

Making Ends Meet: Role of the LARP1 La-Module in TOP mRNA Recognition

by

Hiba A. Al-Ashtal

B.S. in Biological Sciences, Carnegie Mellon University, 2013

Submitted to the Graduate Faculty of the
Dietrich School of Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Hiba A. Al-Ashtal

It was defended on

March 2, 2020

and approved by

Jeffrey Brodsky, Ph.D., Professor

Paula Grabowski, Ph.D., Professor

Andrew VanDemark, Ph.D., Associate Professor

Jayakrishnan Nandakumar, Ph.D., Associate Professor

Thesis Advisor/Dissertation Director: Andrea Berman, Ph.D., Associate Professor

Copyright © by Hiba A. Al-Ashtal

2020

Making Ends Meet: Role of the LARP1 La-Module in TOP mRNA Recognition

Hiba A. Al-Ashtal, Ph.D.

University of Pittsburgh, 2020

The La-related protein (LARP) superfamily is a diverse family of RNA-binding proteins that are characterized by an N-terminal La-Module. The La-Module is comprised of two RNA-binding domains, a La Motif (LAM) and an RNA recognition motif (RRM), which synergistically engage RNA. Sequence and structural changes to the La-Module lead to unique RNA binding specificity and biological function. For example, the La-Module of Genuine La, the prototypical LARP, binds the 3'UUU-OH of pre-tRNAs to facilitate their folding and maturation. In contrast, the LARP6 La-Module binds a stem-loop within the 5'UTR of collagen mRNAs to regulate their translation. LARP1 is the most divergent member of the LARP superfamily, and has been implicated in the stability and translation of mRNAs encoding the translation machinery.

The RNA binding partners and function of the LARP1 La-Module remain elusive. Because it associates with Poly(A)-binding protein, we hypothesized that the LARP1 La-Module has evolved to bind poly(A) RNA; we found that the La-Module indeed binds poly(A) RNA. Interestingly, we discovered that it simultaneously engages the 5' terminal oligopyrimidine (TOP) motif characteristic of mRNAs that encode the translation machinery. We hypothesize that through binding to features at the 5' and 3' end of TOP mRNAs, the 5' TOP motif and poly(A) tails, the La-Module may aid mRNA circularization to regulate TOP mRNA translation regulation. In addition, multivalent interactions between the La-Module and mRNAs may aid sequestration of TOP mRNAs to stress granules and processing-bodies during translation repression in response to stress.

Table of Contents

Preface.....	xv
1.0 Introduction.....	1
1.1 Translation in eukaryotes	2
1.1.1 Overview of translation.....	2
1.1.2 Initiation	2
1.1.3 Elongation	4
1.1.4 Termination.....	5
1.2 mRNAs: agents of their own lives	6
1.2.1 5' cap.....	7
1.2.2 Poly(A) tails and poly(A)-binding protein (PABP).....	7
1.2.3 5' and 3' untranslated regions in mRNA stability and translation.....	8
1.2.4 5' UTRs.....	9
1.2.5 3' UTRs.....	10
1.3 Translation regulation.....	10
1.3.1 Enhancing translation in polysomes: mRNA circularization theory.....	11
1.3.2 mRNA stability and turnover	12
1.3.3 Compartmentalized stress: P-bodies and stress granules in translation repression.....	13
1.3.4 Stress granules (SGs).....	14
1.3.5 Processing-bodies (P-bodies)	15
1.3.6 Phase separation and the formation of SGs and P-bodies.....	15

1.3.7 mTORC1: linking extracellular events with translation output.....	16
1.3.8 Changing protein synthesis in response to various stimuli.....	17
1.3.9 Ribosome biogenesis and synthesis of the translation machinery.....	18
1.4 La and La-Related Proteins (LARPs).....	19
1.4.1 Conserved and divergent features of La and LARPs.....	19
1.4.2 Classification and roles in RNA metabolism.....	19
1.4.3 La-Modules	20
1.4.4 La-Motif (LAM).....	21
1.4.5 RNA recognition motif (RRM).....	21
1.4.6 Conserved and divergent features of RNA recognition by LARPs	22
1.4.7 C-terminal RNA binding domains	23
1.4.8 Protein binding motifs.....	24
1.4.9 LARPs in disease	24
1.5 La-Related Protein 1 (LARP1).....	25
1.5.1 Association with poly(A) RNA and PABP.....	25
1.5.2 LARP1 regulates TOP mRNA stability and translation.....	26
1.6 Summary of goals and discoveries	27
2.0 The LARP1 La-Module Recognizes Both Ends of TOP mRNAs.....	29
2.1 Introduction	29
2.2 Materials and methods.....	31
2.2.1 La-Module (amino acids 310-540) cloning, expression, and purification.....	31
2.2.2 RNA oligonucleotides	32
2.2.3 La-Module electrophoretic mobility assays	33

2.2.4 La-Module competition assays	33
2.2.5 La-Module biotin pull-downs	33
2.2.6 Isothermal titration calorimetry	34
2.2.7 Circular dichroism spectroscopy	34
2.2.8 WT and REYA LARP1 cloning, expression, and purification from Expi293T cells	35
2.2.9 WT and REYA LARP1 electrophoretic mobility shift assays.....	35
2.2.10 WT and REYA LAR1 biotin pull-downs	36
2.3 Results.....	36
2.3.1 The LARP1 La-Module binds poly(A) RNA.....	36
2.3.2 The LARP1 La-Module binds various TOP motifs.....	40
2.3.3 LARP1 La-Module simultaneously engages poly(A) RNA and RPS6 TOP motif	42
2.3.4 Simultaneous binding of poly(A) and TOP motif RNA is faithful in the context of full-length LARP1.....	47
2.4 Discussion	49
3.0 Identification of RNA-Binding Surfaces and Optimization of La-Module Construct ...	54
3.1 Introduction	54
3.2 Materials and methods.....	55
3.2.1 La-Module (310-540) mutants cloning, expression, and purification.....	55
3.2.2 UV cross-linking of La-Module (310-540).....	55
3.2.3 Size exclusion chromatography of La-Module (310-540)	55
3.2.4 La-Module (amino acids 310-647), cloning, expression, and purification.....	56

3.2.5 LAM cloning, expression, and purification.....	57
3.2.6 RRM cloning, expression, and purification	57
3.2.7 La-Module (310-647) electrophoretic mobility shift assays.....	58
3.2.8 La-Module (310-647) competition assays	58
3.2.9 RRM (440-647) electrophoretic mobility shift assays	58
3.2.10 RRM (440-647) static light scattering (SLS)	59
3.2.11 La-Module (310-540) and RRM (440-647) native mass spectrometry	59
3.3 Results.....	59
3.3.1 Attempts to identify the RNA binding surfaces of the LARP1 La-Module (310-540)	59
3.3.2 Attempts to identify the stoichiometry of the LARP1 La-Module (310-540)..	61
3.3.3 Identification of La-Module (310-647) as the optimal construct.....	62
3.3.4 Contributions of the LAM and RRM to RNA binding	65
3.3.5 La-Module (310-647) and RRM (440-647) can simultaneously bind poly(A) and TOP motif RNA.....	67
3.3.6 Attempts to identify the stoichiometry of the RRM (440-647) during simultaneous binding of poly(A) and TOP motif RNA	68
3.3.7 Attempts to identify the RNA binding surfaces of the LARP1 La-Module	70
3.4 Discussion	72
4.0 Identification of LARP1-PABPC1 Interaction Interface	75
4.1 Introduction	75
4.2 Materials and methods.....	78
4.2.1 Cloning, expression, and purification of GST-tagged PABPC1 constructs	78

4.2.2 LARP1 cloning, expression, and purification from <i>E. coli</i>	79
4.2.3 GST pull-downs	80
4.2.4 MLLE cloning, expression, and purification	80
4.2.5 MLLE and PAM2 co-crystallization	81
4.3 Results.....	82
4.3.1 LARP1 La-Module binds PABP MLLE.....	82
4.3.2 LARP1 La-Module binds the PABP RRM.....	83
4.3.3 Crystallization of PABPC MLLE domain and LARP1 PAM2.....	84
4.4 Discussion	85
5.0 Discussion.....	89
5.1 Role of the LARP1 La-Module in TOP mRNA recognition.....	89
5.1.1 Future direction: How does the LARP1 La-Module affect TOP mRNA translation?	91
5.2 Canonical faces of LAM and RRM for RNA-binding activity.....	93
5.2.1 Future direction: How does the La-Module bind two RNAs simultaneously?	94
5.3 Contribution of disordered regions to LARP1 RNA-binding activity	94
5.3.1 Future direction: How do unstructured regions facilitate LARP1 interactions with RNA and proteins? Do binding partners stimulate disorder-to-order changes?	97
5.4 Concluding remarks.....	98
Appendix A Structural Predictions Used to Optimize La-Module Construct	99
Appendix A.1 Introduction.....	99

Appendix A.2 Materials and methods	99
Appendix A.2.1 Protein sequence alignment and secondary structure prediction .99	
Appendix A.2.2 Ab initio protein structure prediction.....100	
Appendix A.2.3 One-to-one threading.....100	
Appendix A.3 Results and discussion	100
Appendix A.3.1 LARP1 RRM region greatly diverges from defined RRM.....100	
Appendix A.3.2 <i>Ab initio</i> homology model	105
Appendix A.3.3 One-to-one threading.....106	
Appendix A.4 Conclusions.....	110
Appendix B Microalgae LARP1	112
Appendix B.1 Introduction	112
Appendix B.2 Materials and methods	113
Appendix B.2.1 Protein cloning, expression, and purification	113
Appendix B.2.2 Secondary structure prediction and homology modeling.....	114
Appendix B.3 Results and discussion.....	114
Appendix B.4 Conclusions	118
Bibliography	120

List of Figures

Figure 1-1 Model of eukaryotic translation initiation.	3
Figure 1-2 Translation elongation and termination in eukaryotes.	5
Figure 1-3 Schematic of an eukaryotic mRNAs.	6
Figure 1-4 Polysome pools are in dynamic equilibrium with stress granules and P-bodies.	14
Figure 1-5 mTORC1 fine-tunes translation in response to extracellular cues.	17
Figure 1-6 Schematic of LARPs and their domain organization.	20
Figure 2-1 LARP1 La-Module binds poly(A) RNA.	37
Figure 2-2 Biophysical characterization of the LARP1 La-Module demonstrates it is folded and interacts with poly(A), but not poly(C) RNA.	38
Figure 2-3 The La-Module poly(A) RNA complex is maintained the presence of different nucleic acid nonspecific competitors.	39
Figure 2-4 LARP1 La-Module binds TOP mRNA 5' UTRs in a TOP motif-dependent manner.	41
Figure 2-5 LARP1 La-Module does not bind all TOP motifs or the 3' UTR of RPS6 mRNA.	42
Figure 2-6 An intermediate complex forms upon the addition of cold RPS6 5'TOP motif or RPL13A 5' UTR to the La-Module poly(A) RNA complex.	43
Figure 2-7 Formation of an intermediate complex is not an artifact of aggregation or RNA basepairing , and is dependent upon an intact TOP motif.	44
Figure 2-8 The intermediate complex forms in the presence of another non-specific competitor.	45

Figure 2-9 The La-Module simultaneously binds to poly(A) RNA and RPS6 TOP motif in biotin pull-down experiments.	46
Figure 2-10 WT and REYA LARP1 bind poly(A) RNA and RPS6 5' UTR.....	48
Figure 2-11 WT and REYA LARP1 simultaneously binds poly(A) RNA and RPS6 TOP motif in biotin pull-down experiments.....	49
Figure 2-12 Model of potential roles for the LARP1 La-Module in TOP mRNA translation and localization to P-bodies and stress granules.....	52
Figure 3-1 La-Module RNA-binding mutants show decreased affinity for RPS6 20-mer and poly(A) RNA.....	60
Figure 3-2 Cross-linking La-Module to poly(A) and RPS6 20-mer RNA was inefficient....	61
Figure 3-3 Constructs generated to identify the optimal La-Module construct.	63
Figure 3-4 Purification of La-Module (amino acids 310-647).....	64
Figure 3-5 LARP1 La-Module (310-647) binds RPS20-mer and poly(A) 25-mer RNAs.....	65
Figure 3-6 RNA-binding is driven by the linker and RRM region.	66
Figure 3-7 Tandem addition of the LAM and RRM suggest that the RRM region drives RNA-binding.	67
Figure 3-8 Revised La-Module and RRM form intermediate complex upon addition of RPS6 TOP motif to pre-bound poly(A) RNA.	68
Figure 3-9 LARP1 RRM might bind poly(A) and TOP motif RNA as a monomer.	69
Figure 3-10 Crystallized LARP1 La-Module, LAM, and RRM constructs.	70
Figure 3-11 Expression of La-Module linker deletion constructs for crystallographic studies.	71
Figure 3-12 Recombinant purified LARP1 RRM is intrinsically disordered.	72

