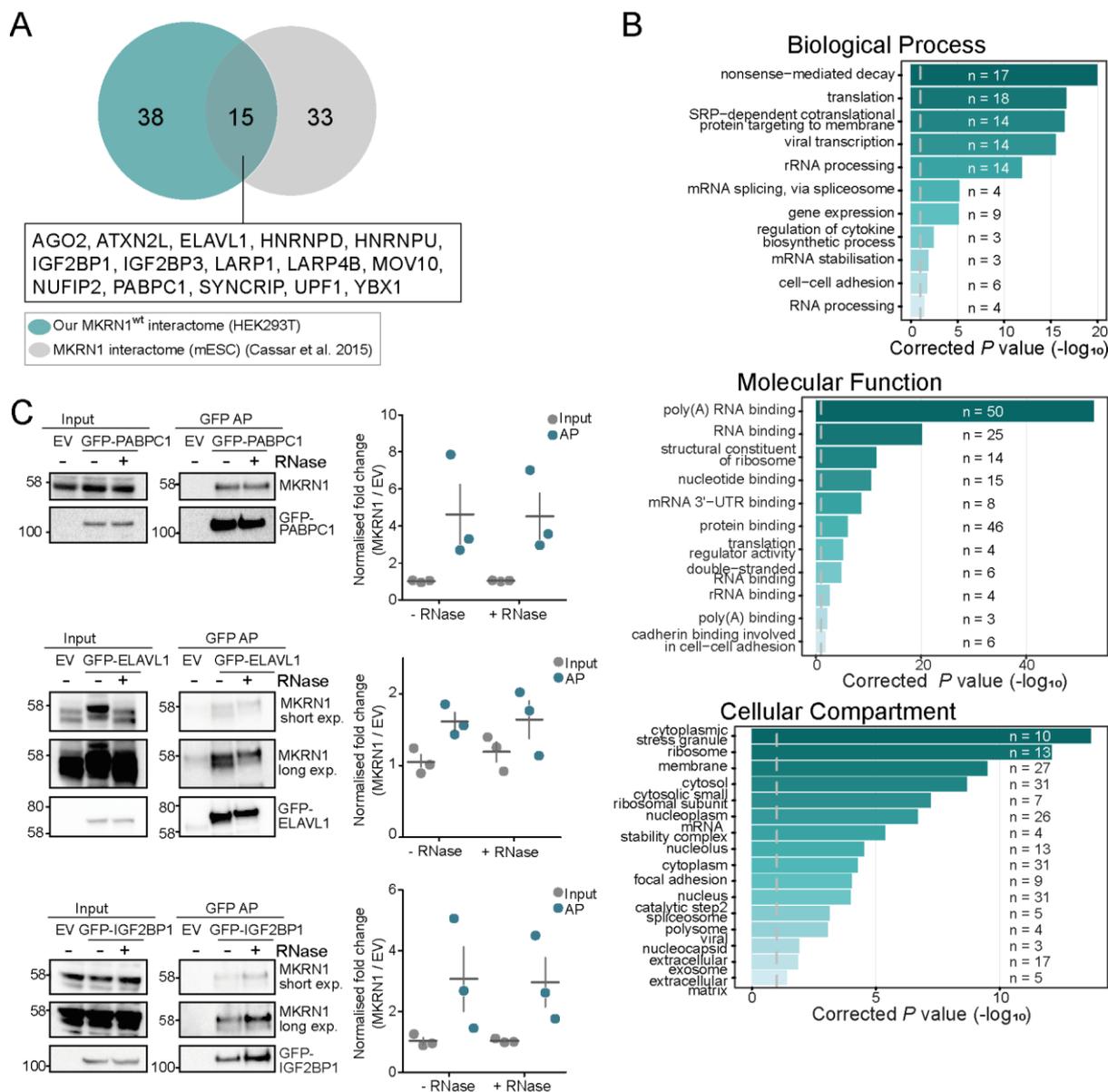
Name	Sequence 5' – 3'	Comment
MKRN1 siRNA1	CAGGCGAAGCUGAGUCAAGAA[dT][dT]	Ko et al. 2012
MKRN1 siRNA2	CGGGAUCCUCUCCAACUGCAA[dT][dT]	Kim et al. 2014
ZNF598 siRNA	CCCUCUAAAGUUGGGAAGA[dT][dT]	Sigma, Rosetta predictions
Control siRNA	UGGUUUACAUGUCGACUAA[dT][dT]	Heidelberger et al. 2018

Supplementary Figures



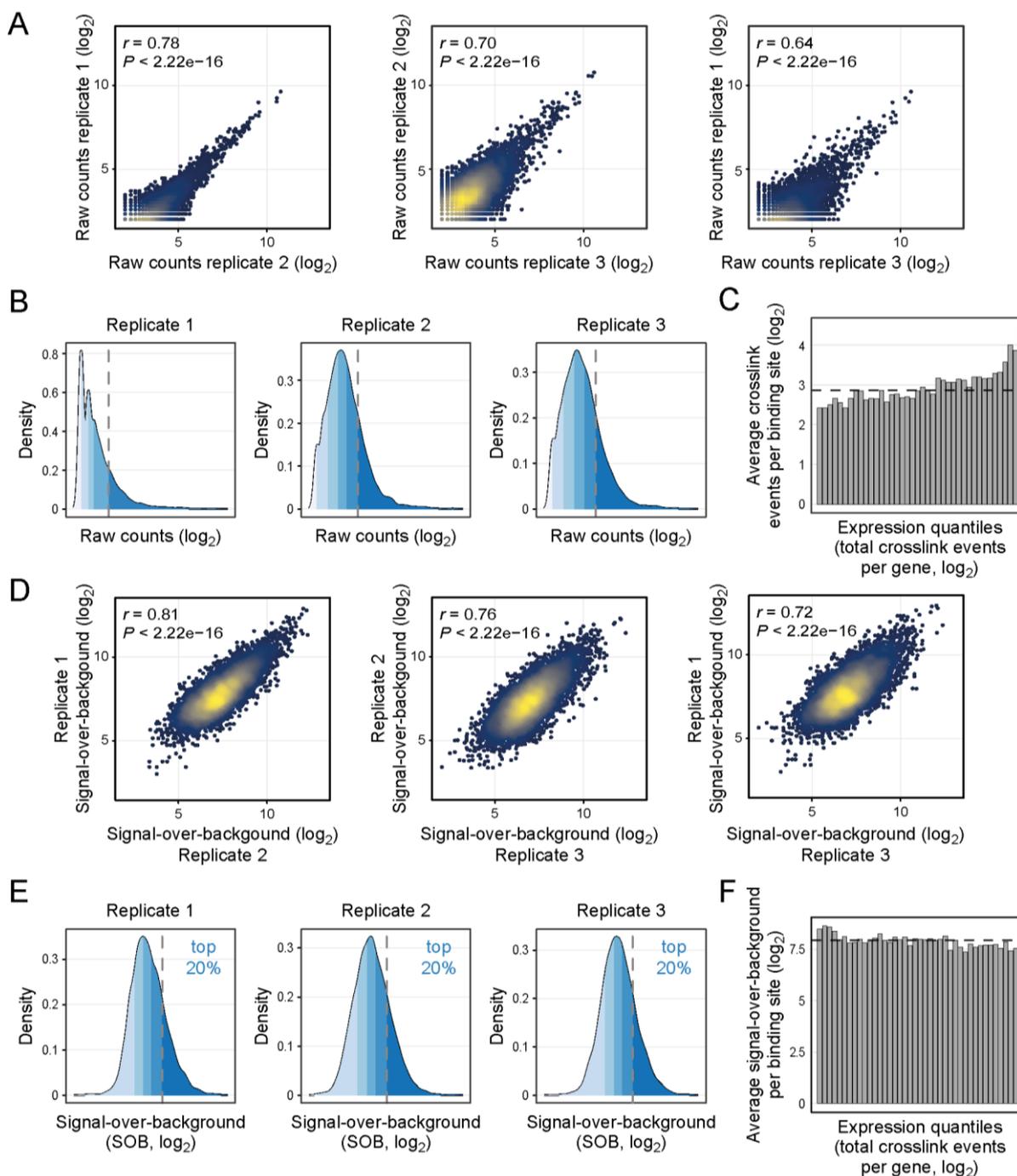
Supplementary Figure IV-1, Maximum likelihood tree of Makorin orthologs with their protein domain architecture. (A) Maximum likelihood tree with 100 bootstrap replicates of selected vertebrate and invertebrate orthologs and *C. elegans* as an outgroup. Bootstrap values at each node indicate the number of replicates (out of 100) that support the local tree structure and thereby serve as confidence estimates. Protein schematics (drawn to scale) on the right depict protein domains corresponding to the following PFAM domains: RING-type zinc finger, PF13445; MKRN1 C-terminus, PF15815; CCCH zinc finger, PF15663, PF14608 and PF00642. PAM2 motifs, predicted to interact with the MLL domain of PABP proteins (Kozlov et al. 2001) as well as the recently reported derivative PAM2L (Pohlmann et al. 2015), were added separately (see Material and methods). Abbreviated and full species names with corresponding UniProt identifiers in order of appearance: ANOGA, *Anopheles gambiae*, Q7QF83; BOVIN, *Bos taurus*, F1MF12, F6QQR5; BRAFL, *Branchiostoma floridae*, C3Y7M0; CAEEL, *Caenorhabditis elegans*, Q9N373; CANLF, *Canis lupus*, J9P921, E2RRA5, E2REH2, J9P9K3; DANRE, *Danio rerio*, Q4VBT5, Q9DFG8, A9C4A6; DROME, *Drosophila melanogaster*, Q9VP20; CHICKEN, *Gallus gallus*, Q9PTI4, F1NI93; GORGO, *Gorilla gorilla*, G3S6Y3, G3QDU4, G3RZ99; HUMAN, *Homo sapiens*, Q9UHC7, Q9H000, Q13064; IXOSC, *Ixodes scapularis*, B7QIJ9, B7Q4B2; LEPOC, *Lepisosteus oculatus*, W5NGW8, W5N9B2, W5LWJ1; MONDO, *Monodelphis domestica*, F6QPR3, F7F013; MOUSE, *Mus musculus*, Q9QXP6, Q9ERV1, Q60764; ORYLA, *Oryzias latipes*, H2MBR3, H2M1P4, H2LQG1; PANTR, *Pan troglodytes*, H2QVH8, H2QM29, H2Q915; RAT, *Rattus norvegicus*, A0A0G2QC40, Q5XI23, D3ZY41; XENTR, *Xenopus tropicalis*, Q6GLD9, B4F720. (B) The PAM2 motif in Makorin proteins from vertebrates (bottom, species abbreviations as in (A)) shows similarities to PAM2 in known PABP-interacting proteins from human (top, protein names given; first amino acid position for all PAM2 motifs indicated on the left in grey). The PAM2 consensus (Kozlov et al. 2001) is given above. Positions 9, 10 and 12 within the aligned regions that are highly consistent between all aligned proteins and important for PAM2 function (Kozlov et al. 2004) are highlighted in brown. Mutations that were introduced to abrogate the function of the PAM2 motif in human MKRN1 (MKRN1PAM2mut) are shown below. The corresponding UniProt identifiers are Q81YD1, Q8NDV7, Q99700, Q9H074, Q9BPZ3 (known PABP-interacting proteins from human), Q9UHC7, H2QVH8, G3S6Y3, J9P921, E2RRA5, F1MF12, Q5XI23, Q9QXP6, Q9PTI4, F6QPR3, W5NGW8, Q4VBT5, H2MBR3 (Makorin orthologs from vertebrates). (C) The closest Makorin orthologs in *Saccharomyces cerevisiae* lack RING domain and PAM2 motif. Domain architecture of Yth1p and Lee1p, which were detected as closest orthologs by HaMStR-OneSeq (Ebersberger et al. 20147), but were not considered as orthologs due to low FAS scores (0,59 and 0,60, respectively). The annotated PFAM domains are CCCH zinc finger, PF15663, PF00642, PF16131.

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



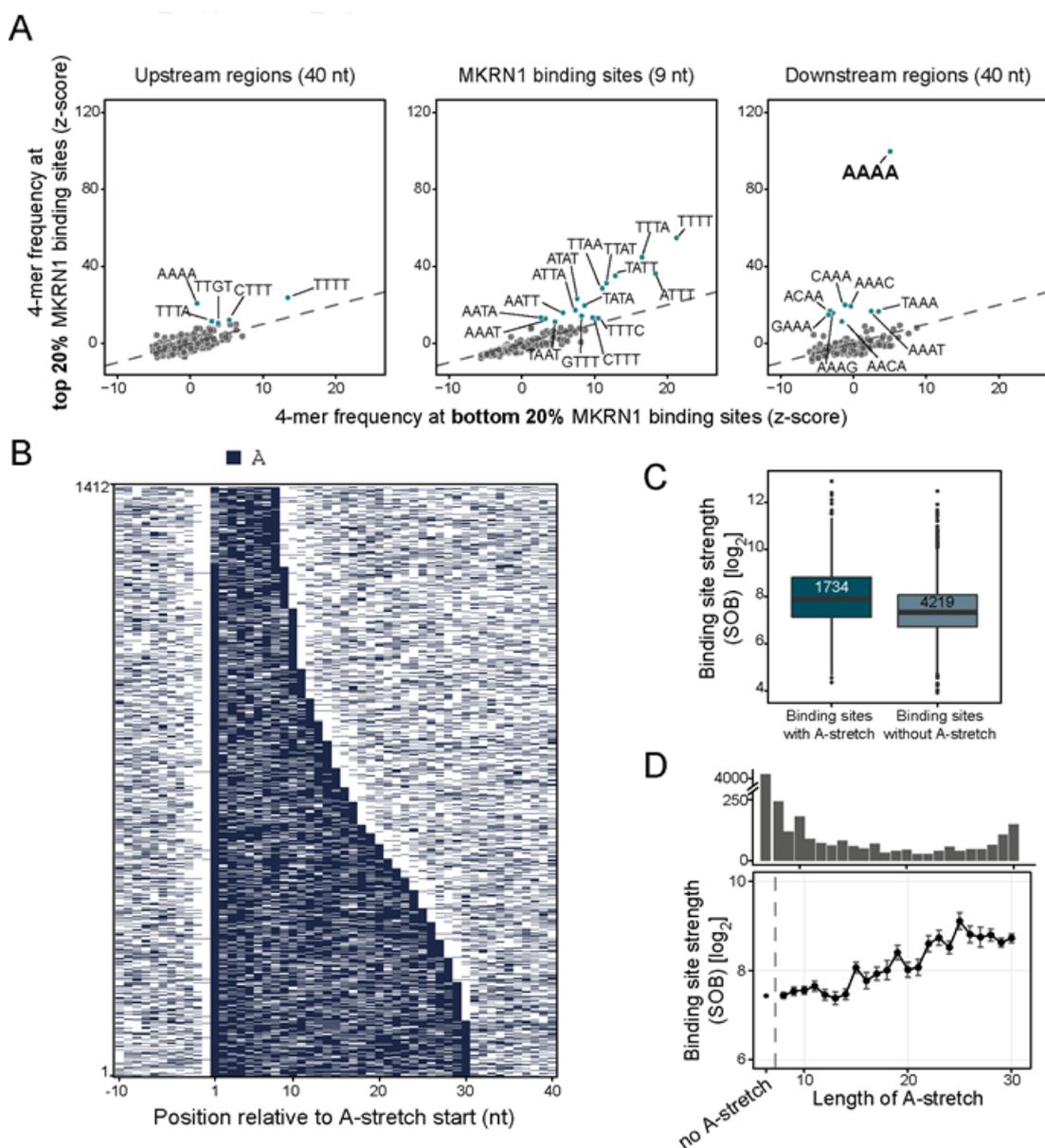
Supplementary Figure IV-2, MKRN1 interacts with translational regulators and other RBPs. (A) Overlap of the 53 significant interaction partners of GFP-MKRN1^{wt} in human HEK293T cells with previously published interactors of MKRN1 in mouse embryonic stem cells (mESC) (Cassar et al. 2015). (B) GO terms enriched for the 53 MKRN1 interactors. P values (modified Fisher exact test, Benjamini-Hochberg correction) are depicted for all significant GO terms (corrected P value < 0.05) for Biological Process, Molecular Function and Cellular Compartment, together with the number of interactors associated with the respective term. (C) Reciprocal APs show that MKRN1 interacts with PABPC1, ELAVL1 and IGF2BP1 independently of RNA. AP with GFP-PABPC1, GFP-ELAVL1 and GFP-IGF2BP1 as baits were performed from HEK293T cells in the presence or absence of RNase A and T1. Bait proteins and endogenous MKRN1 were detected by Western blots (replicate 1). Different exposure times (exp.) for MKRN1 are shown for GFP-ELAVL1 and GFP-IGF2BP1 APs. Quantifications (fold changes of the MKRN1 signal over empty vector (EV)) of three replicates are shown on the right. Replicates 2 and 3, and uncropped gel images are shown in Supplemental Fig. S9D-F.

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

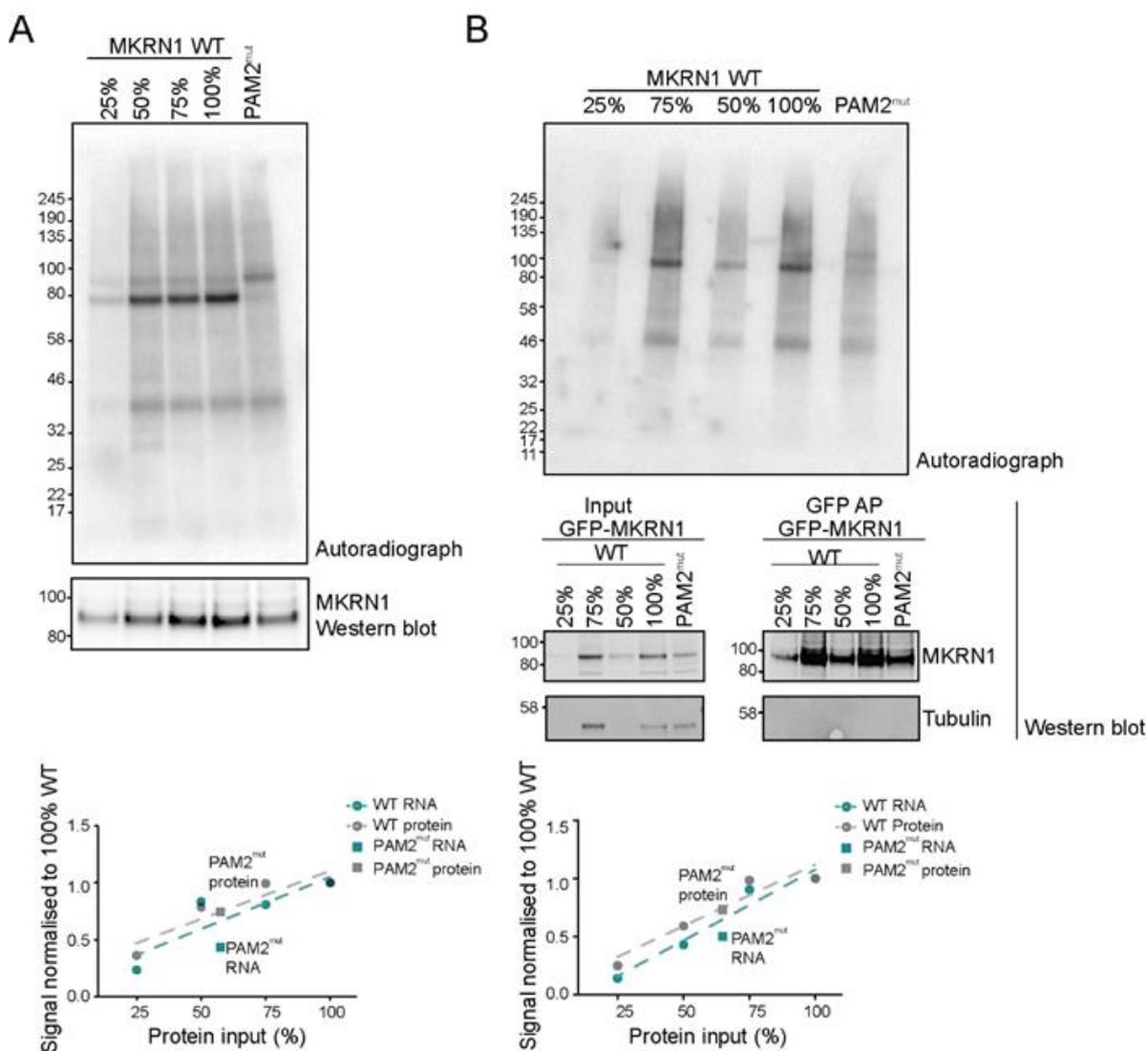


Supplementary Figure IV-3, Signal-over-background transformation allows to estimate MKRN1 binding strength. (A-C) Raw iCLIP signal before signal-over-background transformation. (A) Scatter plots show pairwise comparisons of crosslink events per binding site in three replicate MKRN1 iCLIP experiments. Pearson correlation coefficients (r) and associated P values are given. (B) Density plots depict the distribution of crosslink events per binding site in the three replicate experiments. Shades of blue indicate 20% quantiles; top 20% of binding sites with highest counts are denoted by a dashed line. (C) Raw iCLIP counts are strongly influenced by the expression level of the underlying gene. MKRN1-bound genes were stratified into 50 bins with increasing expression (using the total number of MKRN1 crosslink events within the 3' UTR as a proxy of a gene's expression level). Shown is the average number of crosslink events per binding site for all binding sites in each bin. Dashed line denotes median across all bins. (D-F) Signal-over-background (SOB) values allow to correct for expression-level differences. (D) Pairwise comparison of SOB values for the three MKRN1 iCLIP replicate experiments. Scatter plots as in (A). (E) Distribution of SOB values in the three replicates. Density plots as in (B). Shades of blue indicate 20% quantiles. Dashed lines denote the top 20% MKRN1 binding sites with strongest binding that were used for the analyses in Fig. 2B and Supplemental Fig. S4A. (F) SOB values are independent of the expression level of the underlying gene. Average SOB values for all binding sites in each expression bin are shown as in (C).

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

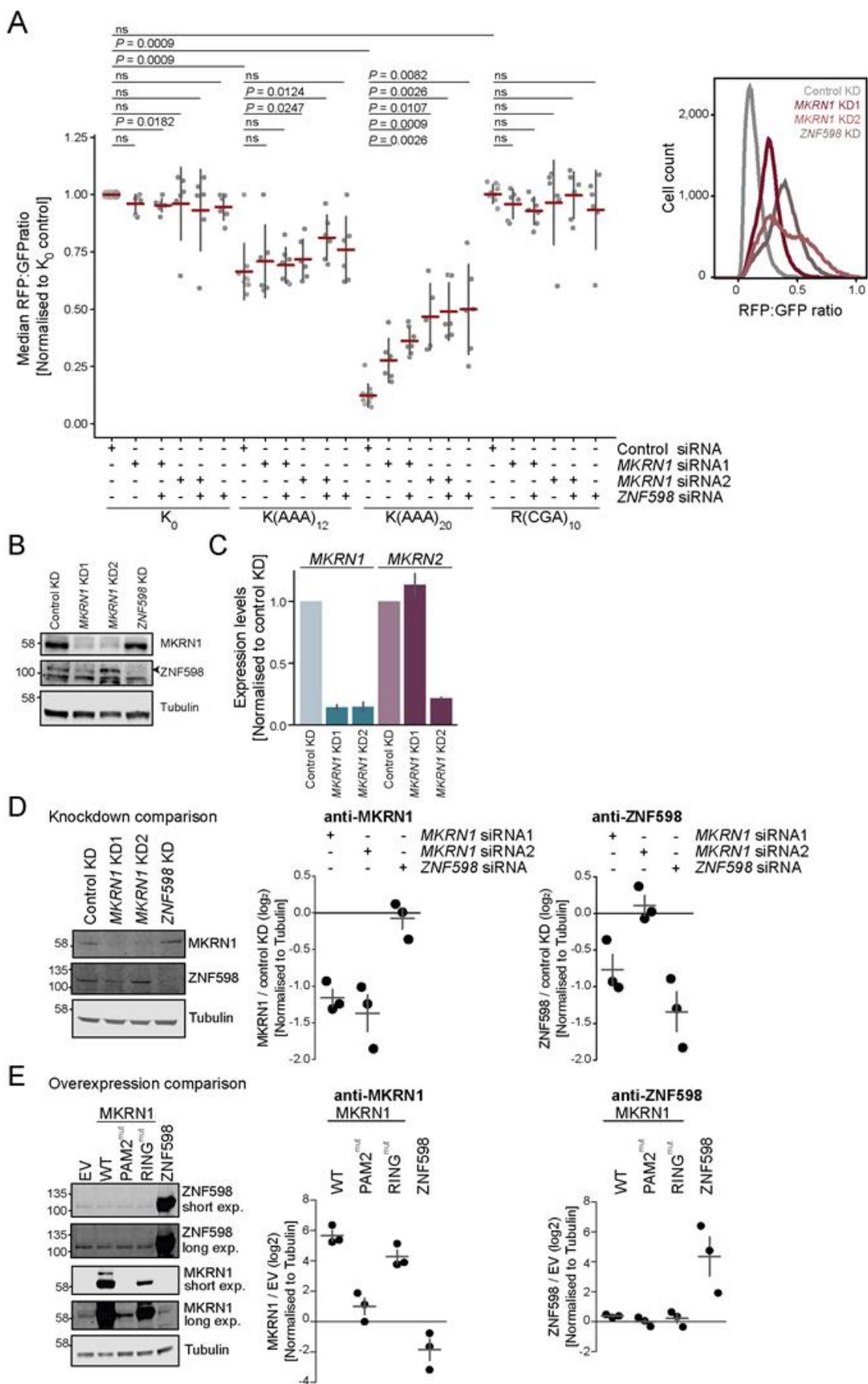


Supplementary Figure IV-4, MKRN1 binds upstream of long A-stretches. (A) Binding sites with associated A-stretches show stronger MKRN1 binding. Boxplot compares the SOB values of MKRN1 binding sites in 3' UTRs with and without associated A-stretches. Number of binding sites indicated inside box. (B) Heatmap representation of 1,412 non-overlapping A-stretches at MKRN1 binding sites, sorted by increasing length (8-30 nt). Only A's are coloured. (C) MKRN1 binding site strength (signal-over-background, SOB) increases with length of associated A-stretch. Mean and standard deviation of MKRN1 binding sites associated with A-stretches of increasing length (x-axis). MKRN1 binding sites without associated A-stretches are shown for comparison on the left. Number of binding sites in each category indicated as bar chart above. (D) The top 20% MKRN1 binding sites show a strong RNA binding preference for AAAA. Scatter plot compares the frequency of 4-mers within the 9-nt MKRN1 binding sites and flanking 40-nt windows for the top 20% and bottom 20% MKRN1 binding sites (according to SOB). 4-mer frequencies are displayed as z-scores based on background distribution from binding site permutations.



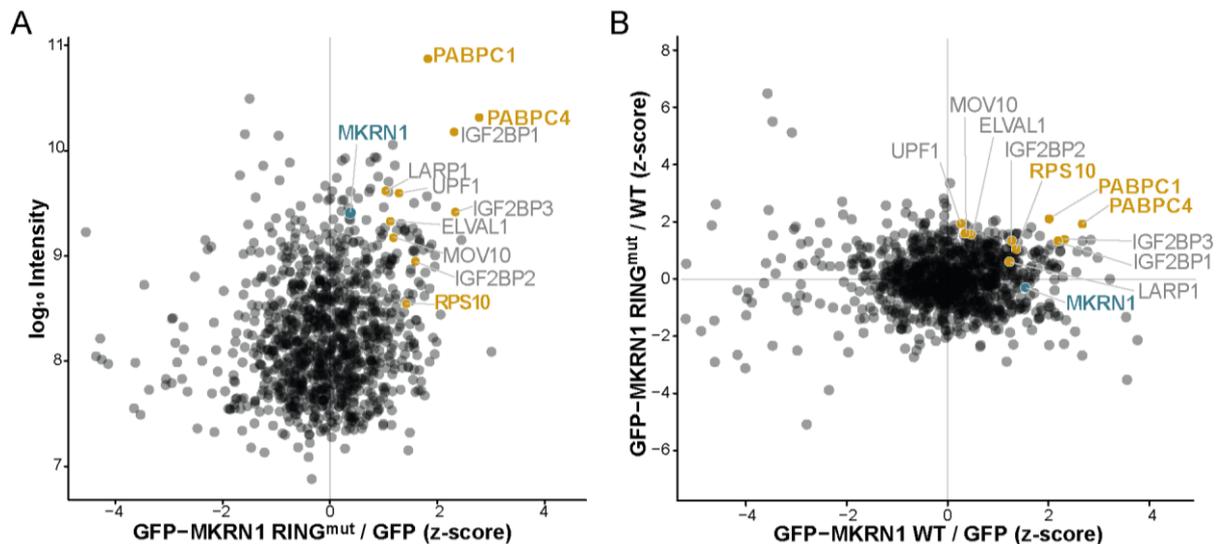
Supplementary Figure IV-5, Interaction with PABP is required for MKRN1 RNA binding. (A,B) UV crosslinking experiments to measure the RNA binding capacity of GFP-MKRN1^{wt} and GFP-MKRN1^{PAM2^{mut}}. Autoradiographs (top) and Western blots (bottom) show GFP-MKRN1/RNA complexes and GFP-MKRN1 protein, respectively, in the eluates from replicates 2 (with 4SU and UV crosslinking at 365 nm) (A) and 3 (with conventional UV crosslinking at 254 nm) (B). For calibration, input samples for GFP-MKRN1^{wt} were diluted to 75%, 50% and 25% prior to GFP AP. Note that samples were loaded in different order in (B). Quantifications are given below. Uncropped gel images are shown in Supplemental Fig. S10.

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



Supplementary Figure IV-6, MKRN1 is required to stall ribosomes at K(AAA)₂₀ in reporter assays.

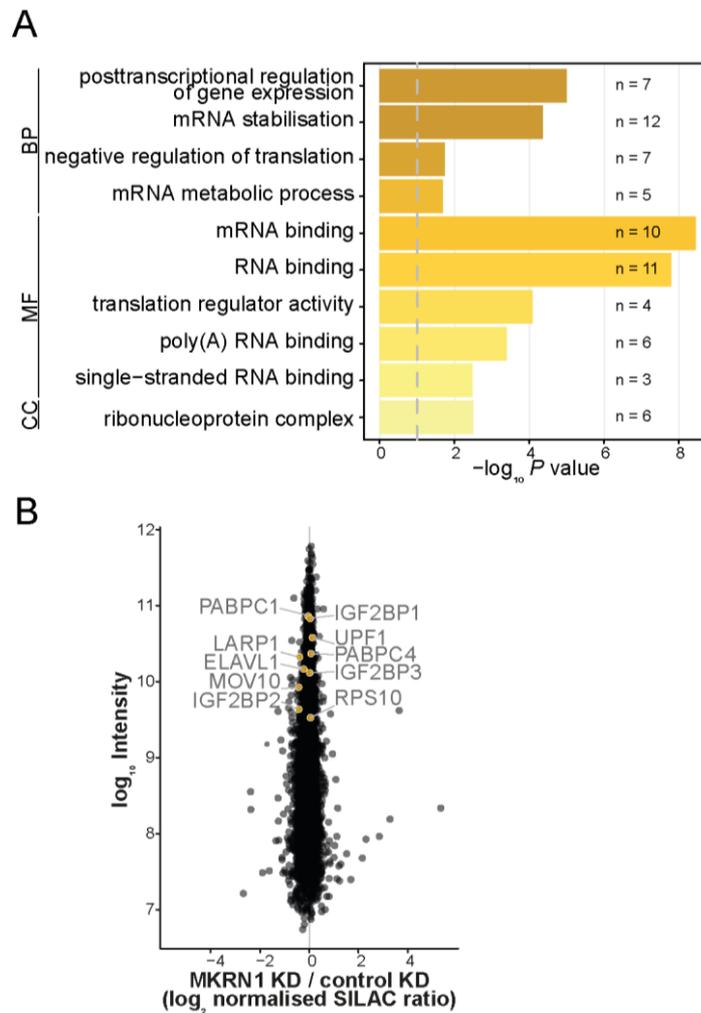
(A) Translation of dual fluorescence reporter plasmids was assessed by flow cytometry upon *MKRN1* and/or *ZNF598* KD. Median RFP:GFP ratios (normalised to K₀ in control KD) are shown for the reporter plasmids K₀, K(AAA)₁₂, K(AAA)₂₀, and R(CG A)₁₀. Error bars represent standard deviation of the mean (s.d.m., n ≥ 6 replicates; paired two-tailed t-test, Benjamini-Hochberg correction). Density plot of median RFP:GFP ratios of one replicate experiment with K(AAA)₂₀ with control or *MKRN1* KD (two independent siRNAs, KD1 and KD2) or *ZNF598* is shown on the right. (B) KDs of *MKRN1* and *ZNF598* were assessed by Western blot (n = 3 replicates). Black arrowhead indicates ZNF598. Replicates 2 and 3, and uncropped gel images are shown in Supplemental Fig. S11A,B. (C) *MKRN1* KD2 also reduces *MKRN2* levels. *MKRN1* KD1 and KD2 were performed for 72 h. Expression levels of *MKRN1* and *MKRN2* were assessed in relation to *β-actin* levels by qPCR in *MKRN1* KD (siRNA 1 and 2) and control KD. Error bars indicate s.d.m. (n = 2 replicates). (D,E) Cross-regulation of *MKRN1* and *ZNF598*. (D) *MKRN1* KD1 reduces endogenous *ZNF598* protein levels. Effect of *MKRN1* KD (KD1, siRNA 1 and KD2, siRNA 2) and *ZNF598* KD for 72 h was assessed by Western blot for endogenous *MKRN1* and *ZNF598*. Quantifications depict *MKRN1* or *ZNF598* expression levels in *MKRN1* or *ZNF598* KD over control KD condition, normalised to tubulin levels (n = 3 replicates). Replicates 2 and 3, and uncropped gel images are shown in Supplemental Fig. S11C,D. (E) *ZNF598* overexpression reduces *MKRN1* protein levels. Effect of *ZNF598* and *MKRN1* (wt and mutants) overexpression was tested after 48 h. Quantification as in (D). Uncropped gel images for all replicates are shown in Supplemental Fig. S11E,F.



Supplementary Figure IV-7, Interactome of GFP-MKRN1^{RINGmut} reveals putative ubiquitylation substrates.

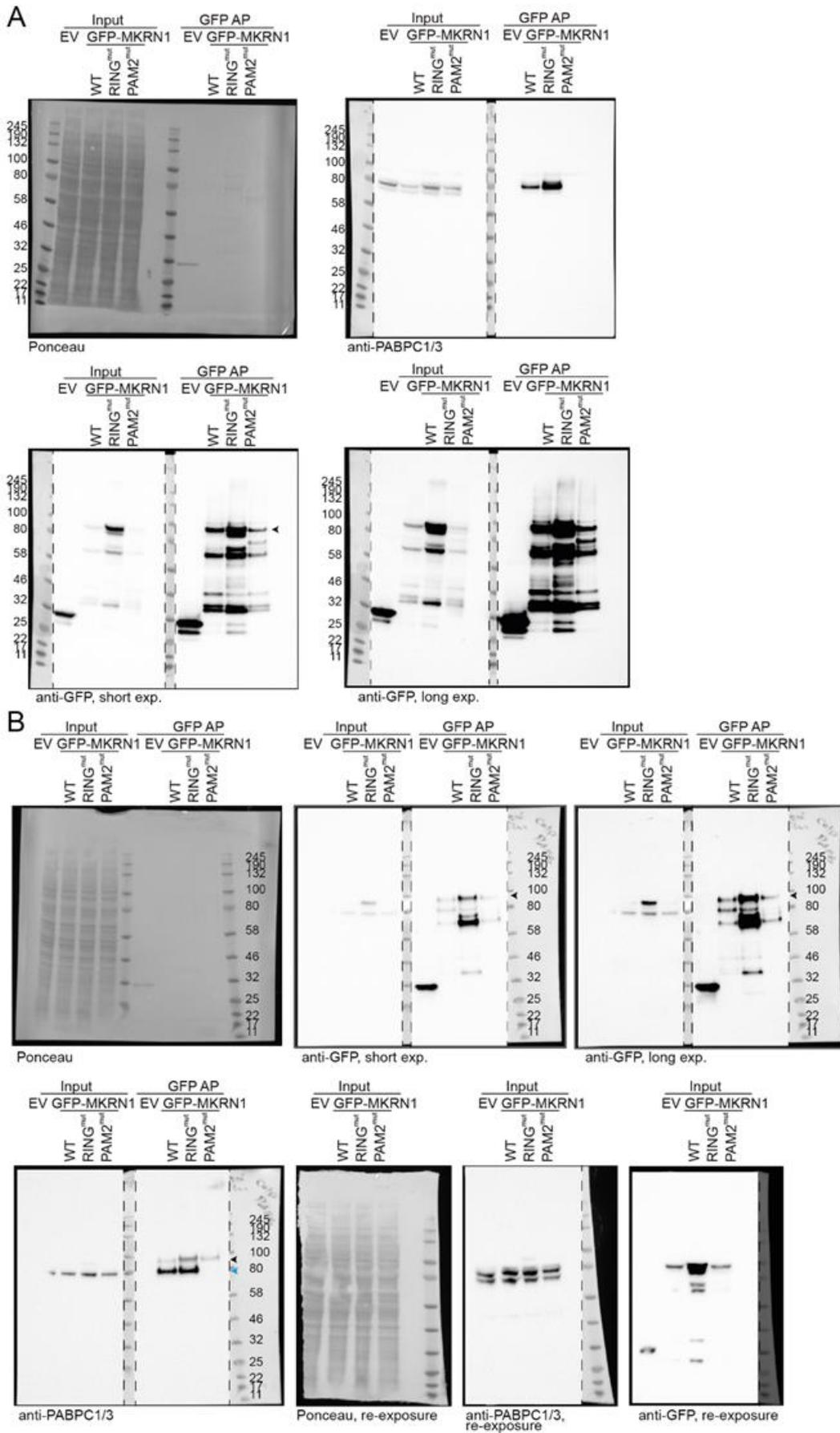
Experiments were performed using SILAC-based MS. Asymmetrical z-scores of combined SILAC ratios (n = 3 replicates) are shown. Proteins are detected in at least two out of three replicates. (A) Protein interactome of GFP-MKRN1^{RINGmut} in HEK293T cells analysed by quantitative mass spectrometry. Combined SILAC ratios (n = 3 replicates) after z-score normalisation are plotted against log₁₀-transformed intensities. 1,097 protein groups were quantified in at least two out of three replicates (Supplemental Table S1). *MKRN1* and interesting candidate ubiquitylation targets are highlighted. (B) Quantitative comparison of the interactome of GFP-MKRN1^{wt} and GFP-MKRN1^{RINGmut} shows that potential ubiquitylation candidates identified in (A) are enriched in GFP-MKRN1^{RINGmut} over GFP-MKRN1^{wt}. Comparison reveals 137 proteins to be significantly enriched (*MKRN1*^{RINGmut} over *MKRN1*^{wt} with FDR < 5% and *MKRN1*^{wt}/GFP z-score > 1). Combined ratios of three replicates are shown in a scatter plot. Only proteins detected in at least two out of three replicates are shown. Highlighting as in (A).

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

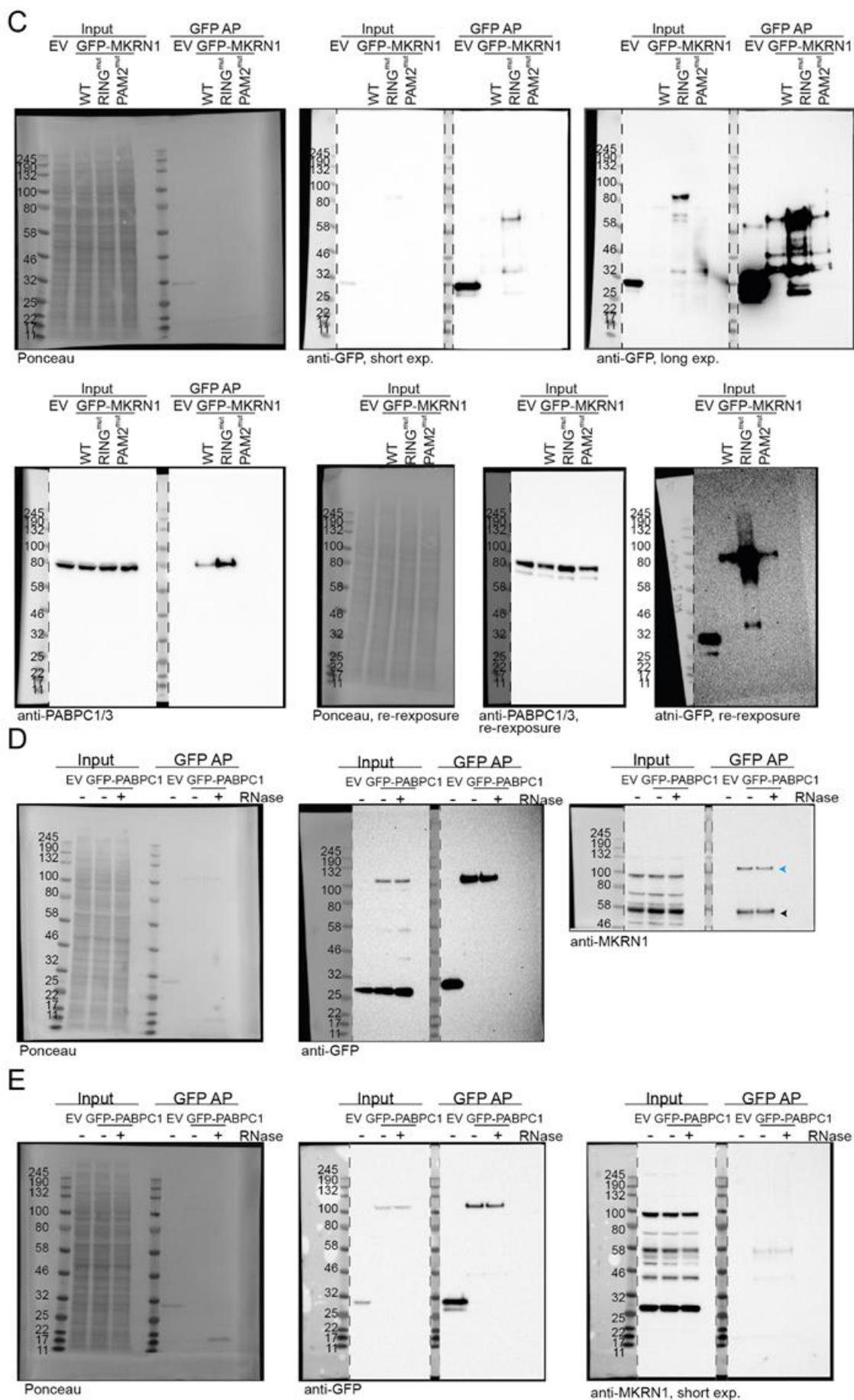


Supplementary Figure IV-8, GO term analysis of MKRN1 ubiquitylation targets and proteome analysis upon MKRN1 KD. (A) GO terms enriched for the 21 MKRN1 ubiquitylation targets. Corrected P values (modified Fisher exact test, Benjamini-Hochberg correction) are depicted for all significant GO terms (corrected P value < 0.05) for Biological Process (BP), Molecular Function (MF) and Cellular Component (CC), together with the number of ubiquitylation targets associated with the respective term. **(B)** Proteome analysis of MKRN1 KD in HEK293T cells analysed by quantitative mass spectrometry. Log₂-transformed, combined normalised SILAC ratios ($n = 3$ replicates) are plotted against log₁₀-transformed intensities. 6,425 protein groups were quantified in at least one out of three replicate experiments (Supplemental Table S4). Selected ubiquitylation targets of MKRN1 are highlighted.

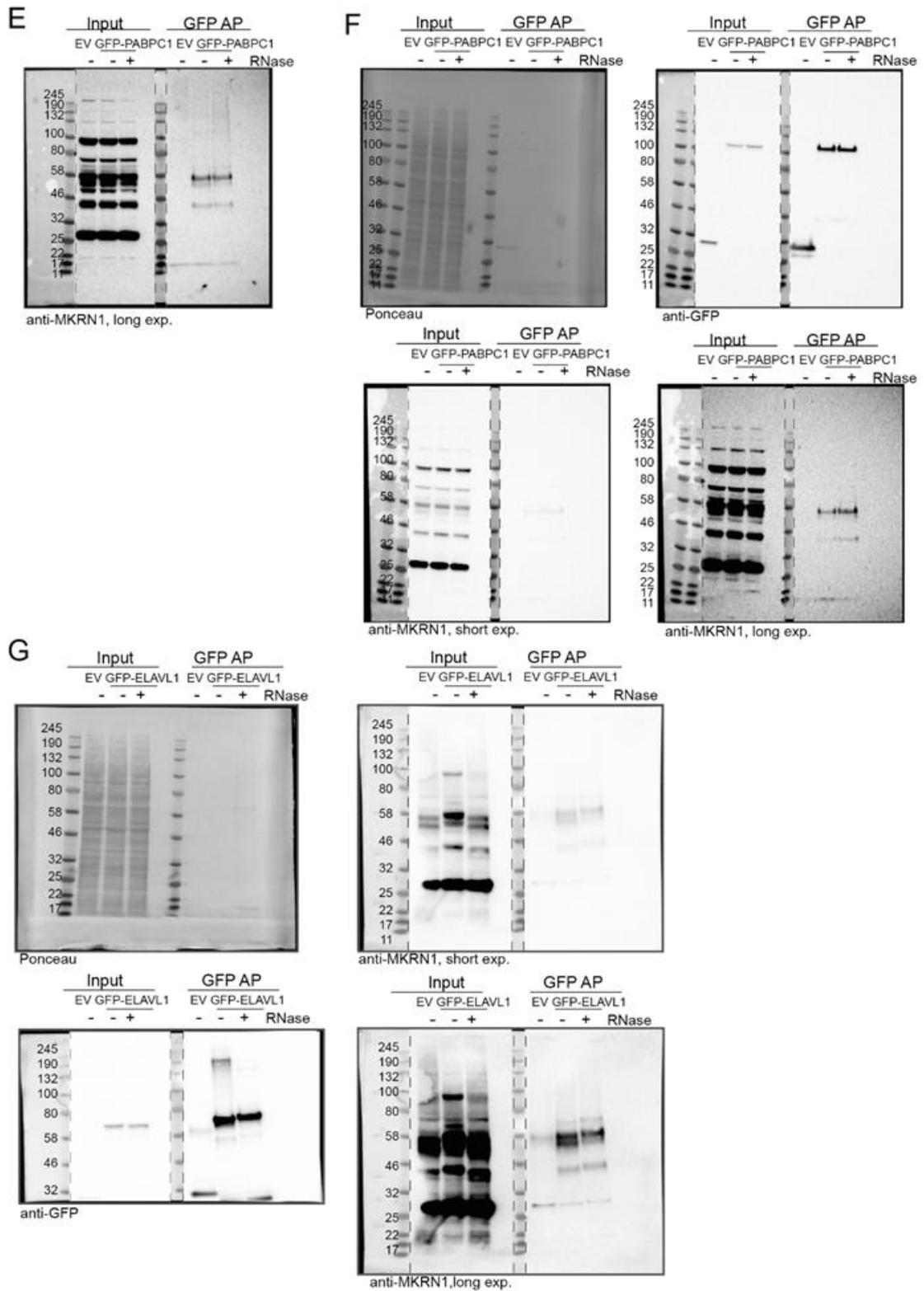
The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



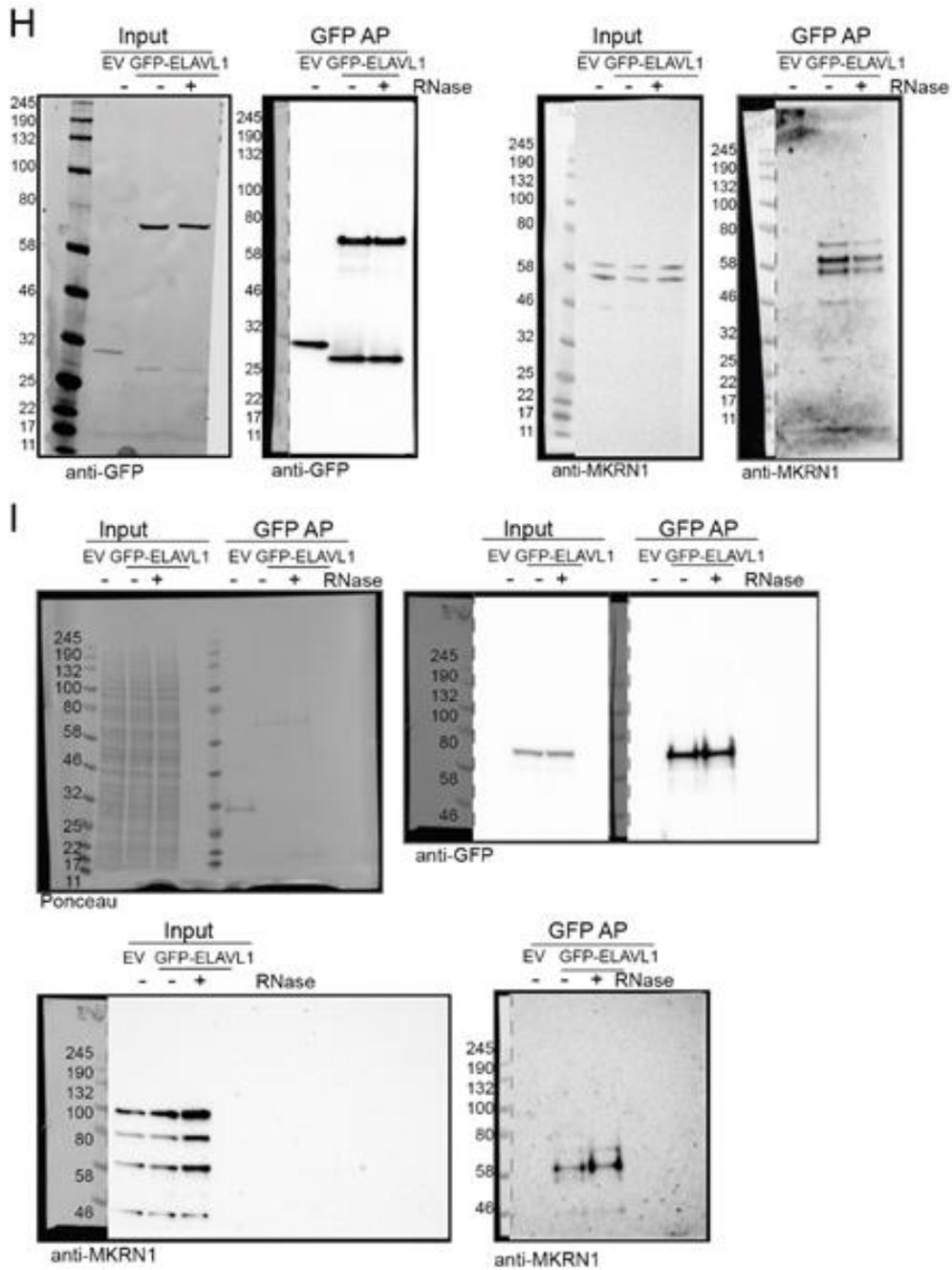
The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



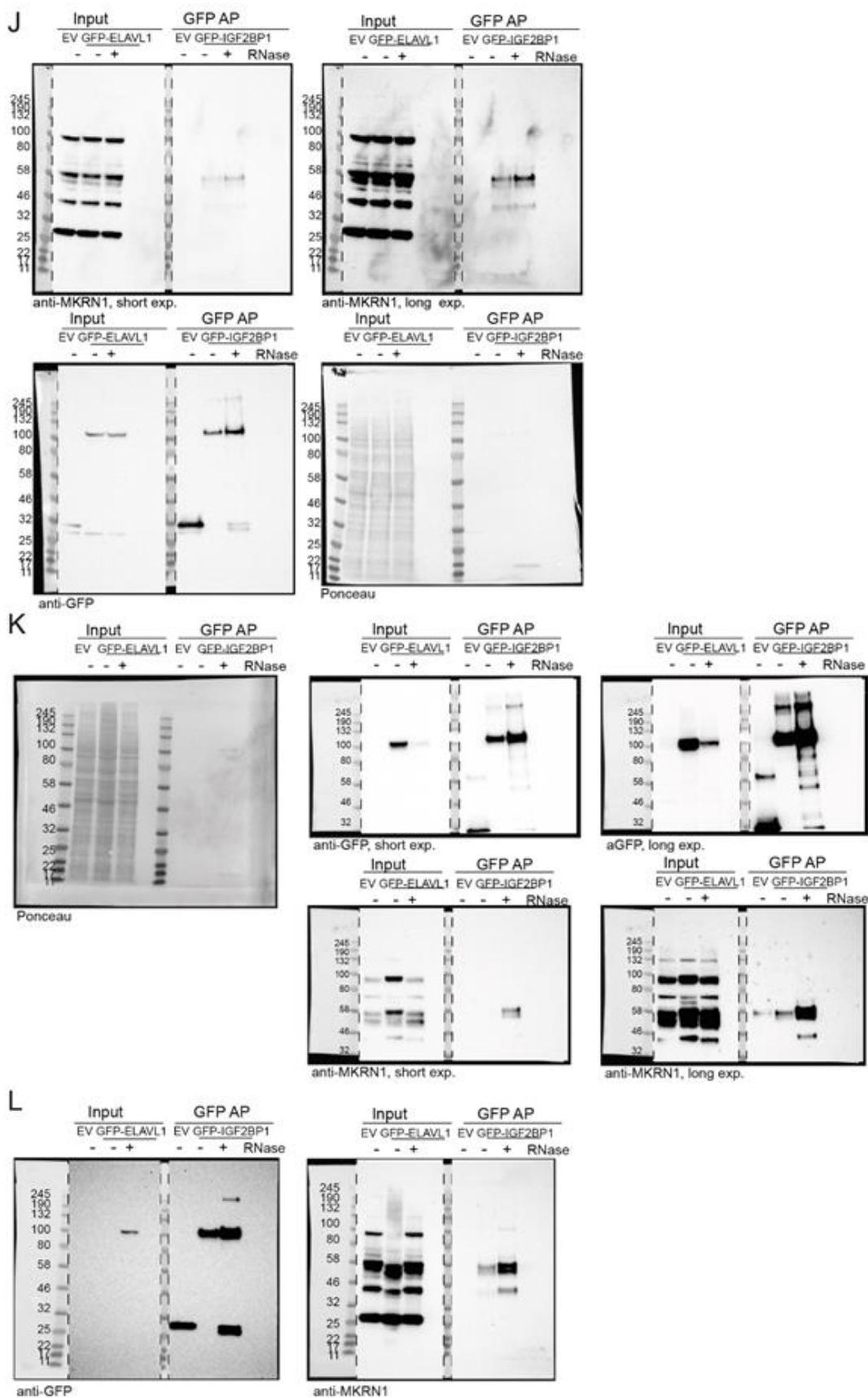
The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

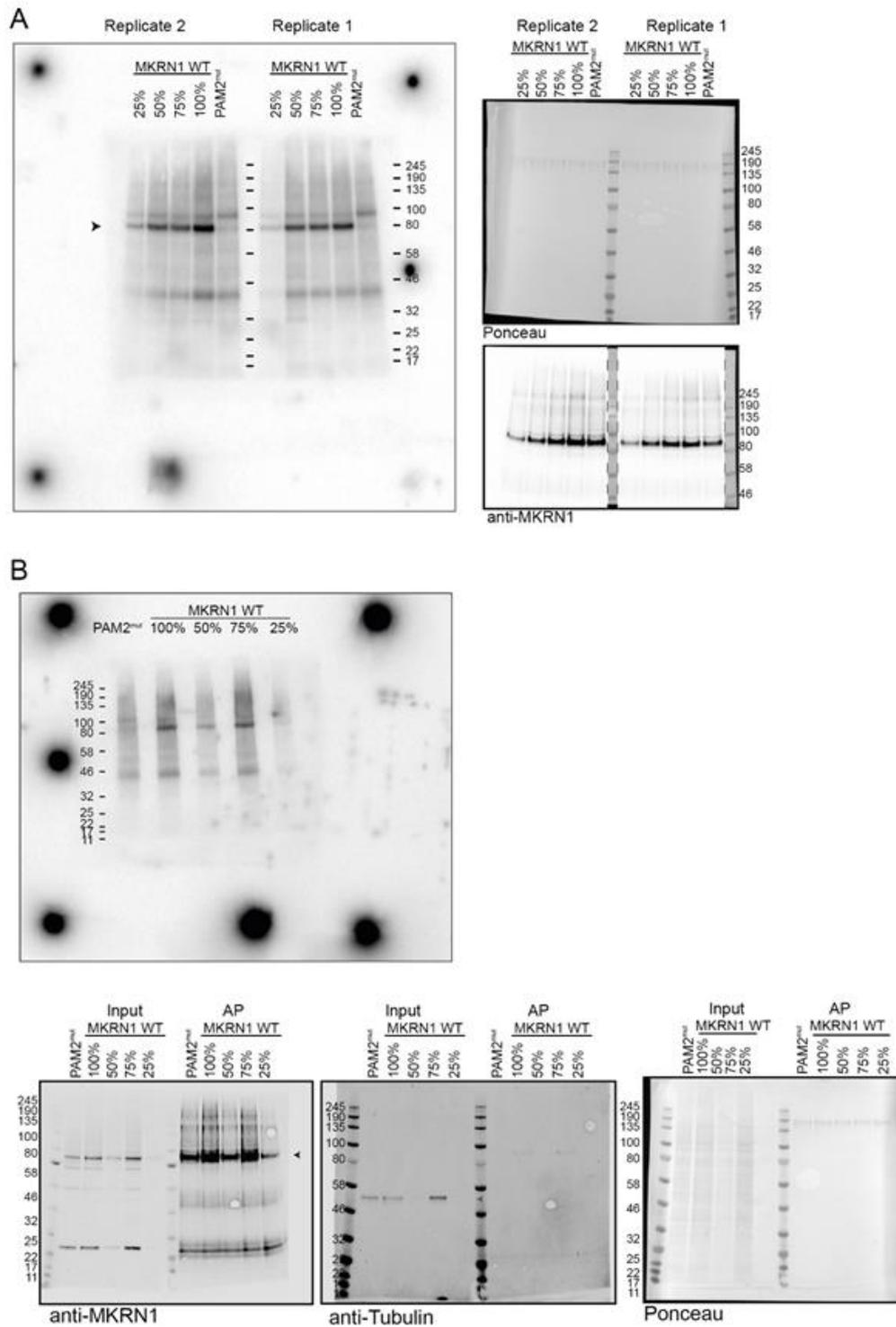


The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



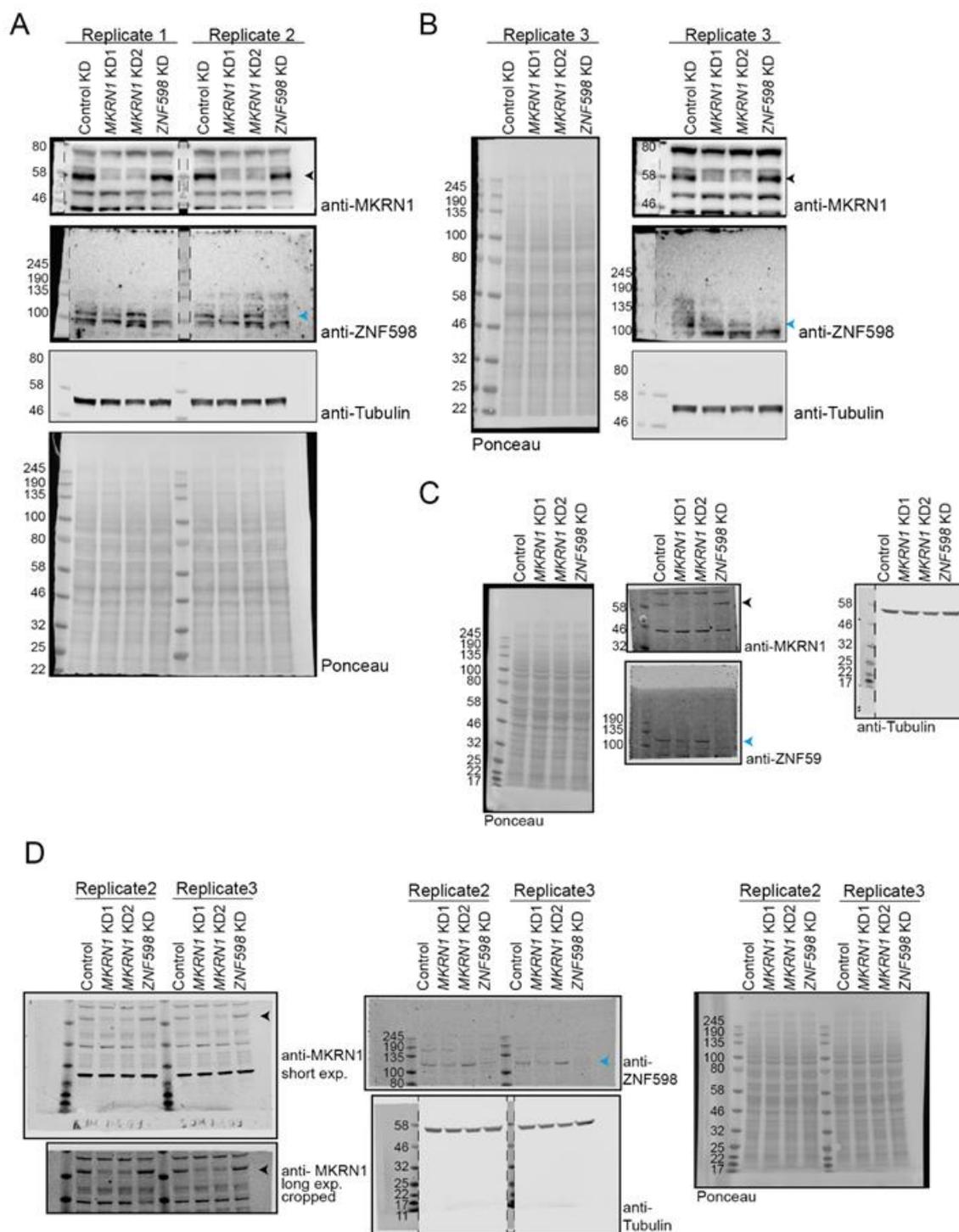
Supplementary Figure IV-9, Images of full membranes and different exposure times (exp.) for Western blot analyses in Fig. 1C and Supplemental Fig. S2C in the presence or absence of RNase A and T1. (A-C) PABP interacts with MKRN1^{wt} and MKRN1^{RINGmut} but not MKRN1^{PAM2mut}. Western blot analysis was performed with antibodies against PABPC1/3 and GFP. Images of full membranes and different exposure (exp.) times for both antibodies are shown for replicate 1 (A) which is presented in Fig. 1C, as well as replicates 2 (B) and 3 (C). Black and blue arrowheads indicate GFP-MKRN1 and PABPC1/3, respectively. **(D-F)** Endogenous MKRN1 interacts with GFP-PABPC1 independent of RNA. Western blot analysis was performed with antibodies against MKRN1 and GFP. Images of full membranes and different exposure times for both antibodies are shown for replicate 1 (D) which is presented in Supplemental Fig. S2C, as well as replicates 2 (E) and 3 (F). Black and blue arrowheads indicate MKRN1 and GFP-PABPC1, replicates. **(G-I)** Endogenous MKRN1 interacts with GFP-ELAVL1 independent of RNA. Western blot analysis was performed with antibodies against MKRN1 and GFP. Images of full membranes and different exposure times for both antibodies are shown for replicate 1 (G) which is presented in Supplemental Fig. S2C, as well as replicates 2 (H) and 3 (I). **(J-L)** Endogenous MKRN1 interacts with GFP-IGF2BP1 independent of RNA. Western blot analysis was performed with antibodies against MKRN1 and GFP. Images of full membranes and different exposure times for both antibodies are shown for replicate 1 (J) which is presented in Supplemental Fig. S2C, as well as replicates 2 (K) and 3 (L).

The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation

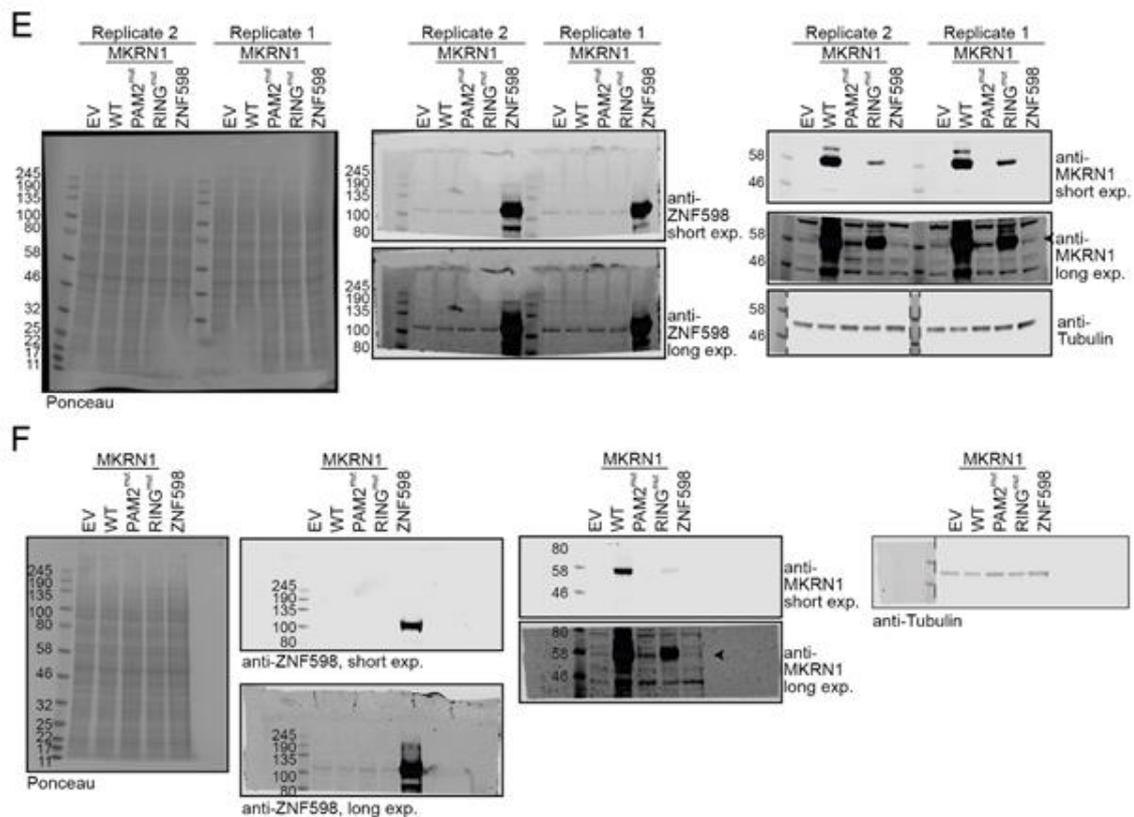


Supplementary Figure IV-10, Images of full membranes of autoradiographs and Western blot analyses in Fig. 3D (replicate 1) and Supplemental Fig. S5 (replicates 2 and 3). UV crosslinking experiments to measure the RNA binding capacity of GFP-MKRN1^{wt} and GFP-MKRN1^{PAM2^{mut}}. Autoradiographs (A, left; B, top) and Western blots (A, right; B, bottom) show GFP-MKRN1/RNA complexes and GFP-MKRN1 protein, respectively, in the eluates from replicates 1 and 2 (with 4SU and UV crosslinking at 365 nm) (A) and 3 (with conventional UV crosslinking at 254 nm) (B). (B) Images of full membranes of Western blot analyses with both antibodies are shown for replicate 3 (B).