Figure 4-1 Schematic of LARP1 and PABP domain organization and PAM2 sequence conservation.....	76
Figure 4-2 LARP1 La-Module binds the PABP MLLE region.....	83
Figure 4-3 LARP1 La-Module binds PABP RRM1+2.....	84
Figure 4-4 Model of LARP1 PAM2 displacement for deadenylation during mRNA decay in P-bodies.....	86
Figure 4-5 Phosphorylation of LARP1 PAM2 and PABP may affect interactions.....	88
Figure 5-1 Model for TOP mRNA translation regulation through LARP1 and outstanding questions.....	90
Figure A-1 Secondary structure prediction of LARP1.....	101
Figure A-2 Sequence conservation of LARP1 LAM region.....	102
Figure A-3 Sequence conservation of LARP1 RRM region.....	103
Figure A-4 Sequence conservation of the LARP1 RRM region amongst different species.....	104
Figure A-5 <i>Ab initio</i> homology model of the LARP1 La-Module (310-647).....	106
Figure A-6 Local and global homology models of the LARP1 RRM generated by threading on the LARP6 RRM.....	108
Figure A-7 Local and global homology models of the LARP1 La-Module (310-647) generated by threading on the La-Protein La-Module.....	109
Figure A-8 Local and global homology model of the LARP1 La-Module (310-647) generated by threading on the LARP4 La-Module.....	110

Figure B-1 Schematic of domain organization in human LARP1 versus microalgae LARP1.
..... 112

Figure B-2 Secondary structure prediction of maLARP1. 115

Figure B-3 Homology model of maLARP1..... 116

Figure B-4 maLARP1 exhibits robust expression and purification..... 117

**Figure B-5 Size exclusion chromatography suggests stability and homogeneity of purified
maLARP1.** 118

Preface

“Once when I was living in the heart of a pomegranate, I heard a seed saying, ‘Someday I shall become a tree, and the wind shall sing in my branches, and the sun will dance on my leaves, and I shall be strong and beautiful through all the seasons.’

Then another seed spoke and said, ‘When I was as young as you, I too held such views; but now that I can weigh and measure things, I see that my hopes were in vain.’

And a third seed spoke also, ‘I see in us nothing that promises so great a future.’

And a fourth said, ‘But what a mockery our life would be, without a greater future!’

Said a fifth, ‘Why dispute what we shall be, when we know not even what we are.’

But a sixth replied, ‘Whatever we are, that we shall continue to be.’

And a seventh said, ‘I have such a clear idea how everything will be, but I cannot put it into words’

Then an eighth spoke – and a ninth – and a tenth – and then many – until all were speaking, and I could not distinguish anything for the many voices.

And so I moved that very day into the heart of quince, where the seeds are fewer and almost silent.”

- The Pomegranate, by Kahlil Gibran

1.0 Introduction

All cellular processes – growth, division, specialization, and even death – are dictated by gene expression. During gene expression, information encoded by DNA is carried by non-coding or messenger RNAs (mRNAs), the latter of which are essential intermediates that instruct protein production during translation. Most cellular processes are heavily dependent upon proteins because they fulfill catalytic, structural, transport, and regulatory roles. But the precise control of protein synthesis presents a formidable challenge for the cell. At any given time, a cell must sense various stimuli and respond by quickly fine-tuning the translation of a myriad of mRNAs; increasing the synthesis of proteins in high demand, while stagnating the synthesis of others.

Throughout evolution, cells have attained efficient mechanisms to adjust mRNA translation with the robust ability to re-adapt. In eukaryotic cells, transcription and translation are compartmentalized processes. As such, multiple opportunities exist within the cytoplasm to regulate the flow of genetic information, largely independent of nuclear mRNA synthesis and processing. Post-transcriptional events in the cytoplasm regulate mRNA translation, with the advantage of bypassing the time and expense of re-synthesizing and processing transcripts in the nucleus. These events are guided by RNA-binding proteins (RBPs), as well as a host of associated proteins, that recognize nascent transcripts and percolate them through various stages of mRNA metabolism¹. Overall, mRNA translation is regulated by: the inherent features of an mRNA that dock RBPs and their associated proteins, the stability and localization of the mRNA, as well as upstream cell signaling events. Packaging mRNAs into ribonucleoprotein complexes (mRNPs) precludes these metabolic processes, with the dynamic interchange of RNP components dictating transcript fate. Notably, RBPs that recognize only a specific class of mRNAs can allow functionally related mRNAs to be organized into distinct RNPs in order to co-regulate their translation²⁻⁵.

1.1 Translation in eukaryotes

Translation is the biosynthetic process in which genetic information carried by mRNAs is used to synthesize proteins. Importantly, translation is a principal opportunity for rapid and reversible adjustments to the proteome in response to cell signaling. For example, activated macrophages de-repress translation of 90 mRNAs encoding anti-inflammatory regulators within one hour of immune stimulation⁶. Translation is also the predominant mode for regulating gene expression⁷ in cell systems with little to no transcription regulation, such as reticulocytes^{8,9} and oocytes^{10,11}. For instance, translation of pre-existing maternal mRNAs in oocytes increases by 30-fold upon fertilization¹¹.

1.1.1 Overview of translation

Translation occurs in three steps: 1) initiation, 2) elongation, and 3) termination. Initiation is the rate-limiting step, with the majority of translation regulation executed at this stage¹². Eukaryotic cap-dependent translation relies upon numerous protein-mRNA and protein-protein interactions that are orchestrated by non-coding features of mRNAs: the 5' cap and poly(A) tail (discussed further in sections 1.2.1 and 1.2.2).

1.1.2 Initiation

Translation starts with the recruitment of various eukaryotic initiation factors (eIFs) to the mRNA^{13,14}. Eukaryotic mRNAs are co-transcriptionally appended with a 5' cap that is recognized by eIF4E during cap-dependent translation¹⁴ (Figure 1-1). eIF4G binds eIF4E and acts as a scaffold (Figure 1-1)¹⁴. An RNA helicase, eIF4A, then binds eIF4G and unwinds the cap-proximal RNA region in preparation of ribosome association¹⁴ (Figure 1-1). Together, eIF4E, eIF4G, and eIF4A form the eIF4F translation initiation complex¹⁴. The 43S pre-initiation complex – consisting of the 40S ribosomal subunit, eIF1, eIF3, and the eIF2-GTP-Met-tRNA_i ternary complex – is recruited to the translation initiation complex *via* eIF3 binding to eIF4G¹⁴ (Figure 1-1). Once associated

with the 5' cap, the 43S then 'scans' the mRNA until it encounters an AUG start codon in appropriate Kozak sequence context¹⁴ (Figure 1-1). Start codon recognition occurs through the base-pairing of the AUG with the Met-tRNA_i anticodon and forms the 48S initiation complex¹⁴. Concomitant GTP hydrolysis triggers the release of associated eIFs and the 60S large ribosomal subunit binds the 48S, forming the 80S ribosome that is competent for elongation¹⁴ (Figure 1-1).

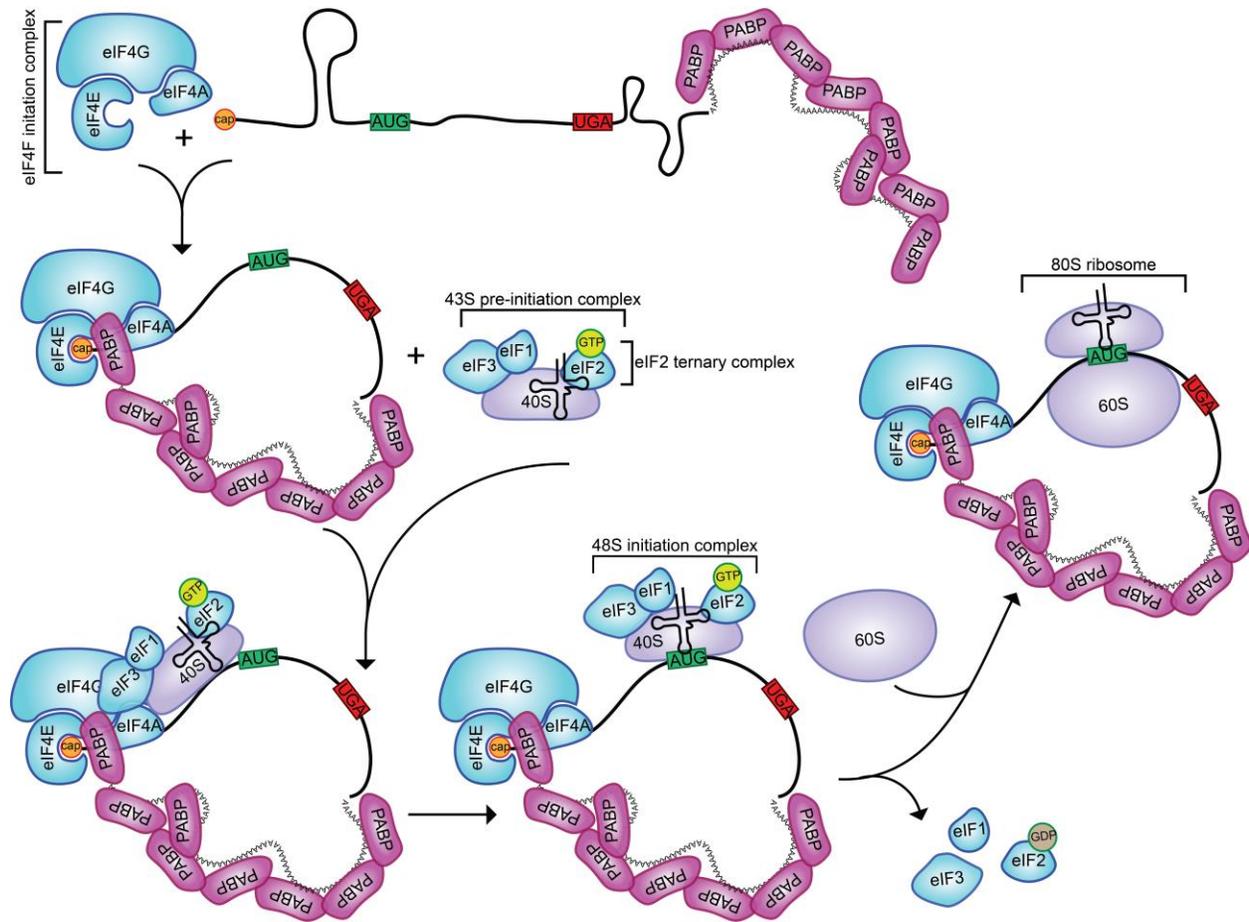


Figure 1-1 Model of eukaryotic translation initiation. Translation commences with the assembly of the eIF4F translation initiation complex on the 5' end of the mRNAs. The 43S pre-initiation complex binds the eIF4F translation initiation complex and scans towards the start codon. Base-pairing between the Met-tRNA_i and start codon forms the 48S initiation complex. The 60S ribosomal subunit joins to form an elongation competent 80S ribosome. The eIF4G-PABP interaction leads to mRNA circularization, which may enhance translation¹⁵ (see 1.1.3).

1.1.3 Elongation

Peptide synthesis and ribosome translocation are mediated by eukaryotic elongation factors (eEFs). After initiation, the 80S ribosome is poised on the AUG codon with the Met-tRNA_i anticodon base-paired to the start codon within the peptidyl site (P-site)¹⁶ (Figure 1-2). The second codon is now positioned in the aminoacyl site (A-site) pending recognition by its cognate aminoacyl-tRNA¹⁶ (Figure 1-2). Eukaryotic elongation factor 1A (eEF1A) binds aminoacyl-tRNAs in a GTP-dependent manner and presents them to the A-site¹⁶ (Figure 1-2). Upon correct codon-anticodon matching, GTP hydrolysis triggers the release of eEF1A, and the aminoacyl-tRNA is accommodated into the A-site¹⁶. After peptide bond formation, eEF2-GTP binds to the ribosome and promotes translocation of the deacylated and peptidyl tRNAs into the E- and P-sites, respectively¹⁶ (Figure 1-2). GTP hydrolysis then triggers release of eEF2 and deacylated tRNA from the E-site¹⁶ (Figure 1-2). With the A-site now empty, the ribosome is primed for the next cycle of elongation¹⁶ (Figure 1-2).