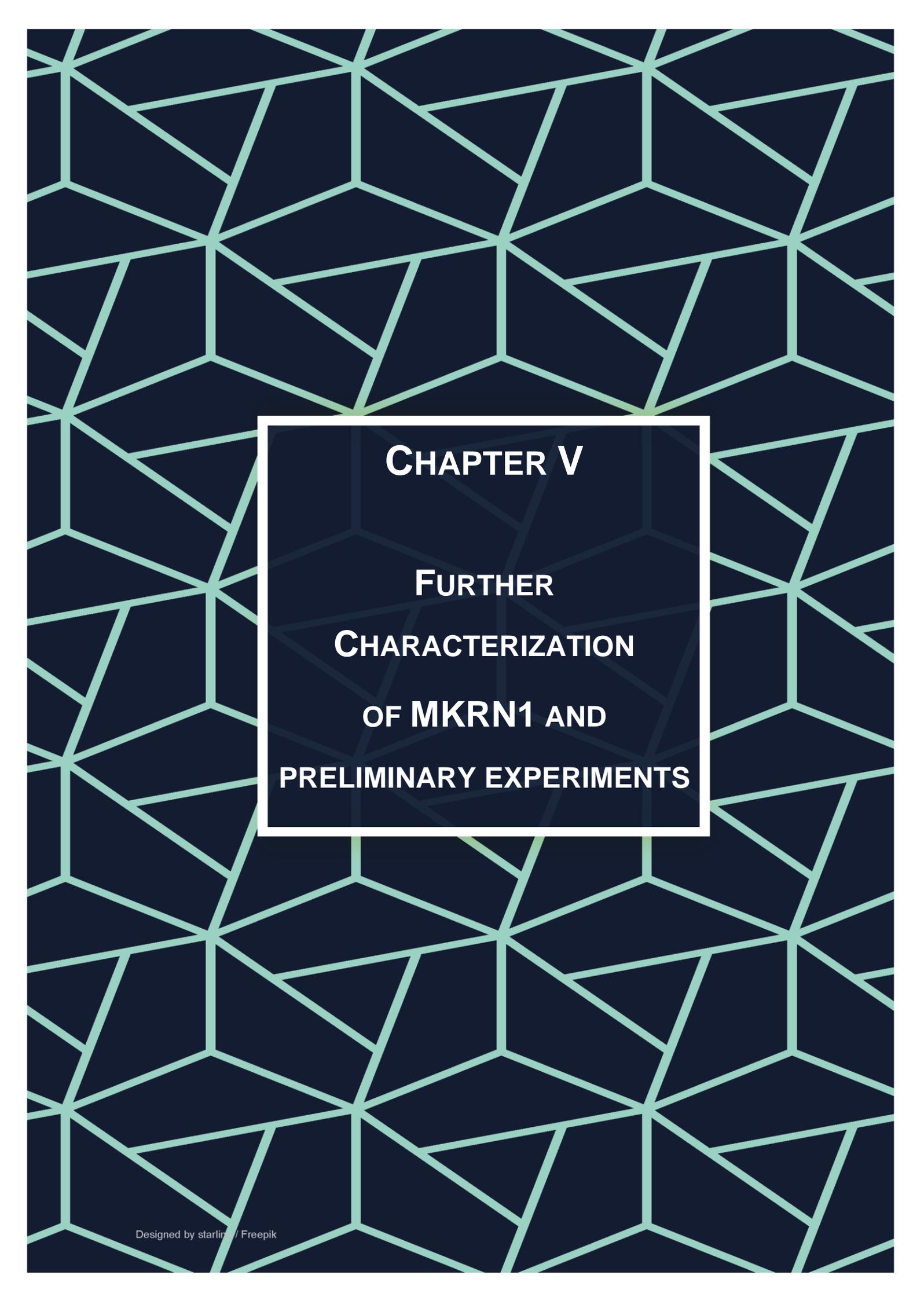
The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



The RNA-binding ubiquitin ligase MKRN1 functions in ribosome-associated quality control of poly(A)-translation



Supplementary Figure IV-11, Images of full membranes and different exposure (exp.) times for Western blot analyses in Supplemental Fig. S6B,D,E. (A,B) KDs of *MKRN1* and *ZNF598* assessed by Western blot ($n = 3$ replicates) from Supplemental Fig. S6B. Western blot analysis was performed with antibodies against *MKRN1*, *ZNF598*, and tubulin. Black and blue arrowheads indicate *MKRN1* (53 kDa) and *ZNF598* (99 kDa), respectively. Uncropped gel images of replicates 1 and 2 (A) and 3 (B). (C,D) Images of full membranes are shown for cross-regulation between *MKRN1* and *ZNF598* KD from Supplemental Fig. S6D. *MKRN1* KD1 reduces endogenous *ZNF598* protein levels. Western blot analysis was performed with antibodies against *MKRN1*, *ZNF598*, and tubulin. Coloured arrowheads as in (A). Uncropped gel images of replicate 1 (C) and replicates 2 and 3 (D). (E,F) Images of full membranes are shown for cross-regulation of *MKRN1* and *ZNF598* overexpression (OE) from Supplemental Fig. S6E. *ZNF598* OE reduces *MKRN1* protein levels. Western blot analysis was performed with antibodies against *MKRN1*, *ZNF598*, and tubulin. Black arrowheads indicate *MKRN1*. Images of full membranes and different exposure times (exp.) for both antibodies are shown for replicates 1, 2 (E), and 3 (F). Note the opposite order of replicates 1 and 2 (2 left, 1 right) in (E).



CHAPTER V

**FURTHER
CHARACTERIZATION
OF MKRN1 AND
PRELIMINARY EXPERIMENTS**

V. FURTHER CHARACTERIZATION OF MKRN1 AND PRELIMINARY EXPERIMENTS

In the previous chapter, I described that MKRN1 is involved in ribosome stalling at poly(A) stretches in the context of RQC. As MKRN1 is a highly conserved protein, I wondered whether there is a more general role for this RBUL. The protein is differentially expressed in mouse tissues and high transcript levels can be found in early embryogenesis of human and mouse, for instance¹⁻⁴. Thus, a tissue-specific or general physiological function of this protein might be only identified in the relevant biological system. Nonetheless, one may be able to identify further molecular functions of MKRN1 in cell culture. To do so, I assessed the influence of MKRN1 on mRNA stability and translation.

In this chapter, I present additional work that is not published in the two preceding manuscripts and pose a resource for future experimental work.

In this chapter, I present additional work that is not published in the two preceding manuscripts and pose a resource for future experimental work.

1. Material and Methods

1.1 Materials

Table V-1, Kits, Enzymes, Antibodies used in this study

Reagent/Antibody/Kit/Enzyme	Catalogue Number	Distributor
Chloroacetamide (CAA)	C0267	Sigma Aldrich/Merck Darmstadt, Germany
Dithiothreitol (DTT)	D5545	Sigma Aldrich/Merck Darmstadt, Germany
LDS Sample Buffer (4x)	NP0007	Thermo Fisher Scientific, MA, USA
NuPage 4-12% BisTris Gels, 1mm, 10 wells	NP0321BOX	Thermo Fisher Scientific, MA, USA
NuPage 4-12% BisTris Gels, 1mm, 12 wells	NP0322BOX	Thermo Fisher Scientific, MA, USA
NuPAGE MES SDS Running Buffer (20X)	NP0002	Thermo Fisher Scientific, MA, USA
NuPAGE MOPS SDS Running Buffer (20X)	NP0001	Thermo Fisher Scientific, MA, USA
NuPAGE Novex 3-8% Tris-Acetate Gels, 1.0 mm, 12 well	EA03752BOX	Thermo Fisher Scientific, MA, USA
NuPAGE Transfer Buffer (20X)	NP0006	Thermo Fisher Scientific, MA, USA
NuPAGE Tris-Acetate SDS Running Buffer (20X)	LA0041	Thermo Fisher Scientific, MA, USA
Pierce BCA Protein Assay Kit	23225	Thermo Fisher Scientific, MA, USA
0.45µm filter	SLHA033SB	Merck KGaA, Darmstadt, Germany
Actinomycin D	A1410-2mg	Sigma Aldrich/Merck Darmstadt, Germany
Bortezomib (CAS 179324-69-7)	sc-217785	Santa Cruz Biotechnology, Inc. Heidelberg, Germany
DMEM	21969035	Thermo Fisher Scientific, MA, USA
Doxycycline Hydrochloride	D3072	Sigma Aldrich/Merck Darmstadt, Germany
DPBS	14190094	Thermo Fisher Scientific, MA, USA
Fetal Bovine Serum (FBS)	10270106	Thermo Fisher Scientific, MA, USA
Gateway LR Clonase II Enzyme mix	11791100	Thermo Fisher Scientific, MA, USA
Lenti-X GoStix	631244	Takara Bio Europe S.A.S, Saint-Germain-en-Laye, France
L-Glutamine	25030024	Thermo Fisher Scientific, MA, USA
Lipofectamine RNAiMAX	13778150	Thermo Fisher Scientific, MA, USA
Penicillin/Streptomycin	15140122	Thermo Fisher Scientific, MA, USA
Polyethylenimine MAX 4000	24885-2	Polysciences, Inc, PA, USA
Puromycin	ANT-PR-1	InvivoGen, CA, USA
Trypsin-EDTA (0.5%)	15400050034	Thermo Fisher Scientific, MA, USA
Antarctic Phosphatase	M0289S	New England Biolabs, Frankfurt, Germany
BamHI	R3136S	New England Biolabs, Frankfurt, Germany
BbsI	R0539S	New England Biolabs, Frankfurt, Germany
BsrGI	R3575S	New England Biolabs, Frankfurt, Germany

Table V-1, Kits, Enzymes, Anitbodies used in this study

Reagent/Antibody/Kit/Enzyme	Catalogue Number	Distributer
Color Prest Protein Standard	P7712S	New England Biolabs, Frankfurt, Germany
DNA ladder 1 kb	N3232S	New England Biolabs, Frankfurt, Germany
DNA ladder 100 bp	N3231S	New England Biolabs, Frankfurt, Germany
dNTP Solution Mix	N0447S	New England Biolabs, Frankfurt, Germany
EcoRI	R3101D	New England Biolabs, Frankfurt, Germany
Human Genomic DNA	G304A	Promega GmbH
Luminaris HiGreen Low ROX qPCR MM	13505260	Thermo Fisher Scientific, MA, USA
MinElute PCR Purification Kit	28004	Qiagen GmbH, Hilden, Germany
Phusion HF DNA Polymerase	M0530S	New England Biolabs, Frankfurt, Germany
Phusion High-Fidelity PCR Kit	M0531 S	New England Biolabs
Q5 High-Fidelity DNA Polymerase	M0491S	New England Biolabs, Frankfurt, Germany
QIAGEN Plasmid Plus Midi Kit (100)	12945	Qiagen GmbH, Hilden, Germany
QIAprep Spin Miniprep Kit	27106	Qiagen GmbH, Hilden, Germany
QIAquick Gel Extraction Kit	28706	Qiagen GmbH, Hilden, Germany
QIAquick PCR Purification Kit	28106	Qiagen GmbH, Hilden, Germany
Quick Ligation Kit	M2200S	New England Biolabs, Frankfurt, Germany
RNeasy Plus Mini Kit	74136	Qiagen GmbH, Hilden, Germany
T4 DNA Ligase	M0202S	New England Biolabs, Frankfurt, Germany
Taq DNA Polymerase	M0273S	New England Biolabs
The Q5 Site-Directed Mutagenesis Kit	E0554	New England Biolabs, Frankfurt, Germany
TOPO XL PCR Cloning Kit	K4700	Life technologies
XbaI	R0145S	New England Biolabs, Frankfurt, Germany
XhoI	R0146S	New England Biolabs, Frankfurt, Germany
Ambion ERCC RNA Spike-In Mix	13499968	Thermo Fisher Scientific, MA, USA
QIAshredder	79656	Qiagen GmbH, Hilden, Germany
Ribo-Zero rRNA Removal Kit (Human/Mouse/Rat)	MRZH11124	Illumina Inc., CA, USA
TruSeq RNA Library Prep Kit v2		Illumina Inc., CA, USA
TURBO DNase	AM2238	Thermo Fisher Scientific, MA, USA
anti-GFP (B-2 clone)	sc-9996	Santa Cruz Biotechnology, Inc. Heidelberg, Germany
anti-Flag(R) M2	F1804	Sigma Aldrich/Merck Darmstadt, Germany
anti-Rabbit IgG (HRP)	7074S	Cell Signaling Technology Inc, MA, USA
anti-Mouse IgG (HRP)	7076S	Cell Signaling Technology Inc, MA, USA

Table V-1, Kits, Enzymes, Antibodies used in this study

Reagent/Antibody/Kit/Enzyme	Catalogue Number	Distributor
anti-Znf598	HPA041760-100 µl	Sigma Aldrich/Merck Darmstadt, Germany
anti-MKRN1	A300-990A	Bethyl Laboratories, Inc, TX, USA
anti-RPS10 (EPR8545)	ab151550	Abcam plc, Cambridge, UK
anti-RPS20 (EPR8716)	ab133776	Abcam plc, Cambridge, UK
anti- HSP70/HSP72 (N15F2-5)	ADI-SPA-8133-F	Enzo Life Sciences Inc, NY, USA
anti- PABP1 / PABP3	4992	Cell Signaling Technology Inc, MA, USA
anti-EEF1A1 (316-328)	E3909	Sigma Aldrich/Merck Darmstadt, Germany
Anti- Alpha-Tubulin	T-5168	Sigma Aldrich/Merck Darmstadt, Germany
Dual-Luciferase Reporter Assay System	E1910	Promega GmbH
Microfluoar - White Plate, flat bottom	10415985	Thermo Fisher Scientific
Tecan Reader Infinite 200 PRO		Tecan Austria GmbH
Typhoon FLA 9500 biomolecular imager	28-9969-43	GE Healthcare Life Sciences, PA, USA
FLA Image Eraser	28-9564-73	GE Healthcare, PA, USA
ViiA 7 Real-Time PCR System	4453536	Thermo Fisher Scientific, MA, USA
Odyssey® CLx Imaging System		Li-cor Biosciences GmbH, Bad Homburg v.d. H., Germany
LSR Fortessa Cell Analyser	649225	Becton Dickinson, NJ, USA
Tecan Austria GmbH		i-control software
FlowJo v10		FlowJo LLC, OR, USA

Table V-2, Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV

Experiment	oA	Name	Sequence	Comment
MKRN1 - stable shRNA KD cell lines	100	MKRN1.479	TGCTGTTGACAGTGAGCGATCCGTATA GTGTAGTGTGCAATAGTGAAGCCACAG ATGTATTGCACACTACACTATACGGACT GCCTACTGCCTCGGA	MKRN1 shRNA
	101	MKRN1.664	TGCTGTTGACAGTGAGCGCCAGGCGAA GCTGAGTCAAGAATAGTGAAGCCACAG ATGTATTCTTGACTCAGCTTCGCCTGTT GCCTACTGCCTCGGA	MKRN1 shRNA
	106	miRE-Xho-F	TGAACTCGAGAAGGTATATTGCTGTTGA CAGTGAGCG	Cloning
	107	miRE-EcoRI-R	TCTCGAATTCTAGCCCCTTGAAGTCCGA GGCAGTAGGC	
	108	miRseq5	TGTTTGAATGAGGCTTCAGTAC	Sanger sequencing
MKRN1 KO by CRISPR/Cas9	256	guide1-fwd	caccGCAGGTAGGACGTTGCGCGG	Cas9 guide RNAs for MKRN1, 5' P
	257	guide1.4-rev	aaacCCGCGCAACGTCCTACCTGC	
	258	guide2-fwd	caccgCCTCAATGCACGACTAGAGA	
	259	guide2.4-rev	aaacTCTCTAGTCGTGCATTGAGG	

Table V-2, Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV

Experiment	oA	Name	Sequence	Comment
	260	guide3-fwd	caccgCAATGCACGACTAGAGAAGG	
	261	guide3.4-rev	aaacCCTTCTCTAGTCGTGCATTG	
	119	U6	GACTATCATATGCTTACCGT	
	120	Cas9-R	CTTCTGGCGGTTCTCTTCAG	Sanger sequencing
	121	Cas9_rev2	cacgcgctAAAacggacta	
	266	M1_g1_fwd	ACAACATCAGGagcaggagc	Sanger sequencing
	267	M1_g1_rev	GCAGAGATCGAGACAACGC	genomic
	268	M1_g2/3_fwd	GTCAGCTTGGGAACATTGTTTT	MKRN1 locus for guide 1 or guides 2/3 (MKRN1 KO)
	269	M1_g2/3_rev	GTTGGCTTTCTCATAGACCACC	
Mutating and analyzing RPS10	246	RPS10 K138R fwd	TGGTGCCGACcggAAAGCCGAGGC	
	247	RPS10 K138R rev	GGTGGCACAGCACTCCGT	
	248	RPS10 K139R fwd	TGCCGACAAGcggGCCGAGGCTG	
	249	RPS10 K139R rev	CCAGGTGGCACAGCACTC	Mutating RPS10
	252	RPS10 K107R fwd	GCCTCGGCCTCGGGTCTGGAGG	
	253	RPS10 K107R rev	CTGCCAGTCTCTGGACGG	
	262	RPS10-siRNA-insens-fwd	gataaaaacGTGCCAACCTTCATGTC	
	263	RPS10-siRNA-insens-rev	ggcgagttcCGGGTGCTTAGGCATGTG	
Luciferase reporter assay	29	YWHAB_fwd	catTCTAGATGTTTCTCGTGCTTTGTGAT	
	30	YWHAB_rev	atgGGATCCGAGAGGTATCAAGTAACGGGTG	
	31	Znf281_fwd	catTCTAGAGGTCCCAAAGTGGCCAGG	
	32	Znf281_rev	atgGGATCCCCTGAGCTCAGGCAATCCACC	
	36	Znf281_LONG-fwd	GCCTAGAACAACGTATCAGAT	
	33	Znf281_LONGrev	TTTTAGGTGCATAGCAATATG	Cloning
	34	PNN_Fwd	catTCTAGATGGAAGAAGCCAGGCTTT	
	35	PNN_rev	atgGGATCCTACTATGATCAGTAGCGCTTAACTG	
	45	Znf281_del_fwd	[P]GTTGCTAGGATTAAGGTTATTC	
	46	Znf281_del_rev	ATTGCTAAACATTGTCACTTTG[P]	
	208	A_insert luc fwd	aaaaaaaaTTCTAGAGTCGGGGCGGC	
	209	A_insert luc rev	tttttttttgACACGGCGATCTTTCCGC	
210	A_stop_luc fwd	aaaaaaaaTAATTCTAGAGTCGGGGCGG		

Table V-2, Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV

Experiment	oA	Name	Sequence	Comment
	211	A_stop_luc rev	ttttttttCACGGCGATCTTTCCGCC	
	215	Luc_seq_utr-F	ATCTTCGACGCAGGTGTCTG	
	37	pGL3p_Seq_fw d	GATTACGTCGCCAGTCAAG	
	38	pGL3p_Seq_rev	CTGTCCTACGAGTTGCATG	
	39	YWHAB-5' -seq	CTAGGCTGAGGCTGTGAAA	
	40	YWHAB-3' - seq	GATGTGTCACCACCATCT	Sanger sequencing
	41	LucN-rev – seq	CCTTATGCAGTTGCTCTCC	
	42	SV40pro_F – seq	TATTTATGCAGAGGCCGAGG	
	43	EBV-rev - seq	GTGGTTTGTCCAAACTCATC	
	44	YWHAB-mid - seq	CTCCCTGTATTGAGGCTAG	
	15	Actin fwd	TCCTCCCTGGAGAAGAGCTAC	
	16	Actin rev	TGGAGTTGAAGGTAGTTCGTG	
	17	MKRN1 P3 fwd	CGATACGGGGAGAAGACTGTGT	
	18	MKRN1 P3 rev	CCTTCTCATGGGCCTCAAT	
	63	CUL7 fwd	TTTGAGGAAGACATTGAGGACAT	
	64	CUL7 rev	AAGGCACAGCATAGAGTTCTGTC	qPCR Actinomycin D treatment
	82	CUL7_rev_intron	acaaggatgtcactagcaactc	
	67	LOX fwd	AACCAAGGGACATCAGATTTCTT	
	68	LOX rev	ATCAAGCAGGTCATAGTGGCTAA	
	90	LOX_rev_intron	aaaatcacctgagaaatgaaaagc	
	73	UNC5B fwd	ATTGCCACGACCACGAAG	
	74	UNC5B rev	CTGCTCGACTCTAAGAACTGCAC	
	83	UNC5B_fwd_intron	CAAGAACATCGTGGCCAAAC	
mRNA stability	84	UNC5B_rev_intron	catcctccctaccaatccag	
	71	ITGA8 fwd	ACGTGGAAAAGCTCACAGTGTA	
	72	ITGA8 rev	CTGGCTGGTGTGGCTTT	
	85	ITGA8_fwd_intron	AGCCGTCTATTACTGTCCTTGG	
	86	ITGA8_rev_intron	cacagcttagaaggacaaatctca	
	59	TNC fwd	GCCTTTGCTGTCTATGACAAGTT	qPCR, pre- mRNA
	60	TNC rev	AATCTGTGTCCTTGTCAAAGGTG	
	87	TNC_rev_intron	ctgtaaagtcaagaaggctcca	
	75	HTRA3 fwd	AGCTACAGAATGGGGACTCCTAT	
	76	HTRA3 rev	AGCAACAACACAGGGAGCTT	
	88	HTRA3_rev_intron	tctgggaagaggccaccat	
	61	ID3 fwd	GTCATCGACTACATTCTCGACCT	
	62	ID3 rev	TTTGTCGTTGGAGATGACAAGTT	
	89	ID3_rev_intron	gtcccgactcaggctta	

Table V-2, Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV

Experiment	oA	Name	Sequence	Comment
	220	MKRN1-sR17-H307E-fwd	taattgtaatgaaCACCTACTGTCTCAAGTGC ATTCGC	
	221	MKRN1-sR17-H307E-rev	gacagaatgccaaaAAGCGGCGCTCACTGG GG	
	222	MKRN1-PAMmut fwd	gccgttgccGGGCAACCCTACTGTGGC	
	223	MKRN1-PAMmut rev	ctcaatagaATTCACCCAGTCCTCTGAACC	
	158	MKRN1-Znf1-mut-F	[Phos]tgccaaggaaggagacaacgccCGCTACT CGCATGACCTC	
	159	MKRN1-Znf1-mut-R	[Phos]accccatgcataaaatacctggcGGTGACC TGTTTAGTCCAG	
	160	MKRN1-Znf2-mut-F	[Phos]cgccattatggagaccgcgccAGATATGA ACATAGCAAACC	
	161	MKRN1-Znf2-mut-R	[Phos]taccctcgctgaaaatactggcCACTACAC TATACGGACTG	
	162	MKRN1-Znf3-mut-F	[Phos]ggcccatacggggagaacgccGTGTATC TCCACGGAGATTC	
	163	MKRN1-Znf3-mut-R	[Phos]tctcccactgcagcatagggggcCAGCTGC TTCTTTGTCTC	
	164	MKRN1-Znf4-mut-F	[Phos]gagcgccccattggagggaacgccTTTTAC AAGCATGCGTACC	
	165	MKRN1-Znf4-mut-R	[Phos]ccacgtccttcatcaaaaatacctggcCGCCTT GTTGCTCATTGC	
Cloning, Mutating, and Sequencing MKRN1	53	MKRN1-sR16-fwd	[P][P]gaatccaggaacTCAAACCTTTGCAACT GTAG	Mutating MKRN1
	54	MKRN1-sR16-rev	[P][P]ggcctctccggtATTCATTTCAACAAGTG GTC	
	55	MKRN1-sR17-fwd	[P][P]tctaattgtaatCACACCTACTGTCTCAA GTGCATTTCG	
	56	MKRN1-sR17-rev	[P][P]cagaatgccaaaGCGGCGCTCACTGG GGTT	
	220	MKRN1-sR17-H307E-fwd	taattgtaatgaaCACCTACTGTCTCAAGTGC ATTCGC	
	221	MKRN1-sR17-H307E-rev	gacagaatgccaaaAAGCGGCGCTCACTGG GG	
	222	MKRN1-PAMmut fwd	gccgttgccGGGCAACCCTACTGTGGC	
	223	MKRN1-PAMmut rev	ctcaatagaATTCACCCAGTCCTCTGAACC	
	7	MKRN1_PAM2_del_fwd	[Phos]CCCTACTGTGGCCGCTACT	
	8	MKRN1_PAM2_del_rev	[Phos]TCCTACAGTTGCAAAGTTTGAATT TC	
9	MKRN1_H307E_fwd	[Phos]CAACTGCAACGAAACCTACTGTCT CAAG		
10	MKRN1_H307E_rev	[Phos]GAGAGGATCCCGAAGCGG		
254	pMX-del-GFP fwd	GAGCTCTACAAAAGCGGTTC		
255	pMX-del-GFP rev	GCTGGCCATGGTGTCTAG		

Table V-2, Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV

Experiment	oA	Name	Sequence	Comment	
	80	MKRN1-HindIII_fwd	TACTttAAGCTTACCATGGCGGAGG	Cloning	
	81	MKRN1-BsrGI_rev	tacttgtagaGGTTCTATAGATCCAAGTCAT		
	225	MKRN1-fwd_2	TGCACTGAAGCACCCCTGCAG	Sager sequencing	
	216	MKRN1-fwd_1	ACAACAGCCACAACATCAGG		
MKRN1 ubiquitylation targets	178	HSPA1A-F	CTGGAGTCCTACGCCTTCAACA		qPCR RPS10
	179	HSPA1A-R	TGTCCAGCACCTTCTTCTGTCC		
	180	IGF2BP1-F	GAAGTGAAGCTGGAGACCCACATAC		
	181	IGF2BP1-R	CGTTTCATCAGGGGTCTGGTC		
	182	EEF1A1-F	GCAAAAATGACCCACCAATGGAAGC		
	183	EEF1A1-R	AGCCGTGTGGCAATCCAATACA		
	264	RPS10-qPCR-fwd	tccgtgattacctcatctgcc		
	265	RPS10-qPCR-rev	tatctctgtcagcttcccctct		

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
Stable cell lines - MKRN1 shRNA KD	pL71	miRE_11_SGE P_Ren713	N' term GFP	SGEP	Ampicillin, Puromycin	Addgene, #111170
	pL72	miRE 18_LT3GEPIR_Ren713	N' term GFP	miRE 18_LT3GEPI R_Ren713	Ampicillin, Puromycin	Addgene, #111177
	pL73	psPAX2	none	psPAX2	Ampicillin	Addgene, #12260
	pL74	pMD2.G	none	pMD2.G	Ampicillin	Addgene, #12259
	pL77	miRE_18_LT3G EPIR_MKRN1.4 79	N' term GFP	miRE 18_LT3GEPI R_Ren713	Ampicillin, Puromycin	
	pL78	miRE_18_LT3G EPIR_MKRN1.6 64	N' term GFP	miRE 18_LT3GEPI R_Ren713	Ampicillin, Puromycin	
	pL79	miRE_11_SGE P_MKRN1.664	N' term GFP	SGEP	Ampicillin, Puromycin	
	pL80	miRE_11_SGE P_MKRN1.479	N' term GFP	SGEP	Ampicillin, Puromycin	
Stable cell lines - MKRN1 mutants	pL27	pMX-Dest53-notag	none	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL52	pMX-Dest53-MKRN1-sR17insens	none	pMX-DEST53-IP	Ampicillin, Puromycin	
	pL53	pMX-Dest53-MKRN1-sR17insens-H307E	none	pMX-DEST53-IP	Ampicillin, Puromycin	

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
	pL76	pMX-Dest53-MKRN1-sR17insens-PAM2mut	none	pMX-DEST53-IP	Ampicillin, Puromycin	
	pL90	pMX-Dest53-MKRN1-sR17insens-PAM2mut-H307E	none	pMX-DEST53-IP	Ampicillin, Puromycin	
CRISPR Cas9	pL75	pSpCas9(BB)-2A-Puro (PX459) V2.0	Flag	pSpCas9(BB)-2A-Puro (PX459) V2.0	Ampicillin, Puromycin	Roukos Lab
	pL81	pSpCas9(BB)-2A-Puro (PX459) V2.0_MKRN1_guide1	Flag	pSpCas9(BB)-2A-Puro (PX459) V2.0	Ampicillin, Puromycin	
	pL82	pSpCas9(BB)-2A-Puro (PX459) V2.0_MKRN1_guide2	Flag	pSpCas9(BB)-2A-Puro (PX459) V2.0	Ampicillin, Puromycin	
	pL83	pSpCas9(BB)-2A-Puro (PX459) V2.0_MKRN1_guide3	Flag	pSpCas9(BB)-2A-Puro (PX459) V2.0	Ampicillin, Puromycin	
Flow cytometry reporter assay	pL131	pM-eGFP-K0-RFP	none	pmGFP-N1 (Clontech)	Kanamycine	Hegde Lab
	pL132	pM-eGFP-R10-RFP	none	pmGFP-N1 (Clontech)	Kanamycine	Hegde Lab
	pL133	pM-eGFP-K20-RFP	none	pmGFP-N1 (Clontech)	Kanamycine	Hegde Lab
	pL134	pM-eGFP-K12-RFP	none	pmGFP-N1 (Clontech)	Kanamycine	Hegde Lab
MKRN1	pL5	pcDNA-DEST53-IP-GFP MKRN1	N' term GFP	pcDNA-DEST53-GFP	Ampicillin, Puromycin	
	pL13	pMX-DEST53-IP-GFP MKRN1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL23	pMX-DEST53-GFP-MKRN1_H307E	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL25	pcDNA-DEST53-GFP-MKRN1- Δ PAM2	N' term GFP	pcDNA-DEST53-GFP	Ampicillin, Puromycin	
	pL26	pcDNA-DEST53-GFP-MKRN1-H307E	N' term GFP	pcDNA-DEST53-GFP	Ampicillin, Puromycin	
	pL29	pMX-DEST53-GFP-MKRN1- Δ PAM2	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL37	pMX-DEST53-IP-GFP vector	N' term GFP		Ampicillin, Chloramphenicol	

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
	pL38	pMX-DEST53-IP-GFP MKRN1 H307E-sR17insens	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL52	pMX-Dest53-MKRN1-sR17insens	none	pMX-DEST53-IP	Ampicillin, Puromycin	
	pL53	pMX-Dest53-MKRN1-sR17insens-H307E	none	pMX-DEST53-IP	Ampicillin, Puromycin	
	pL55	pMX-DEST53-IP-GFP MKRN1 PAM2mut	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL59	pMX-DEST53-GFP-MKRN1-sR17insens	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL63	pMX-DEST53-GFP-MKRN1_sR16insens	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL64	pcDNA5_MKRN1-sR17insens-PAM2mut	N' term GFP	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL65	pcDNA5_MKRN1-sR17insens-H307E	N' term GFP	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL66	pcDNA5nt-MKRN1-WT	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL67	pcDNA5nt-MKRN1-PAM2mut	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL68	pcDNA5_not tag	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL69	pcDNA5nt_MKRN1_sR17insens	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL69	pcDNA5nt_MKRN1_sR17insens	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL70	pcDNA5nt_MKRN1_sR16insens	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL87	pcDNA5-MKRN1-sR16insens-ZNF1+2	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL88	pcDNA5-MKRN1-sR16insens-ZNF1+3	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	
	pL89	pcDNA5-MKRN1-sR16insens-ZNF1+4	none	pcDNA5 (Reymond pL18)	Ampicillin, Hygromycin	

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
	pL95	pcDNA5-MKRN1-WT-ZNF1-mut	none	none (pL66)	Ampicillin, Hygromycin	
	pL96	pcDNA5-MKRN1-WT-ZNF2-mut	none	none (pL66)	Ampicillin, Hygromycin	
	pL97	pcDNA5-MKRN1-WT-ZNF3-mut	none	none (pL66)	Ampicillin, Hygromycin	
	pL98	pcDNA5-MKRN1-WT-ZNF4-mut	none	none (pL66)	Ampicillin, Hygromycin	
	pL99	pcDNA5-MKRN1-sR16insens-ZNF1-mut	none	none (pL70)	Ampicillin, Hygromycin	
	pL100	pcDNA5-MKRN1-sR16insens-ZNF2-mut	none	none (pL70)	Ampicillin, Hygromycin	
	pL101	pcDNA5-MKRN1-sR16insens-ZNF3-mut	none	none (pL70)	Ampicillin, Hygromycin	
	pL102	pcDNA5-MKRN1-sR16insens-ZNF4-mut	none	none (pL70)	Ampicillin, Hygromycin	
	pL107	pcDNA5-MKRN1-H307E	none	none (pL66)	Ampicillin, Hygromycin	
	pL108	pcDNA5-MKRN1- Δ PAM2	none	none (pL66)	Ampicillin, Hygromycin	
	pL109	pcDNA5-MKRN1-sR16insens-H307E	none	none (pL70)	Ampicillin, Hygromycin	
	pL118	pMX-DEST53-IP-GFP MKRN1 ZNF1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL119	pMX-DEST53-IP-GFP MKRN1 ZNF2	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL120	pMX-DEST53-IP-GFP MKRN1 ZNF3	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL121	pMX-DEST53-IP-GFP MKRN1 ZNF4	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL150	pcDNA-DEST53-IP-GFP EV	N' term GFP	pcDNA-DEST53-GFP	Ampicillin, Puromycin	
RSP10	pL57	pMX-Dest53-RPS10	none	pMX-Dest53	Ampicillin, Puromycin	
	pL130	pMX-DEST53-GFP RPS10	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
	pL143	pMX-Dest53-RPS10 K107R	none	pMX-Dest53	Ampicillin, Puromycin	
	pL144	pMX-Dest53-RPS10 K138R	none	pMX-Dest53	Ampicillin, Puromycin	
	pL145	pMX-Dest53-RPS10 K139R	none	pMX-Dest53	Ampicillin, Puromycin	
	pL146	pMX-Dest53-RPS10 sR79-insens	none	pMX-Dest53	Ampicillin, Puromycin	
	pL147	pMX-Dest53-RPS10 K107R sR79-insens	none	pMX-Dest53	Ampicillin, Puromycin	
	pL148	pMX-Dest53-RPS10 K138R sR79-insens	none	pMX-Dest53	Ampicillin, Puromycin	
	pL149	pMX-Dest53-RPS10 K139R sR79-insens	none	pMX-Dest53	Ampicillin, Puromycin	
	pL123	pENTR223-RPS10	none	pENTR223	Spectinomycine	Harvard medical school
Luciferase reporter assay	pL10	pRL-TK	none	pRL-TK	Ampicillin	Niehrs Lab
	pL11	pGL3-promoter	none	pGL3-pro	Ampicillin	Niehrs Lab
	pL112	phPRF SST2 IRES	none	none	Ampicillin	Dr. Tobias Schmid
	pL135	pGL3pro-HIST1H1C	none	pGL3-pro (pL11)	Ampicillin	Invitrogen/ Geneart
	pL136	pGL3pro-HIST1H4B	none	pGL3-pro pL(11)	Ampicillin	
	pL140	pGL3pro-Ai	none	pGL3-pro (pL11)	Ampicillin	
	pL141	pGL3pro-As	none	pGL3-pro pL(11)	Ampicillin	
	pL18	PNN-3'UTR-pGL3-pro	none	pGL3pro	Ampicillin	
	pL24	ZNF281-3'UTR-pGL3-pro	none	pGL3pro	Ampicillin	
	pL30	pGL3pro-ZNF281-3'UTR-mut	none	pGL3pro		
	pL32	pGL3pro-YWHAB-3'UTR	none	pGL3pro	Ampicillin	
	Putative MRKN1 ubiquitylation targets	pL125	pMX-DEST53-GFP YWHAB	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin
pL126		pMX-DEST53-GFP HSPA8	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
pL127		pMX-DEST53-GFP LARP1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
pL128		pMX-DEST53-GFP RPS20	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	

Table V-3, Plasmids used in this study

Experiment	internal name	Plasmid	Tag	Backbone	Resistance	Origin
	pL129	pMX-DEST53-GFP HSP90AA1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL130	pMX-DEST53-GFP RPS10	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL103	pMX-DEST53-GFP-IGF2BP1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL104	pMX-DEST53-GFP-HSPA1A	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL105	pMX-DEST53-GFP-HIST1H1C	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL106	pMX-DEST53-GFP-EEF1A1	N' term GFP	pMX-DEST53-IP-GFP	Ampicillin, Puromycin	
	pL91	pENTR223.1-IGF2BP1 (no stop)	none	pENTR223.1	Spectinomycine	
	pL92	pENTR221-HSPA1A	none	pENTR221	Kanamycine	IMB Orfeome Clones
	pL93	pENTR221-EEF1A1	none	pENTR221	Kanamycine	
	pL94	pENTR221-HIST1H1C	none	pENTR221	Kanamycine	
	pL122	pENTR223-HSP90AA1	none	pENTR223	Spectinomycine	
	pL123	pENTR223-RPS10	none	pENTR223	Spectinomycine	
	pL114	pENTR221-YWHAB entry	none	pENTR221	Kanamycine	
	pL115	pENTR221-HSPA8	none	pENTR221	Kanamycine	IMB Orfeome Clones
	pL116	pENTR221-LARP1	none	pENTR221	Kanamycine	
	pL117	pENTR221-RPS20	none	pENTR221	Kanamycine	

Table V-4, siRNAs used in this study additionally to Suppl. Table S6 in Chapter IV. For MKRN1 siRNA1 (sR16) and siRNA2 (sR17), as well as control siRNA (sR18), see Table S6 in Chapter IV

Target	Internal number	Sequence	Origin
RPS10	sR79	GAGCUGGCAGACAAGAAU[dT][dT]	Ren J, J Biol Chem 2010, PMID: 20159986; siRNA RPS10A
On Target plus non targeting pool	sR12	ugguuuacaugucgacuaa	Fisher / Dharmacon (D-001810-10-05)
		ugguuuacauguuguguga	
		ugguuuacauguuuucuga	
		ugguuuacauguuuuccua	
MKRN1	sR13	UGACUUGGAUCUAUAGCAA	Dharmacon/ GE Healthcare (L-006959-00)
		UCAAGUCUCUCAUCGAUAG	
		GAGUGGGACUUGUUUCAUG	
		GCAAGUGGAGGAGUGCUAA	
MKRN1	sR19	GUGUAUCUCCACGGAGAAU[dT][dT]	Sigma, Rosetta predicitons (SASI_Hs01_00102106)
HSPA70-1	sR54	UUUCUCUUGAACUCCUCAUU [dT] [dT]	Kishor et al. 2013
IFG2BP1-1.3	sR55	CCAUCCGCAACAUCACAAA[dT] [dT]	Zirkel et al. 2013
IFG2BP1-1.2	sR56	UGAAUGGCCACCAGUUGGA[dT] [dT]	Zirkel et al. 2013
EEF1A1	sR57	GUACUAUGUGACUAUCAUU[dT][dT]	Sigma, Rosetta predicitons (SASI_Hs02_00331771)
MKRN1	sR70	CAGAUCCGGGCCCAACGA[dT][dT]	Sigma, Rosetta predicitons (SASI_Hs01_00102105)

1.2 Methods

1.2.1 MKRN1 mutants

1.2.1.1 Non-tagged MKRN1

MKRN1 was cloned into the pcDNA5 vector (pL68) by extending MKRN1 (mutant) ORFs (from pL13, pL63, and pL59) with sequences recognized by restriction enzymes (oA80/oA81) using the Phusion High-Fidelity PCR Kit according to the manufacturer's recommendations. The pcDNA5 vector and *MKRN1* ORF PCR products were digested with HindIII and BsrGI. Purified inserts were ligated into the dephosphorylated vector using the Quick Ligation Kit according to the manufacturer's recommendations.

To express non-tagged MKRN1 from the pMX-Dest53 reporter, the GFP-coding sequence was deleted from the pMX-Dest53-GFP vector (pL37) using the primers oA254/oA255 and the Q5 Site-Directed Mutagenesis Kit according to the manufacturer's recommendations. Additional information on the used oligonucleotides and plasmids can be found in the Table V-2Table V-3 of this Chapter.

1.2.1.2 Mutations in MKRN1

To enhance GFP-MKRN1 expression, the MKRN1 ORF was cloned into the pcDNA-DEST53-GFP destination vector (pL105), harboring a CMV promoter, using the Gateway LR Clonase II Enzyme mix according to the manufacturer's recommendations, as described for pMX-DEST53-GFP in Chapter IV.

All mutations were introduced into MKRN1 plasmids using the Q5 Site-Directed Mutagenesis Kit according to the manufacturer's recommendations. RING^{mut} and PAM2^{mut} mutations were introduced into non-tagged MKRN1 vectors as described above (Chapter IV), using the oligonucleotides oA09/oA10 (H307E) and oA222/ oA223 (PAM2^{mut}). The PAM2 motif was deleted (Δ PAM2) with the oligonucleotides oA07/oA08. MKRN1 was rendered insensitive to MKRN1 siRNA1 and MKRN1 siRNA2 using the oligonucleotides oA53/oA54 (MKRN1 siRNA1) and oA55/oA56 (MKRN1 siRNA2). For a MKRN1 mutant, harboring the H307E mutation and being insensitive to MKRN1 siRNA2, the oligonucleotides oA220/oA221 were utilized. In order to mutate the cysteine residues within the ZNF domains of MKRN1 to alanine, the following oligonucleotides were applied: oA158/ oA159 (ZNF1), oA160/161 (ZNF2), oA162/163 (ZNF3), and oA164/ oA165 (ZNF4). All plasmids were amplified in *E. coli* DH5 α cells. The primers oA216 and oA225 were used for Sanger sequencing to confirm the plasmid identities. Additional information on the used oligonucleotides and plasmids can be found in the Table V-2Table V-3 of this Chapter.

1.2.2 Stable *MKRN1* knock down cell lines

To create stable *MKRN1* KD cell lines, restriction sites were added to oligonucleotides of predicted shRNA sequences for MKRN1 (MKRN1.479 (oA100) and MKRN1.664 (oA101))⁵. For this, Q5 Polymerase-based PCR was performed, using the primers oA106 and oA107, according to the manufacturer's recommendations. After restriction digest of the PCR products and the lentiviral plasmid backbones miRE11_SEGP_ren713 (pL71) and miRE18_LTGEPIR_Ren713 (pL72) with XhoI and EcoRI, vectors and inserts were ligated using the Quick ligation kit. All plasmids were amplified in *E. coli* DH5 α cells and their sequence was verified by Sanger sequencing. Additional information on the used oligonucleotides and plasmids can be found in the Table V-2Table V-3 of this Chapter.

To establish stable *MKRN1* KD cell lines, lentiviral particles were produced under biosafety level 2 conditions. HEK293T cells were transfected with the shRNA-coding

plasmids (pL77, pL78, pL79, and pL60), the psPAX plasmid (pL73), and the pMD2.G plasmid (pL74). After 48 hours (h), cell supernatants were harvested, filtered through a 0.45 μm filter and diluted in DMEM. For transduction, HEK293T cells were first incubated with viral-particle-containing medium overnight. Then, the virus-particle-containing medium was exchanged for new virus particle containing-medium and the cells were spun at 1200 revolutions per minute (rpm) for 90 minutes (min). This procedure was repeated three times. After 48 h, cells were split and kept in normal cell culture medium as described in Chapter IV. After another 48 h, MKRN1-shRNA-positive cells were selected using puromycin. Following single cell dilutions, MKRN1-shRNA-positive cells were grown in 1.5 $\mu\text{g}/\text{ml}$ puromycin and checked for MKRN1 depletion by Western blot.

1.2.3 Stable *MKRN1* mutant expressing cell lines

To obtain stable cell lines expressing MKRN1 siRNA2-insensitive MKRN1 protein, HEK293T cells were transfected with plasmids coding for *MKRN1* wt or mutants (pL52, pL53, pL76, pL90) or an empty vector (EV) control (pL27) using PEI as described above. Additional information on the used plasmids can be found in Table V-3 of this Chapter. 48 h after transfection, puromycin selection was started and stable, puromycin-resistant cell lines were kept in DMEM containing 1.5 $\mu\text{g}/\text{ml}$ puromycin. Expression of MKRN1 mutant proteins was assessed by Western blot.

1.2.4 Stable Knockout cell lines

Based on the protocol from the Zhang lab (<https://www.addgene.org/crispr/zhang>), the CRISPR-SpCas9 backbone (pL75; Addgene #62988), was digested with BbsI, dephosphorylated and gel purified using the Qiagen gel extraction kit according to the manufacturer's recommendations. *MKRN1* guide RNAs (guide 1: oA256/oA257, guide2: oA258/oA259, guide 3: oA260/oA261) were designed using the ChopChop Website (<http://chopchop.cbu.uib.no>)^{6,7}. Phosphorylated oligonucleotides were annealed according to the "Zhang Lab General Cloning Protocol" (https://media.addgene.org/cms/filer_public/e6/5a/e65a9ef8-c8ac-4f88-98da-3b7d7960394c/zhang-lab-general-cloning-protocol.pdf)^{8,9}. Using T4 DNA ligase at 16°C overnight, the oligonucleotides were ligated into the digested vector. For Sanger sequencing, to confirm the plasmid's sequences, the primer oA119 was used. Additional information on the used oligonucleotides and plasmids can be found in the Table V-2 and Table V-3 of this

Chapter. CRISPR-SpCas9 plasmids, containing guide RNAs for MKRN1 (guide1: pL81, guide2: pL82, guide3: pL83), were transfected into HEK293T cells. After 48 h, CRISPR-SpCas9 positive cells were selected using puromycin (1.5 µg/ml) for five days. MKRN1 KO cells were subsequently cultured in DMEM containing with 10% fetal bovine serum, 1% penicillin/streptomycin, and 1% L-glutamine. MKRN1 protein levels of these cells were assessed by Western Blot. *MKRN1* mRNA levels were analyzed by qPCR.

1.2.5 Actinomycin D treatment to assess mRNA stability

To assess the influence of MKRN1 on mRNA stability, HEK293T cells were transfected with MKRN1 siRNA1 or control siRNA (sR18) for 48 h. Next, the cells were treated with 5 µg/ml actinomycin D for 0, 1, 2, 4, or 6 h. mRNA was isolated using the Qiagen RNeasy Plus Mini Kit according to the manufacturer's recommendations and was subsequently analyzed by qPCR for the mRNA targets (*MKRN1* (oA17/18), *ACTB* (oA15/16), *LOX* (oA67/68), *UNC5B* (oA73/74), and *CUL7* (oA63/64)) using the Luminaris HiGreen Low ROX qPCR Master Mix according to the manufacturer's recommendations. Information on the used oligonucleotides can be found in Table V-2 of this Chapter.

1.2.6 Pre-mRNA assessment

To determine the influence of MKRN1 on mRNA stability in HEK293T cells, pre-mRNA and mRNA levels of different target mRNAs were assessed by qPCR analysis upon *MKRN1* KD2 or control KD after 72 h. For this, the mRNA target genes *LOX* (oA67/68/90), *HTRA3* (oA75/76/88), *CUL7* (oA63/64/82), *UNC5B* (oA73/73/83/84), *ID3* (oA61/62/89), *ITGA8* (oA71/72/85/86), *TNC* (oA59/60/87), *MKRN1* (oA17/18), and *ACTB* (oA15/16), were analyzed by qPCR using the Luminaris HiGreen Low ROX qPCR Master Mix according to the manufacturer's recommendations. Information on the used oligonucleotides can be found in Table V-2 of this Chapter.

1.2.7 RNA sequencing

For transcriptome analysis, untagged MKRN1 (pL66) or EV control (pL68) were ectopically expressed for 48 h in HEK293T cells. The cells were lysed in RLT buffer using Qias shredder columns. Subsequently, RNA was isolated using the Qiagen RNeasy Plus Mini Kit according to the manufacturer's recommendations in two rounds. After the

first round of RNA isolation, residual genomic DNA was removed by incubating the eluted RNA with Turbo DNase at 37°C for 10 min. Afterwards, the second round of RNA isolation was performed. RNA sequencing (RNA-Seq) libraries were prepared by the IMB Genomics Core Facility. To deplete rRNA from the RNA pool, RNA-Seq libraries were prepared using the Ribo Zero Magnetic Kit and the TruSeq RNA Sample Prep Kit v2 according to the manufacturer's recommendations. For normalization, the Ambion ERCC RNA Spike-In Mix was added prior to ribozero depletion to the RNA-Seq libraries according to the manufacturer's recommendations. RNA-Seq libraries were sequenced on an Illumina HiSeq 2500 as 51 bp single-end reads, yielding 50 - 55 million reads per sample. Experiments were carried out in triplicates.

Data analysis was performed by Dr. Anke Busch as follows; Basic sequencing quality checks were applied to all reads using FastQC (version 0.11.3) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Reads were mapped to the human genome (assembly version GRCh37) and its annotation based on GENCODE release 19 as well as ERCC control spike-in genes using STAR (version 2.5.1b)¹⁰⁻¹². When running STAR, up to two mismatches were allowed and only uniquely mapping reads were kept for further analysis. Uniquely mapped reads were summarized per gene using featureCounts from the subread package (version 1.4.6-p2)¹³. Differential gene expression analysis was done using DESeq2¹⁴. Spike-in genes were used to normalize the data. Log₂ fold changes were calculated using shrinkage for low coverage genes. Subsequently, all genes with a p-value corrected for multiple testing of < 0.01 were considered significantly differentially expressed.

1.2.8 Pulsed SILAC

Pulsed SILAC experiments were performed as described before¹⁵. Briefly, HEK293T-LT3GEPiR-MKRN1-664-#10 cells (Chapter V.2.1, Figure V-2A) were cultivated in light SILAC medium. *MKRN1* KD was induced by doxycycline treatment (2 µg/ml). DMSO was used as a control treatment. After 48 h in light SILAC medium, *MKRN1*-deficient cells were incubated with medium SILAC medium containing doxycycline, while DMSO-treated cells were cultured in heavy SILAC medium. After 24 h, the cells were washed with DPBS and subsequently, equal numbers of cells were combined from the medium and the heavy SILAC fraction (i.e. a 1:1 ratio). The cell mix was then lysed in mRIPA buffer, supplemented with inhibitors for proteases, phosphatases and DUBs as described in Chapter IV. Protein concentrations were determined using the Pierce BCA

Protein Assay Kit. 50 µg total protein was heated in LDS sample buffer, supplemented with 1 mM DTT for 10 min at 70°C and alkylated using 5.5 mM CAA for 30 min at room temperature (RT) in the dark¹⁶. Samples were prepared for MS and analyzed in a 4 h gradient as described in Chapter IV. Raw data files were analyzed and peptides were identified using the MaxQuant software (version 1.5.28) as described in Chapter IV¹⁷. Experiments were carried out in quadruplicates.

Analysis of the pulsed SILAC data was performed by Dr. Anke Busch as follows; Proteins that were categorized as “only identified by site”, potential contaminants and reverse hits were removed. Only proteins identified with at least two peptides (including at least one unique peptide) and a SILAC ratio count of at least two in any replicate were used for analysis. SILAC ratios were log₂ transformed and an asymmetric z-score was calculated on these ratios as described in Chapter IV¹⁸. The asymmetric z-score was used to perform a differential protein abundance analysis between two conditions using the R package limma (see Chapter IV)¹⁹. For all genes, the length of the longest A-stretch was extracted.

1.2.9 Luciferase reporter assay

The luciferase reporter plasmids pGL3promoter (pGL3pro, pL11; coding for Firefly luciferase) and pRL-TK (pL10; coding for Renilla luciferase) were a kind gift from Dr. Michael Musheev (Niehrs lab, IMB). MKRN1-bound 3' UTRs including the poly(A) site of *PNN*, *YWHAB*, and *ZNF281* were amplified from genomic DNA using the Phusion High-Fidelity PCR Kit according to the manufacturer's recommendations. For *PNN*, the primers oA35/oA37 and for *YWHAB* the primers oA30/oA37 were used. The 3' UTR of *ZNF281* was amplified in two steps, first with the primers oA33/oA36, which gained a longer product. The final *ZNF281* product was obtained with the primers oA31/oA32. The 3' UTR PCR products were ligated into the pGL3pro vector either directly by restriction digest (BamHI and XbaI) and ligation, or indirectly, by TOPO cloning. Using the TOPO XL PCR Cloning Kit according to the manufacturer's instructions, *YWHAB* PCR product was first inserted into the pCR2.1 vector and then transferred into the pGFL3pro vector by restriction digest (BamHI, XbaI). To mutate the pGL3pro vector, Q5 site-directed mutagenesis kit was used according to the manufacturer's recommendations. The poly(A) stretch and the MKRN1-binding site within the *ZNF281* 3' UTR were removed from the *ZNF281*-3' UTR-pGL3pro plasmid using the primers oA45/oA46. To assess the effect of A-tracts on MKRN1-dependent translation, 21 A

nucleotides were inserted into the pGL3pro vector either before the stop codon (As, pL140) or by replacing the stop codon (Ai, pL141) using the oligonucleotides oA210/oA211 or oA208/oA209, respectively. A luciferase reporter plasmid containing the SST2 5' UTR (phpRF-SST2-IRES, pL112) to assess the role of translation initiation was a kind gift from Dr. Tobias Schmid (Institut für Biochemie I, Goethe-Universität Frankfurt)²⁰. Control luciferase reporter plasmids, harboring histone stem loops instead of polyadenylation sites, were ordered from Thermo Scientific via Invitrogen GeneArt Gene Synthesis (pGL3pro-HIST1H4B, pL136 and pGL3pro-HIST1H1C, pL135). All vectors were amplified in *E. coli* DH5 α cells. By Sanger sequencing, the sequences of the luciferase reporter plasmids were confirmed using the primers oA37, oA38, oA39, oA40, oA41, oA42, oA43, oA44, and oA215. Additional information on the used oligonucleotides and plasmids can be found in the Table V-2 and Table V-3 of this Chapter.

For luciferase reporter assays the Dual-Luciferase Reporter Assay System was used according to the manufacturer's recommendations. Briefly, HEK293T cells were seeded in 6 well plates. KDs and MKRN1 (mutant) overexpression (OE) were performed for 24 h as described above. Cells were harvested and diluted to 1×10^6 cells/ml in their own medium. Subsequently 5×10^4 cells/well were seeded into a 96-well plate in 100 μ l (50 μ l fresh DMEM with 50 μ l cell suspension). The cells were allowed to settle for two to three hours and were then transfected with 50 ng Firefly luciferase plasmid and 1 ng Renilla luciferase plasmid (pRL-TK) using PEI as described above. With a medium change after 24 h, Luciferase luminescence measurements were performed according to the manufacturer's recommendations after 48 h. Briefly, the medium of the cells was exchanged for lysis buffer. After mixing, the cells were incubated at -80°C for 30 min. Subsequently, 20 μ l of the thawed cell suspension were transferred into a white microtiter plate. The luminescence signals were analyzed using a Tecan Reader Infinite 200 PRO.

1.2.10 Ubiquitylation targets of MKRN1

Di-glycine remnant profiling of HEK293T cells transfected with *MKRN1* siRNA1 or control siRNA was performed as described for *MKRN1* siRNA2 in Chapter IV.

1.2.11 Knock down and overexpression of putative MKRN1 ubiquitylation targets

1.2.11.1 Knock down of target proteins

To knock down the putative MKRN1 ubiquitylation targets, HEK293T cells were transfected with siRNAs targeting *HSPA1A* (sR54), *IGF2BP1* (KD1: sR56, KD2: sR55), *EEF1A1* (sR57), or *RPS10* (sR79) for 72 h. RNA was extracted using the Qiagen RNeasy Plus Mini Kit according to the manufacturer's recommendations. RT-PCR and qPCR were performed as described in Chapter 0. The following primers were used to assess target transcript levels in duplicates (for *RPS10* only one replicate was measured): *HSPA1A* (oA178/179), *IGF2BP1* (oA180/181), *EEF1A1* (oA182/183), and *RPS10* (oA264/265). Expression levels were normalized to *ACTB*. Additional information on the used primers and siRNAs can be found in the Table V-2 and Table V-4 of this Chapter.

1.2.11.2 Overexpression of putative MKRN1 target proteins

Gateway cloning using the Gateway LR Clonase II Enzyme mix was performed as described above (Chapter IV). To insert the ORFs of *HSPA1A*, *HSP90AA1*, *HSPA8*, *HIST1H1C*, *LARP1*, *YWHAB*, *EEF1A1*, *IGF2BP1*, *RPS10*, and *RPS20* into the destination vector pMX-DEST53-GFP (pL37), the following entry vectors were obtained from the IMB Core Facility ORFeome Collection²¹: pENTR221-*HSPA1A* (pL92), pENTR221-*HSPA8* (pL115), pENTR221-*HIST1H1C* (pL94), pENTR221-*LARP1* (pL116), pENTR221-*YWHAB* (pL114), pENTR221-*EEF1A1* (pL93), pENTR223.1-*IGF2BP1* (pL91), and pENTR221-*RPS20* (pL117). The entry vectors pENTR223-*RPS10* (pL123) and pENTR223-*HSP90AA1* (pL122) were obtained from the Harvard PlasmID Repository (<https://plasmid.med.harvard.edu/PLASMID/>). Additional information on the used plasmids can be found in Table V-3 of this Chapter. Ectopic expression of GFP-tagged proteins was assessed by Western blot.

1.2.12 RPS10 mutants and stable cell lines

All mutations were introduced into the *RPS10* coding sequence using the Q5 Site-Directed Mutagenesis Kit according to the manufacturer's recommendations. From the pMX-Dest53-GFP-*RPS10* plasmid (pL130), the GFP-tag was removed with the oligonucleotides oA254/255. To mutate ubiquitylated lysines within *RPS10* to arginine, the following oligonucleotides were utilized: oA252/253 (K107R), oA246/247 (K138R), and oA248/249 (K139R). To render *RPS10* insensitive to the siRNA sR79, *RPS10* was

mutated using the oligonucleotides oA262/263. Additional information on the oligonucleotides and plasmids used here can be found in the Table V-2 and Table V-3 of this Chapter. Expression of RPS10 mutant proteins was assessed by Western blot as described above. For SDS-PAGE, NuPAGE MES SDS Running Buffer was used.

1.2.13 MKRN1 response to stress

HEK293T cells were treated with either 2 nM H₂O₂ for 2 h or exposed to heat shock. 42°C for 30 min or 60 min were considered as very mild heat shock, while 44°C for 2 h were considered as severe heat shock²². After stress treatment, Western blot analysis was performed as described above.

2. Results

2.1 MKRN1 can be efficiently depleted in HEK293T cells

In order to have a small interfering RNA (siRNA)-based system to reduce MKRN1 levels, I tested different siRNAs targeting *MKRN1* mRNA for potent KD in HEK293T cells. As assessed by Western blot, only MKRN1 siRNA1 and siRNA2 (see Chapter IV) were efficiently depleting *MKRN1*. Opposed to this, the *MKRN1* targeting siRNAs sR13, sR19, and sR70 did not change MKRN1 expression levels compared to control KD (Figure V-1A). Thus, MKRN1 siRNA1 or siRNA2 were used for all MKRN1 siRNA-based experiments.

Another *MKRN1* KD approach based on the small hairpin RNA (shRNA) method was used to create stable *MKRN1* KD cell lines. For this, HEK293T cell lines that express two different shRNAs targeting *MKRN1* were created. While expression from the SGEP vector (pL71, pL79, pL80) leads to stable *MKRN1* KD (Figure V-2C and D), *MKRN1* KD from the LT3GEIR vector (pL72, pL77, pL78) is inducible by doxycycline (dox) treatment (Figure V-2A and B)^{5,23}. Single cell dilutions were performed and in different clones, *MKRN1* KD was assayed by Western blot (Figure V-2A – D). Most of the *MKRN1* KD cell lines (stable and inducible) showed a decrease of MKRN1 levels. Selected cell lines for *MKRN1* KD and control cell lines were chosen for storage and further experiments (highlighted with 'x', Figure V-2A – D).

To knock out (KO) *MKRN1* in HEK239T cells, the CRISPR SpCas9 System was employed. CRISPR SpCas9 plasmids harboring target sequences against *MKRN1* were created (pL 81, guide1; pL82, guide2; pL82, guide3)^{8,9}. CRISPR SpCas9 transfected

HEK293T cells were selected using puromycin. Single cell-derived cell lines were analyzed for MKRN1 expression by Western Blot and for *MKRN1* and *MKRN2* expression by qPCR (Figure V-2E and F). MKRN2 levels seem to be slightly elevated in *MKRN1* KO cell lines (Figure V-2F). The *MKRN1* KO cell line clones #3.3 and #1.16 display reduced MKRN1 mRNA levels, which hints towards a heterogenous genotype. Only one of the tested cell lines, clone #3.13, displayed decreased *MKRN1* mRNA as well as reduced MKRN1 protein levels (Figure V-2F). This cell line (clone #3.13) will be used for further experiments.

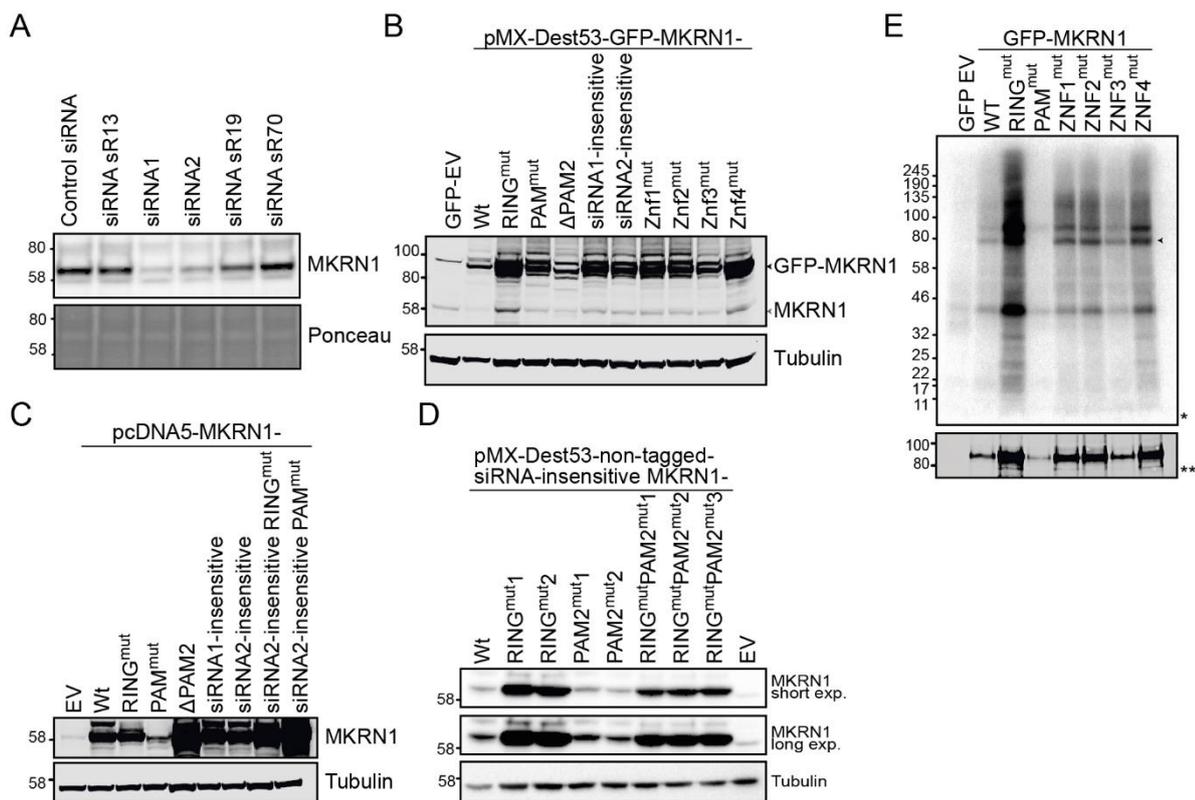


Figure V-1, In depth characterization of MKRN1. **A**, Different siRNAs were tested for *MKRN1* KD. Upon transfection of HEK293T cells, MKRN1 expression as assessed by Western Blot. Only siRNA1 and siRNA2 were found to successfully deplete MKRN1. **B – D**, Ectopic MKRN1 (mutant) levels in HEK293T cells expressed from the pMX-Dest53-GFP backbone (**B**), the pcDNA5 vector (**C**), or the non-tag pMX-Dest53 backbone (**D**) were determined by Western Blot. MKRN1 was detected after 48 h with an antibody against endogenous MKRN1. All overexpressed proteins showed higher levels of expression compared to endogenous MKRN1. **E**, Mutating any ZNF domain of MKRN1 singularly does not affect the RNA-binding behavior of MKRN1. The three cysteines within a ZNF domain of MKRN1 (ZNF1, ZNF2, ZNF3, and ZNF4) were mutated for alanine. The autoradiograph (*) and Western blot (**) show GFP-MKRN1 (mutant)/RNA complexes and GFP-MKRN1 protein, respectively. Crosslinking was performed at 254 nm after ectopic expression for 48 h. Further information on plasmids and siRNAs can be found in Suppl. Table S6 of Chapter IV and in the Table V-3 and Table V-4 of this Chapter.

2.2 MKRN1 proteins are ectopically or stably expressed in HEK293T cells

Besides GFP-tagged MKRN1^{wt} (pL13), MKRN1^{RINGmut} (pL23), and MKRN1^{PAM2mut} (pL55) (Chapter IV), several other plasmids coding for MKRN1 (mutants) were created

(Chapter V.1). A list of all MKRN1 plasmids can be found in Table V-3 in this Chapter. The ectopic expression of most MKRN1 (mutants) was assessed by Western blot (Figure V-1B – D). MKRN1 (mutant) proteins were expressed from all tested plasmids to a higher extent than endogenous MKRN1 (see EV or GFP-EV controls). MKRN1^{RINGmut}, the ligase-dead mutant, is generally stronger expressed than MKRN1^{wt} (Figure V-1B – D), indicating that MKRN1^{wt} levels are usually regulated by proteasomal degradation due to auto-ubiquitylation of MKRN1²⁴. The pcDNA5 (pL68) backbone harbors a CMV promoter, resulting in higher MKRN1 levels compared to MKRN1 expression from the pMX-Dest53-GFP (pL37) backbone, which harbors a MoMuLV LTR (Figure V-1B – D). Non-tagged MKRN1 siRNA2-insensitive MKRN1 (mutants) (pL52, pL53, pL76, and pL90) expressed from the non-tagged pMX-Dest53 backbone (pL27) were used to create stable cell lines. The stable expression of MKRN1^{mut} proteins was confirmed by Western blot upon *MKRN1* KD2 (to deplete endogenous MKRN1) in puromycin resistant cells (Figure V-3). Single cell dilutions of these cells were made to gain populations with even expression of MKRN1 (mutant) proteins. All siRNA2-insensitive *MKRN1* sequences were stably incorporated in most of the tested cell lines and displayed resistance to *MKRN1* siRNA2 (Figure V-3A). The expression of MKRN1 in the selected stable cell lines that were selected for further experiments is shown in Figure V-3B.

Of note, *MKRN1* KD2 affected cell growth in puromycin-containing medium. Puromycin is a translational inhibitor that causes premature chain termination during translation. The observed growth defect could only be rescued by the stable expression of MKRN1^{wt} but not by expressing MKRN1 mutant proteins (RING^{mut}, PAM2^{mut}, or RING^{mut}PAM2^{mut}). *MKRN1* KD2 does not have any negative effects on cellular growth in normal medium (data not shown). As I found MKRN1 to play a role in stalling ribosomes, these findings might indicate a role of MKRN1 in general translation regulation.

2.3 MKRN1 binds to mRNA with more than one ZNF domain

To find out, which MKRN1 ZNF domain is responsible for MKRN1's RNA-binding capability, the three cysteines of each ZNF domain of MKRN1 were mutated into alanine and UV crosslinking experiments were performed as described in Chapter IV. GFP-MKRN1 ZNF-mutant proteins (pL118, pL119, pL120, and pL121) were expressed to a higher extent compared to GFP-MKRN1^{wt} (Figure V-1E, bottom). When compared to protein levels, no changes in RNA-binding behavior of the GFP-MKRN1 single

ZNF^{mut} proteins could be observed relative to MKRN1^{wt} RNA-binding (Figure V-1E). These results indicate that the MKRN1 RNA-binding capability does not depend on a single ZNF domain but might be mediated by a combinatorial action of two or more ZNF domains. To test this, all four ZNF domains of MKRN1 could be mutated simultaneously for new UV crosslinking experiments. Additionally, the RNA-binding affinity might be increased by the interaction of MKRN1 with other RBPs, such as PABP, IGF2BP1, or ELAVL1. This hypothesis is backed up by the finding that the RNA-binding ability of GFP- MKRN1^{PAM2mut} was reduced compared to GFP-MKRN1^{wt} (Chapter IV, Figure IV-3D).

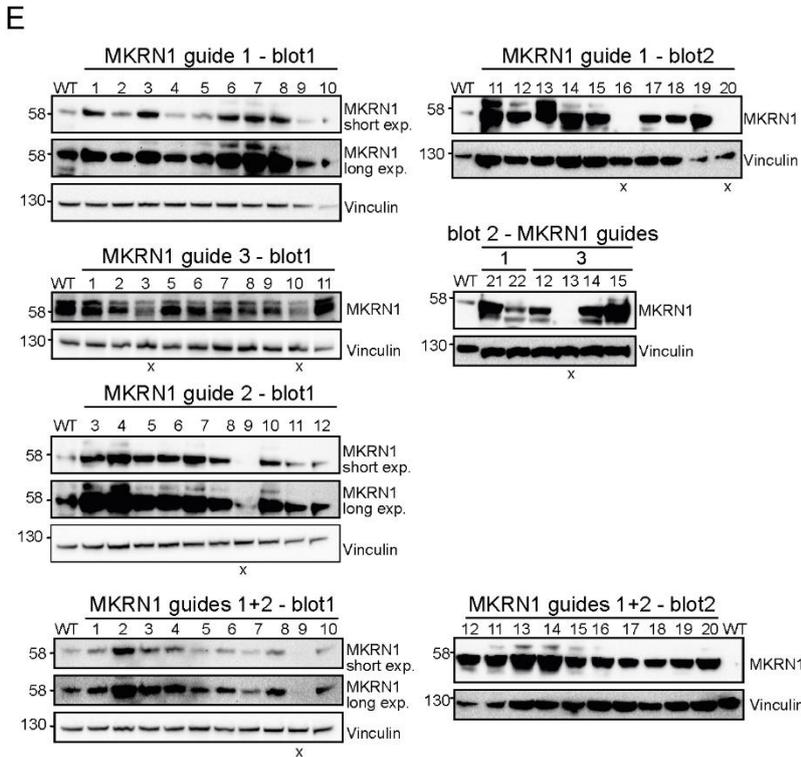
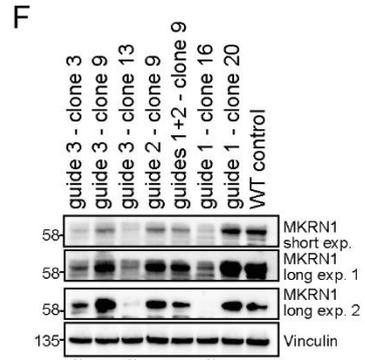
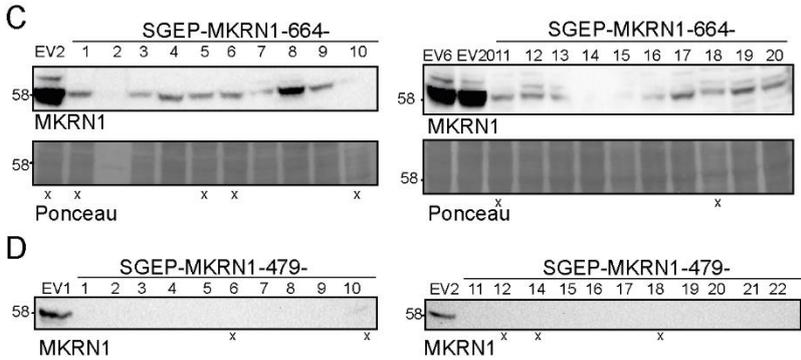
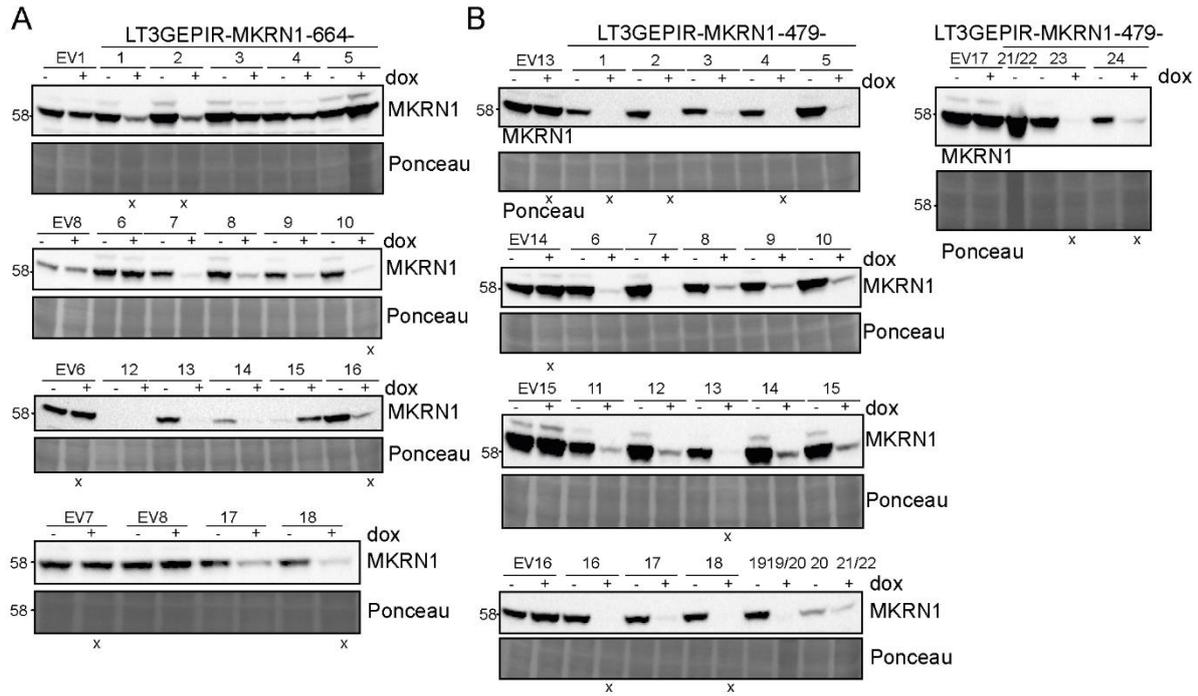
Even though the reduced RNA- as well as protein-binding capacity of MKRN1^{PAM2mut} could be due to aberrant protein folding of MKRN1^{PAM2mut}, it might also be the case that MKRN1 needs the PAM2 motif to interact with other RBPs that guide its binding to RNA.