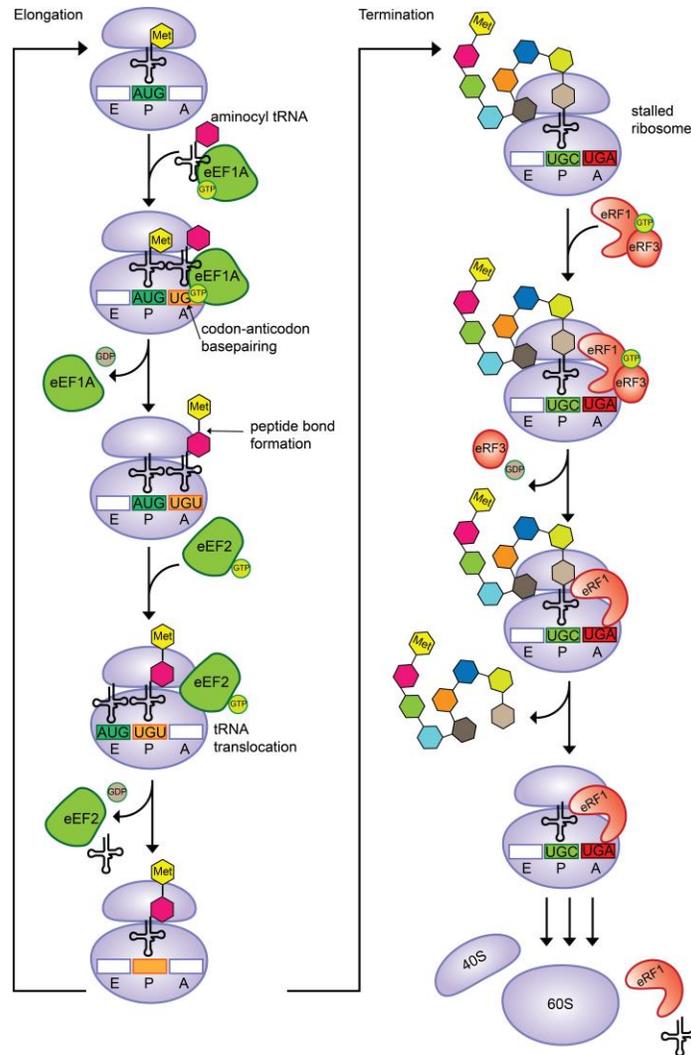


Figure 1-2 Translation elongation and termination in eukaryotes. Elongation is a cyclic process that occurs in the large ribosomal subunit. Amino-acyl tRNAs bind to the A-site through recognition of the mRNA by the tRNA anticodon. Peptide bond formation is followed by translocation into the P-site. The deacetylated tRNA is released and the A-site is emptied for the next amino-acyl tRNA. Upon occupation of the A-site by the stop codon, eukaryotic release factors stimulate release of the nascent polypeptide from the P-site.

1.1.4 Termination

Translation ends when the ribosome traverses the coding sequence and encounters a stop codon (UAA, UGA, or UAG) in the A-site^{16,17}(Figure 1-2). This process is catalyzed by the collaboration of two eukaryotic release factors (eRFs) – eRF1 and eRF3^{16,17}. When a stop codon occupies the ribosome A-site, the ribosome stalls as there is no tRNA with a corresponding

anticodon^{16,17}(Figure 1-2). The preformed eRF1/eRF3-GTP ternary complex binds to the A-site^{16,17}(Figure 1-2). GTP hydrolysis by eRF3 causes conformational changes in eRF1 that allow it to enter the pre-termination complex and induce the hydrolysis of the nascent polypeptide from the peptidyl-tRNA situated in the P-site^{16,17}.

1.2 mRNAs: agents of their own lives

During translation, the ribosome reads the coding sequence of mRNAs and produces the encoded protein. However, mRNAs contain non-coding features that flank both ends of the coding sequences, namely: the 5' cap, 3' poly(A) tail, as well as the 5' and 3' UTRs (Figure 1-3). Although these features do not encode proteins themselves, they are critical components of mRNAs that regulate their own stability and translation of transcripts throughout their lifetimes.

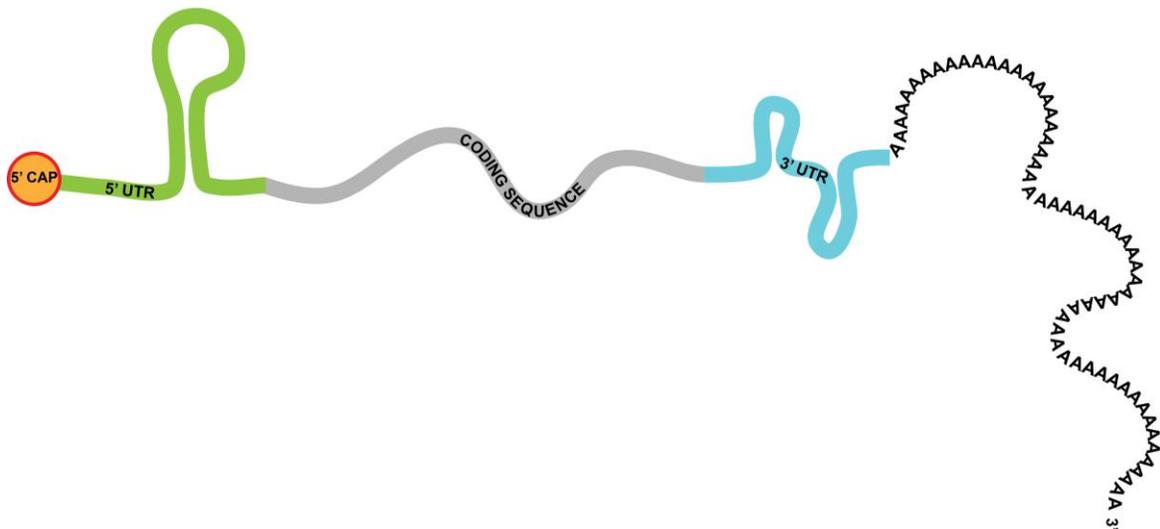


Figure 1-3 Schematic of an eukaryotic mRNAs. The coding sequence is embedded within untranslated features. The 5' and 3' untranslated regions (UTRs) flank either end of the coding sequence. These UTRs can have various, lengths, sequences and secondary structures that dock RNA-binding proteins. mRNAs are also co-transcriptionally capped and polyadenylated within the nucleus.

1.2.1 5' cap

The 5' m⁷G cap is the first modification placed upon eukaryotic mRNAs and accompanies transcripts throughout their lifetimes^{18,19} (Figure 1-3). The cap is co-transcriptionally incorporated in three steps after the first 20-25 nts of a nascent transcript emerge from RNA polymerase II²⁰. First, RNA 5' triphosphatase removes the γ -phosphate from the 5'-triphosphate of the first nucleotide, leaving a 5'-diphosphate. RNA guanylyltransferase then transfers GMP from GTP to the 5'-diphosphate, adding a guanine cap *via* a 5'-5' linkage that is inverted relative to the 3'-5' phosphodiester bonds throughout the rest of the transcript. Finally, guanine-N7-methyltransferase transfers a methyl group to the N7 amine of the guanine cap, resulting in a cap 0 structure. The 5' cap protects transcripts from 5' \rightarrow 3' degradation²¹, acts an identifier to recruit protein factors for pre-mRNA splicing^{22,23}, polyadenylation, and nuclear export²⁴, and anchors translation initiation factors during cap-dependent translation¹⁸.

The 5' cap is nearly always bound by proteins. In the nucleus, the cap is bound by nuclear cap-binding complex (CBC), which consists of CBP20 and CBP80²³. Once in the cytoplasm, CBC is exchanged for eukaryotic initiation factor 4E (eIF4E)²⁵. Modifications to the cap and the identity of cap-proximal nucleotides can affect recognition by cap-binding proteins and facilitate different cellular functions^{19,26}. For instance, eIF4E has the highest affinity for transcripts with a cap 1 structure and a purine in the +1 position²⁷; a +1C confers greater translation inhibition during stress than +1G or +1A²⁶. Methylation of the 2'OH of the +1 and +2 ribose generate cap 1²⁸ and cap 2²⁹, respectively. Cap 1 is also critical for non-self discrimination of the innate immune response to foreign RNAs²⁸. Thus, the cap can mediate different cellular processes by recruiting unique proteins that preferentially recognize unique modifications¹⁹.

1.2.2 Poly(A) tails and poly(A)-binding protein (PABP)

Perhaps the most well-studied non-coding feature of mRNAs is the poly(A) tail, which has been an active area of research since the early 1970s^{30,31}. Nearly all eukaryotic mRNAs are post-transcriptionally amended with a 3' poly(A) tail³⁰ (Figure 1-3). Polyadenylation occurs in two steps. Briefly, the mRNA undergoes endonuclease cleavage by the cleavage and polyadenylation

specificity factor (CPSF) complex. CPSF recognizes an AAUAAA motif located 20-30 nucleotides upstream of the cleavage site³². After cleavage, polyadenylation is catalyzed by polyadenylate polymerase (PAP). While PAP can catalyze polyadenylation independently, processing is very slow and is substantially enhanced by CPSF³³ and poly(A)-binding protein (PABP)^{33,34}. Typical poly(A) tails are 200-250 nucleotides long³⁴, though their lengths are highly dynamic, even within the lifetime of a single transcript³⁵⁻³⁷; poly(A) tails facilitate the export, localization, decay, and translation of their corresponding mRNAs³⁸ through PABP³⁹.

PABP was first discovered in 1973 as the predominant poly(A)-binding factor in cytoplasmic mRNPs⁴⁰. PABP is an abundant protein, with HeLa cells containing six PABP molecules for every ribosome⁴¹. It was initially postulated that PABP simply protects bound transcripts from 3' → 5' exonuclease activity, but has been shown to participate, and even guide, numerous mRNA-dependent events³⁹. With several known protein-binding factors and the ability to bind virtually any polyadenylated transcript, PABP affects all stages of mRNA metabolism³⁹. PABP binds poly(A) tails using four tandem RRM domains and interacts with proteins *via* a C-terminal Mademoiselle (MLLE) domain.

Nuclear PABP (PABPN1) controls poly(A) tail length³⁴; consecutive binding of PABPN1 forms a 21 nm-diameter filament that serves as a “molecular ruler” for poly(A) tail length⁴². Cytoplasmic PABP (PABPC1; hitherto referred to as PABP), mediates mRNA localization⁴³, and regulates decay⁴⁴ and translation⁴⁵. PABP regulates translation by interacting with various effector proteins that can enhance or inhibit the global translation of mRNAs or specifically tune a subclass of transcripts⁴⁶. For instance, PAIP1⁴⁷ and PAIP2⁴⁸ compete⁴⁸ for binding to the PABP MLLE domain to enhance and inhibit mRNA translation, respectively. In a unique autoregulatory fashion, PABP represses the translation of its own mRNA *via* binding to adenosine tracts within the 5' UTR to block ribosome scanning^{49,50}.

1.2.3 5' and 3' untranslated regions in mRNA stability and translation

Non-coding regions of mRNAs can drive the regulation of gene expression. Not surprisingly, only 1.5% of the human genome encodes proteins^{51,52}. Genome size increased throughout evolution from invertebrates to humans⁵³. The 5' and 3' untranslated regions (UTRs)

exhibited marked expansion in length, creating a ‘playground’ for mRNA regulation⁵³ (Figure 1-3). The 5’ and 3’ UTRs harbor sequence determinants that control mRNA localization, decay, and translation either directly *via cis*-acting sequence and structural elements or by docking *trans*-acting factors. Accordingly, sequence and structural features within the UTRs can directly or indirectly block or recruit ribosomes and other translation regulatory factors, allowing rapid control of translation in response to diverse and fastidious cellular conditions.

1.2.4 5’ UTRs

5’ UTRs are generally much shorter than 3’UTRs, with the average length being 100-200 nts⁵³. However, 5’ UTR lengths can vary dramatically among different genes in eukaryotes, ranging from a few to thousands of base pairs⁵³. This provides unique platforms for regulating specific mRNA subsets. Features of the 5’ UTR that affect transcript stability or translation can be non-structural or structural. Non-structural features include a strong Kozak sequence, which improves start codon recognition in highly translated mRNAs⁵⁴ or upstream open reading frames (uORFs) that regulate translation of the downstream ORF^{55,56}.

Generally, 5’ UTR secondary structures repress translation⁵⁷. A classic example is the iron responsive element (IRE), which is a stem-loop near the 5’ cap of mRNAs required for iron homeostasis. During iron depletion, iron regulatory protein 1 (IRP1) or IRP2 binds to the IRE and prevents ribosome scanning^{58,59}. Cap-proximal hairpins or G-quadruplexes can inhibit formation of the 43S preinitiation complex or 48S scanning⁶⁰. Additionally, RNA tertiary structures like pseudoknots regulate translation⁶¹ in response to downstream encoded proteins or metabolites⁶².

Some 5’ UTR structures can promote translation. For instance, internal ribosome entry sites (IRESs) enable cellular cap-independent translation during global repression of cap-dependent translation as a result of stress⁶³, mitosis⁶⁴, or apoptosis^{65,66}. Viruses use IRESs to hijack host translation machinery in response to repressed translation during viral infection⁶⁷⁻⁶⁹. Intermolecular interactions with non-coding RNAs such as long non-coding and circular RNAs can enhance also translation⁷⁰.

1.2.5 3' UTRs

Genome expansion correlates with 3' UTR length, with the human transcriptome having an average 3' UTR length of 1,200-nt⁷¹. 3' UTRs regulate mRNA stability and translation largely through AU-rich elements⁷² and binding sites for microRNAs⁷³. AU-rich elements promote deadenylation-mediated decay and are preferentially found in genes with short half-lives whose expression requires tight regulation, such as proto-oncogenes⁷⁴. MicroRNAs decrease protein output by destabilizing mRNAs or repressing their translation⁷³.

Structural elements within 3' UTRs dock RBPs that facilitate temporal and localized translation regulation⁷⁵. For instance, during embryo development in *Drosophila*⁷⁶ and *Xenopus*⁷⁷, various mRNAs are translationally repressed until they are localized to the anterior or posterior pole. Histone mRNAs lack poly(A) tails and undergo cell-cycle dependent decay mediated by a 3' terminal stem loop motif⁷⁸. Stem-loop binding protein (SLBP) binds the 3' end of histone mRNAs and is required for pre-mRNA processing and accompanies mature histone mRNAs to the cytoplasm⁷⁹.

At least 54% of human genes contain alternative polyadenylation sites within their 3' UTRs⁸⁰, which can produce mRNAs with shorter 3'UTRs or encode different protein isoforms. Cancer cell lines tend to have mRNAs with shorter 3' UTRs resulting from alternative cleavage and polyadenylation⁸¹. These transcripts show increased stability and produce 10-fold more protein due to the loss of miRNA-mediated repression⁸¹.

1.3 Translation regulation

Translation is interconnected with and controlled by various cytoplasmic events within eukaryotic cells. Translation is fine-tuned by upstream cell signaling events to regulate global protein synthesis or the translation of a specific class of mRNAs⁸². Various quality control systems monitor mRNA translation and, when aberrant translation is detected, sequester the underlying mRNAs into specialized decay pathways⁸³. Furthermore, translationally repressed mRNAs can also accumulate into P-bodies and stress granules when translation initiation is inhibited during

cellular stress⁸⁴. These mRNAs can re-enter the translation pool once stress has lifted or targeted for degradation during prolonged stress⁸⁴.

1.3.1 Enhancing translation in polysomes: mRNA circularization theory

The closed-loop model for mRNA translation was first proposed in the early 1980s and has been reiterated several times since⁸⁵. The idea of mRNA circularization began with the observation that poly(A) tails enhance translation. Capped poly(A)⁺ mRNAs increase polysome association, translation initiation, and yield 2.5-fold greater translation than capped poly(A)⁻ mRNAs^{85,86}. Furthermore, poly(A) RNA competitively inhibits the translation of poly(A)⁺ mRNAs in rabbit reticulocyte extracts as compared to other ribopolymers⁸⁷. Together these data established the importance of the poly(A) tail for eukaryotic cap-dependent translation initiation.

Subsequent studies showed that the 5' cap and poly(A) tail synergistically increase mRNA translation in a stability-independent manner⁸⁸. Capped poly(A)⁺ mRNAs undergo more efficient translation than mRNAs possessing neither or only one moiety^{88,89}. PABP and eIF4G interact in yeast⁹⁰, plants⁹¹, and humans⁹²⁻⁹⁴, highlighting its functional significance. The eIF4G-PABP complex increases the affinity of eIF4E for the cap⁹⁵ and translation efficiency⁹⁶. Additionally, eIF4G promotes poly(A) tail-dependent translation *in vitro*⁹⁷. Expression of an eIF4G mutant that cannot bind PABP inhibits translation of polyadenylated mRNA in *Xenopus* oocytes⁹⁸. These data suggest communication between mRNA ends, as the poly(A) tail affects 5' activities and *vice versa*.

Proximity of 5' and 3' ends through mRNA circularization has also been observed. Electron micrographs of rat pituitary cells revealed circular polysomes on the surface of the rough endoplasmic reticulum⁹⁹. Atomic force microscopy and cryoelectron tomography showed the formation of circular polysomes when recombinant yeast eIF4E, eIF4G, and PABP were mixed with capped and polyadenylated mRNAs^{15,100}. Furthermore, omitting any protein factors abrogated mRNA circularization as did eIF4G mutants defective in PABP binding¹⁵.

But what is the advantage of mRNA circularization? Circularization of mRNAs may promote ribosome recycling for more efficient translation; terminating ribosomes at the 3' end can quickly reinitiate at the proximal 5' end¹⁰¹. PABP may promote ribosome recycling as it also

associates with ribosome release factors¹⁰². Some data also suggest that PABP enhances translation by promoting 60S ribosomal subunit joining⁸⁵ and/or recruiting the 40S^{102,103}.

The closed-loop model of mRNA translation has recently been scrutinized¹⁰⁴. This single-molecule study suggested that translationally repressed transcripts in stress granules (SGs) showed proximal 5' and 3' ends, while translationally active transcripts rarely showed co-localized ends¹⁰⁵. However, the probe used to observe the 3' end of the reporter mRNA targeted the junction between the 3'UTR and poly(A) tail¹⁰⁵. As such, the distance measured between the 5' and 3' probes is actually from the 3' UTR-poly(A) tail junction to the cap. The average poly(A) tail is 250 nts, a length that could bridge the gap. Due to the crowded nature of SGs – intra- and inter- mRNA interactions¹⁰⁶, as well as multimeric RNA-protein and protein-protein interactions¹⁰⁷ – it is unsurprising that the 5' end and the 3'UTR-poly(A) tail junction are proximal. Nonetheless, a more recent study suggested mRNA circularization promotes oncogenic translation and malignancies¹⁰⁸.

In reality, mRNA circularization may not be a global occurrence; transcript-specific circularization may be guided by the length of the poly(A) tail, polymer chain flexibility, thermodynamics, and associated proteins. Additionally, mRNA circularization may occur only in short bursts. Some studies suggest that post-transcriptional regulators, such as microRNAs, affect PABP and eIF transcript occupancy without changing poly(A) tail length^{109,110}. Consistent with this, other studies report that mRNA translation is stochastic, with rapid switches between active and inactive translation states¹¹¹⁻¹¹³, and <20% of translating mRNPs show 5' and 3' distances consistent with the closed-loop model of translation¹¹⁴. Thus, changes to mRNP composition before, during, and after translation can affect the proximity of mRNA ends.

1.3.2 mRNA stability and turnover

Cellular mRNA levels are the net result of mRNA synthesis in the nucleus and degradation in the cytoplasm. Differential mRNA turnover in the cytoplasm expedites translation changes in response to environmental and developmental cues¹¹⁵. Stable mRNAs undergo more translation¹¹⁶⁻¹¹⁸. Proteins required in high demand are typically encoded by relatively stable transcripts, as is the case for mRNAs encoding β -globin¹¹⁹ and housekeeping proteins such as HSP70¹²⁰. By

contrast, mRNAs encoding proteins required for brief timeframes, for example proto-oncogenic mRNAs¹²¹, tend to have shorter half-lives.

A number of mRNA decay pathways exist in the cytoplasm of mammalian cells¹²². The rates of mRNA decay are dictated by *cis*- and *trans*-acting factors¹²³. The degradation of a specific mRNA may be limited to a distinct pathway or may be targeted to various, seemingly redundant pathways, depending upon the cellular condition¹²². Two integral determinants of mRNA stability exist in the majority of eukaryotic mRNAs, namely, the 5' cap and poly(A) tail. Either one of these features can be compromised to initiate directional decay, or the mRNA can be cleaved endonucleolytic attack¹²².

The bulk of eukaryotic mRNA degradation commences with deadenylation of the poly(A) tail. After this, the mRNA is subject to 3' → 5' degradation by the exosome or 5' → 3' degradation by exonuclease XRN1 following decapping¹²². These two methods do not seem mutually exclusive, allowing flexibility in the means of decay¹²². Endonuclease-mediated decay initiates with the internal cleavage of an mRNA. This generates two fragments of RNA with an unprotected 5' or 3' end that can then be degraded by XRN1 and the exosome, respectively¹²². Endonuclease cleavage sites vary in length and secondary structure, and have been identified in the coding region and in UTRs¹²⁴. Endonucleolytic cleavage of specific mRNAs is regulated by RBPs that bind at or in the vicinity of cleavage sites to block accessibility to cleavage sites¹²⁴.

1.3.3 Compartmentalized stress: P-bodies and stress granules in translation repression

In eukaryotes, mRNPs travel to different membrane-enclosed compartments both as part of mRNA biogenesis and in response to changing conditions. However, high concentrations of protein and RNA also cause mRNPs to coalesce into membrane-less subcellular compartments called RNA granules. These granules can be found in the nucleus, for example the nucleolus or paraspeckles, and in the cytoplasm, such as stress granules (SGs) or processing-bodies (P-bodies)¹²⁵. SGs and P-bodies help regulate mRNA stability and translation by sequestering mRNPs that are stalled at translation initiation due to stress^{84,126,127}. This allows rapid re-entry into the translation pool once the stress has lifted, which permits a faster response than re-transcribing, processing, and exporting new mRNAs. Although SGs and P-bodies share some properties and

components, their proposed functions, behavior, and compositions differ. Ultimately, both SGs and P-bodies respond to translation control events that modulate the proteome and influence cell fate in ways that are not yet completely understood.

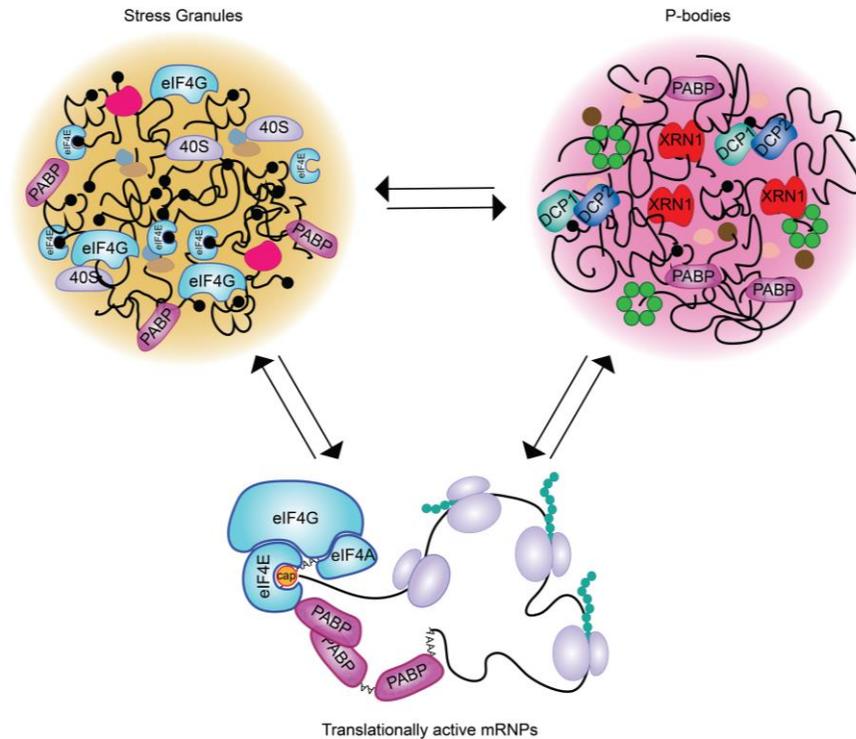


Figure 1-4 Polysome pools are in dynamic equilibrium with stress granules and P-bodies. mRNAs and their associated proteins can be sequestered into SGs and P-bodies during translation repression in response to stress. mRNPs can be exchanged between SGs, P-bodies, and translationally active mRNPs.