2.4 Changes in MKRN1 levels redound to global alterations of mRNA levels

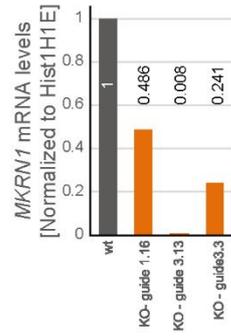
As discussed in Chapter IV, GFP-MKRN1 binds to hundreds of mRNAs, mainly at the 3' UTR and in front of poly(A) tails. To investigate, whether MKRN1 has an effect on the levels of its bound mRNAs, RNA-Seq experiments were performed in MKRN1 or EV OE in HEK293T cells. Spike-ins were added to the RNA-Seq libraries as a normalization means. Upon MKRN1 OE, an overall remote downregulation of polyadenylated mRNAs could be observed relative to spike-in levels (Figure V-4A). Replication-dependent histone mRNAs, harboring 3' stem loops instead of poly(A) tails, behaved similarly to ERCC spike-in probes. In contrast, poly(A) tail-possessing histone mRNAs levels were reduced, equally to the general polyadenylated mRNA pool (Figure V-4A). These findings hint towards a general, albeit modest, response of polyadenylated mRNA levels towards changes in MKRN1 expression. Thus, further experiments are necessary to elucidate the putative influence of MKRN1 on poly(A)-mRNAs. Additionally, it would be interesting to see whether the observed downregulation of poly(A)-mRNAs is dependent on MKRN1's ubiquitylation function. This could be done in experiments with MKRN1^{RINGmut} OE upon depletion of endogenous *MKRN1*. It might be the case that the changes observed on mRNA levels upon MKRN1 OE are not due to a direct effect of MKRN1 on mRNA but might be a response to MKRN1-mediated alteration in translational regulation. In RQC, aberrant mRNAs, which ribosomes stall on, are degraded by NGD or NSD. Thus, the reduction in mRNA levels upon MKRN1 OE might result from extensive ribosome stalling, which triggers mRNA decay.

I wondered whether MKRN1 could affect mRNA levels in terms of regulating mRNA stability. To address this question, cells transfected with MKRN1 siRNA1 or control siRNA were treated with actinomycin D. In a first time course experiment, mRNA abundances of selected mRNAs were monitored by qPCR. Figure V-4B illustrates that the inhibition of transcription by actinomycin D in *MKRN1* KD1 conditions did not result in the depletion of the monitored mRNA pools over time, relative to control KD. Thus, MKRN1 does not seem to affect the mRNA stability of the analyzed mRNAs. In general, lower levels of mRNAs

Further characterization of MKRN1 and preliminary experiments



Expression of *MKRN1*



Expression of *MKRN2*

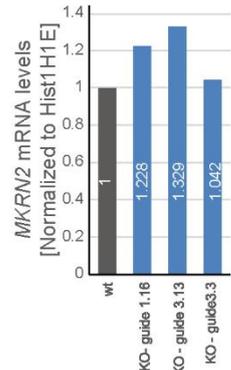


Figure V-2, Stable cell lines. **A & B,** *MKRN1* KD could be induced by doxycycline. Stable cell lines carrying constructs for two different shRNAs targeting *MKRN1* (shRNA 664, A) and (shRNA 479, B) were created and *MKRN1* KD efficiency was assessed in comparison to empty vector (EV) control cells⁵. **C & D,** The stable KD of *MKRN1* with two different shRNAs (shRNA 664, C) and (shRNA 479, D) was assessed by Western Blot⁵. Cell lines selected for further experiments are marked with an 'x'. **E & F,** HEK293T cells were transfected with plasmids coding for SpCRISPRCas9 and three different guide RNAs targeting endogenous *MKRN1*. Upon puromycin selection, cell lines, derived from single cell colonies, were assessed for *MKRN1* protein levels by Western blot. Cell lines selected for further experiments are marked with an 'x'. **E,** Single cell-derived HEK293T clones were tested for *MKRN1* expression. **F,** Selected cell lines from (E) were analyzed for *MKRN1* protein levels by Western blot (top) and *MKRN1* mRNA levels by qPCR (bottom). The clone guide 3.13 was chosen for further experiments.

in *MKRN1* KD1 were found. All expression data was normalized to *ACTB* levels. As *ACTB* is a polyadenylated mRNA also this mRNA, which was used for normalization, might be affected by altered *MKRN1* levels. Consequently, the normalization might not have worked. Thus a control is needed that is independent of potential *MKRN1*-driven effects. Replication-dependent histone mRNAs do not contain a poly(A) tail but a stem loop at their 3' end. Thus, one could utilize histone mRNAs for normalization. To assess whether *MKRN1* exhibits a function in transcriptional regulation as suggested by Omwancha *et al.*, further experiments need to be conducted using an unaffected normalization control²⁵. Besides its mainly cytoplasmic localization, I found a subpopulation of *MKRN1* protein to

A

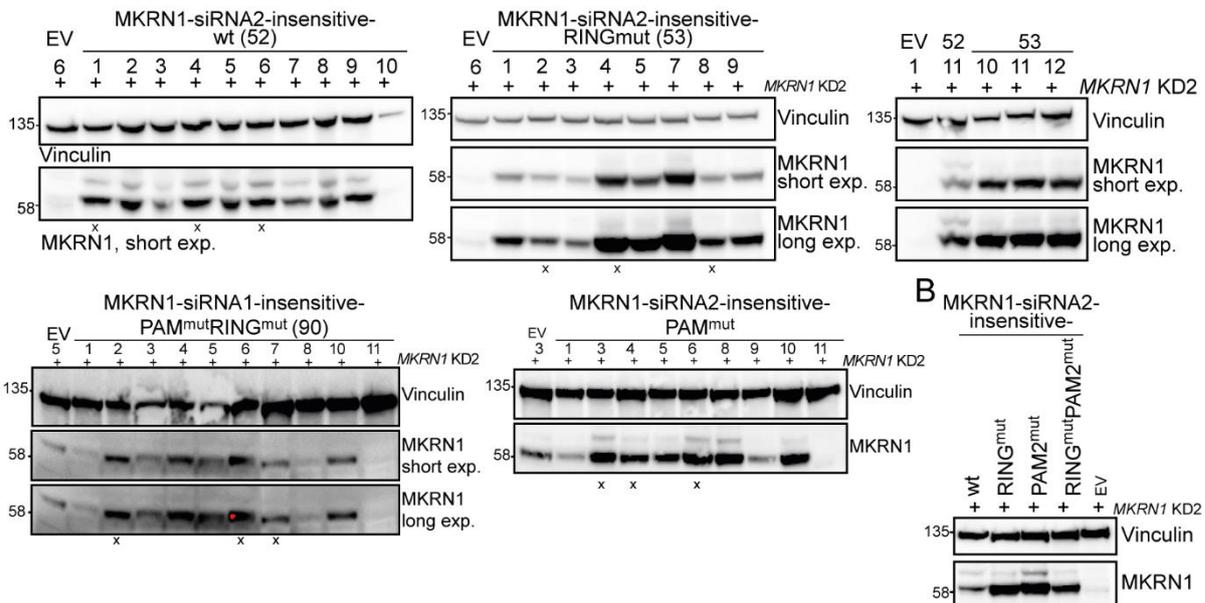


Figure V-3, MKRN1 mutants are stably expressed in HEK293T cells. Endogenous *MKRN1* was knocked down with in HEK293T cells for 72 h and the expression of the stably expressed, siRNA2-insensitive *MKRN1* was assessed by Western Blot. **A,** Stable cell lines derived from single cells were tested for stable insertion of *MKRN1* coding sequences. *MKRN1* KD2 did not affect the expression of stably inserted *MKRN1*^{wt} and *MKRN1*^{RINGmut} in all tested cell lines. Some of the cell lines tested for stable insertion of *MKRN1*^{PAM2mut} and *MKRN1*^{PAM2mutRINGmut} were affected by *MKRN1* KD2, while others had stably inserted the coding sequences for *MKRN1*^{PAM2mut} and *MKRN1*^{PAM2mutRINGmut}. Cell lines selected for further experiments are marked with an 'x'. **B,** *MKRN1* KD2 targeting endogenous *MKRN1* did not affect the expression of stably inserted *MKRN1* wt or mutants in a pool of transfected HEK293T cells.

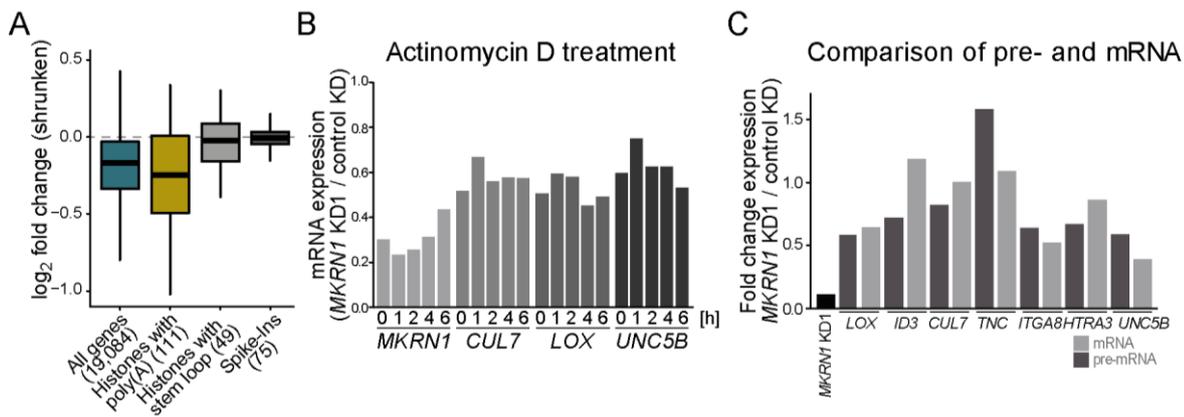


Figure V-4, Changes in MKRN1 levels alter mRNA abundances. **A**, MKRN1 might regulate the levels of polyadenylated mRNAs. Slight, albeit global, changes in mRNA abundances were observed, when mRNA levels were compared to Spike-ins and replication-dependent histone transcripts. RNA-Seq experiments were performed upon rRNA depletion in conditions of ectopic overexpression of MKRN1 vs. empty vector controls. RNA-Seq libraries were sequenced on an Illumina HiSeq 2500 as 51 bp single-end reads, yielding 50 - 55 million reads per sample. Spike-in genes were used to normalize the data. Log₂ fold changes were calculated using shrinkage for low coverage genes; n= 3. **B**, mRNA levels are decreased in *MKRN1* KD relative to *ACTB*. HEK293T cells were treated with Actinomycin D for 0, 2, 4, or 6 h in *MKRN1* KD1 and control KD conditions. mRNA levels were assessed by qPCR and normalized to *ACTB*; n = 1. **C**, General mRNA levels (pre- and mRNA) are reduced in *MKRN1* KD compared to *ACTB*. Pre-mRNA levels were compared to mRNA levels by qPCR analysis. *MKRN1* levels upon *MKRN1* KD1 are shown in black; n = 1.

be present within the nucleus (Chapter III, Figure III-2E and Suppl. Table III-3-6). This goes in line with the putative function of MKRN1 in transcriptional regulation. Moreover, MKRN1 is associated with HELZ2, HNRNPD, PURA, AKAP8, and YBX1 (Chapter IV). Those proteins have been linked to transcription regulation, further supporting a nuclear function of MKRN1. Consequently, the function of nuclear MKRN1 would need to be assessed in further experiments. In a second approach, the effect of MKRN1 on mRNA was assessed by comparing pre-mRNA to mRNA levels of selected targets under *MKRN1* KD1 conditions by qPCR analysis. Levels of pre-mRNAs and mRNAs were normalized to *ACTB*. In an initial experiment, pre-mRNA and mRNA levels seem to be generally slightly reduced in *MKRN1* KD1 conditions compared to control KD. For *LOX*, *CUL7*, *ITGA8*, *HTRA3*, and *UNC5B* no major changes between pre- and mRNA levels could be observed (Figure V-5C). *ID3* expression might be regulated on the transcriptional level in a MKRN1-dependent manner. A role in transcriptional regulation has been assigned to MKRN1 before²⁵. *TNC* is the only mRNA, whose mRNA stability might be negatively affected by *MKRN1* KD (Figure V-4C). However, as generally (pre-) mRNA levels seem to be decreased in *MKRN1* KD, this might also be the case for *ACTB* levels that were used to normalize expression values. Thus, these experiments addressing mRNA stability would need to be repeated using a normalization control, like spike-ins, that is not affected by *MKRN1* KD.

2.5 MKRN1 affects translational regulation

2.5.1 MKRN1 controls protein levels of newly translated proteins

As global protein expression was not altered upon *MKRN1* KD1 (see Chapter IV, Suppl. Fig. IV-8), I set out to assess the effect of *MKRN1* KD on newly translated proteins. With pulsed SILAC experiments, differentially expressed proteins in *MKRN1* KD and control conditions could be quantitatively compared^{15,26}. By quantitative LC-MS/MS, protein abundances of proteins translated within 24 h were assessed for *MKRN1* KD and control conditions^{15,26}. For analysis, quantified proteins were classified into the groups replication-dependent histones (with stem loop), replication-independent histones (with poly(A) tail), and ‘all other quantified proteins’. The mean expression of the group of ‘all other quantified proteins’ was not altered upon *MKRN1* KD. In comparison, the levels of histones with stem loops (i.e. replication-dependent histones) were slightly reduced in *MKRN1* depleted cells (Figure V-5A). However, in the RNA-Seq experiments I found that MKRN1 did not influence replication-dependent histone mRNAs but affected the polyadenylated mRNA pool (Figure V-4A). This prompted me to assume that, due to normalization, the data was skewed towards a global ‘no change’ effect, masking the actual influence of *MKRN1* KD on protein expression. Under this premise, newly translated proteins would be upregulated in *MKRN1* KD compared to unaltered levels of replication-dependent histones as a control.

This indicates that MKRN1 might be indeed involved in translational regulation. RBP-binding to sequences within 3' UTRs is a common means to regulate translation. With regard to the role of MKRN1 in RQC and its binding to poly(A) tracts within 3' UTRs (Chapter IV), all quantified proteins were classified according to their internal ‘longest A stretch’ within the ORF to assess whether MKRN1 has an effect on mRNAs with long internal A stretches. The expression changes of the proteins within the ‘longest A stretch’ groups were compared between the *MKRN1* KD and the control condition. No major differences were observed between proteins harboring short (2 – 6 As; 4188 proteins) or longer (> 6 As; 1892 proteins) poly(A) sequences (Figure V-5B). Thus, MKRN1 might not regulate protein levels via internal A stretches but rather control translation via the poly(A) tail. This hypothesis is supported by the finding that MKRN1 also binds to mRNA in front of poly(A) tails (Chapter IV, Figure IV-3A – C) and stalls ribosomes on poly(A) stretches consisting of 60 A nucleotides (Chapter IV, Figure IV-4C).

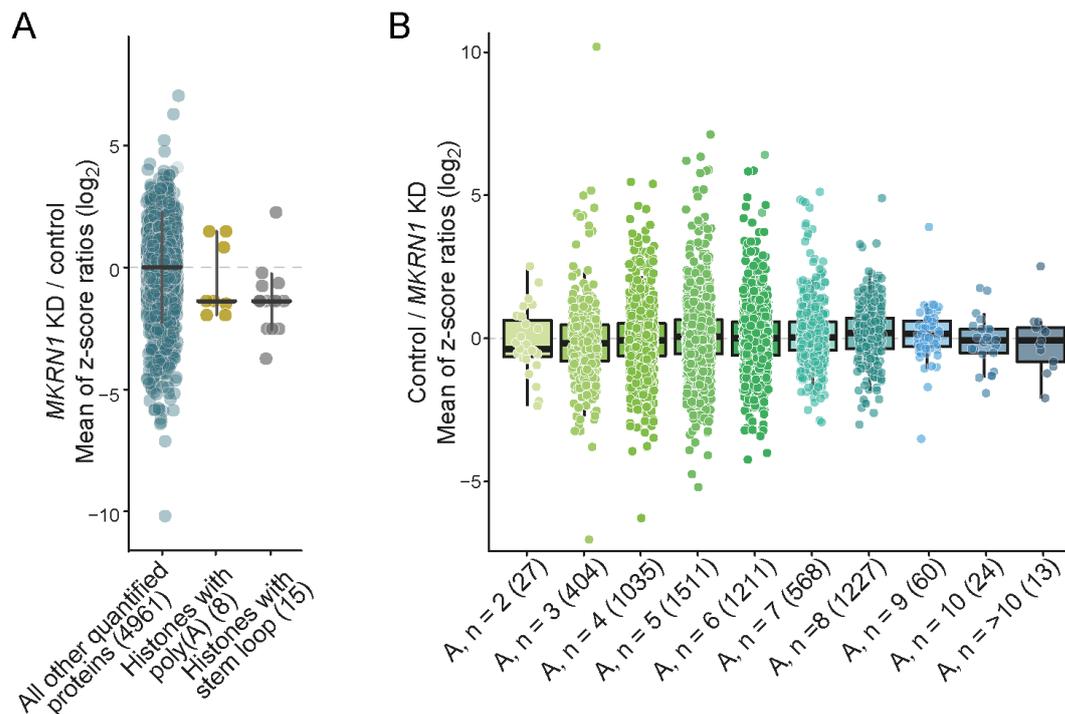


Figure V-5, MKRN1 controls protein levels of newly translated proteins. Proteins synthesized within 24 h were analyzed in *MKRN1* KD conditions in a pulsed SILAC experiment; $n = 4$. **A**, The dotplot shows asymmetric z-scores of log₂ ratios of *MKRN1* KD vs. control. Means of asymmetric z-scores are indicated. Replication dependent and replication independent histones are compared to all other quantified proteins. **B**, The longest A-stretch was determined for all quantified proteins. Boxplots indicate of mean asymmetric z-scores of log₂ SILAC ratios are shown. Control conditions are compared to *MKRN1* KD.

2.5.2 MKRN1 might regulate translation via the poly(A) tail and does not control translation initiation

To further assess the role of MKRN1 in translation, luciferase reporter assays were performed. 3' UTRs of three genes (*YWHAB* (pL32), *PNN* (pL18), and *ZNF281* (pL24)) harboring MKRN1-binding sites were cloned into Firefly luciferase vectors (Figure V-6A). The Firefly luminescence of the different reporters was compared to the luminescence of a control Renilla vector. In HEK293T cells, MKRN1 OE resulted in the downregulation of Firefly as well as Renilla luminescence signal, independent of the Firefly luciferase 3' UTRs (normalized to EV control) (Figure V-6B, left). In line, *MKRN1* KD led to increased Firefly as well as Renilla luminescence signals compared to control KD (Figure V-6B, right). Again, this effect was not dependent on Firefly luciferase 3' UTRs, i.e. MKRN1 binding to 3' UTRs, as Renilla luminescence signals changed concomitantly with Firefly luminescence. Hence, all reporters seemed to be regulated by MKRN1, independently of 3' UTR binding sites. This observation hinted towards a general role of MKRN1 in translation via the poly(A) tail. To assess the importance of MKRN1 binding sites within mRNA 3' UTRs, the MKRN1 binding site within the *ZNF281*

3' UTR (including the poly(A) stretch immediately downstream of this binding site) was deleted (*ZNF281^{mut}*; pL30). When compared to the luminescence of the Firefly *ZNF281^{wt}* reporter, expression of the Firefly *ZNF281^{mut}* vector did not exhibit significant alterations in Firefly luminescence (Figure V-6B and C). This result supported the finding that MKRN1 seems to regulate translation via the poly(A) tail and strengthens the notion that MKRN1 binding within the 3' UTR might not be the major regulation site for MKRN1 in translational control. Hence, Renilla luminescence signals could not be utilized as a normalization means for Firefly luminescence, as Renilla luminescence behaved directional to Firefly luminescence (Figure V-6B and C).

To assess whether MKRN1 regulates translation by rendering translation initiation, I determined the effect of MKRN1 on cap- and IRES (internal ribosome entry site)-dependent translation. To do so, I evaluated Firefly and Renilla translation from the phpRF-SST2-IRES vector (phpRF; pL112) under MKRN1 OE or *MKRN1* KD conditions. In this vector, Renilla luciferase is translated in a cap-dependent manner. However, translation efficiency is reduced due to a stem loop that is inserted upstream of the Renilla luciferase coding region. Firefly luciferase translation is driven by the SST2 IRES, inserted intercistronically upstream of the Firefly luciferase coding sequence (Figure V-6A)^{20,27}. As observed for the reporter plasmids tested before, MKRN1 OE resulted in decreased protein levels of Firefly luciferase, whereas MKRN1 KD increased Firefly luciferase levels; independent of cap- or IRES-driven translation (Figure V-6C).

In comparison to the Firefly-*ZNF281^{wt}* or Firefly-*ZNF281^{mut}* luminescence, the IRES-driven expression of Firefly from the phpRF-SST2-IRES plasmid was not altered in MKRN1 OE conditions (Figure V-6C). Therefore, MKRN1 did not seem to regulate translation initiation but rather via binding in front of poly(A) tails.

To find out more about the molecular mechanism behind MKRN1-regulated translation, a suitable normalization control is needed, as Renilla Firefly is equally affected by rendered MKRN1 levels. Therefore, I utilized Firefly luciferase plasmids harboring a histone stem loop (stem loop sequence from *HIST1H1C*, pL135 and *HIST1H4B*, pL136) instead of the usual SV40 late poly(A) signal (Figure V-6A). The expression of 3' stem-loop containing Firefly luciferase was assayed compared to Firefly-*ZNF281^{wt}* luciferase and control Firefly luciferase (pGL3pro). Firefly luminescence was normalized to Renilla luminescence. Firefly levels expressed from the vectors containing histone 3' stem-loops were only detected at background levels and no conclusions could be drawn from this experiment (Figure V-6D). Thus, the vectors harboring Firefly-histone 3' stem-loop

luciferase sequences were not suitable for normalization in luciferase reporter experiments at this state. Histones with 3' stem loops are expressed in a replication-dependent manner and this might also be the case for Firefly-histone 3' stem-loop luciferase²⁸.

Further experiments are necessary to optimize Firefly-histone 3' stem-loop luciferase expression. Synchronizing cells might help to overcome this issue, as expression of Firefly-histone 3' stem-loop luciferase from a synchronous cell population may increase signal strength. Altering the expression vector in a way that Firefly-histone 3' stem-loop luciferase expression is forced is another option to increase Firefly luminescence. Moreover, driving histone 3' stem-loop luciferase expression from the Renilla-coding vector would enable 'one-well' signal normalization for cell density, for example.

To sum up, these initial experiments imply a general role of MKRN1 in translation regulation, rather via the poly(A) tail than by controlling translation initiation. However, the molecular mechanism of this translational control mode still needs to be deciphered. Furthermore, the role of the ubiquitylation function and the RNA-binding capacity of MKRN1 in this pathway needs to be determined. For this purpose, MKRN1 mutant-expressing stable cell lines or *MKRN1*-depleted cell lines can be utilized for further analyses where luciferase reporter constructs with and without poly(A) tails are compared, for example. Additionally, translation could be inhibited with cycloheximide in *MKRN1* KD or KO conditions and the stability of target proteins could be monitored by Western blot in order to assess the effect of MKRN1 on the translation and stability of specific proteins.

2.5.3 Utilizing the Luciferase reporter assay to analyze the role of MKRN1 in RQC

To obtain a tool for assessing the effect of poly(A) stretches on MKRN1-mediated regulation of RQC responses, 21 A nucleotides were inserted either before the Firefly luciferase stop codon (As, pL141) or replaced the Firefly luciferase stop codon (Ai, pL140) (Figure V-6A). In an initial experiment, the translation of a non-stop Firefly luciferase construct with a 3' terminal A-stretch ('Firefly-Ai') was assessed in relation to Renilla luminescence. Firefly-Ai seems to be expressed at lower levels compared to Firefly luciferase translated from a control plasmid (pGL3pro) or the plasmid that harbors the *ZNF281* 3' UTR (Figure V-6D). Whether MKRN1 regulates the translational stalling on this non-stop reporter or which effect the insertion of 21 A nucleotides before

the stop codon has, still needs to be determined. As the effect of MKRN1 on ribosome stalling at

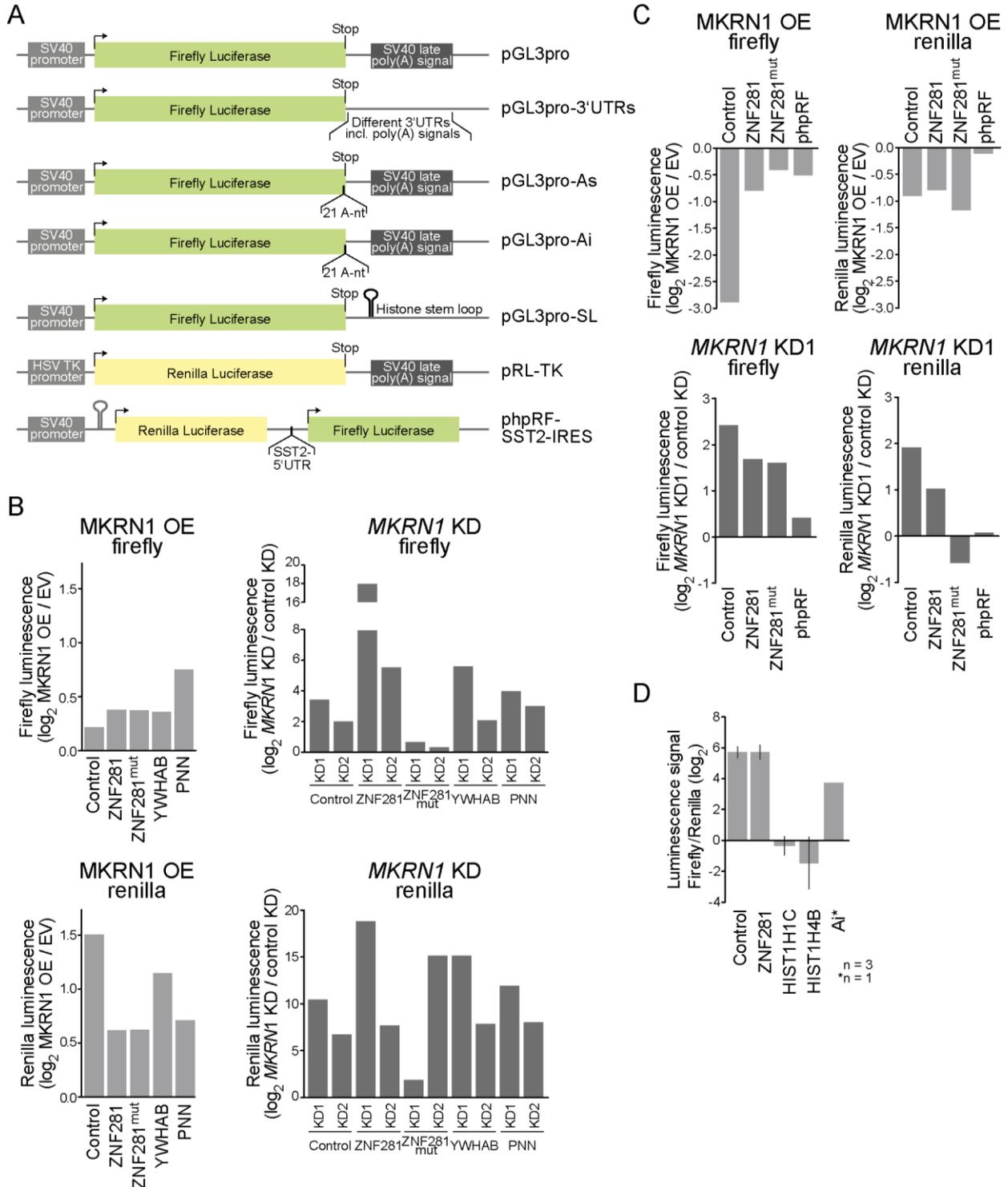


Figure V-6, MKRN1 might regulate translation via the poly(A) tail and does not control translation initiation. **A**, Luciferase plasmids used in this study. **B**, MKRN1 overexpression decreases firefly and renilla luciferase levels independently of 3' UTR binding sites. *MKRN1* KD enhances luciferase protein levels. 3' UTRs (including MKRN1 binding sites and PASs) of different mRNAs (ZNF281, YWHAB, and PNN) were inserted into firefly luciferase vectors. Firefly and Renilla luminescence was determined in three technical replicates. Mean luminescence signals of MKRN1 OE or *MKRN1* KD compared to control conditions are depicted. **C**, MKRN1 does not regulate translation initiation. Luminescence signals of firefly luciferase with different 3' UTRs were compared to the luminescence of the firefly luciferase encoded from the phpRF-SST2-IRES plasmid. In the latter, translation initiation is under the control of an IRES. Independent of translation initiation, MKRN1 OE decreased firefly and renilla protein levels, while *MKRN1* KD had the opposite effect. Experiments were performed in three technical replicates. Mean luminescence signals of MKRN1 OE or *MKRN1* KD compared to control conditions are depicted. **D**, Firefly luciferase transcripts harboring histone stem loops instead of canonical 3' UTRs are only minimally translated. Luminescence of these transcripts could only be detected at background levels. Mean luminescence signals of MKRN1 OE or *MKRN1* KD compared to control conditions from three biological replicates and standard deviations of the mean are depicted. Substitution of the stop codon with 21 A nt did not alter protein expression. Mean luminescence signals from three technical replicates of MKRN1 OE or *MKRN1* KD compared to control conditions are depicted. * indicates n = one biological replicate.

poly(A) stretches only becomes apparent when 60 A nucleotides are inserted into dual fluorescence reporter (showing a minor effect with 36 A nt; see Chapter IV, Figure IV-4C), the insertion of 21 A nucleotides into luciferase vectors might not result in significant changes in Firefly luciferase translation. Thus, longer A-stretches could be inserted into the reporter plasmids to assess the role of MKRN1 on long A-tract containing Firefly luciferase translation.

2.6 MKRN1 ubiquitylates proteins found within the translational elongating complex

2.6.1 MKRN1 depletion slightly reduces ubiquitylation of heat shock proteins and histones

While MKRN1 siRNA2 affects both *MKRN1* and *MKRN2* transcript abundances, MKRN1 siRNA1 also reduces ZNF598 levels (see Chapter IV, Suppl. Fig. IV-6B). By comparing relative abundances of ubiquitylation sites between both knockdown conditions (*MKRN1* KD1 vs. KD2) relative to a control KD, MKRN1-dependent ubiquitylation targets, additional to the ones described in Chapter IV, might be identified. Thus, I searched for peptides exhibiting decreased abundances in *MKRN1* KD1 and KD2 conditions, in two replicate experiments. For most detected ubiquitylation sites, the effect of *MKRN1* KD1 was stronger than KD2 (Figure V-7A and B) and hence, it cannot be fully excluded that the detected changes in ubiquitylation abundance might also result from altered levels of ZNF598. The identified presumptive MKRN1 ubiquitylation targets included EEF1A1, a protein that delivers amino acid-charged tRNAs to the ribosomal A site, and the heat shock proteins HSPA1A, HSP90AA1, and HSPA8 (Figure V-7A)²⁹. Heat shock proteins can be found associated with nascent proteins to assist folding and release from the ribosomes^{30,31}. These findings indicate that MKRN1

might not only ubiquitylate RPS10 and RPS20 (see Chapter IV) to induce ribosomal stalling, but also other proteins within the translational hub, such as PABPC1/4, EEF1A1, and heat shock proteins associated with the nascent peptide. As MKRN1 interacts with VCP (Chapters III and IV) and ubiquitylates HSP70 (Chapter IV, Figure V-7A), one could speculate that those proteins are also involved in the MKRN1 mRNP.

Moreover, reduced abundances of ubiquitylation on UBE2O could be observed upon *MKRN1* KD (Figure V-7B). In search of a collaborative E2 ligase, UBE2O, an E2/E3 hybrid ubiquitin-ligase mediating monoubiquitylation, could be exploited in more detail in relation to MKRN1^{32,33}. Of note, MKRN1 also binds to the 3' UTR of *UBE2O* mRNA (data not shown). Another point of interest is the type of ubiquitylation mediated by MKRN1. The fact that MKRN1 autoubiquitylates itself, which results in MKRN1 degradation, hints towards the formation of K48 chains by MKRN1²⁴. In the *MKRN1* KD1 and KD2 comparison, a minor decrease in ubiquitylation abundance could be observed on K27-linked ubiquitin (*UBC*) (Figure V-7B). Therefore, it might be interesting to elucidate the E2 ligases interacting with MKRN1 to identify the respective MKRN1 linkage specificity. In addition to the 14-3-3 proteins YWHAB and YWHAЕ, several histones (HIST1H1E, HIST1H2A, HIST2H2BE, and HIST1H1C) could be identified as putative MKRN1 ubiquitylation targets. Especially HIST1H1C displayed strongly decreased ubiquitylation abundances in both *MKRN1* KD conditions compared to control KD (Figure V-7B). The functional relevance of this ubiquitylation, however, still needs to be elucidated. Moreover, the ubiquitylation of the here and in Chapter IV identified ubiquitylation targets of MKRN1 needs to be validated in further experiments.

2.6.2 MKRN1 ubiquitylation targets can be efficiently depleted from HEK293T cells

A next step would be to assess the interplay of putative ubiquitylation targets with MKRN1 and to find out whether there is a link to RQC. To do so, initial assays were performed to set up an experimental system to address those questions. Therefore, *EEF1A1*, *IGF2BP1*, *HSPA1A*, and *RPS10* knockdowns were tested by qPCR. Using the siRNA approach, transcript levels of the aforementioned target mRNAs could be reduced by at least 80% (Figure V-7C) Thus, the selected siRNAs can be employed in further experiments. Additionally, the effect of the KD of one of these putative ubiquitylation substrates on *MKRN1* transcript abundance relative to *ACTB* levels was measured.

Generally, *MKRN1* transcript levels were not altered upon KD of *EEF1A1*, *IGF2BP1*, or *HSPA1A*. Only *RPS10* KD seemed to negatively affect *MKRN1* expression (Figure V-7D). For convincing evidence, however, these experiments would need to be repeated to have a good number of replicates.

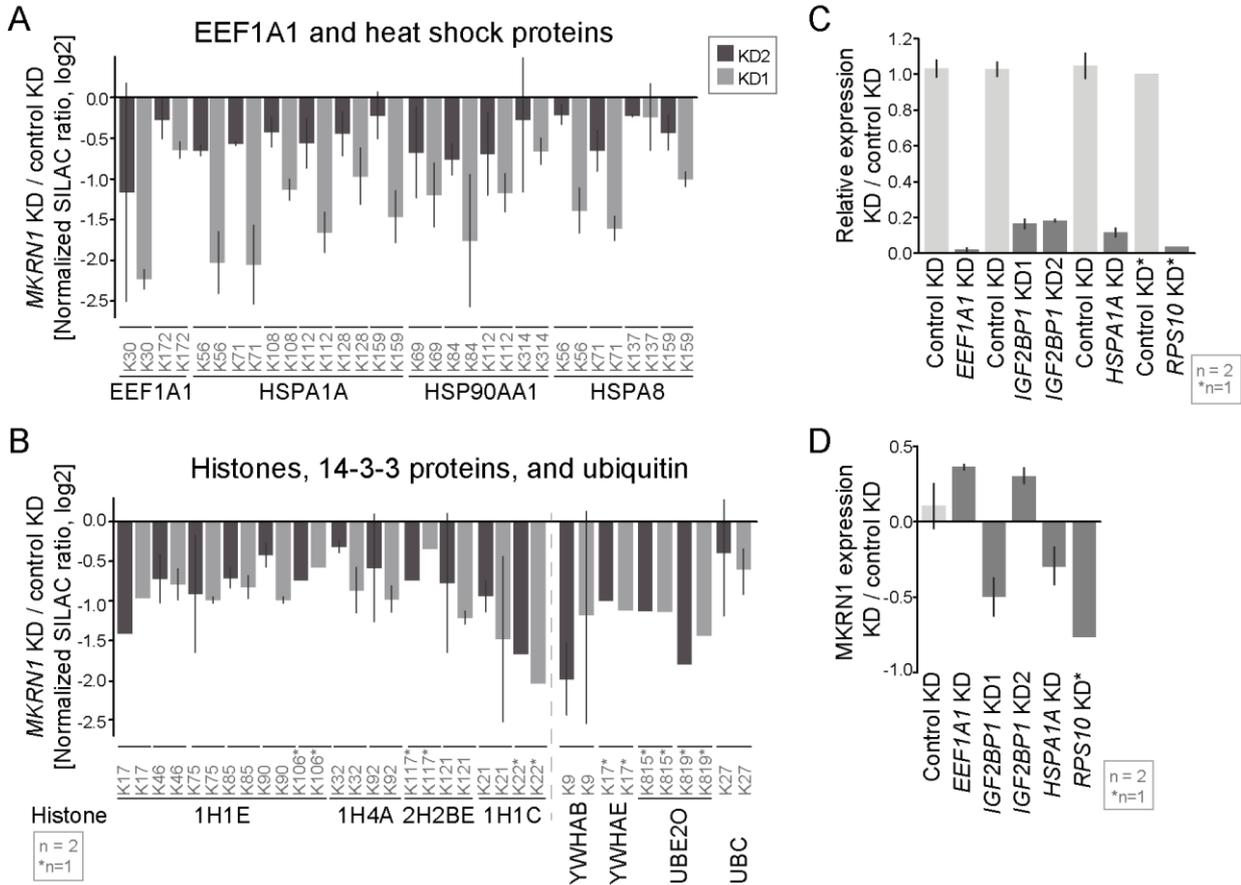


Figure V-7, MKRN1 ubiquitylates regulators of RNA stability and translation. **A & B**, Ubiquitin remnant profiling results for *MKRN1* KD1 and KD2. Mean and standard deviation of the mean (s.d.m., error bars) are given for ubiquitylation sites detected in two replicates. * indicates sites that were only identified in one replicate (B). Legend in (A) defines the color coding for *MKRN1* KD1 and KD2 for (A) and (B). **C**, Ubiquitylation targets of *MKRN1* could be knocked down efficiently using specific siRNAs. Expression levels were normalized to *ACTB* levels and KD conditions were compared to control KDs. Mean and standard deviation of the mean (s.d.m., error bars) are given for two replicates. * indicate experiments with one replicate. **D**, *MKRN1* levels were not affected by the KD of *EEF1A1*, *IGF2BP1*, or *HSPA1A*. *RPS10* KD resulted in reduced *MKRN1* expression. The effect of the KD of *MKRN1* ubiquitylation substrate on *MKRN1* transcript levels was assessed by qPCR. Expression levels were normalized to *ACTB* levels and KD conditions were compared to control KDs. Mean and standard deviation of the mean (s.d.m., error bars) are given for two replicates. * indicate experiments with one replicate.

2.6.3 MKRN1 ubiquitylation targets can be overexpressed as a GFP-tagged protein

For experiments that would confirm the *MKRN1*-mediated ubiquitylation of substrates, identified putative *MKRN1* ubiquitylation target proteins were GFP-tagged for OE. The ectopical expression of some of these proteins was assayed by Western blot. As illustrated in Figure V-8A, all tested tagged proteins, namely GFP-IGF2BP1 (pL103), GFP-HSPA1A (pL104), GFP-EEF1A1 (pL106), GFP-HIST1H1C (pL105), GFP-

HSP90AA1 (pL129), and GFP-LARP1 (pL127), were expressed in HEK293T cells and could be detected by Western blot. The expression of GFP-YWHAB (pL125), GFP-RPS20 (pL128), and GFP-HSPA8 (pL126) still remains to be assessed.

In order to assess the role of RPS10 K107 ubiquitylation in RQC, the lysines K107, K138, and K139 were mutated to arginine. Moreover, the RPS10 encoded on a non-tag vector was rendered insensitive towards RPS10 siRNA (sR79). These untagged vectors and GFP-RPS10 were transfected into HEK293T cells and ectopic expression of (mutated) or GFP-tagged RPS10 was assessed by Western Blot (Figure V-8B, green and blue arrow). As ribosomal protein levels are highly regulated, only minor levels of untagged RPS10 and GFP-RPS10 levels could be observed (Figure V-8B). This suggests that extrinsic RPS10 is indeed expressed in HEK293T cells and that excessive protein levels are reduced by K48-ubiquitylation and proteasomal degradation.

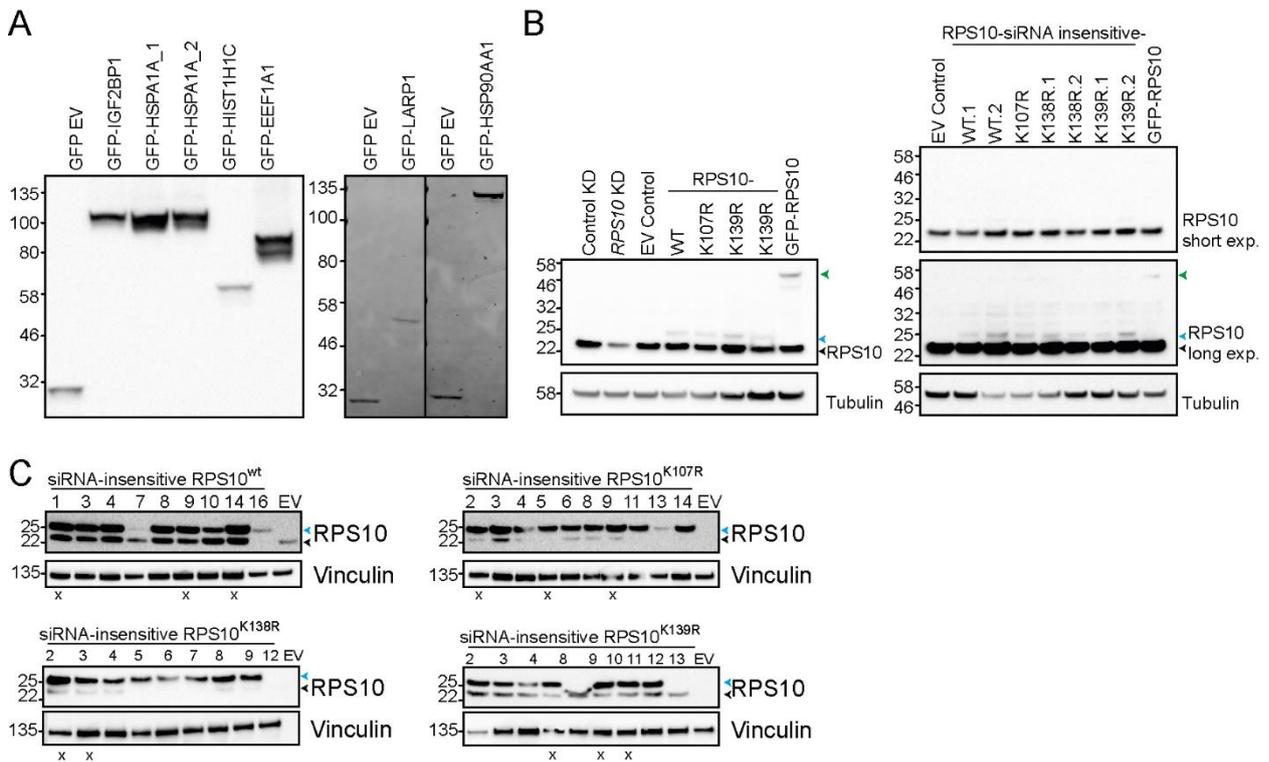


Figure V-8, Ubiquitylation targets of MKRN1 can be ectopically expressed in HEK293T cells. A & B, Potential MKRN1-ubiquitylation substrates were cloned into expression vectors and their expression was assessed by Western Blot 48 h after transfection. **B,** RPS10 expression. RPS10 KD was evaluated after 48 h. Known RPS10 ubiquitylation sites (K107, K138, and K139) were mutated into arginine and RPS10 was rendered insensitive to the used siRNA (right). The expression of RPS10 mutant proteins is shown. The black arrow indicates endogenous RPS10. The blue arrow indicates untagged, ectopically expressed RPS10. The green arrow indicates GFP-RPS10. **C,** Cell lines stably expressing siRNA-insensitive RPS10 and that either encode for the wild type coding sequence of RPS10 (wt) or that harbor lysines mutated to arginines (K107R, K138R, or K139R), are were created. Using siRNA KD of RPS10, cell lines that stably expressed siRNA-insensitive RPS10 were selected. The black arrow indicates endogenous RPS10. The blue arrow indicates non-tagged, ectopically expressed RPS10. Selected cell lines for further experiments are marked with an 'x'.

In addition, the RPS10-coding vectors were utilized to create HEK293T cells that stably express siRNA-insensitive RPS10 with or without mutated ubiquitylation sites (RPS10-WT, -K107R, -K138R, and -K139R), as described for stable MKRN1-expressing cell lines above (see Chapter V.2.2). The single cell derived cell lines were tested for stable expression of siRNA-insensitive RPS10 upon KD of the endogenous *RPS10* mRNA (Figure V-8C). The selected stable cell lines can be used in the flow cytometry reporter assay (Chapter IV, Figure IV-4A) in future experiments.

Moreover, the effect of the siRNA targeting endogenous *RPS10* was assessed on *RPS10* mRNA and protein levels, by qPCR and Western blot, respectively. *RPS10* transcript levels were reduced to 3.4%, and concomitantly, RPS10 protein levels were reduced under KD conditions (Figure V-7C and B, left).

2.7 MKRN1 levels are reduced during stress conditions

To elucidate the physiological function of MKRN1, I assessed the behavior of MKRN1 in cells exposed to stress conditions. Previously, MKRN1 has been reported to accumulate in stress granules upon treatment with sodium arsenite and ER stress in mESCs³⁴. When HEK293T cells were treated with H₂O₂ for 2 h, ectopic MKRN1^{wt}, MKRN1^{RINGmut}, and MKRN1^{ΔPAM2} protein levels were found to be decreased (Figure V-9A). The same was observed for endogenous MKRN1 under oxidative stress (Figure V-9B). As MKRN1 expression was greatly downregulated after 2 h, I explored endogenous MKRN1 protein levels upon H₂O₂ exposure over time. Already after 60 min, MKRN1 depletion could be observed. No further reduction of MKRN1 was detected over a time course of 6 h. Protein levels remained low even after HEK293T cells were allowed to recover for 4 h after 2 h of H₂O₂ treatment (Figure V-9C). To figure out the role of MKRN1 in stress and whether MKRN1 is also found in stress granules after H₂O₂ treatment, as reported for sodium arsenite treatment before, further experiments need to be conducted. During heat shock, global translation is transiently inhibited and HSP70 is involved in regulating translation elongation during heat shock²². As HSP70 might be an ubiquitylation target of MKRN1, I analyzed the behavior of MKRN1 in mild and severe heat shock. Levels of overexpressed MKRN1^{wt}, MKRN1^{RINGmut}, and MKRN1^{ΔPAM2} were reduced after 30 and 60 min of mild heat shock (Figure V-9D, left).

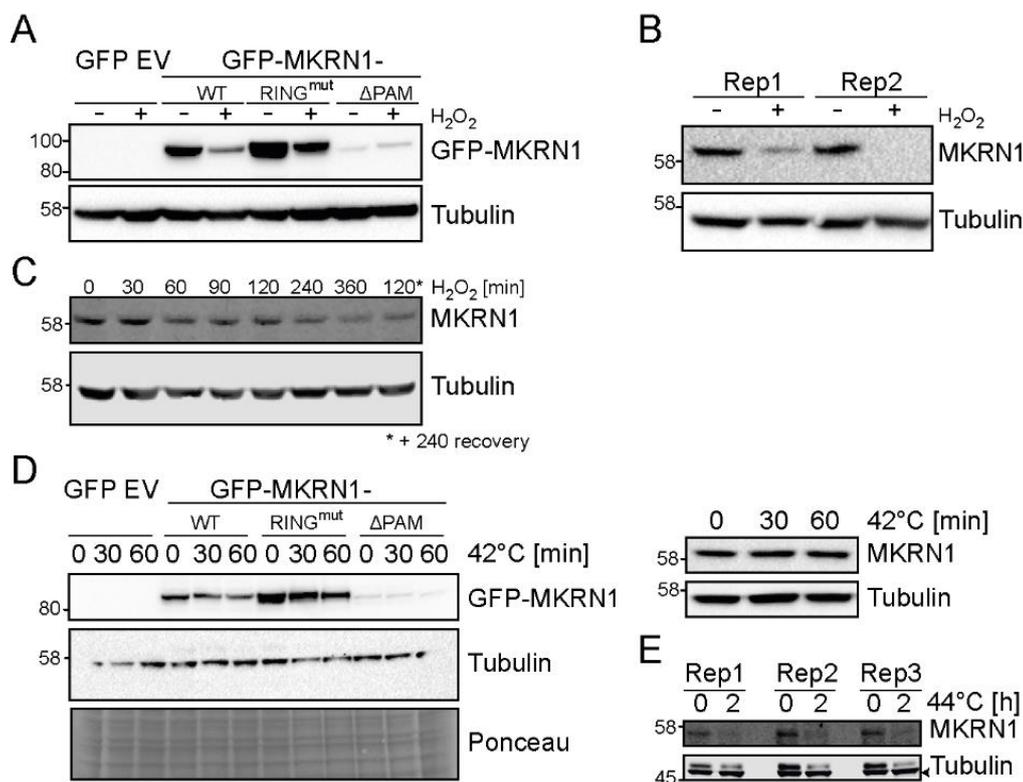


Figure V-9, MKRN1 levels are reduced during stress. **A**, Treatment of HEK293T cells with 2 nM H₂O₂ for 2 h reduced levels of overexpressed MKRN1 (mutant) proteins. Protein levels were assessed by Western Blot using an anti-GFP antibody. **B**, Endogenous MKRN1 levels declined after 2 h of H₂O₂ treatment. Protein levels were assessed by Western Blot using an anti-MKRN1 antibody. **C**, A time course experiment indicated that MKRN1 levels were diminished after 60 min of H₂O₂ treatment. The protein levels did not further drop within 360 min of treatment. After 2 h of H₂O₂ treatment and 4 h recovery time (indicated by *), MKRN1 levels did not recover. **D**, In mild heat shock (42°C), overexpressed MKRN1 levels were downregulated after 30 and 60 min (left). Endogenous MKRN1 protein levels were not altered after 30 or 60 min of mild heat shock (right). **E**, MKRN1 levels were decreased after exposure of HEK293T cells to 2 h of severe heat shock (44°C)²².

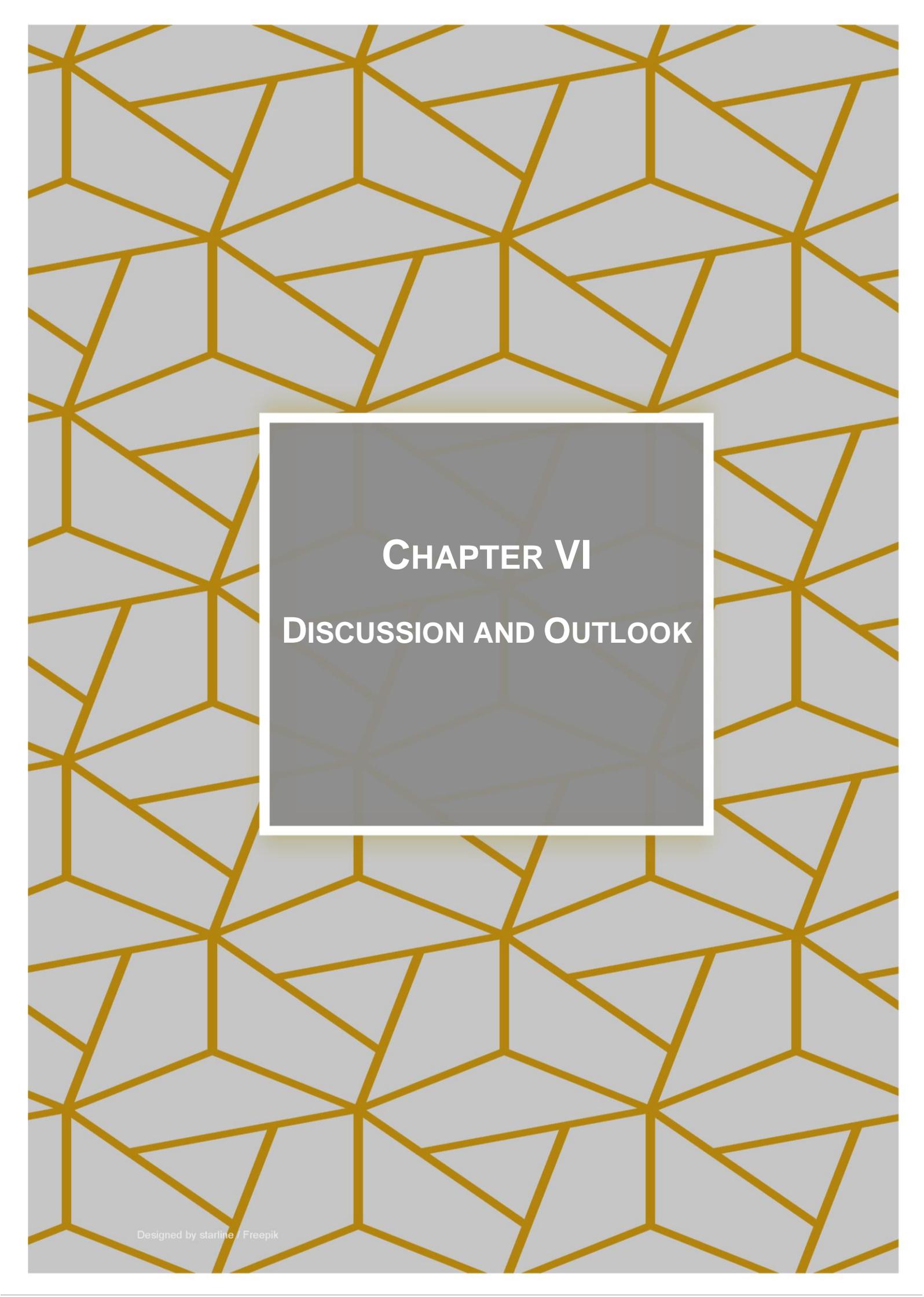
However, endogenous MKRN1 levels were not altered after exposure of HEK293T cells to 42°C for 30 or 60 min (Figure V-9D, right). Of note, endogenous MKRN1 levels were reduced after 2 h of severe heat shock (Figure V-9E). While ectopically expressed MKRN1 protein levels are already affected by 30 min of mild heat shock, endogenous MKRN1 levels only respond to severe heat shock. This phenomenon might be explained by the hypothesis that MKRN1 levels need to be tightly regulated within the cell during stress conditions and that excess MKRN1 is already removed in mild heat shock, while endogenous MKRN1 levels are only controlled upon severe heat shock.

Downregulation of MKRN1 might alter ubiquitylation levels of MKRN1 target proteins, possibly including HSP70. Thus, this level of regulation of HSP70 might be involved in altered HSP70 levels observed in severe heat shock by Shalgi *et al.*²². Whether MKRN1 indeed is involved in the regulation of HSP70 and whether it might be involved in translational regulation upon stress, still remains to be deciphered.

3. References

1. Böhne, A. *et al.* The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* 11, 721 (2010).
2. Gray, T. A., Wilson, A., Fortin, P. J. & Nicholls, R. D. The putatively functional Mkrn1-p1 pseudogene is neither expressed nor imprinted, nor does it regulate its source gene in trans. *Proc. Natl. Acad. Sci.* 103, 12039–12044 (2006).
3. Miroci, H. *et al.* Makorin ring zinc finger protein 1 (MKRN1), a novel poly(A)-binding protein-interacting protein, stimulates translation in nerve cells. *J. Biol. Chem. Biol. Chem.* 287, 1322–1334 (2012).
4. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500, 593–597 (2013).
5. Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* 5, 1704–1713 (2013).
6. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* 44, W272–W276 (2016).
7. Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* 42, 401–407 (2014).
8. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* (80-.). 339, 819–823 (2013).
9. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc.* 8, 2281–2308 (2013).
10. Harrow, J. *et al.* GENCODE: The reference human genome annotation for the ENCODE project. *Genome Res.* 22, 1760–1774 (2012).
11. National Institute of Standards and Technology. External RNA Controls Consortium. (2015).
12. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).
13. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930 (2014).
14. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550 (2014).
15. Lebedeva, S. *et al.* Transcriptome-wide Analysis of Regulatory Interactions of the RNA-Binding Protein HuR. *Mol. Cell* 43, 340–352 (2011).
16. Nielsen, M. L. *et al.* Iodoacetamide-induced artifact mimics ubiquitination in mass spectrometry. *Nat. Methods* 5, 459–460 (2008).
17. Cox, J. *et al.* A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* 4, 698–705 (2009).
18. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–72 (2008).
19. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47 (2015).
20. Kunze, M. M. *et al.* sST2 translation is regulated by FGF2 via an hnRNP A1-mediated IRES-dependent mechanism. *Biochim. Biophys. Acta* 1859, 848–859 (2016).
21. ORFeome Collaboration. The ORFeome Collaboration: A genome-scale human ORF-clone resource. *Nat. Methods* 13, 191–192 (2016).
22. Shalgi, R. *et al.* Widespread Regulation of Translation by Elongation Pausing in Heat Shock. *Mol. Cell* 49, 439–452 (2013).
23. Fellmann, C. *et al.* Functional Identification of Optimized RNAi Triggers Using a Massively Parallel Sensor Assay. *Mol. Cell* 41, 733–746 (2011).
24. Kim, J. H. *et al.* Ubiquitin ligase MKRN1 modulates telomere length homeostasis through a proteolysis of hTERT. *Genes Dev.* 19, 776–81 (2005).
25. Omwancha, J. *et al.* Makorin RING finger protein 1 (MKRN1) has negative and positive effects on RNA polymerase II-dependent transcription. *Endocrine* 29, 363–73 (2006).
26. Schwanhäusser, B., Gossen, M., Dittmar, G. & Selbach, M. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* 9, 205–209 (2009).
27. Lampe, S. *et al.* Identification of the TXNIP IRES and characterization of the impact of regulatory IRES trans-acting factors. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1861, 147–157 (2018).
28. Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. The human and mouse replication-dependent histone genes. *Genomics* 80, 487–498 (2002).

29. Negrutskii, B. S. & El'skaya, A. V. Eukaryotic translation elongation factor 1 alpha: structure, expression, functions, and possible role in aminoacyl-tRNA channeling. *Prog Nucleic Acid Res Mol Biol* 60, 47–78 (1998).
30. Beckmann, R. P., Mizzen, L. A. & Welch, W. J. Interaction of HSP 70 with newly synthesized proteins: Implications for protein folding and assembly. *Science* (80-.). 248, 850–854 (1990).
31. Nelson, R. J., Ziegelhoffer, T., Nicolet, C., Werner-Washburne, M. & Craig, E. A. The translation machinery and 70 kd heat shock protein cooperate in protein synthesis. *Cell* 71, 97–105 (1992).
32. Zhang, X. *et al.* Fine-tuning BMP7 signalling in adipogenesis by UBE2O/E2-230K-mediated monoubiquitination of SMAD6. *EMBO J.* 32, 996–1007 (2013).
33. Mashtalir, N. *et al.* Autodeubiquitination protects the tumor suppressor BAP1 from cytoplasmic sequestration mediated by the atypical ubiquitin ligase UBE2O. *Mol. Cell* 54, 392–406 (2014).
34. Cassar, P. A. *et al.* Integrative genomics positions MKRN1 as a novel ribonucleoprotein within the embryonic stem cell gene regulatory network. *EMBO Rep.* 16, 1334–57 (2015).



CHAPTER VI
DISCUSSION AND OUTLOOK

VI. DISCUSSION AND OUTLOOK

Bifunctional RBPs are involved in many cellular processes and have the potential to connect different, seemingly unrelated, pathways. One class of these RBPs are RBULs that are able to ubiquitylate substrate proteins besides their RNA-binding capability. Only little is known about the general biological function of this protein group. Hence, I set out to decipher the role of RBULs in human cells. To do so, I applied several genome- and proteome-wide approaches and complemented them with biochemical assays to get hints on the global role of six RBULs. Later on, I examined the molecular function of one selected RBUL, namely MKRN1.

1. Functional interaction network of RBULs

To elucidate the cellular and molecular function of RBULs, I established the adapted AP protocol. With the help of Prof. Miguel Andrade and Dr. Gregorio Alanís Lobato, we were able to set up an experimental approach that recovers known and new stable interactors of RBULs by combining it with computational evaluation based on GO similarities (Chapter III, Figure III-1b and Figure III-2a). As RBPs tend to commit to non-physiological interactions *in vitro*, our adapted AP approach neglects interactions established only after lysis and helps to stably recover 'true' cellular interactors^{1,2}. As expected, the interaction network of six selected RBULs provided hints towards their biological function. Besides proteins of the core ubiquitylation pathways, e.g. VCP, hardly any overlap of interactors could be observed between the tested bifunctional RBPs. This indicates that each RBUL holds specific cellular roles and that these ligases are not functionally redundant. Furthermore, the interaction profiles of the RBULs highlighted their cellular functions in transcriptional and translational processes. Exemplarily, MEX3B is a protein that seems to bridge ubiquitylation to transcription and post-transcriptional processes. The GO analysis of preys linked MEX3B to transcription (subunits of RNA Pol I, II and III, ZGPAT, and histone-associated proteins). Concerning post-transcriptional control, the GO terms 'regulation of mRNA stability' and 'poly(A) RNA binding' were significantly recovered among MEX3B interactors (Chapter III, Figure III-3 and Figure III-4b). The other RBULs could be linked to protein transport (e.g. ACTR2 and TUBB4A), cell cycle, or DNA damage repair. Some of the interacting proteins (PGM1, HK2, PRKCI, APPL1, GNE, and PYGB) of ARIH2 connected this RBUL to glycolysis and monosaccharide-related pathways, for instance (Chapter III, Figure III-3). In summary, the adapted AP approach allows to recover physiological

interactors of RBPs, whose biophysical behavior might impede the identification of 'true' interactors in conventional AP approaches. Our protein-protein interaction profiling study presented here provides clues on the functional relevance of several uncharacterized RBULs. Additional in-depth studies are necessary to determine the molecular links between ubiquitylation and RNA-related processes. To unravel the molecular function of one protein of this ubiquitin ligase family, the RBUL MKRN1 was chosen for further studies.

2. MKRN1 is involved in detecting aberrant mRNAs, in translational regulation and has an effect on mRNA stability

2.1 MKRN1 is a mRNP member

MKRN1 is a highly conserved RBUL. Several ubiquitylation substrates have been identified but no definite cellular role could be defined for MKRN1 so far. Different activities have been ascribed to this RBUL in several organisms. First, the MKRN1 orthologue *lep-2* helps to regulate developmental switches in *Caenorhabditis elegans* (*C. elegans*)³. Second, MKRN1 has been linked to embryonic patterning and germ cell formation in *Drosophila melanogaster* (referred to as *Drosophila* below), where it positively regulates *osk* translation^{4,5}. Third, in rat neurons, it colocalizes with PABP and is involved in local translation⁶. And fourth, MKRN1 can be found in stress granules and binds to hundreds of mRNAs in mESCs⁷. Lastly, it seems to play a negative role in transcriptional regulation and has a role in tuning the activity of telomerase in human cells⁸⁻¹⁰. To functionally characterize MKRN1 in HEK293T cells, I first performed quantitative MS-based interactome studies of MKRN1. With these, I could confirm the reported interaction of MKRN1 with PABPC1/4^{6,7}. Additionally, I found MKRN1 to interact with proteins that are involved in the regulation of translation and RNA stability (Chapter IV, Figure IV-1A). These findings place MKRN1 within a mRNP.

Besides PABPC1, 14 other proteins (e.g. YBX1, MOV10, or HNRNPD) were found to be conserved MKRN1 interactors between human and murine cells (Chapter IV, Suppl. Fig. IV-2A)⁷. To further determine whether the interactions between MKRN1 and its binding partners are conserved among other species, I compared the human MKRN1 interactome with Mkrn1-bound proteins in *Drosophila* S2R+ cells. Eight protein interactions of Mkrn1 were found to be identical with the human MKRN1 interactome, with PABPC1, IGF2BP1, and LARP1 being among the *Drosophila* Mkrn1 interactors⁴. Interestingly, all eight proteins that interacted with Mkrn1 in *Drosophila* also were also

found to bind to MKRN1 in murine ESCs⁷. As MKRN orthologues are found in *C. elegans* and zebrafish, a next step would be to compare the interactomes MKRN proteins of these organisms to human. Miroci *et al.* identified the MKRN1 PAM2 motif to be responsible for the MKRN1-PABP interaction. When I mutated this motif, most MKRN1 interactors did not bind to MKRN1 anymore (Chapter IV, Figure IV-1D). A reason for this could be that most proteins interact with MKRN1 via the PAM2 motif and lose binding affinity upon PAM2 mutation. Furthermore, PABP might bring along its binding partners (e.g. LARP1 or IGF2BP1) and thus bridge their interaction to MKRN1. Consequently, MKRN1^{PAM2mut} would shed the ‘secondary’ interactors by being deprived of the PABP-interaction. Lastly, another option why most interactors do not bind to MKRN1^{PAM2mut} is that the three introduced point mutations alter the structure of the protein, thereby rendering the protein-binding site within MKRN1 inaccessible. To elucidate the details of MKRN1 mRNP structure and formation, as well as the role of the MKRN1 PAM2 motif in mediating interactions with many RBPs, further experiments need to be conducted.

IGF2BP1 and ELAVL1 are members of different RBP families. While ELAVL1 is ubiquitously expressed, its paralogues ELAVL3 and ELAVL4 display tissue-specific expression. The same specificity in expression patterns can be found for the IGF2BP family, where IGF2BP1 and IGF2BP3 expression are specific, whereas IGF2BP2 is omnipresent¹¹. MKRN1 also has up to four orthologous proteins and the expression of different *Mkrn* transcript versions varies for example in mouse tissues and during development. Additionally, there are high transcript levels of *MKRN/Mkrn* in early embryogenesis of human and mouse, for instance^{12–14}. As I found MKRN1 to interact with IGF2BP1, IGF2BP2, and IGF2BP3, it could be important to analyze the MKRN1 protein interactomes in a tissue-specific manner. In summary, MKRN1 seems to be a component of an mRNP together with many other RBPs, such as PABP1, LARP1, and IGF2BP1. This mRNP composition seems to be conserved in several organisms. Functionally, these mRNPs might be involved in the control of translation or mRNA stability, possibly in different developmental stages or tissues.

2.2 Poly(A) tails and 3' UTRs of mRNAs are bound by MKRN1

Accompanying to the identification of the MKRN1-bound proteins, I determined the RNA interactome of MKRN1 by iCLIP. We found poly(A) sequences enriched downstream of the MKRN1 binding sites (Chapter IV, Figure IV-2, Figure IV-3, and Suppl. Fig. IV-4). But other than that no further binding motif of MKRN1 could be determined. It is a

possibility that MKRN1 itself recognizes poly(A) stretches but as we find A nucleotides accumulating only after the cross link site, we consider the following scenario more likely. As MKRN1 binds to PABP, which binds to poly(A) sequences, PABP might recruit MKRN1 to mRNA-binding sites via their interaction. This could suggest that MKRN1 itself does not recognize mRNA in a sequence-specific manner but that the specificity is brought about by its interactors.

RRMs are formed of a surface of four β -strands backed by two α -helices. The surface of the β -sheet recognizes four nucleotides by stacking interactions, electrostatics, and hydrogen bonding. Opposed to this, proteins harboring C₃H ZNF domains harbor little secondary structure. The RNA is contacted by forming hydrogen bonds between the RNA bases and aromatic side chains, which form hydrophobic binding pockets. In other words, nucleotide bases are bound by stacking interactions with aromatic residues of the protein backbone. Hence, the protein backbone, whose structure is established by the protein shape, determines the sequence specificity of RNA recognition¹⁵. Thus, specific RNA recognition by the CCCH ZNF-containing MKRN1 may be supported by PTMs on MKRN1, which might have an influence on the protein conformation. Moreover, interacting proteins that specially form the protein shape of MKRN1 might help the RBUL to recognize particular mRNAs. As mentioned above, PABPC1 might position MKRN1 on the mRNA. We found that 30% of mRNA binding sites of MKRN1 can be found in front of poly(A) segments within 3' UTRs or in front of poly(A) tails (Chapter IV, Figure IV-2 and Figure IV-3). As PABPC1 binds to poly(A) tails, it could be the case that MKRN1 is placed in front of poly(A) tails with the help of PABPC1. By binding to or in front of the poly(A) tail, MKRN1 could regulate mRNA stability and translation in cooperation with other RBPs.