1.3.4 Stress granules (SGs)

SGs assemble when stress-activated pathways stall translation initiation^{107,128}. Mechanistically, stress-activated serine/threonine kinases phosphorylate eIF2 α in response to external stimuli^{128,129}. For instance, GCN2 and PKR phosphorylate eIF2 α in response to nutritional starvation and ER perturbations, respectively^{128,129}. Phosphorylation of eIF2 α , depletes the eIF2-GTP-Met-tRNA_i complex needed to load initiating Met-tRNA_i in the 40S^{128,129}. Not surprisingly, SGs contain ribosomal pre-initiation complexes and some translation initiation factors^{84,130}, in addition to RBPs involved in translation repression¹⁰⁷, PABP^{84,130,131}, protein chaperones^{84,130}, and mRNAs¹³². Ultimately, SGs facilitate survival and adaptation in response to changing

environments by sequestering translationally inactive mRNP pools in response to cell signaling¹²⁶. SGs are disassembled upon adaptation to or removal of stress, restoring translation equilibrium¹²⁶.

1.3.5 Processing-bodies (P-bodies)

P-bodies also assemble in response to cellular stress and sequester mRNPs that are translationally inactive^{84,133}. While SGs generally house proteins involved in translation initiation, P-bodies mostly harbor mRNA decay factors^{84,133}. Like SGs, P-bodies contain translation repressors and mRNAs, although the latter are mostly deadenylated^{84,133}. P-bodies were first described as “XRN1 foci” due to the localization of exoribonuclease XRN1 to distinct cytoplasmic foci⁸⁴. Subsequently, a conserved core of proteins functioning in mRNA decay or silencing have been observed to co-localize to P-bodies from yeast to humans, including: decapping enzymes DCP1/DCP2 and their activators EDC3/EDC4, deadenylase complex CCR4/CAF/NOT and its enhancer TOB2^{84,133-135}. In metazoans, P-bodies are also home to RNAi machinery, such as GW182 and Argonaut^{84,133,135}. While P-bodies are generally thought of as sites of mRNA decay and silencing¹³⁴, some studies have shown that mRNPs can exit P-bodies and reenter the translationally active pool of mRNAs¹³⁵⁻¹³⁷.

1.3.6 Phase separation and the formation of SGs and P-bodies

Protein-protein interactions between RBPs drive SG and P-body assembly by linking various mRNP populations together^{107,133}. Post-translational modifications that alter protein-protein interactions can also influence granule formation¹⁰⁷. For example, phosphorylation of endoribonuclease G3BP prevents its ability to multimerize, which decreases SG assembly¹³⁸. In addition, multivalent interactions in which a single RBP binds more than one mRNA enhances SG and P-body assembly¹⁰⁷. This creates a concentrated network of molecules that phase separate, a process known as liquid-liquid phase separation (LLPS)¹³⁹. These multivalent interactions are particularly driven by RBPs enriched with intrinsically disordered regions (IDRs), which provide proteins with the flexibility to undergo dynamic and promiscuous interactions¹³⁹⁻¹⁴³.

Studies of SGs have shown they possess a highly concentrated and stable core surrounded by a more dilute and dynamic shell¹³². These features permit dynamic assemblies; mammalian SGs undergo fusion, fission, and flow in the cytosol¹²⁷. Most SG components are exchanged rapidly, sometimes within seconds¹²⁷. However, SGs also contain an relatively immobile pool of proteins that exchange very slowly, if at all¹²⁷.

1.3.7 mTORC1: linking extracellular events with translation output

The mammalian target of rapamycin complex 1 (mTORC1) is a critical serine/threonine kinase complex that integrates environmental cues and regulates protein synthesis accordingly⁸². mTORC1 functions at the convergence point of numerous signaling pathways that sense changes in extracellular signals, such as secreted growth factors, and intracellular signals, like nutrient levels (amino acid, glucose, and oxygen) and ATP availability⁸². Importantly, mTORC1 relays this information to downstream effector proteins by means of phosphorylation⁸². These phosphorylated enzymes and metabolic regulators then adjust anabolic and catabolic processes downstream of mTORC1¹⁴⁴. By signaling to downstream effector proteins, mTORC1 not only regulates global protein synthesis, but also the translation of specific subclasses of mRNAs¹⁴⁴.

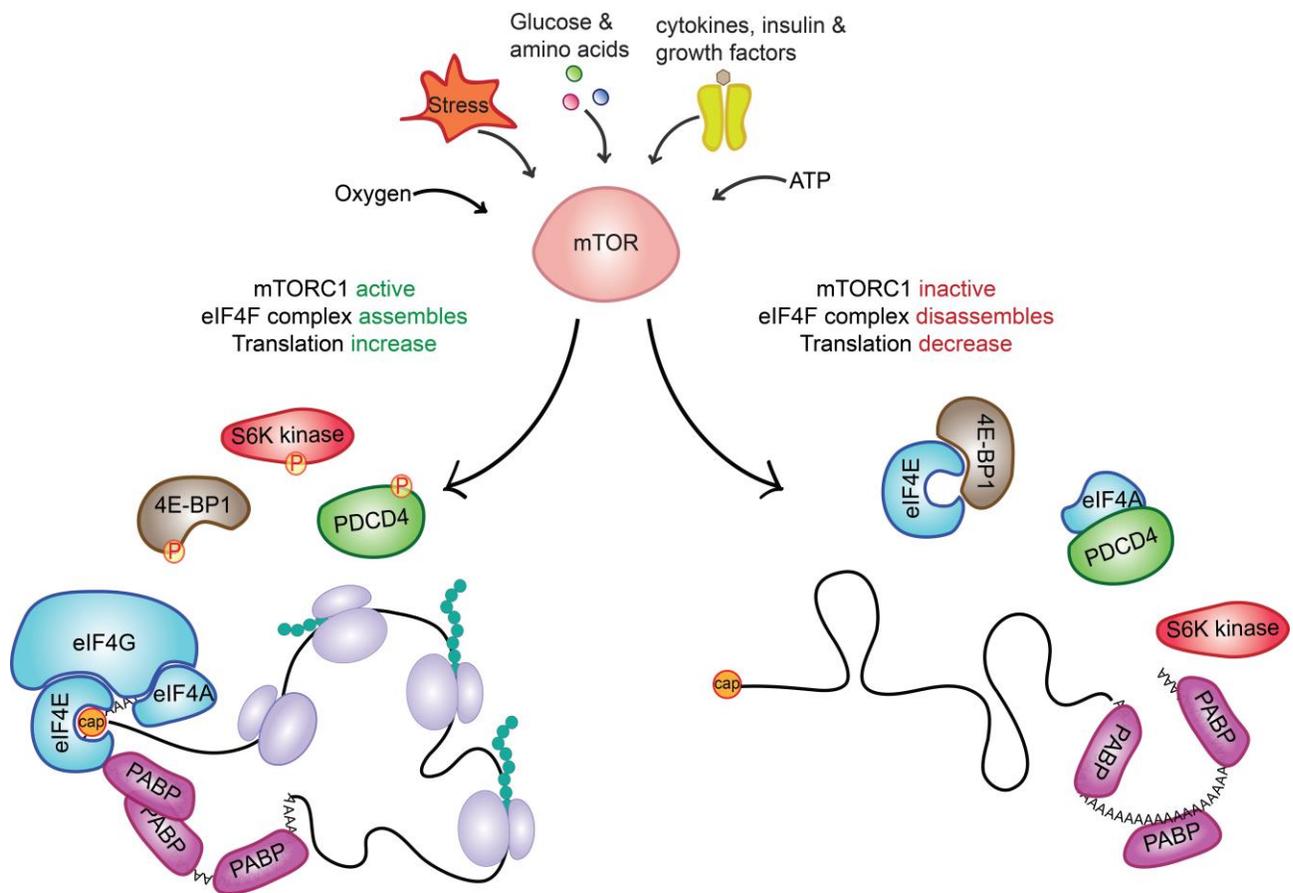


Figure 1-5 mTORC1 fine-tunes translation in response to extracellular cues. Model depicting downstream targets of mTORC1 and their phosphorylation state in response to various conditions. Phosphorylation of proteins downstream of mTORC1 permits translation initiation during favorable conditions and *vice versa*.

1.3.8 Changing protein synthesis in response to various stimuli

After binding to their respective receptors, cues from growth factors, amino acids, oxygen, and insulin converge upon the TSC1/TSC2 (Tuberosclerosis complex) heterodimer⁸². This signal is then delivered to mTORC1 *via* Rheb (Ras-homolog enriched in brain)⁸². Activated mTORC1 increases global mRNA translation by promoting the availability of translation initiation factors. mTORC1 phosphorylates 4E-BP1/2 to release eIF4E, which can then bind the cap and dock the remaining eIFs to the mRNA¹⁴⁴. The eIF4A RNA helicase is de-repressed by mTORC1 upon phosphorylation of PDCD4¹⁴⁴. mTORC1 also phosphorylates S6K kinase, which phosphorylates eIF4B and ribosomal subunit 6, allowing them to also join the translation initiation

complex¹⁴⁴. Thus, mTORC1 enhances translation by permitting assembly of the eIF4F translation initiation complex and the ribosomal pre-initiation complex through downstream phosphorylation¹⁴⁴.

1.3.9 Ribosome biogenesis and synthesis of the translation machinery

Cells decrease protein synthesis to conserve energy and resources in response to hostile conditions, such as infection or nutrient depletion, until favorable conditions are restored. A fundamental means to that end is to halt the production of new translation machinery, such as ribosomal proteins and translation factors. This is achieved by repressing the translation of mRNAs that encode components of the translation machinery, known as TOP mRNAs¹⁴⁵. These transcripts are characterized by a 5' terminal oligopyrimidine (TOP) motif within their 5' UTRs and account for 30% of total cellular mRNAs in actively growing cells¹⁴⁶. The TOP motif is found in 79 of the 80 transcripts encoding the ribosomal proteins¹⁴⁷, as well as those encoding translation initiation and elongation factors¹⁴⁵, and some RBPs such as hnRNPA1 involved in IRES-mediated translation¹⁴⁵. The TOP motif consists of an invariant +1C, followed by a stretch of 4-14 pyrimidines (with equal proportion of Cs and Us) and an adjacent GC-rich region^{148,149}. The TOP motif hallmarks are conserved in all vertebrates and some ribosomal mRNAs of *Drosophila melanogaster*¹⁵⁰. Importantly, the TOP motif is required for the coordinated and temporal repression of transcripts that encode the translation machinery in response to unfavorable environmental conditions¹⁴⁵.

The predominant subclass of mRNAs that is specifically regulated by mTORC1 is TOP mRNAs⁵. Indeed, TOP mRNA translation displays hypersensitivity to mTORC1 signaling as compared to non-TOP mRNAs with similar 5' UTR length and complexity⁵. TOP mRNAs are translated in an “all-or-nothing” manner; translation is either inactive or active, with active translation occurring at maximal efficiency¹⁵¹. This bimodal nature of TOP mRNA translation led to the hypothesis that the translation regulation of TOP mRNA is conferred at the initiation step, whereby a repressor must be relieved in order to permit TOP mRNA cap-dependent translation¹⁴⁵. Because mTORC1 does not bind RNAs, it was speculated that an mTORC1 has an RBP target¹⁴⁵; this RBP would have specificity for the TOP motif to regulate TOP mRNA translation downstream

of mTORC1. This was supported by the observation that titrating a pyrimidine competitor RNA into cell-free translation systems and mammalian cells relieves TOP mRNA translation repression¹⁵². More recently, La-Related protein 1 (LARP1) has emerged as the RBP that links mTORC1 signaling to TOP mRNA translation¹⁵³.

1.4 La and La-Related Proteins (LARPs)

La protein was first described in 1974 as an autoantigen in sera of patients with systemic lupus erythematosus and Sjögren's syndrome^{154,155}. Since its discovery, six other subfamilies of related proteins have been identified, forming the La-Related Protein (LARP) superfamily (Figure 1-6). The LARP superfamily contains over 250 eukaryotic RBPs that function in nearly all stages of RNA metabolism¹⁵⁶.

1.4.1 Conserved and divergent features of La and LARPs

Although LARPs share some core RNA-binding domains (RBDs), LARP-specific sequence and structural variations within these RBDs, as well as additional unique RBDs and motifs, contribute to the functionalization of each LARP^{156,157}. The targets of LARPs include non-coding, messenger, and viral RNAs, with LARPs executing functions ranging from RNA chaperone activity to the regulation of mRNA stability and translation¹⁵⁷. Despite their relation, each LARP is unique in its biological role, target recognition, and mechanism of RNA binding¹⁵⁷.

1.4.2 Classification and roles in RNA metabolism

All LARPs are characterized by an N-terminal La-Motif (LAM) and have been classified into seven subfamilies based on sequence conservation of the LAM¹⁵⁶: LARP1, LARP2, LARP3 (La protein or Genuine La), LARP4, LARP5, LARP6, and LARP7 (Figure 1-6). Despite their common LAM, each LARP is specialized to recognize an RNA subclass and facilitate its

metabolism at a particular stage. Genuine La, the prototypical LARP, transiently binds the UUU_{OH} 3' of premature tRNAs to facilitate their folding and maturation. LARP7 also recognizes poly(U) sequences, but instead stably binds the 3' terminus of the 7SK noncoding RNA to mediate transcription regulation in metazoans¹⁵⁸⁻¹⁶⁰. In stark contrast, LARP6 has evolved to engage a stem-loop within the 5' UTR of collagen mRNAs to enhance their translation¹⁶¹⁻¹⁶³. LARP4 and LARP5, the most closely related to LARP1, bind poly(A) and AU-rich mRNA sequences, respectively, and also enhance transcript stability and translation¹⁶⁴⁻¹⁶⁶. LARP1 regulates the stability and translation of TOP transcripts by recognition of the mRNA cap and TOP motif¹⁶⁷⁻¹⁷¹. Thus, despite seemingly convergent domain features,LARPs exhibit specified RNA target selection and function.

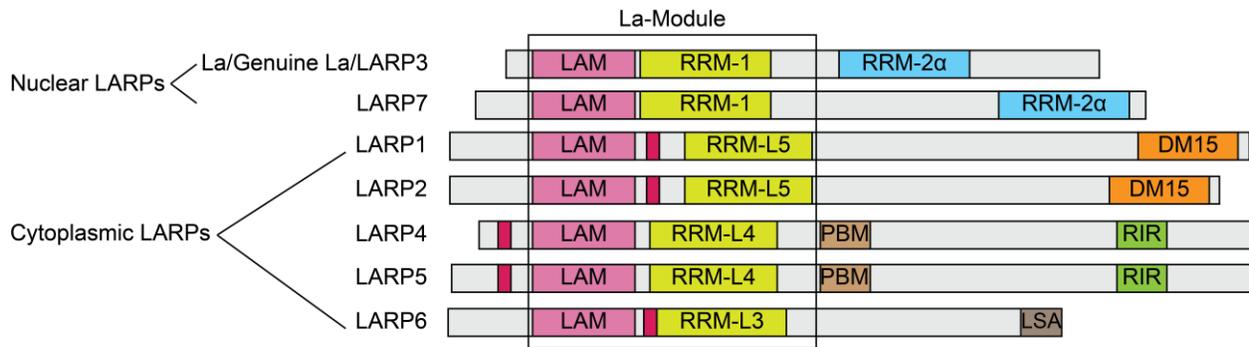


Figure 1-6 Schematic of LARPs and their domain organization. LAM, La-Motif; RRM, RNA recognition motif; RRM-1, the n-terminal RRM in La and LARP7; RRM-2α, the c-terminal RRM in La and LARP7; L5, like 5; L4, like 4; L3, like 3; Dark pink box, PAM2; PBM, PABP interaction motif; DM15, *Drosophila melanogaster* repeat; RIR, RACK-1 interacting region; LSA, LAM and S1-associated motif.

1.4.3 La-Modules

In La and LARPs, the LAM is followed by an RRM or RRM-like domain (Figure 1-6), separated by an interdomain linker¹⁵⁶. Phylogenetic analysis reveals that the LAM co-evolved with the RRM¹⁵⁶ in LARPs. Together, the LAM and RRM comprise the ‘La-Module’.

1.4.4 La-Motif (LAM)

The LAM is a well-conserved 90-amino acid structural motif that adopts a variation of the winged helix-turn-helix fold¹⁷². LAM-containing proteins are found in all eukaryotes suggesting its arrival soon after the archaea-eukarya separation¹⁵⁶. The LAM was first identified in La protein, but has since been found in all LARPs and other non-LARP proteins¹⁵⁶. The winged helix-turn-helix fold is used for interactions with nucleic acids, and sometimes proteins, and is enriched in eukaryotic and prokaryotic transcription factors¹⁷³. Classically, the winged helix-turn-helix motif consists of three α -helices and β -strands in the canonical $\beta\alpha\beta\beta\alpha\alpha$ topology, although many variations exist¹⁷³. The second helix, or recognition helix, typically docks transcription factors in the major groove of DNA¹⁷³. However, in LARPs, a conserved aromatic patch lined by basic residues is used to recognize RNAs¹⁷⁴.

1.4.5 RNA recognition motif (RRM)

An RNA recognition motif (RRM) or RRM-like domain immediately follows the LAM of LARPs (Figure 1-6)¹⁵⁶. The RRM is the most common RNA-binding domain, existing in over 200 human RBPs, with 44% containing two to six RRMs^{175,176}. RRMs are found in all domains of life, including prokaryotes and viruses^{175,176}. A typical RRM contains 80-90 residues with a $\beta_1\alpha_1\beta_2\beta\alpha_2\beta$ topology that folds into four antiparallel β -strands buttressed against two α -helices. Canonical RNA recognition is mediated by two conserved sequence motifs, RNP1 (eight amino acids, [K/R]-G-[F/Y]-[G/A]-[F/Y]-[I/L/V]-X-[F/Y]) and RNP2 (six amino acids, [I/L/V]-[F/Y]-[I/L/V]-X-N-L), located in the central two strands of the RRM^{175,176}. The RNPs are enriched with basic and aromatic residues that hydrogen bond and base stack with RNA nucleotides, respectively^{175,176}. However, RRMs vary in structure and RNA recognition modes^{175,176}. The secondary structure elements and connecting loops can have different lengths that modulate RNA recognition¹⁷⁶. Furthermore, the connecting loops may contain secondary structural elements that enhance RNA specificity¹⁷⁵. For instance, an extra β -strand between β_1 and β_2 in the RRM of T-cell intracellular antigen 1 (TIA-1) enhances specificity for uridine-rich sequences¹⁷⁷. Furthermore, RRM termini can sometimes be conjoined to structural motifs that also enhance RNA

recognition or modulate function^{175,178}. The C-terminal RRM of Genuine La is joined to a helix that serves as a nuclear retention element¹⁷⁹. Some LARPs exhibit more extensive changes to their RRMs, and are thus termed RRM-like¹⁵⁶ domains (Figure 1-6). These include: an absence of RNP1 and RNP2^{163,165} (all LARPs), changes to the conformation and length of loops^{163,165} (all LARPs), as well as helix insertions¹⁶³ (La, LARP6).

1.4.6 Conserved and divergent features of RNA recognition by LARPs

The pairing of the LAM and RRM to form the La-Module is an arrangement unique to LARPs. Initial structural studies of the La-Modules of La¹⁸⁰⁻¹⁸² and LARP7¹⁶⁰ bound to RNA showed that the canonical binding surfaces of the LAM and RRM do not participate in RNA binding. Instead, the La-Module adopts a V-shape, with RNAs making intimate contacts with a hydrophobic patch in the LAM and the edge of RRM β_2 ^{160,180-182}. High specificity for the 3' hydroxyl and penultimate uridylate is a result of interactions with residues from both the LAM and RRM^{180,182}. Importantly, the flexible interdomain linker between the LAM and RRM adopts a helical conformation upon RNA-binding which orients the domains into a V-shape^{181,182}.

The LAM and RRM in LARP6 also bind RNA synergistically¹⁶³. Surprisingly, the LARP6 LAM and RRM adopt an elongated tandem domain orientation rather than a V-shape¹⁶¹. This may be due to the shorter interdomain linker, which is only two amino acids in length, as compared to the eight and nine amino acids in La and LARP7, respectively¹⁶¹. Substitution of the LARP6 linker for that of Genuine La obliterates RNA binding, suggesting strong functional significance of the linker length¹⁶³. In addition to the short linker, the LARP6 La-Module differs in that its RRM contains a helix between α_1 and α_2 that obscures the canonical RNA-binding surface¹⁶³. This helix aids in the recognition of the stem-loop within the 5'UTR of collagen mRNAs¹⁶³. Studies of LARP6 demonstrated that the LAM and RRM can synergistically bind RNA using varied topological arrangements¹⁶³.

Recently the LARP4 La-Module was also shown to tandemly orient the LAM and RRM, which may also be attributed to the short interdomain linker of four amino acids¹⁶⁵. However, the LARP4 La-Module was shown to play a minor role in RNA recognition. Instead an unstructured region N-terminal (NTR) to the LAM has high affinity for poly(A) RNA, with some binding also

contributed by the *canonical* face of the RRM¹⁶⁵. The NTR is highly flexible, undergoing conformational changes on the picosecond timescale, and toggles between an open and closed conformation relative to the La-Module¹⁶⁵. The NTR also binds PABP, suggesting potential interplay between LARP4, poly(A) RNA, and PABP¹⁶⁴.

Thus, the interdomain linker appears to govern relative domain orientations within the La-Module that guide functional topologies. Different domain orientations of the LAM and RRM might exist within each LARP to allow binding to distinct RNA targets. In addition, novel secondary structural elements and unstructured regions enhance RNA affinity and specificity, and can also mediate protein-protein interactions¹⁵⁷. The exploitation of conformational plasticity combined with unique structural motifs may allow each LARP to bind and regulate the metabolism of unprecedented RNAs.

1.4.7 C-terminal RNA binding domains

While the La-Module is proposed to be the main locus of RNA recognition, additional motifs and domains within each LARP enhance specificity and function (Figure 1-6). La and LARP7 contain a second RRM towards their C-termini dubbed RRM2 α ^{183,184}. RRM2 α contains a third helix that lies across the face of the β -sheet¹⁸³. In La, RRM2 α does not independently bind RNA, but will bind internal stem-loop structures in synergy with the La-Module, such as in microRNAs and the HCV IRES¹⁸⁵. Addition of ~20 basic residues downstream of RRM2 α increases binding¹⁸⁴. The LARP7 RRM2 α is required for binding of LARP7 to 7SK RNA, and recognizes unpaired and base paired nucleotides at the apical loop of hairpin 4¹⁸⁶.

LARP6 contains a unique LSA (LAM and S1 associated) motif at the C-terminus¹⁵⁷. Although its function is not characterized, the LSA consists of 20-30 amino acids that are appended to cold-shock domains in cold-shock response protein 1 (CSP1) and are putative nucleic acid binding motifs¹⁵⁷.

LARP1 and LARP2 contain an ~150 amino acid long C-terminal DM15 region that is unique in the human proteome and displays high sequence conservation¹⁵⁶. The LARP1 DM15 adopts a HEAT-like fold containing three helix-turn-helix repeats, which bind the cap and TOP

motif of mRNAs encoding the translation machinery to promote their stability and translation^{168,171}.

1.4.8 Protein binding motifs

Although LARPs are RBPs, each one contains at least one protein-binding motif¹⁵⁷. Protein-protein interactions allow LARPs to shuttle between the nucleus and cytoplasm, bind PABP, and associate with the ribosome and other translation factors¹⁵⁷. Interactions between LARPs and other proteins may be critical to changing mRNP composition and dynamics in order to regulate localization, stability, and translation.