The hypothesis that MKRN1 is positioned on the mRNA by PABPC1 is supported by the observation that MKRN1^{PAM2mut}, which loses the interaction with PABPC1 shows decreased RNA-binding behavior. RBPs often cooperate for RNA contacts¹⁶. Hence, MKRN1 might also bind to mRNA within the context of its mRNP. A hint that this might be true results from the observation that in *Drosophila*, *pAbp* depletion results in reduced binding of Mkrn1 to *osk* mRNA⁴. To elucidate the question of collaborative mRNA binding of the mRNP components, one could compare mRNA-binding site profiles of MKRN1 and its binding partners. Overlaps of side-by-side binding sites might indicate whether the binding of MKRN1 to mRNA on specific sites is mediated by the specific RBP. For example, computationally comparing RNA-binding sites of MKRN1

with those of IGF2BP1, YBX1, or PABPC1 could be a first step to validate this hypothesis¹⁷⁻²¹. iCLIP experiments in cells that are devoid of specific MKRN1 interactors (e.g. IGF2BP1, YBX1, or PABPC1/4) are a possibility to find out whether the MKRN1 mRNA-binding specificity is mediated by its binders. In addition, MKRN1 *in vitro* iCLIP experiments, to which interactor proteins would be added singularly or as a pool, could be performed²².

From the MKRN1 iCLIP data, we could extract 1734 binding sites that are followed by a poly(A)-stretch. These binding sites represent 30% of all identified binding sites and they display a higher binding site strength compared to the binding sites that are not preceding a poly(A) tract (4219 binding sites) (Chapter IV, Suppl. Fig. IV-4C). Computational analyses are needed to examine the similarities and properties of the 'non-poly(A) mRNA binding sites' of MKRN1. By removing the 'poly(A) binding sites' from the set, it might be possible to retrieve a binding motif, which might be masked by the presence of the strong poly(A) binding sites. In addition, determining the protein structure of MKRN1 by cryogenic electron microscopy or crystallization experiments could provide insights into the mechanisms of MKRN1 binding to mRNA and proteins. Taken together, we could elucidate that approximately one third of mRNA binding sites of MKRN1 can be found in front of poly(A) segments. I suggest that the attachment to or the recognition of mRNA elements, especially the poly(A) tail, is supported by PABPC1.

2.3 MKRN1 mediates ribosome stalling

We found that MKRN1 binds in front of poly(A) tracts within the 3' UTR and, importantly, poly(A) tails. We propose that within the RQC pathway, PABPC1 helps MKRN1 to bind upstream of (premature) poly(A) tails. Actively translating ribosomes are stalled by MKRN1, which poses a roadblock to these ribosomes and thereby inhibits the translation of the poly(A) tail. Upon contact with the translating ribosome, MKRN1 ubiquitylates the RPS10, a member of the 40S ribosomal SU. Ubiquitylation of RPS10 leads to ribosome stalling. The trailing ribosomes collide with the first stalled ribosome and are then recognized and ubiquitylated by ZNF598, a process that promotes RQC.

Premature poly(A) tails arises due to premature cleavage and polyadenylation, which results in a truncated, but polyadenylated, mRNA without a stop codon²³⁻²⁵. Besides several other RBPs, U1 snRNP plays a role in regulation of APA and the usage of internal PASs. To avoid premature cleavage and polyadenylation, U1 snRNP blocks the recognition of internal PASs by binding to the canonical 5' splice site²⁵. To avoid the

production of aberrant proteins, these prematurely polyadenylated mRNAs are co-translationally recognized and degraded²⁴.

Ribosomes might translate into the poly(A) tail of these mRNAs due to a lack of stop codons. While these mRNAs are usually subject to NGD and NSD, occasionally proteins are produced from non-stop mRNAs. We discovered that 84% of the MKRN1 mRNA-binding sites are situated in 3' UTRs (Chapter IV, Figure IV-2A). This imposes the question for the rationale of this specific mRNA-binding behavior of MKRN1. Interestingly, defective ribosome recycling in yeast increases the abundance of ribosomes, some of them actively translating, at the stop codon and in 3' UTRs. Rli1 is involved in ribosome recycling as it helps to clear ribosomes that accumulate at the stop codon and can reinitiate translation in 3' UTRs. Dom34 rescues the ribosomes that are present within 3' UTRs or translate into the poly(A) tail by mistake as they escaped ribosome recycling²⁶. Consequently, it is a possibility that MKRN1 sits within 3' UTRs and in front of poly(A) tails to stop the ribosomes that got lost during translation termination or that overlooked a stop codon. In this way, MKRN1 could aid proteins, such as Dom34 in yeast, in recognizing and recycling translating ribosomes within 3' UTRs.

The translation of poly(A) tails can be detrimental to the cell. For example, poly-lysine sequences represent nuclear localization signals. The translation of poly(A) stretches resulting in a poly-lysine sequence might thus affect the localization or stability of the respective protein^{27,28}. Unspecific poly(A) translation should therefore be avoided. Hence, as we propose that MKRN1 poses a roadblock for translating ribosomes in front of poly(A) tails, it helps to avoid the production of aberrant poly-lysine tagged proteins that could be mislocated and thus have detrimental effects on the cell²⁹.

In conclusion, in case of prematurely polyadenylated mRNAs or stop codon read-through, MKRN1 might be involved in the rescue of ribosomes that are falsely translating within 3' UTRs or the poly(A) tail.

2.4 MKRN1 is involved in RQC and ubiquitylates RPS10

2.4.1 The role of MKRN1 in RQC

MKRN1 is involved in ribosome stalling, which initiates RQC activation (Chapter IV, Figure IV-4C). By performing ubiquitin remnant profiling, I found that MKRN1 ubiquitylates RPS10 (Chapter IV, Figure IV-5). As this protein is also an ubiquitylation target of ZNF598, it was interesting to further investigate the interplay between MKRN1

and ZNF598 in ribosome stalling. The finding that *MKRN1* and *ZNF598* KD do not have an additive effect on ribosome stalling supports the notion that the two proteins act within the same pathway (Chapter IV, Figure IV-4C and Suppl. Fig. IV-6A). Consequently, *MKRN1* and *ZNF598* could be related or be functional homologous protein. As we do not observe additive effects upon knocking down *ZNF598* and *MKRN1* together, the latter option can be excluded, due to the fact that functional homologous proteins would be redundant and the KD of proteins would show an additive effect. Hel2 and ZNF598 share partial homology with approximately 15% sequence identity and fulfill similar cellular roles in yeast and human, respectively²⁷. ZNF598 and *MKRN1* are both RING E3 ligases but their ZNF RBDs are different. While, *MKRN1* possesses C₃H ZNFs, ZNF598 has five C₂H₂ domains (Figure VI-1). Hence, the proteins do not share a similar domain structure. In line with this, both proteins only share 6.6% protein sequence identity (<https://www.uniprot.org/align/>)³⁰.

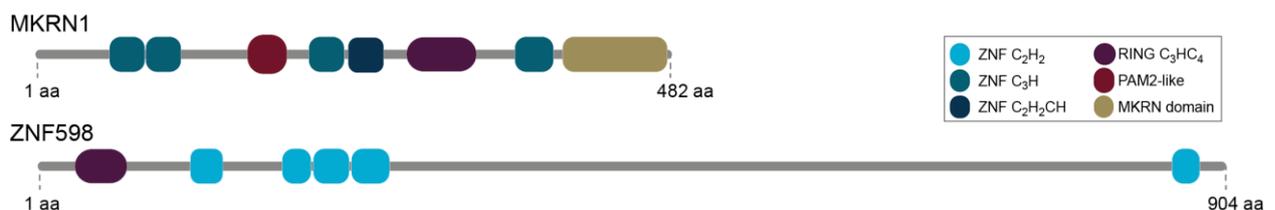


Figure VI-1, Domain comparison of *MKRN1* and *ZNF598*. *MKRN1* possesses four C₃H ZNF domains for RNA binding and an unusual C₂H₂CH ZNF domain. The RING C₃HC₄ domain harbors the ubiquitylation function. *MKRN1* interacts with PABP proteins via the PAM2 motif. A *MKRN1* specific domain can be found in the C' term. ZNF598 also possesses a RING C₃HC₄ domain, which is responsible for its ubiquitylation activity. Moreover, ZNF598 carries five C₂H₂ ZNF domains. Protein and domain sizes are depicted to scale.

As no additive effect was observed upon the combined KD of *MKRN1* and *ZNF598* and as the proteins share little protein sequence identity, the proteins seem to exhibit a nonredundant function. After the first ribosome is stalled at poly(A) stretches, trailing ribosomes will hit the initially stalled ribosome. It was recently shown that ZNF598 recognizes the interface formed by two or more collided ribosomes and thus is not involved in ribosome stalling per se. Upon recognition of the newly formed ribosomal surface, ZNF598 ubiquitylates RPS10 of the collided but not the initially stalled ribosome³¹. In agreement, the collision interface of two ribosomes is recognized by Hel2, which subsequently ubiquitylates its ribosomal targets in yeast^{32,33}. Moreover, it was proposed that Asc1/RACK1 steadies the ribosomal interface and stabilizes or fosters Hel2/ZNF598 binding to it^{31,33,34}. Upon binding the ribosomal interface, Hel2 and ZNF598 ubiquitylate their ribosomal protein targets (RPS20/RPS3 and RPS10, respectively) and, with this, trigger the RQC and NGD/NSD responses (see

Chapters I.1.6.1 and I.1.6.2)^{31–33,35,36}. The recognition of ribosomal clashes by Hel2/ZNF598 could occur after several types of ribosomal stalls, e.g. at mRNA ends or at secondary structures³¹. While this could pose a more global recognition mode to induce RQC, MKRN1-mediated stalling could be specialized for stalled ribosomes at poly(A) stretches. RPS10 ubiquitylation by Hel2 induces NGD but not NSD^{33,34}. Prematurely polyadenylated mRNAs are degraded by NSD²⁴. On the basis of these findings, one could hypothesize that RPS20 ubiquitylation by Hel2 triggers NGD, while RPS10 ubiquitylation, mediated by ZNF598 and MKRN1, induces NSD. In human, two E3 ligases would be responsible for activating NSD and RQC in this case, which might allow for a more delicate signaling response or provide two checkpoints that would need to be passed before the RQC and NSD responses are mounted. In line, Not4-mediated monoubiquitylation of RPS7A is topped up by ubiquitylation of Hel2 in yeast³³. This is seen as backup mechanism of mRNA QC and could mirror fine-tuning of signaling.

Including MKRN1, there are four human ubiquitin ligases involved in RQC signaling (ZNF598, LTN1, and CNOT4)³⁷. The reason for involving several E3 ligases in this pathway is that efficiency, speed, and specificity of the QC response can be improved. The adding of particular ubiquitin chains to target proteins in a timely defined manner can trigger specific and rapid responses, such as the degradation of nascent polypeptides by the proteasome in a VCP-dependent manner^{37–39}. The molecular mechanisms and effects of ribosomal ubiquitylation are largely elusive. The same is true for the identity of ribosome ubiquitylating and deubiquitylating enzymes. By ubiquitylating RPS10, MKRN1 joins the list of ribosome-ubiquitylating ligases^{37,38,40,41}.

Of note, all four RQC-related E3 ligases are RING ligases. The chain specificity of this protein family is mediated by the respective E2 ligase⁴². Thus, it would be compelling to elucidate the identity of the E2 ligases involved in RQC signaling. For example, the formation of K48 ubiquitin chains by the Ube2D family and LTN1 is fostered by TCF25 (Rqc1 in yeast), which also constrains the assembly of other types of ubiquitin chains⁴³. Additionally, ZNF598 has been shown to collaborate with UBE2D3 in ribosome stalling⁴⁴. In our case, especially the MKRN1-interacting E2 ligases would be of interest. UBAITs (ubiquitin-activated interaction traps) could be used to trap transient interaction partners, such as E2 ligases, of MKRN1⁴⁵.

2.4.2 MKRN1 ubiquitylates lysine K107 of RPS10

To initiate RQC, ZNF598 ubiquitylates the lysine residues K138 and K139 of RPS10 (Chapter IV, Figure IV-5E)^{27,44,46}. I found that MKRN1 mainly ubiquitylates K107 on

RPS10. The ubiquitylation on the residues K138, K139, K38, and K53 on RPS10 also slightly, albeit not significantly, decreased in abundance upon *MKRN1* KD2 (Chapter IV, Figure IV-5D). The RPS10 ubiquitylation sites of ZNF598 (K138 and K139) are located in close proximity to K107, which is ubiquitylated by MKRN1. All three lysines are localized within the unstructured C' terminus of the protein, which has not been structurally determined yet^{47,48}. The local concentration of ubiquitylated lysines linked to RQC responses indicates a central role for the C' terminus of RPS10 in signaling responses. Juszkievicz and Hegde suggest that the exact localization of ubiquitylation on either RPS10 or RPS20 is not essential. They claim that any ubiquitylation event in this part of the 40S ribosomal SU would block translation²⁷. While ZNF598 only ubiquitylates RPS10 K139/139 of collided ribosomes, MKRN1 administers ubiquitin to RPS10 K107 of initially stalled ribosomes³¹. This suggests that ubiquitylation at different lysines on ribosomal proteins might have different effects.

To test the effect of MKRN1 on stalling ribosomes by ubiquitylation of RPS10, *MKRN1* deprived cells could be treated with hygromycin B, which leads to increased read-through of stop codons and impairs translational accuracy. If MKRN1 is involved in clearing ribosomes from stop codon read-throughs, *MKRN1* KD/KO cells will be more sensitive to this drug compared to other, not read-through-mediating but translation elongation drugs, such as cycloheximide or anisomycin⁴⁹. Moreover, one could trawl through ribosome profiling data from HEK293T cells and analyze whether ribosomes accumulate before MKRN1 3' UTR binding sites upstream of poly(A) segments identified by iCLIP. Assuming that MKRN1 might leave a footprint of 10 to 20 nt upstream of poly(A) sequences, a peak for ribosomes would be expected 5' to this imprint.

Ubiquitylation of ribosomal proteins by MKRN1 might be a means to stall ribosomes and initiate ribosome recycling. By mutating RPS10 K107 into arginine, it will be possible to test whether ribosome stalling indeed depends on this ubiquitylation event by using the RQC reporter assay. To validate the involvement of the ubiquitylation function of MKRN1 in RQC, rescue experiments with MKRN1^{RINGmut} protein in *MKRN1* KD/KO cells could be performed. In terms of RQC initiation, I propose a model in which PABPC1/4 and maybe other interaction partners position MKRN1 in front of premature poly(A) tails, where MKRN1 poses a roadblock to ribosomes and stalls these by ubiquitylation of K107. Moreover, I suggest that MKRN1 is positioned at the beginning of the signaling cascade that triggers RQC upon ribosome stalling. The notion that MKRN1 sits

upstream of poly(A) stretches and stalls ribosomes, helps to understand how poly(A) sequences are initially recognized. In this aspect it would be interesting to assess the importance of the MKRN1-PABPC1/4 interaction in RQC and to test whether other MKRN1 interactors can mediate the same outcome by positioning MKRN1.

Ribosomes release the fully translated polypeptides only in collaboration with chaperones, which are involved in protein folding, processing, and localization. To prevent misfolding or degradation of nascent chains, that might occur due to translational errors, the emerging peptides are bound by heat shock proteins (e.g. HSP70). This protein family assures correct folding of 'entwining-challenged proteins' by ATP hydrolysis. Hence, ribosome-associated chaperones are part of a complex co-translational protein QC system that cells maintain to avoid toxic protein aggregations⁵⁰⁻⁵⁵. Moreover, mRNA-bound protein domains have been identified in HSP70 and HSP90 proteins, suggesting that HSPs can bind to mRNA⁵⁶. An interpretation of this scenario is that HSP70 is bound to the nascent peptide and is in contact with the translating ribosome and maybe the mRNA. It is possible that MKRN1 ubiquitylates also other proteins within the translational hub in addition to the identified targets HSP70, EEF1A1, and mRNP components.

To summarize, MKRN1 is involved in ribosome stalling. RPS10 ubiquitylation has been implied in ribosome stalling previously^{27,44,46}. I propose that by ubiquitylating K107 on RPS10, MKRN1 stalls elongating ribosomes, which leads to ribosomal collision. In a subsequent step, ZNF598 would then reinforce the signaling cascade elicited by ribosome stalling by ubiquitylating K138/139 on RPS10 within collided ribosomes, resulting in RQC initiation³¹.

2.5 The role of MKRN1 in RQC might be conserved in other organisms

The RQC process is highly similar between yeast and human. As MKRN1 is also a highly conserved protein, I investigated whether the role of MKRN1 in protein QC pathways might also be conserved in different organisms. Hints for a conserved involvement of MKRN1 in RNA surveillance and protein QC mechanisms come even from *Arabidopsis thaliana* (*Arabidopsis*). In these plants, NGD is conserved and only occurs on poly(A) stretches, with a preference for long poly(A) tracts⁵⁷. Interestingly, MKRN orthologues are found in *Arabidopsis* as well (uniprot identifier / gene name: Q6IDS6 / MKRN and E0X9N4 / At3g63550). In *C. elegans*, the NSD pathway is present and so is the MKRN1 orthologue *lep-2*^{3,58}. It would be interesting to examine whether

there is a poly(A)-triggered RQC pathway in *C. elegans* and *Arabidopsis* and whether MKRN orthologous proteins play a role in NSD, RQC, or ribosome ubiquitylation in these eukaryotes. Also in *Drosophila*, *Mkrn1* is conserved and similarly to human MKRN1, binds in front of a poly(A) tract within the *osk* 3' UTR in S2R+ cells⁴. As the RBUL is found in fruit flies and seems to exhibit a similar mRNA-binding behavior, one could use an adapted RQC fluorescence reporter assay to determine whether *Mkrn1* exhibits the same functionality in *Drosophila* and human cells. Moreover, it would be interesting to determine the ubiquitylation targets of *Mkrn1* in these animals.

As mentioned above, the RQC pathway is highly conserved in yeast and human. However, differences can be observed in sequences leading to ribosome stalling. Ribosome stalling in yeast is mediated by poly-basic amino acids rather than by a poly(A) sequence as observed in human cells^{28,36,59–62}. Of note, a homologous protein of MKRN1 is missing in baker's yeast. The proteins Yth1p and Lee1p in yeast are similar to MKRN1 but are deprived of the RING domain and the PAM2 motif (Chapter IV) and were hence not considered as MKRN1 functional orthologues in *Saccharomyces cerevisiae* (*S. cerevisiae*). In contrast, a homologous protein of human ZNF598 is found in yeast (Hel2) but the two proteins exhibit different ubiquitylation preferences. ZNF598 ubiquitylates the C' terminus of RPS10, whereas Hel2 ubiquitylates RPS20 and RPS3. The Tollervey lab explains the different ubiquitylation substrates between ZNF598 and Hel2 by the conservation of the respective target proteins RPS10 and RPS20. In *S. cerevisiae* the C' terminus of RPS10 is absent and hence the protein does not possess the ubiquitylation targets of human ZNF598. Winz *et al.* argue that Hel2 ubiquitylates RPS20, while ZNF598 has evolved to ubiquitylate the C' terminus of RPS10, which is absent in yeast. Moreover, they point out that RPS10 of *Schizosaccharomyces pombe* (*S. pombe*) contains lysines that resemble the human ZNF598 ubiquitylation sites, which are ubiquitylated by Hel2 in *S. pombe* (Figure VI-2)³⁴.

```

P46783|RS10_HUMAN  MLMPKKNRIAIYELLFKEGVMVAKKDVHMPKHPPELADKNVPNLHVMKAMQSLKSRGYVKE  60
Q08745|RS10A_YEAST MLMPKEDRNKIHQYLFQEGVVAKKDFNQAKHEEID---TKNLYVIKALQSLTSKGYVKT  57
O14112|RS10A_SCHPO MLIPKENRKAIHQALFSQGVLVAKKDFNLPKHPEVG---VPNLQVIKACQSLDSRGYLKT  57
      **::**::*  *:: **::**::*****:  ** *:  .  ** *::**  ***  *::**::*

P46783|RS10_HUMAN  QFAWRHFFYWYLTNEGIQYLRDYLHLPPEIVPATLRRSRPETGRPREKSGLEGERPARLTRG  120
Q08745|RS10A_YEAST QFSWQYQYTYTLTEEGVEYLREYLNLPPEHIVPGTYIQERNPTQRPPQRY-----  105
O14112|RS10A_SCHPO RYNWGWFFYYTLTNEGVEYLREYLNLPPEVVPATHKRQVRPTAPRAGRPEPRE-----RA  111
      :: *  *:: **::**::*****:***:***  .:***.*  :.  *  :

P46783|RS10_HUMAN  EADRDTYRRSAVPPGAIKKAEAGAGSATEFQFRGGFGRGRGQPPQ  165
Q08745|RS10A_YEAST -----  -----
O14112|RS10A_SCHPO SAD-AGYRR-----AFKKDEGAAPSGFAPSGFRGGFGRPQ-----  144

```

hK107
hK138/K139

Figure VI-2, The C' term of hRPS10 is not conserved between *Homo sapiens*, *S. cerevisiae*, and *S. pombe*. RPS10 protein sequences of *Homo sapiens*, *S. cerevisiae*, and *S. pombe* were aligned and compared. Human K107 is substituted by arginine in *S. cerevisiae* and *S. pombe* (indicated in red). The C' terminus of human RPS10 is not present in *S. cerevisiae*. RPS10 of *S. pombe* possesses a C' term domain, which, however, does not resemble the C' terminus of human RPS10. K138 and K139, found in human RPS10 are also found in *S. pombe*.

Notably, K107, the major RPS10 ubiquitylation target of MKRN1 is interchanged with an arginine in *S. cerevisiae* and *S. pombe* (Figure VI-2). We searched for MKRN orthologues in *S. pombe* (using HaMStR-OneSeq) but only found one potential MKRN orthologue (Cps3). However, this protein does not harbor a RING finger domain and was subsequently neglected. We therefore concluded that MKRN1 is not conserved in *S. pombe*. While RPS10 K138 and K139, the ubiquitylation sites of ZNF598, are conserved in *S. pombe*, the human RPS10 K107, the ubiquitylation substrate of MKRN1, is replaced by an arginine as it is the case in yeast (Chapter IV, Figure VI-2). With the loss of MKRN1, no conservation pressure exists for its ubiquitylation substrate lysine RPS10 K107. This is why this lysine might not be conserved in *S. cerevisiae* and *S. pombe*. As the stalling mechanism in yeast is different from human, MKRN1 might not be needed to recognize poly(A) sequences. Hence, while the RQC pathway is highly conserved between yeast and human, the signaling to initiate it might differ^{28,37,38,62}.

To summarize, the RQC and NSD/NGD pathways are conserved in human, animals, plants, and fungi. Similarly, MKRN proteins are found in many, however not all, of the organisms that maintain those surveillance pathways. In *S. cerevisiae* and *S. pombe*, the protein is not found, possibly due to another mode of recognition of aberrant mRNAs. In contrast, MKRN1 could have a functional role in QC pathways in *Arabidopsis*, *C. elegans*, and *Drosophila*. Nevertheless, this hypothesis would need to be assessed experimentally.

2.6 MKRN1 ubiquitylates regulators of mRNA stability and translation

Besides RPS10, I found several MKRN1 interactors, namely PABPC1/4, IGF2BP proteins, ELAVL1, LARP1, MOV10, and SYNCRIP to be putatively ubiquitylated by MKRN1 (Chapter IV, Figure IV-5A - C). The ubiquitylation by MKRN1 does not seem to mediate degradation of those proteins, as their protein levels were not affected by *MKRN1* KD (Chapter IV, Suppl. Fig. 8B). This is in line with the finding that MKRN1 interactors are probably not destabilized by MKRN1 in mESCs⁷. Of note, MKRN1 autoubiquitylation triggers MKRN1 degradation⁹, indicating that MKRN1 is capable of

mediating the addition of different ubiquitin modules (monoubiquitylation or specific chains) with the help of several E2 ligases⁴². To identify the ubiquitin linkage specificity of MKRN1, the ubiquitylation chain topology could be identified by using chain-specific antibodies for Western blot analysis in *MKRN1* KD/KO conditions. Another option is the use of affimers or TUBEs (tandem-repeated ubiquitin-binding entities) to elucidate the ubiquitylation chain specificity^{63,64}.

The functionality of the ubiquitylation of MKRN1 interactors still remains unclear. I propose that, when MKRN1 ubiquitylates RPS10 and stalls the ribosome, RBPs within the MKRN1 mRNP, like PABP and IGF2BP proteins, are simultaneously ubiquitylated. However, I cannot exclude that the ubiquitylation of these proteins by MKRN1 might take place in a completely unrelated context. The ubiquitylation of MKRN1-interactors might entail conformational changes or alterations in the RNA-binding behavior of the respective RBPs. A possible outcome would be the displacement of the MKRN1 mRNP from the bound mRNA, making room at the ribosomal stall site for the binding of ZNF598 and ribosome recycling proteins (Pelota, HBS1, and ABCE1). Diverse PTMs (e.g. phosphorylation, acetylation, or methylation) have previously been described to change the protein conformation, which can result in reduced RNA-binding ability or altered protein-protein interactions. For example, methylation of FMRP influences its ability to form dimers with FXR1P and has an effect on the RNA-binding behavior of FMRP⁶⁵. In addition, phosphorylation of p40^{AUF1}, PCBP1, or TTP reduced their affinity for mRNA^{66–70}. On a side note, depending on its phosphorylation, plant PABP seems to exhibit different affinities towards poly(A) RNA as well as to its protein interactors⁷¹. These outcomes might also be the case for ubiquitylation. IGF2BP and PABP protein ubiquitylation sites lie within their RBDs or within the linkers connecting RBDs (Chapter IV, Figure IV-5E). This modification may reduce RNA-binding capacity of those proteins and release them from mRNAs. Conformational changes induced by ubiquitylation have previously been reported. For instance, the conformation of DIO2 is reshaped upon ubiquitylation, which inhibits its enzymatic activity⁷². Furthermore, ubiquitylation can destabilize the structure of proteins, of FKBP12 or FABP4 for example, and result in altered protein conformations⁷³. Of note, RBDs, for example RRM, pose a hub for PTMs^{56,74}. Flexible linkers between RBDs can aid RNA-binding by fostering specific conformational arrangements of RBDs. This is the case for IGF2BP1, KSRP, or PTB, for instance, and has also been suggested for PABP^{75–78}. Hence, ubiquitylation of the RBD or the interdomain linker region, as I observed for PABP and IGF2BP proteins here (Chapter IV, Figure IV-5E), might influence protein

conformation and eventually have an effect on the conformation of the RBP or its RNA-binding capability.

MKRN1 ubiquitylates proteins that are involved in regulating mRNA stability (e.g. ELAVL1 or MOV10) and translation (e.g. IGF2BP1, PABPC1, or LARP1). The functional relevance of these ubiquitylation events, however, remains enigmatic. The trigger of MKRN1 ubiquitylation is unknown as well. It might be the case that mRNA binding induces the ubiquitylation activity of MKRN1. All four single MKRN1^{ZNFmut} proteins showed high levels of ubiquitylation, which could be a result of autoubiquitylation, and are still capable of mRNA binding (Chapter V, Figure V-1E)⁹. To investigate whether mRNA binding influences the MKRN1 ubiquitylation ability, as it has been shown for the RBUL TRIM25 for example, all four ZNF domains of MKRN1 could be mutated simultaneously⁷⁹. Moreover, the ubiquitylation ability of MKRN1^{PAM2mut} could also be explored, as this protein displays reduced RNA binding capability (Chapter IV). In addition, *in vitro* ubiquitylation assays with recombinant MKRN1 proteins and its recombinant targets could be performed in the absence and presence of RNA to elucidate whether there is a dependency of MKRN1 on RNA-binding for ubiquitylation. Another option is that the ubiquitylation activity of MKRN1 is induced by the interaction with other proteins, such as PABPC1 or RPS10. Lastly, MKRN1 could be permanently catalytically active, without any stimulation, and ubiquitylate proteins in close proximity. In this case, regulation of ubiquitylation could be provided by the respective E2. In conclusion, I propose that by ubiquitylation of its mRNP components and proteins within the translational complex, MKRN1 might not only stall ribosomes but also dissolve the assembly of proteins involved in translational regulation in 3' UTRs or at (premature) poly(A) tails.

2.7 MKRN1 might regulate translation via the 3' UTR or the poly(A) tail

With the help of a luciferase reporter assay, I found that MKRN1 might have a general effect on the translation of polyadenylated mRNAs but is not involved in translation initiation (Chapter V). The majority of the MKRN1 mRNA-binding sites are situated within 3' UTRs, where also many other RBPs bind to regulate translation. Thus MKRN1 might have an effect on translation elongation or termination. This hypothesis is backed up by the finding that it interacts with proteins that repress, regulate, or terminate translation (GSPT1, YBX1, YBX3, LARP1, LARP4B, AGO2, MOV10, PABPC1, and PABPC4; Chapter IV).

PCBP1 binds to TGF-beta-activated translational elements (BAT elements) in mRNA 3' UTRs where it forms a complex with EEF1A1, which is in contact with the elongating ribosome. Hussey *et al.* suggest that the interaction between EEF1A1, bound to the A site of the translating ribosome, and PCBP1, bound to the BAT element in the 3' UTR, bridges across the circularized mRNA to inhibit translation elongation by trapping EEF1A1 in the ribosomal A site. When PCBP1 is phosphorylated, it is released from mRNA and translation can proceed^{68,69}. While MKRN1 does neither interact with nor ubiquitylate PCBP1, it might ubiquitylate EEF1A1. When considering how this ubiquitylation is taking place, one could imagine that MKRN1 sits in the 3' UTR and contacts EEF1A1 in a way similar to PCBP1. Thus, MKRN1 could facilitate ubiquitylation of EEF1A1 and other translational regulators by bridging the mRNA loop structure to regulate translation elongation.

As MKRN1 interacts with PABPC1/4 and is found upstream of poly(A) tails, it could also play a role in translation termination. PABPC1 has already been linked to translation termination. It helps to summon ERF1 and ERF3 to the translating ribosomes and fosters weak translation termination when being overexpressed^{80–82}. Schuller and Green suggest that besides PABP, other RBPs, such as PTBP1, might be involved in this process²⁹. One of them might be MKRN1. To assess whether MKRN1 plays a general role in translation, ribosome profiling experiments could be performed. By tethering MKRN1 to specific locations within the mRNA or to mRNAs with short and long 3' UTRs, one could test translation efficiencies with the luciferase reporter assay. Another hint towards a function of MKRN1 in translational control comes from the finding that Mkrn1 regulates *osk* mRNA localization and translation in *Drosophila* ovaries. By binding 5' to A-rich regions within the 3' UTR of *osk mRNA* and by outcompeting Bru, Mkrn1 fosters *osk* translation⁴.

The ribosome constitution can be described as heterogeneous depending on its associated factors. Among the ribosomal interactors, several kinases and ubiquitin ligases have been found, suggesting that PTM of ribosomes adds another layer of translational regulation. Besides being phosphorylated, ribosomal proteins have been shown to be ubiquitylated^{41,83}. I found MKRN1 to interact with several ribosomal proteins, which is in agreement with previous findings⁸³. Thus, MKRN1 seems to be one of the ribosome associated ubiquitin ligases and might ubiquitylate ribosomal proteins also on occasions other than RQC initiation. Ribosome ubiquitylation can act as a means for rapid responses to external stimuli. For example, Rpl28 is ubiquitylated in a

cell cycle dependent manner in yeast⁸⁴. In addition, ubiquitylation of RPS2 and RPS3 has been described upon endoplasmic reticulum (ER) stress, rendering the cell more prone to cell death⁸⁵. Monoubiquitylation is involved in protein-protein interactions and might help to accumulate the correct regulatory proteins on the ribosomes to control translation elongation and termination^{41,86}. Taken together, I found that MKRN1 might have an effect on translation elongation or termination. One possibility is that MKRN1 influences the mRNP composition on mRNAs 3' UTRs or poly(A) tails by competing or collaborating with other RBPs. The other prospect would be that MKRN1 regulates translation via ribosome ubiquitylation. However, the exact mechanism of how MKRN1 might influence translation remains elusive.

2.8 MKRN1 levels affect mRNA abundances

I found that upon MKRN1 OE there are slight, albeit global, changes in mRNA abundance. On the other hand, MKRN1 did not seem to affect mRNA stability directly as indicated by initial experiments (Chapter V, Figure V-4A and B). On the other hand, MKRN1 interacts with regulators of RNA stability (e.g. UPF1, ELAVL1, or PUM1). Hence, it is not clear whether MKRN1 has an effect on mRNA stability. The molecular mechanism on how MKRN1 would affect mRNA levels, however, is unknown as well. A new canonical mRNA decay mechanism, namely ribothrypsis, which is mechanistically related to NGD/NSD, has recently been discovered. In ribothrypsis, transient ribosome pausing results in endonucleolytic cleavage 5' to the ribosome⁸⁷. I hereby hypothesize that MKRN1 might be involved in this decay pathway. In line, MKRN1 interacts with SYNCRIP and HNRNPD, two proteins that have been proposed to play a role in translation-coupled RNA decay^{88,89}. A possible function of MKRN1 might be that it fosters transient ribosome pauses and intensifies ribosome stalling by ribosome ubiquitylation. The stalls would then trigger mRNA cleavage. This might explain reduced protein levels and slightly lower RNA abundances in MKRN1 OE, as RNA cleavage in ribothrypsis might simultaneously reduce translation and RNA levels. As the effect of MKRN1 on ribosome stalling has only been observed on long poly(A) stretches, I speculate that MKRN1 efficiency or activity might depend on the respective interaction partner or its current mRNP composition. To examine a putative role of MKRN1 in ribothrypsis, Akron-Seq experiments could be performed in MKRN1-depleted cells⁸⁷. To summarize, I found MKRN1 to interact with regulators of RNA stability (Chapter IV, Figure IV-1A). Moreover, I found mRNA levels to be altered upon manipulating MKRN1

expression. Nonetheless, how mRNA levels are influenced by MKRN1 on a molecular level still needs to be elucidated.