La and LARP7 predominantly function in the nucleus, but can shuttle to the cytoplasm to regulate cellular and viral mRNA translation¹⁵⁷. La protein contains a nuclear export element within the RRM, as well as a nuclear retention element and nuclear localization signal (NLS) towards its C-terminus¹⁵⁷. Differential phosphorylation of these sequence motifs regulates nuclear-cytoplasmic shuttling of La¹⁵⁷. While LARP6 is predominantly cytoplasmic, it also contains a nuclear localization sequence, although the function of nuclear LARP6 is uncharacterized¹⁵⁷.

The cytoplasmic LARPs, namely 1, 2, 4, 5, 6, contain a PABP-interacting motif 2 (PAM2) motif that binds the MLE domain of PABP¹⁵⁷ (Figure 1-6). LARPs 4 and 5 contain an additional PABP-Interacting Motif (PBM) that aids in PABP binding¹⁵⁷ (Figure 1-6). Also exclusive to LARPs 4 and 5 is the RACK1-Interacting Motif (RIR) towards their C-termini¹⁵⁷ (Figure 1-6). The RIR interacts with the RACK1 scaffolding protein, an integral component of the 40S ribosomal subunit with roles in cap-dependent and cap-independent translation, RNA decay, ribosome quality control, and more¹⁵⁷.

1.4.9 LARPs in disease

La protein was originally identified as an autoantigen in patients with systemic lupus and Sjogren's syndrome. Since then, the discovery of other LARPs has been concomitant with our understanding of their role in various diseases. Loss of function mutations in LARP7 cause

impaired telomere maintenance, severe intellectual disability, and facial dysmorphism symptoms in patients suffering from Alzami syndrome¹⁸⁷.

Thus far, the cytoplasmic LARPs have been correlated with cancer phenotypes in various cancers¹⁸⁸. LARP6 enhances proliferation, lamellipodia formation, and invasion in breast cancer cells¹⁸⁹. These functions appear to be linked to uncharacterized nuclear functions, as deletion of the LARP6 NLS inhibits these effects¹⁸⁹. LARP4 appears to regulate cancer cell migration and invasion; LARP4 knockdown increases migration and invasion in prostate cancer cells^{190,191}. Mutational analysis suggested that the anti-invasion properties of LARP4 are linked to its interaction with PABP¹⁹¹. LARP5 acts as a tumor suppressor¹⁹² and an oncoprotein^{193,194}, depending on the cell type. Overexpression of LARP5 induces mitotic arrest and apoptosis that appears to be dependent upon the La-Module in some cancer lines¹⁹². In contrast, LARP5 appears to be an oncoprotein in acute myeloid leukemia mouse models¹⁹⁴.

LARP1 expression levels are correlated with cancer progression by increasing cell proliferation, migration, invasion, and tumorigenicity in mice¹⁹⁵. LARP1 levels are increased in epithelial malignancies as compared to adjacent normal tissues¹⁸⁸. In epithelial cancers, such as cervical and lung cancer, LARP1 is linked to disease progression and is an independent marker of patient prognosis¹⁹⁶. RIP-chip studies suggest that the LARP1 interactome in HeLa cells is particularly enriched for oncogenic mRNAs, including anti-apoptotic, focal adhesion, and actin remodeling factors¹⁹⁷. In ovarian cancer, LARP1 increases chemotherapeutic resistance perhaps by stabilizing certain oncogenic transcripts through their 3' UTRs, although further studies are needed to confirm a direct interaction that enhances transcript stability¹⁹⁸.

1.5 La-Related Protein 1 (LARP1)

1.5.1 Association with poly(A) RNA and PABP

LARP1 from human cell extracts immunoprecipitates with poly(A), but not poly(U), (G) or (C) RNAs¹⁹⁹. LARP1 associates with PABP^{167,170} *via* a putative PAM2 motif located between the LAM and RRM²⁰⁰, and co-sediments with PABP in polysome profiling gradients¹⁶⁷. More

recently, LARP1 has emerged as a key regulator of TOP mRNA translation downstream of mTORC1. For over 20 years, TOP mRNA was known to display hypersensitivity to mTORC1 signaling, but the mediator between mTORC1 and TOP mRNA translation was missing¹⁴⁵. The identification of LARP1 as a regulator of TOP mRNA translation downstream of mTORC1 is critical to our understanding of ribosome biogenesis and the synthesis of other translation machinery that are encoded by TOP mRNAs.

1.5.2 LARP1 regulates TOP mRNA stability and translation

LARP1 is a direct substrate of mTORC1^{167,201-203} that regulates the stability and translation of TOP mRNAs, although its role in translation has been debated^{167-170,195,199,201}. A few studies suggest that LARP1 increases mRNA translation, as shown by polysome profiling gradients^{170,195}. Yet other studies report an inhibitory role for LARP1 in TOP mRNA translation^{167-169,201}. The LARP1 DM15 region recognizes the cap^{168,169} and TOP motif^{168,169,171}, and outcompetes eIF4E for binding capped TOP motif sequences *in vitro*¹⁶⁸. These data are consistent with the preference of eIF4E for +1G or +1A transcripts, and its lowest affinity for +1C²⁶. Occlusion of eIF4E from TOP mRNAs would inhibit cap-dependent translation by preventing the assembly of the eIF4F translation initiation complex. Consistent with this, RNA-binding mutations to the DM15 region decreases the association of LARP1 with TOP mRNAs in cells¹⁶⁸ and LARP1 represses the translation of TOP mRNAs *in vitro* through the combined recognition of both the 5' cap and TOP motif¹⁶⁹. Interestingly, phosphorylation by mTORC1 decreases the affinity of LARP1 for TOP mRNA 5' UTRs and permits TOP mRNA translation^{167,169,201}. Taken together, the current data indicate a role for LARP1 in regulating TOP mRNA translation downstream of mTORC1, with mTORC1-mediated phosphorylation governing whether LARP1 represses or permits TOP mRNA translation.

Indeed, LARP1 contains up to 26 serine and threonine residues, whose phosphorylation states are controlled by the mTORC1 pathway²⁰⁴. These residues lay within seven clusters dispersed throughout LARP1²⁰⁴. Mutations to residues within clusters 4 and 5 – which reside within the linker between the RRM and DM15 – impair the association of LARP1 with mTORC1. Furthermore, RNA-binding assays show that an extended DM15 region encompassing these

residues has higher affinity for TOP motif sequences as compared to its phosphomimetic counterpart²⁰⁴. These data may indicate that mTORC1-mediated phosphorylation of LARP1 dictates whether the DM15 region binds to the 5' cap and adjacent TOP motif. In this case, eIF4E can now bind and initiate cap-dependent translation of TOP mRNAs that are freed from the LARP1 DM15 region.

1.6 Summary of goals and discoveries

Prior to the work presented in this thesis, the role of the LARP1 C-terminal DM15 region in TOP mRNA translation was an active area of research. However, the contribution of the N-terminal La-Module was unknown. Herein we provide the first characterization of the LARP1 La-Module.

We began our study by identifying the LARP1 La-Module RNA targets. Consistent with the association of LARP1 with PABP and poly(A) RNA, we found that the LARP1 La-Module binds poly(A) RNA. We further show that the La-Module also binds some TOP motifs, and is able to simultaneously engage both poly(A) and TOP motif RNA. We next optimized the LARP1 La-Module construct from that which was suggested in the literature¹⁵⁶, due to difficulties in its expression, purification, and RNA-binding activity¹⁵⁶. We worked towards identifying the RNA-binding surfaces of La-Module, as well as its stoichiometry when bound to both poly(A) and TOP motif RNA simultaneously.

We also investigated the LARP-PABP interaction. We found that LARP1 and PABP directly bind through the La-Module and MLLE domains, respectively, and that the interaction is RNA-independent. The La-Module binds MLLE through a PAM2 motif located between the LAM and RRM. In addition, we identified a potential interaction between the La-Module and PABP RRM regions through a putative PAM1 sequence.

Finally, in addition to our investigation of human LARP1, we also identified a microalgae homologue of LARP1 for use in crystallographic studies to better understand interdomain interactions between the N- and C- termini. Changes to the overall conformation of LARP1 through intramolecular interactions may provide LARP1 with the ability to engage various TOP

mRNAs and regulate translation. With these studies, we also hope to identify druggable pockets that are unique to LARP1 for chemotherapeutic treatments.

We present the first characterization of the LARP1 La-Module in TOP mRNA recognition. Our findings that the LARP1 La-Module binds poly(A) RNA, TOP motifs, and PABP, build upon our current model for how LARP1 may recognize and regulate the translation of TOP mRNAs. Our discovery that the La-Module can simultaneously engage poly(A) and TOP motif RNAs expands the known RNA binding modes in LARPs.

2.0 The LARP1 La-Module Recognizes Both Ends of TOP mRNAs

This is the **Author's Version of Record** of an article published by Taylor & Francis in *RNA Biology* on **October, 10 2019** available at the Taylor & Francis Ltd web site²⁰⁵. Link to the article: <https://www.tandfonline.com/doi/full/10.1080/15476286.2019.1669404>.

2.1 Introduction

La-related Proteins (LARPs) are a diverse family of RNA-binding proteins that are conserved throughout eukaryotic evolution and function in nearly all stages of RNA metabolism^{156,206,207}. Each LARP subfamily has evolved to recognize and regulate the metabolism of particular RNAs. The target RNAs of these subfamilies range from non-coding RNAs to messenger and viral RNAs.

LARPs are characterized by an N-terminal La-Motif (LAM) and are classified into subfamilies based on sequence conservation of the LAM: LARP1, LARP2, Genuine La (LARP3), LARP4, LARP5, LARP6, and LARP7¹⁵⁶. The LAM adopts a winged helix-turn-helix fold¹⁷² and is followed by an RNA recognition motif (RRM). The RRM has a canonical $\beta\alpha\beta\alpha\beta$ structure or an RRM-like (RRM_L) fold, predicted to fold similarly to an RRM, but lacking the consensus RNP motifs^{156,175,208}. Together, the LAM and RRM comprise the 'La-Module.' In addition to the La-Module, most LARPs have a C-terminal RNA-binding domain that enhances RNA recognition and biological function^{156,206,207}.

Phylogenetic analyses reveal that the LAM and RRM co-evolved in LARPs, suggesting the importance of the entirety of the La-Module in RNA recognition¹⁵⁶. Consistent with this, initial structural and biochemical studies in Genuine La and LARP7 show that the canonical RNA-binding surfaces of the LAM and RRM do not engage RNA^{160,180,181,207}. Instead, the LAM and RRM function synergistically; RNAs make intimate contacts with a conserved patch of hydrophobic residues in the LAM, and are supported by contacts with the edge of the RRM β 2 strand^{160,180,181}. Similarly, in LARP6 the LAM and RRM synergistically bind RNA, but also

require participation of the interdomain linker¹⁶³. However, in LARP4 the La-Module plays a minor role in RNA recognition¹⁶⁵. Instead, intrinsically disordered regions N-terminal to the La-Module drive RNA-binding¹⁶⁵. Thus, while the main RNA-binding unit is conserved, additional unstructured features and conformational plasticity within the La-Module coordinate RNA recognition.

Although all LARPs share a La-Module, sequence and structural variations give rise to distinct RNA binding specificities and roles in RNA metabolism^{156,157,207}. In Genuine La, the prototypical LARP, the La-Module binds the UUU_{OH} 3' termini of premature tRNAs to promote their folding and maturation^{180,181,209}. The LARP7 La-Module also recognizes UUU_{OH}, however it binds the 3' terminus of 7SK non-coding RNA to facilitate transcription regulation^{160,210}. The LARP6 La-Module evolved to engage a stem-loop within the 5'UTR of collagen mRNAs to increase their translation^{161,162}. The La-Modules of LARP4 and LARP5, the most closely related to LARP1, bind poly(A) and AU-rich regions of mRNAs, respectively, and also stimulate mRNA stability and translation^{164,165,211,212}. Thus, despite seemingly convergent features, La-Modules display adaptable and specified RNA target selection.

LARP1 regulates the stability and translation of mRNAs that encode components of the translation machinery, such as ribosomal proteins and translation factors^{153,167-170,199,201}. These transcripts, known as TOP mRNAs, are characterized by a terminal oligopyrimidine (TOP) motif in the 5'UTR, immediately after the 5'cap¹⁴⁵. The TOP motif, comprised of 4-14 pyrimidines followed by a GC-rich region¹⁴⁵, allows for the coordinated translation of TOP mRNAs downstream of mTORC1⁵. Under conditions of mTORC1 inhibition, the LARP1 C-terminal DM15 region binds the 5' cap and TOP motif^{168,169}. This obstructs the formation of the translation initiation complex and thereby represses TOP mRNA translation during metabolically unfavorable conditions^{168,169}.