2.9 MKRN1 levels are downregulated in response to stress

To delineate the physiological role of MKRN1, I started to assess the behavior of this protein during oxidative stress and heat shock. I found that endogenous MKRN1 levels are reduced during oxidative stress and severe heat shock (Chapter V, Figure V-9). During environmental stress, MKRN1 localizes to stress granules (SGs) in mESCs⁷. But the role of MKRN1 in stress conditions still needs to be determined. As my results suggest that MKRN1 is associated with the translational machinery, it might be the case that it is involved in the transit of the translating polysome to SGs during stress. The interaction of MKRN1 and ATXNL2, a protein involved in regulation of stress granules, supports this assumption⁹⁰. O-linked N-acetyl glucosamine (O-GlcNAc)-modified ribosomal proteins and RACK1 are present in SGs upon stress. Ohn *et al.* suggest that upon arsenite treatment, translational components are modified with O-GlcNAc, which leads to polysome disassembly and the recruitment of untranslated mRNPs into SGs⁹¹. MKRN1 might be a part of these SG-resident mRNPs, which could be the reason why it is detected in SGs upon sodium arsenite treatment in mESCs⁷. In addition, HSP70 was shown to regulate translational pausing during severe heat shock. Moreover, translational pausing can be induced by inhibiting HSP70 activity⁹². MKRN1 seems to ubiquitylate HSP1A1 within the catalytic ATPase domain (K56, K71, K108, K112, and K128; Chapter V, Figure V-7A)⁹³. It remains to be tested whether ubiquitylation of the ATPase domain of HSPA1A inhibits its enzymatic activity. In case this would be proven to be correct, MKRN1 might support translational pausing during severe heat shock by ubiquitylation of HSP70. It is still unknown, however, whether the ligase function of MKRN1 is active during heat shock. Taken together, MKRN1 levels are downregulated in response to stress. Whether MKRN1 plays an active role in the stress response or whether it is tethered to SGs for storage, remains to be worked out. MKRN1 is down regulated in response to different kinds of stress. Moreover, the protein is localized to SGs upon oxidative stress in mESCs⁷. Whether MKRN1 holds an active role during stress responses or whether it is passively localized to SGs together with the translational machinery still needs to be determined.

2.10 Why MKRN1 is important

In RQC initiation MKRN1 seems to pose a roadblock for ribosomes translating in 3' UTRs or premature poly(A) tails. With this, the production of proteins bearing polylysine tracts would be avoided. Especially neurons are very sensitive to disturbances and the avoidance of malfunctioning proteostasis is vital in these cells⁹⁴. Peptides containing polylysine or polybasic sequences, e.g. nonstop proteins, are prone to agglomeration and are target to QC pathways. As MKRN1 is involved in ribosome stalling at poly(A) tracts, it probably supports the QC pathways in avoiding the translation of aberrant mRNAs and the production of aggregation-prone proteins. Mutations in or loss of several QC components have been implicated in protein aggregation and malignancies^{94–96}. In yeast, depletion of *ltn1* results in the formation of protein aggregates and inclusions that trigger chronic proteotoxic stress⁹⁵. Furthermore, a loss-of-function mutation in *Ltn1* causes severe motor dysfunction and neurodegeneration in mice⁹⁷. As MKRN1 is involved in a QC mechanism, mutations in MKRN1 might also be involved in disease development. In *Drosophila*, depletion of *Mkrn1* was found to negatively affect oogenesis, germ cell specification, and embryonic patterning. Moreover, it has been reported to modulate developmental timing and sexual maturation^{4,5,98}. In line, male fertility is impaired in *Mkrn2* deficient mice due to abnormal spermatogenesis⁹⁹. In *Xenopus laevis*, *Mkrn2* KD is embryonically lethal. When the KD is performed at a later developmental stage, neurogenesis and axis formation are battered¹⁰⁰. In summary, mutations in MKRN proteins might indeed pose threats to germ cell and embryonic development as well as to neurogenesis.

2.11 Summary

In my PhD thesis, I could set up an experimental pipeline to identify stable interaction partners of RNA-binding ubiquitin ligases. After benchmarking the setup on PRPF19, I determined the interaction profiles of selected RBULs. With this approach, I found novel protein-protein interaction partners of six human RBULs. The interaction network provided hints towards new cellular functions of those RBULs, linking ubiquitylation to transcription and post-transcriptional pathways. Furthermore, I characterized the molecular function of the RBUL MKRN1 in detail. I found that MKRN1 interacts with PABPC1 and other RBPs, which suggested that MKRN1 is a member of an mRNP. In addition, the RBUL MKRN1 binds to mRNA poly(A) tails and at internal A-stretches of 3'UTRs. MKRN1 binds to mRNA with more than one ZNF domain and is possibly recruited to its binding sites by PABPC1 and/or other interaction partners. Binding sites of MKRN1 directly upstream of poly(A) stretches also include poly(A) tails of prematurely polyadenylated mRNAs. This finding hinted towards a function of MKRN1 in recognizing aberrant mRNAs. Indeed, I found that MKRN1 plays a role in the human RQC pathway by stalling ribosomes at poly(A) stretches. Moreover, MKRN1 ubiquitylates RPS10, which might result in ribosome stalling. Additionally, the mRNA binding sites of MKRN1 and its interactome suggested that the RBUL plays a role in regulating mRNA levels or translation. Indeed, in preliminary experiments I saw an effect of altered MKRN1 abundance on newly synthesized proteins and mRNA levels. Moreover, I could show that MKRN1 ubiquitylates translational regulators. However, the exact role of MKRN1 in the regulation of translational and mRNA still needs to be elucidated. In summary, I could establish MKRN1 as a new player in the co-translational recognition of aberrant mRNAs, where it stalls ribosomes. I propose the mechanistic model that MKRN1 is positioned on mRNAs with the help of PABP. In this way, prematurely polyadenylated mRNAs can be detected in a co-translational manner with the help of MKRN1 posing a roadblock to ribosomes. By ubiquitylation of RPS10 and other components of the ribosomal complex, MKRN1 helps to initiate an RQC response, which in turn helps to avoid the accumulation of toxic protein products and stress responses.

3. References

1. Mili, S. & Steitz, J. A. Evidence for reassociation of RNA-binding proteins after cell lysis: Implications for the interpretation of immunoprecipitation analyses Evidence for reassociation of RNA-binding proteins after cell lysis: Implications for the interpretation of immunopr. *RNA* 10, 1692–1694 (2004).
2. Kittur, N., Darzacq, X., Roy, S., Singer, R. H. & Meier, T. U. Dynamic association and localization of human H/ACA RNP proteins. *RNA* 12, 2057–2062 (2006).
3. Herrera, R. A., Kiontke, K. & Fitch, D. H. A. Makorin ortholog LEP-2 regulates LIN-28 stability to promote the juvenile-to-adult transition in *Caenorhabditis elegans*. *Development* 143, 799–809 (2016).
4. Dold, A. *et al.* Makorin1 controls embryonic patterning by alleviating Bruno-mediated repression of oskar translation. *bioRxiv* (2018). doi:10.1101/501643
5. Liu, N. & Lasko, P. Analysis of RNA Interference Lines Identifies New Functions of Maternally-Expressed Genes Involved in Embryonic Patterning in *Drosophila melanogaster*. *G3 - Genes/Genomes/Genetics* 5, 1025–1034 (2015).
6. Miroci, H. *et al.* Makorin ring zinc finger protein 1 (MKRN1), a novel poly(A)-binding protein-interacting protein, stimulates translation in nerve cells. *J. Biol. Chem. Biol. Chem.* 287, 1322–1334 (2012).
7. Cassar, P. A. *et al.* Integrative genomics positions MKRN1 as a novel ribonucleoprotein within the embryonic stem cell gene regulatory network. *EMBO Rep.* 16, 1334–57 (2015).
8. Omwancha, J. *et al.* Makorin RING finger protein 1 (MKRN1) has negative and positive effects on RNA polymerase II-dependent transcription. *Endocrine* 29, 363–73 (2006).
9. Kim, J. H. *et al.* Ubiquitin ligase MKRN1 modulates telomere length homeostasis through a proteolysis of hTERT. *Genes Dev.* 19, 776–81 (2005).
10. Salvatico, J., Kim, J. H., Chung, I. K. & Muller, M. T. Differentiation linked regulation of telomerase activity by Makorin-1. *Mol. Cell. Biochem.* 342, 241–250 (2010).
11. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15, 829–845 (2014).
12. Böhne, A. *et al.* The vertebrate makorin ubiquitin ligase gene family has been shaped by large-scale duplication and retroposition from an ancestral gonad-specific, maternal-effect gene. *BMC Genomics* 11, 721 (2010).
13. Gray, T. A. *et al.* The ancient source of a distinct gene family encoding proteins featuring RING and C(3)H zinc-finger motifs with abundant expression in developing brain and nervous system. *Genomics* 66, 76–86 (2000).
14. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* 500, 593–597 (2013).
15. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat. Rev. Mol. Cell Biol.* 8, 479–490 (2007).
16. Dassi, E. Handshakes and Fights: The Regulatory Interplay of RNA-Binding Proteins. *Front. Mol. Biosci.* 4, (2017).
17. Hosono, Y. *et al.* Oncogenic Role of THOR, a Conserved Cancer/Testis Long Non-coding RNA. *Cell* 171, 1559-1572.e20 (2017).
18. Kini, H. K., Silverman, I. M., Ji, X., Gregory, B. D. & Liebhaber, S. A. Cytoplasmic poly(A) binding protein-1 binds to genomically encoded sequences within mammalian mRNAs. *RNA* 22, 61–74 (2016).
19. Conway, A. E. *et al.* Enhanced CLIP Uncovers IMP Protein-RNA Targets in Human Pluripotent Stem Cells Important for Cell Adhesion and Survival. *Cell Rep.* 15, 666–679 (2016).
20. Hinze, F. *et al.* Expanding the map of protein-RNA interaction sites via cell fusion followed by PAR-CLIP. *RNA Biol.* 15, 359–368 (2018).
21. Hwang, H. W. *et al.* cTag-PAPERCLIP Reveals Alternative Polyadenylation Promotes Cell-Type Specific Protein Diversity and Shifts Araf Isoforms with Microglia Activation. *Neuron* 95, 1334-1349.e5 (2017).
22. Sutandy, F. R. *et al.* In vitro iCLIP-based modeling uncovers how the splicing factor U2AF2 relies on regulation by cofactors. *Genome Res.* 28, 699–713 (2018).
23. Oszolak, F. *et al.* Resource Comprehensive Polyadenylation Site Maps in Yeast and Human Reveal Pervasive Alternative Polyadenylation. *Cell* 143, 1018–1029 (2010).
24. Frischmeyer, P. A. *et al.* An mRNA Surveillance Mechanism That Eliminates Transcripts Lacking Termination. *Science (80-.)*. 295, 2258–2262 (2002).
25. Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* 468, 664–668 (2010).

26. Young, D. J., Guydosh, N. R., Zhang, F., Hinnebusch, A. G. & Green, R. Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. *Cell* 162, 872–884 (2015).
27. Juszkiwicz, S. & Hegde, R. S. Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol. Cell* 65, 743-750.e4 (2017).
28. Ito-Harashima, S., Kuroha, K., Tatematsu, T. & Inada, T. Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev.* 21, 519–524 (2007).
29. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* 19, 526–541 (2018).
30. Pundir, S., Martin, M. J., O'Donovan, C. & The UniProt Consortium. UniProt Tools. *Curr. Protoc. Bioinforma.* 53, 29.1–15 (2016).
31. Juszkiwicz, S. *et al.* ZNF598 Is a Quality Control Sensor of Collided Ribosomes. *Mol. Cell* 72, 469-481.e7 (2018).
32. Simms, C. L., Yan, L. L. & Zaher, H. S. Ribosome Collision Is Critical for Quality Control during No-Go Decay. *Mol. Cell* 68, 361-373.e5 (2017).
33. Ikeuchi, K. *et al.* Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J.* e100276 (2019). doi:10.15252/embj.2018100276
34. Winz, M.-L., Peil, L., Turowski, T. W., Rappsilber, J. & Tollervey, D. Molecular interactions between Hel2 and RNA supporting ribosome-associated quality control. *Nat. Commun.* 10, 563 (2019).
35. Saito, K., Horikawa, W. & Ito, K. Inhibiting K63 Polyubiquitination Abolishes No-Go Type Stalled Translation Surveillance in *Saccharomyces cerevisiae*. *PLoS Genet.* 11, 1–18 (2015).
36. Matsuo, Y. *et al.* Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat. Commun.* 8, 159 (2017).
37. Joazeiro, C. A. P. Ribosomal Stalling During Translation: Providing Substrates for Ribosome-Associated Protein Quality Control. *Annu Rev Cell Dev Biol.* 33, 343–368 (2017).
38. Brandman, O. & Hegde, R. S. Ribosome-associated protein quality control. *Nat Struct Mol Biol.* 23, 7–15 (2016).
39. Oh, E., Akopian, D. & Rape, M. Principles of Ubiquitin- Dependent Signaling. *Annu. Rev. Cell Dev. Biol.* 34, 18.1-18.26 (2018).
40. Dimitrova, L. N., Kuroha, K., Tatematsu, T. & Inada, T. Nascent Peptide-dependent Translation Arrest Leads to Not4p-mediated Protein Degradation by the Proteasome. *J. Biol. Chem.* 284, 10343–10352 (2009).
41. Simsek, D. & Barna, M. An emerging role for the ribosome as a nexus for post-translational modifications. *Curr. Opin. Cell Biol.* 45, 92–101 (2017).
42. Deshaies, R. J. & Joazeiro, C. A. P. RING Domain E3 Ubiquitin Ligases. *Annu Rev Biochem.* 78, 399–434 (2009).
43. Kuroha, K., Zinoviev, A., Hellen, C. U. T. & Pestova, T. V. Release of Ubiquitinated and Non-ubiquitinated Nascent Chains from Stalled Mammalian Ribosomal Complexes by ANKZF1 and Pth1. *Mol. Cell* 72, 1–17 (2018).
44. Garzia, A. *et al.* The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat Commun.* 8, (2017).
45. O'Connor, H. F. *et al.* Ubiquitin-Activated Interaction Traps (UBAITs) identify E3 ligase binding partners. *EMBO Reportseports* 16, 1699–1712 (2015).
46. Sundaramoorthy, E. *et al.* ZNF598 and RACK1 Regulate Mammalian Ribosome-Associated Quality Control Function by Mediating Regulatory 40S Ribosomal Ubiquitylation. *Mol. Cell* 65, 751-760.e4 (2017).
47. Anger, A. M. *et al.* Structures of the human and *Drosophila* 80S ribosome. *Nature* 497, 80–85 (2013).
48. Khatler, H., Myasnikov, A. G., Natchiar, S. K. & Klaholz, B. P. Structure of the human 80S ribosome. *Nature* 520, 640–645 (2015).
49. Bengtson, M. H. & Joazeiro, C. A. P. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467, 470–473 (2010).
50. Paknia, E., Chari, A., Stark, H. & Fischer, U. The Ribosome Cooperates with the Assembly Chaperone pICln to Initiate Formation of snRNPs. *Cell Rep.* 16, 3103–3112 (2016).
51. Inada, T. The Ribosome as a Platform for mRNA and Nascent Polypeptide Quality Control. *Trends Biochem. Sci.* 42, 5–15 (2017).
52. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332 (2011).
53. Pechmann, S., Willmund, F. & Frydman, J. The Ribosome as a Hub for Protein Quality Control. *Molecular Cell* 49, 411–421 (2013).

54. Beckmann, R. P., Mizzen, L. A. & Welch, W. J. Interaction of HSP 70 with newly synthesized proteins: Implications for protein folding and assembly. *Science* (80-). 248, 850–854 (1990).
55. Nelson, R. J., Ziegelhoffer, T., Nicolet, C., Werner-Washburne, M. & Craig, E. A. The translation machinery and 70 kd heat shock protein cooperate in protein synthesis. *Cell* 71, 97–105 (1992).
56. Castello, A. *et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol. Cell* 63, 696–710 (2016).
57. Szádeczky-kardoss, I., Gál, L., Auber, A., Taller, J. & Silhavy, D. Plant Science The No-go decay system degrades plant mRNAs that contain a long A-stretch in the coding region. *Plant Sci.* 275, 19–27 (2018).
58. Arribere, J. A. & Fire, A. Z. Nonsense mRNA suppression via nonstop decay. *Elife* 7, 1–23 (2018).
59. Sitron, C. S., Park, J. H. & Brandman, O. Asc1, Hel2, and Slh1 couple translation arrest to nascent chain degradation. *RNA* 23, 798–810 (2017).
60. Brandman, O. *et al.* A Ribosome-Bound Quality Control Complex Triggers Degradation of Nascent Peptides and Signals Translation Stress. *Cell* 151, 1042–1054 (2012).
61. Kuroha, K. *et al.* Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Rep.* 11, 956–961 (2010).
62. Letzring, D. P., Wolf, A. S., Brule, C. E. & Grayhack, E. J. Translation of CGA codon repeats in yeast involves quality control components and ribosomal protein L1. *RNA* 19, 1208–1217 (2013).
63. Heidelberger, J. B., Wagner, S. A. & Beli, P. Mass Spectrometry-Based Proteomics for Investigating DNA Damage-Associated Protein Ubiquitylation. *Front Genet.* 7, (2016).
64. Hjerpe, R. *et al.* Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Rep.* 10, 1250–8 (2009).
65. Dolzhanskaya, N., Merz, G., Aletta, J. M. & Denman, R. B. Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP. *J. Cell Sci.* 119, 1933–1946 (2006).
66. Wilson, G. M. *et al.* Phosphorylation of p40AUF1 regulates binding to A + U-rich mRNA-destabilizing elements and protein-induced changes in ribonucleoprotein structure. *J. Biol. Chem.* 278, 33039–33048 (2003).
67. Carballo, E. *et al.* Decreased Sensitivity of Tristetraprolin-deficient Cells to p38 Inhibitors Suggests the Involvement of Tristetraprolin in the p38 Signaling Pathway. *J. Biol. Chem.* 276, 42580–42587 (2001).
68. Hussey, G. S. *et al.* Identification of an mRNP Complex Regulating Tumorigenesis at the Translational Elongation Step. *Mol. Cell* 41, 419–431 (2011).
69. Chaudhury, A. *et al.* TGF- β -mediated phosphorylation of hnRNP E1 induces EMT via transcript-selective translational induction of Dab2 and ILE1. *Nat. Cell Biol.* 12, 286–293 (2010).
70. García-Mauriño, S. M. *et al.* RNA Binding Protein Regulation and Cross-Talk in the Control of AU-rich mRNA Fate. *Front. Mol. Biosci.* 4, (2017).
71. Le, H., Browning, K. S. & Gallie, D. R. The phosphorylation state of poly(A)-binding protein specifies its binding to poly(A) RNA and its interaction with eukaryotic initiation factor (eIF) 4F, eIFiso4F, and eIF4B. *J. Biol. Chem.* 275, 17452–17462 (2000).
72. Sagar, G. D. V. *et al.* Ubiquitination-Induced Conformational Change within the Deiodinase Dimer Is a Switch Regulating Enzyme Activity. *Mol. Cell Biol.* 27, 4774–4783 (2007).
73. Morimoto, D., Walinda, E., Fukada, H., Sugase, K. & Shirakawa, M. Ubiquitylation Directly Induces Fold Destabilization of Proteins. *Sci. Rep.* 6, (2016).
74. Sloutsky, R. & Naegle, K. M. Proteome-Level Analysis Indicates Global Mechanisms for Post-Translational Regulation of RRM Domains. *J. Mol. Biol.* 430, 41–44 (2018).
75. Dagil, R. *et al.* IMP1 KH1 and KH2 domains create a structural platform with unique RNA recognition and re-modelling properties. *Nucleic Acids Res.* (2019). doi:10.1093/nar/gkz136
76. Järvelin, A. I., Noerenberg, M., Davis, I. & Castello, A. The new (dis)order in RNA regulation. *Cell Commun. Signal.* 14, (2016).
77. Díaz-Moreno, I. *et al.* Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. *Nucleic Acids Res.* 38, 5193–5205 (2010).
78. Oberstrass, F. C. *et al.* Structure of PTB Bound to RNA: Specific Binding and Implications for Splicing Regulation Downloaded from. *Science* (80-). 309, 2054–2057 (2005).
79. Choudhury, N. R. *et al.* RNA-binding activity of TRIM25 is mediated by its PRY/SPRY domain and is required for ubiquitination. *BMC Biol.* 15, (2017).
80. Ivanov, A. *et al.* PABP enhances release factor recruitment and stop codon recognition during translation termination. *Nucleic Acids Res.* 44, 7766–7776 (2016).
81. Ivanov, P. V., Gehring, N. H., Kunz, J. B., Hentze, M. W. & Kulozik, A. E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO J.* 27, 736–747 (2008).
82. Cosson, B. *et al.* Poly (A) -Binding Protein Acts in Translation Termination via Eukaryotic Release Factor 3 Interaction and Does Not Influence [PSI +] Propagation Poly (A) -Binding

- Protein Acts in Translation Termination via Eukaryotic Release Factor 3 Interaction an. *Mol. Cell. Biol. Biol.* 22, 3301–3315 (2002).
83. Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* 169, 1051–1057 (2017).
 84. Spence, J. *et al.* Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* 102, 67–76 (2000).
 85. Higgins, R. *et al.* The Unfolded Protein Response Triggers Site-Specific Regulatory Ubiquitylation of 40S Ribosomal Proteins. *Mol. Cell* 59, 35–49 (2015).
 86. Komander, D. & Rape, M. The Ubiquitin Code. *Annu. Rev. Biochem.* 81, 203–229 (2012).
 87. Ibrahim, F., Maragkakis, M., Alexiou, P. & Mourelatos, Z. Ribothrypsis, a novel process of canonical mRNA decay, mediates ribosome-phased mRNA endonucleolysis. *Nat. Struct. Mol. Biol.* 25, 302–310 (2018).
 88. Grosset, C. *et al.* A mechanism for translationally coupled mRNA turnover: Interaction between the poly(A) tail and a c-fos RNA coding determinant via a protein complex. *Cell* 103, 29–40 (2000).
 89. Lee, K.-H. *et al.* AUF1 contributes to Cryptochrome1 mRNA degradation and rhythmic translation. *Nucleic Acids Res.* 42, 3590–3606 (2014).
 90. Kaehler, C. *et al.* Ataxin-2-Like Is a Regulator of Stress Granules and Processing Bodies. *PLoS One* 7, e50134 (2012).
 91. Ohn, T., Kedersha, N., Hickman, T., Tisdale, S. & Anderson, P. A functional RNAi screen links O-GlcNAc modification of ribosomal proteins to stress granule and processing body assembly. *Nat. Cell Biol.* 10, 1224–1231 (2008).
 92. Shalgi, R. *et al.* Widespread Regulation of Translation by Elongation Pausing in Heat Shock. *Mol. Cell* 49, 439–452 (2013).
 93. Flaherty, K. M., DeLuca-Flaherty, C. & McKay, D. B. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* 346, 623–628 (1990).
 94. Shao, S. & Hegde, R. S. Target Selection during Protein Quality Control. *Trends Biochem. Sci.* 41, 124–137 (2016).
 95. Choe, Y.-J. *et al.* Failure of RQC machinery causes protein aggregation and proteotoxic stress. *Nature* 531, 191–195 (2016).
 96. Chiti, F. & Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* 75, 333–366 (2006).
 97. Chu, J. *et al.* A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2097–2103 (2009).
 98. Tran, H. T. *et al.* Makorin1 Regulates Developmental Timing in Drosophila. *Mol. Cells* 41, 1024–1032 (2018).
 99. Qian, X. *et al.* Deficiency of Mkrn2 causes abnormal spermiogenesis and spermiation, and impairs male fertility. *Sci. Rep.* 6, 39318 (2016).
 100. Yang, P. H. *et al.* Makorin-2 is a neurogenesis inhibitor downstream of phosphatidylinositol 3-kinase/Akt (PI3K/Akt) signal. *J. Biol. Chem.* 283, 8486–8495 (2008).

VII. APPENDIX

Abbreviations

4SU	4-thiouridine
µm	Micrometer
A	Adenosine
A site	Aminoacyl site
AP	Affinity purification
APA	Alternative polyadenylation
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>
Arg / R	Arginine
A-stretches	Adenosine-rich stretches
ATP	Adenosine triphosphate
BAT	TGF-beta-activated translational element
BP	Biological Process
C' term	carboxy terminal
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CAA	Chloroacetamide
CAT tail	C-terminal alanine- and threonine-containing tails
CC	Cellular Compartment
CF	Cleavage factor
cor	Correlation
CPSF	Cleavage and polyadenylation specificity factor
CstF	Cleavage stimulation factor
CTD	Carboxy terminal domain
Cys / C	Cysteine
diGly	Diglycine
DNA	Desoxyribonucleic acid
dox	Doxycycline
<i>Drosophila</i>	<i>Drosophila melanogaster</i>
DSE	Downstream element
DTT	Dithiothreitol
DUB	Deubiquitylase
E site	Exit site
e.g.	Exemplī grātiā, for example
EEF	Eukaryotic elongation factor
EIF	Eukaryotic initiation factor
EJC	Exon junction complex
ER	Endoplasmic reticulum

Appendix

ERF	Eukaryotic release factor
ESI	Electrospray ionization
EV	Empty vector
FDR	False discovery rate
Fig.	Figure
GFP	Green fluorescent protein
Gly / G	Glycine
GlyGly	Glycine-glycine
GO	Gene ontology
GTP	Guanosine-5'-triphosphate
h	Hours
HCD	Higher-energy collisional dissociation
HECT	Homologous to the E6-AP carboxyl terminus
HEK	Human embryonic kidney
HIPPIE	Human Integrated Protein-Protein Interaction Reference
His / H	Histidine
HSP	Heat shock protein
iCLIP	Individual-nucleotide resolution UV crosslinking and immunoprecipitation
iCLIP	Individual nucleotide resolution UV crosslinking and immunoprecipitation
IDR	Intrinsically disordered regions
IP	Immunoprecipitation
IRES	Internal ribosome entry site
iTRAQ	Isobaric tags for relative and absolute quantification
KD	Knock down
KD	Knockdown
kDa	Kilo Dalton
kDa	Kilo Dalton
KH	K homology
KO	Knock out
KO	Knockout
LC	Liquid chromatography
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
lnc	Long non-coding
Lys / K	Lysine
m ⁷ G	7-methylguanosine
mESCs	Mouse embryonic stem cells
Met / M	Methionine
MF	Molecular Function
min	Minutes
MKRN	Makorin
MKRN1 ^{PAM2mut}	MKRN1 variant with point mutations in the PAM2 motif
MKRN1 ^{RINGmut}	MKRN1 variant with point mutation in RING domain

Appendix

MLLE	Methionine-leucine-leucine-glutamic acid
mRIPA	Modified RIPA
mRNA	Messenger RNA
mRNP	Messenger ribonucleoprotein particle
MS	Mass spectrometry
N' term	amino terminus/ amino terminal
NEM	N-ethylmaleimide
NMD	Nonsense mediated decay
nt	Nucleotide
NTC	Nineteen complex
NTC	Nineteen complex
OE	Overexpression
O-GlcNAc	O-linked N-acetylglucosamine
ORF	Open reading frame
P bodies	Processing bodies
P site	Peptidly site
PABP	Poly(A) binding protein
PAM2	PABP-interacting motif 2
PAP	Poly(A) polymerase
PAS	polyadenylation site
PCV2	Porcine circovirus type 2
PEI	Polyethylenimine
pGL3pro	pGL3promoter
PIC	Preinitiation complex
POI	Protein of interest
Pol	Polymerase
Pro / P	Proline
PTM	Post-translational modification
QC	Quality control
<i>r</i>	Pearson correlation coefficients
RBD	RNA-binding domain
RBP	RNA-binding protein
RBR	Ring between RING
RBUL	RNA-binding ubiquitin ligase
RING	Really interesting new gene
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
RNP	Ribonucleoprotein particle
RPL	Ribosomal protein of the large subunit
rpm	Revolutions per minute
RPS	Ribosomal protein of the small subunit
RQC	Ribosome associated quality control
RRM	RNA recognition motif
rRNA	Ribosomal RNA
RT	Room temperature
S	Svedberg unit

Appendix

<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i> / SCHPO	<i>Schizosaccharomyces pombe</i>
SCX	Strong-cation exchange 32 chromatography
Ser / S	Serine
SGs	Stress granules
shRNA	Small hairpin RNA
SILAC	Stable isotope labelling with amino acids in cell culture
siRNA	Small interfering RNA
SOB	Signal-over-background
SS	Splice site
SU	Subunit
SUMO	Small ubiquitin-related modifier
Suppl.	Supplementary
Tab.	Table
TAP	Tandem affinity purifications
TMT	Tandem mass tag
tRNA	Transfer RNA
TUBEs	Tandem-repeated ubiquitin-binding entities
Tyr / Y	Tyrosine
UBAITS	Ubiquitin-Activated Interaction Traps
UPR	Unfolded protein response
USE	Upstream auxiliary element
UTR	Untranslated region
WI	Wang's Index
wt	Wild type
ZNF	Zinc finger

List of Figures

Figure I-1 , RNA binding proteins often harbor several RBDs that collaboratively mediate RNA binding. _____	15
Figure I-2 , mRNA (alternative) polyadenylation. _____	17
Figure I-3 , Co-translational non-stop decay. _____	23
Figure I-4 , Steps of the ribosome-associated quality control pathway. _____	25
Figure I-5 , Ubiquitylation. _____	27
Figure III-1 , The adapted SILAC-based AP approach efficiently recovers known interactors of PRPF19. _____	70
Figure III-2 , Experiments following the adapted AP approach and GO similarity analyses for six RBULs. _____	77
Figure III-3 , The core interactomes of six RBULs. The core interactomes of the six RBULs ARIH2, MEX3B, MKRN1, MKRN2, RNF17 and PRPF19 were identified using the adapted AP. _____	78
Figure III-4 , The RBULs link posttranscriptional processes to the ubiquitin system. _____	80
Figure IV-1 , MKRN1 interacts with PABP and other regulators of translation and RNA stability. _____	116
Figure IV-2 , MKRN1 binds upstream of A-stretches in 3' UTRs. _____	119
Figure IV-3 , MKRN1 binds at poly(A) tails. _____	121
Figure IV-4 , MKRN1 stalls ribosomes at poly(A) sequences. _____	122
Figure IV-5 , MKRN1 ubiquitylates ribosomal protein RPS10 and translational regulators. _____	125
Figure IV-6 , MKRN1 is a sensor for poly(A) sequences that stalls ribosomes to initiate ribosome-associated quality control. _____	126
Figure V-1 , In depth characterization of MKRN1. _____	186
Figure V-2 , Stable cell lines. _____	191
Figure V-3 , MKRN1 mutants are stably expressed in HEK293T cells. _____	191
Figure V-4 , Changes in MKRN1 levels alter mRNA abundances. _____	192
Figure V-5 , MKRN1 controls protein levels of newly translated proteins. _____	194
Figure V-6 , MKRN1 might regulate translation via the poly(A) tail and does not control translation initiation. _____	198
Figure V-7 , MKRN1 ubiquitylates regulators of RNA stability and translation. _____	200
Figure V-8 , Ubiquitylation targets of MKRN1 can be ectopically expressed in HEK293T cells. _____	201
Figure V-9 , MKRN1 levels are reduced during stress. _____	203
Figure VI-1 , Domain comparison of MKRN1 and ZNF598. _____	213
Figure VI-2 , The C' term of hRPS10 is not conserved between <i>Homo sapiens</i> , <i>S. cerevisiae</i> , and <i>S. pombe</i> . _____	218
Supplementary Figure III-1 , RNase optimization for the conventional SILAC-AP workflow. _____	95
Supplementary Figure III-2 , Conventional SILAC-AP in combination with RNase digest enables differentiation between RNA-dependent and RNA-independent protein interactions of the six studied RBULs. _____	96
Supplementary Figure III-3 , Validation of RBUL interaction partners by pulldowns and Western blot (corresponding to Figure 4c, d and 2d). _____	97
Supplementary Figure III-4 , Images of full membranes and different exposure times for Western blot analyses in Figure 4c, d, Supplementary Figure 3a and b. _____	103
Supplementary Figure IV-1 , Maximum likelihood tree of Makorin orthologs with their protein domain architecture. _____	145

Supplementary Figure IV-2 , MKRN1 interacts with translational regulators and other RBPs. _____	146
Supplementary Figure IV-3 , Signal-over-background transformation allows to estimate MKRN1 binding strength. _____	147
Supplementary Figure IV-4 , MKRN1 binds upstream of long A-stretches. _____	148
Supplementary Figure IV-5 , Interaction with PABP is required for MKRN1 RNA binding. (A,B) UV crosslinking experiments to measure the RNA binding capacity of GFP-MKRN1 ^{wt} and GFP-MKRN1 ^{PAM2mut} . _____	149
Supplementary Figure IV-6 , MKRN1 is required to stall ribosomes at K(AAA) ₂₀ in reporter assays. _____	151
Supplementary Figure IV-7 , Interactome of GFP-MKRN1 ^{RINGmut} reveals putative ubiquitylation substrates. ____	151
Supplementary Figure IV-8 , GO term analysis of MKRN1 ubiquitylation targets and proteome analysis upon <i>MKRN1</i> KD. _____	152
Supplementary Figure IV-9 , Images of full membranes and different exposure times (exp.) for Western blot analyses in Fig. 1C and Supplemental Fig. S2C in the presence or absence of RNase A and T1. _____	158
Supplementary Figure IV-10 , Images of full membranes of autoradiographs and Western blot analyses in Fig. 3D (replicate 1) and Supplemental Fig. S5 (replicates 2 and 3). _____	159
Supplementary Figure IV-11 , Images of full membranes and different exposure (exp.) times for Western blot analyses in Supplemental Fig. S6B,D,E. (A,B) KDs of <i>MKRN1</i> and <i>ZNF598</i> assessed by Western blot (n = 3 replicates) from Supplemental Fig. S6B. _____	161

List of Tables

Table I-3-1 , Possible controls during iCLIP library preparation. _____	37
Table I-3-2 , Troubleshooting for common problems during iCLIP library preparation. _____	41
Table V-1 , Kits, Enzymes, Antibodies used in this study. _____	165
Table V-2 , Oligonucleotides used in this study additionally to Suppl. Table S5 in Chapter IV. _____	167
Table V-3 , Plasmids used in this study. _____	171
Table V-4 , siRNAs used in this study additionally to Suppl. Table S6 in Chapter IV. For MKRN1 siRNA1 (sR16) and siRNA2 (sR17), as well as control siRNA (sR18), see Table S6 in Chapter IV. _____	177
Supplementary Table III-1 , MaxQuant analysis of MS data from the conventional SILAC-AP and the adapted AP for the six RBULS. _____	103
Supplementary Table III-2 , Summary of adapted AP interactome analyses for the six RBULS. _____	103
Supplementary Table III-3 , Comparison of numbers of interaction partners in the conventional SILAC-AP to the adapted AP experiments. _____	104
Supplementary Table III-4 , DAVID GO ontology enrichment for GFP-MEX3B interactors for all three GO domains. _____	104
Supplementary Table III-5 , Known and novel protein-protein interactions identified from the adapted APs of the six RBULS. _____	106
Supplementary Table III-6 , A comparison of the cellular localization of the RBULS. _____	108

Appendix

Supplementary Table IV-1 , provided as an Excel file with the publication	
Supplementary Table IV-2 , Summary of MKRN1 iCLIP experiments. iCLIP experiments with GFP-MKRN1 were performed in three independent replicates.	142
Supplementary Table IV-3 , provided as an Excel file with the publication	
Supplementary Table IV-4 , provided as an Excel file with the publication	
Supplementary Table IV-5 , Primers used in this study.	143
Supplementary Table IV-6 , siRNAs used in this study.	143