However, thus far, the role of the LARP1 La-Module in recognizing TOP mRNAs and regulating their translation is unknown. A few lines of evidence led us to hypothesize that the La-Module binds the poly(A) tails of TOP mRNAs. First, LARP1 associates with Poly(A)-Binding Protein (PABP) *via* a putative PABP-interacting motif 2 (PAM2) located between the LAM and RRM^{167,170}. Second, LARP1 co-sediments with PABP through polysome gradients with TOP mRNAs^{167,170,195}. Finally, LARP1 from human cell extracts immunoprecipitates with poly(A) RNA, but not with poly (U), poly(C), or poly(G)¹⁹⁹.

Here we show that the LARP1 La-Module directly engages poly(A) RNA. Unexpectedly, we find that the La-Module also binds some TOP motifs, in a cap-independent manner, with similar affinity to poly(A) RNA. We also present evidence that the LARP1 La-Module can simultaneously bind poly(A) and the TOP motif RNA. Integration of these data with the established roles of the DM15 region allow us to assemble a model for the role of LARP1 in TOP mRNA recognition that reconciles many of the observations of LARP1 published hitherto. We establish a direct role for LARP1 at both the 5' and 3' ends of TOP mRNAs, with the La-Module recognizing distinguishing features at either end of TOP transcripts.

2.2 Materials and methods

2.2.1 La-Module (amino acids 310-540) cloning, expression, and purification

The La-Module region of the LARP1 coding sequence (amino acids 310-540 from 1019-amino acid isoform LARP1a) (Integrated DNA Technologies) was PCR amplified and inserted into a pET28a vector (Novagen Inc) using NdeI and BamHI sites. The resulting construct expresses the La-Module with an N-terminal 6XHis tag followed by a Tobacco Etch Virus (TEV) protease cleavage site. The 6XHis-La-Module fusion protein was expressed in *E. coli* BL21(DE3) and cultured at 37°C for two hours, shifting to 17.5°C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

Cells were resuspended in lysis buffer [50 mM Tris-HCl, pH 8.0, 400 mM NaCl, 10 mM imidazole, 10% v/v glycerol, one cOmplete EDTA-free Protease Inhibitor™ tablet (Roche)]. Cells were lysed using homogenization and lysate was cleared *via* centrifugation. The 6XHis-La-Module was purified in batch using nickel agarose affinity chromatography (ThermoScientific) and eluted with 50 mM Tris-HCl, pH 8.0, 400 mM NaCl, 300 mM imidazole, 10% v/v glycerol. The 6XHis tag was removed by cleavage with 0.5 mg TEV per 10 mL eluate overnight at 4°C in dialysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 0.5 mM EDTA, 0.5 mM DTT). Nucleic acid and protein contaminants were removed using HiTrap Heparin followed by tandem HiTrap S and HiTrap QP (GE Healthcare Lifesciences) chromatography with an NaCl

gradient (150 mM-1M), with La-Module eluting from the Heparin and Q columns. Fractions containing the La-Module were collected, concentrated, dialyzed into storage buffer (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25% glycerol, 4 mM DTT), frozen in liquid nitrogen, and stored at -80°C. Samples for ITC were exchanged in 50 mM sodium phosphate pH 7.6, via 10K MWCO Centriprep concentrator and assessed for folding via circular dichroism spectroscopy (Supp. Fig. 2) prior to use.

2.2.2 RNA oligonucleotides

RNA oligonucleotides (Sigma-Aldrich) have the following sequences: Poly(A)
 RNA: 5'-AAAAAAAAAAAAAAAAAAAAAAAAA_{OH} Poly(G) RNA:
 5'-GGGGGGGGGGGGGGGGGGGG_{OH}
 Poly(C) RNA: 5'-CCCCCCCCCCCCCCCCC_{OH}
 Poly(U) RNA: 5'-UUUUUUUUUUUUUUUUUUU_{OH}
 Poly(A) RNA-PO₄: 5'-AAAAAAAAAAAAAAAAAAAAAAAAA_{PO₄}
 RPS6 TOP 20-mer: 5'-CCUCUUUCCGUGGCGCUC
 RPS6 5'UTR: 5'-CCUCUUUCCGUGGCGCCUCGGAGGCGUUCAGCUGCUUCAAG
 RPS6 5'UTR ΔTOP: 5'-GUGGCGCCUCGGAGGCGUUCAGCUGCUUCAAG
 RPS18 5'UTR:
 5'-CUCUCUUCCACAGGAGGCCUACACGCCGCCGCUUGUGCUGCAGCC
 PABPC1 42-mer: 5'-CCUUCUCCCCGGCGGUUAGUGCUGAGAGUGCGGAGUGUGUG
 RPL13A 5'UTR: 5'-CCUUUCCAAGCGGCUGCCGAAG
 RPL13A 5'UTR ΔTOP: 5'-AAGCGGCUGCCGAAG
 RPS6 3'UTR: 5'-AAGAUUUUUUGAGUAACAAAU
 Loading Control: 5'-CCAGUCAUGCUAGCCAUAUGCCUGGUCCGCCUGUUGC
 Recovery Control:
 5'-UCCUGAAUGCUACGUUAAUCGGUAUCCAGCAGUUCUUUCAGUUU
 CG

2.2.3 La-Module electrophoretic mobility assays

RNA oligonucleotides were 5'-end labeled with [γ - 32 P]-ATP (Perkin Elmer) using T4-polymerase kinase (New England Biolabs) and gel purified. Capped RPS6 20-mer was prepared as previously described¹⁶⁸. La-Module 5X stocks were prepared in dilution buffer (50 mM Tris-HCl pH, 7.5, 100 mM NaCl, 25% glycerol, 4 mM DTT). 10 μ L binding reactions contained final concentrations of 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 7.5 % v/v glycerol, 1mM DTT, 1 μ M BSA (Thermo Fisher Scientific), \leq 2 nM radiolabeled RNA, and either: 10 U/mL poly(dI-dC) (Sigma-Aldrich), 5 μ M Yeast tRNA (Ambion) or 3 μ M salmon sperm DNA (Thermo Fisher Scientific). La-Module was titrated at 0, 0.01, 0.03, 0.1, 0.3, 1, 3, 10 μ M. Reactions were incubated on ice for 30 min then run on 7% polyacrylamide (29:1) native 0.5X TBE gels at 125 V for 45 min at 4°C. Exposed phosphor screens (GE Healthcare Lifesciences) were imaged on a Typhoon FLA plate reader (GE Healthcare Lifesciences) and quantitated using Imagequant TL (GE Healthcare Lifesciences). Dissociation constants were determined by plotting (KaleidaGraph) the fraction of shifted RNA versus the concentration of protein after band intensities were corrected for background (ImageQuant)TL.

2.2.4 La-Module competition assays

La-Module was pre-bound to radiolabeled RNA on ice in 8 μ L reactions as described above. 2 μ L 5X cold competitor RNA was titrated to yield final concentrations of 0, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5, 10 μ M. Competition reactions were incubated on ice for an additional 30 min prior to loading and run as described above. Gels were dried, exposed overnight, and analyzed as described for EMSAs.

2.2.5 La-Module biotin pull-downs

3' Biotinylated RNA oligonucleotides (IDT) were pre-bound to streptavidin magnetic beads and 1 μ M La-Module was pulled down as per kit instructions (Pierce # 20164), with the

exception that the La-Module was incubated in 1X binding buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 7.5% glycerol, 1 mM DTT) in the presence of 1 μ M BSA and 10 U/mL poly(dI-dC) (Sigma-Aldrich) as non-specific competitors. Beads were washed twice with 1X binding buffer to remove excess La-Module prior to the addition of ≤ 25 nM 5' radiolabelled RNA in 100 μ L 1X binding buffer in the presence of non-specific competitors as described above. Reactions were incubated for an additional 30 min at 4°C, then washed with 100 μ L 1X binding buffer to remove unbound radiolabeled RNA. Flow through and washes were ethanol precipitated and resuspended in 10 μ L 47.5% formamide, 0.1% bromophenol blue. Beads were resuspended in 2X formamide to a final volume of 10 μ L. All samples were loaded onto 10% polyacrylamide (29:1) 1X TBE 7M urea sequencing gels and run at 90 W for 45 min at room temperature. Gels were dried, exposed overnight, and analyzed as described for EMSAs above.

2.2.6 Isothermal titration calorimetry

Purified La-module was titrated with poly(A) or poly(C) RNA in a TA instruments nanoITC to confirm selectivity. Representative traces (Supp. Fig. 2) used 300 μ M polyC RNA titrated into 110 μ M La-module or 200 μ M polyA RNA into 85 μ M La-Module with both titrant and titrand in 50 mM sodium phosphate, pH 7.6, at 25°C. Initial injection was 0.18 μ L and subsequent 33 injections were 1.49 μ L each into an active cell volume of 200 μ L (overfilled to 300 μ L). Stirring speed was 300 rpm with a 100 second injection interval and 1 second data collection interval.

2.2.7 Circular dichroism spectroscopy

La-module folding was assessed in a JASCO J1500 circular dichroism spectrophotometer using 20 μ M La-module in 50 mM sodium phosphate, pH 7.6. Spectra were collected in a 0.5 mm cuvette from 280 to 180 nm at 25°C with 0.5 nm steps at a rate of 50 nm/min. Eight repetitions were averaged and smoothed with a 5 nm Savitzky-Golay filter²¹³. Ellipticity was normalized to mean residue molar ellipticity (Supp. Fig. 2).

2.2.8 WT and REYA LARP1 cloning, expression, and purification from Expi293T cells

The construct encoding 6XHis-TEV-6XGlycine- WT or REYA LARP1 sequence (LARP1 isoform 2) was PCR amplified and inserted into a pCMV6 vector using and Mlu1 and AsiSI/SfaAI sites for expression in human cells. Expi293T cells were cultured and transfected as per instructions (Thermo Fisher Scientific #14527). Cells were harvested, and the resulting pellet was frozen in liquid nitrogen then stored at -80°C .

For purification, the cell pellet from a 30 mL culture was resuspended in lysis buffer [25 mM HEPES, pH 8.0, 600 mM NaCl, 5 mM imidazole, 5 % v/v glycerol, protease inhibitors (10 μM leupeptin, aprotinin, 10 μM bestatin, 1 μM pepstatin, 10 μM PMSF)] by pipetting. Cells were lysed by passing twice through a medium gauge needle and the lysate was cleared by centrifugation. The WT or REYA LARP1 fusion proteins was then purified by nickel affinity chromatography using a HiTrap Nickel FF (GE Healthcare Lifesciences). The lysate was loaded manually at 2 mL/min and washed with 10 mL of each wash buffer [lysis buffer supplemented with: 10, 15, 20, and 30 mM imidazole]. The protein was eluted with two 10 mL elutions [25 mM HEPES, pH 8.0, 600 mM NaCl, 250 mM imidazole, 5 % v/v glycerol, (10 μM leupeptin, aprotinin, 10 μM bestatin, 1 μM pepstatin, 10 μM PMSF) protease inhibitors]. The eluate was then diluted to 10 mM imidazole in lysis buffer, then manually loaded, washed, and eluted once more as described. The eluate was buffer exchanged [25 mM HEPES, pH 8.0, 600 mM NaCl, 5 mM imidazole, 10 % v/v glycerol, 1mM DTT] and concentrated to 5-6.5 μM using a Vivaspin 50K MWCO Centrifugal Concentrator (Sartorius), frozen in liquid nitrogen, and stored at -80°C until use.

2.2.9 WT and REYA LARP1 electrophoretic mobility shift assays

Gel shift assays were performed as described for the La-Module (2.2.3) with the following exceptions: 1) reaction buffers contained a final concentration of 20 mM HEPES, pH 7.5, 120 mM NaCl, 7.5% glycerol, and 2) final protein concentrations were 0, 0.01, 0.03, 0.1, 0.3, 1 μM .

2.2.10 WT and REYA LARP1 biotin pull-downs

WT and REYA LARP1 pull-downs were conducted as described above for the La-Module, with the exception that the 1X binding buffer contained 20 mM HEPES, pH 7.5, 120 mM NaCl, 7.5% glycerol.

2.3 Results

2.3.1 The LARP1 La-Module binds poly(A) RNA

To identify the RNA targets of the LARP1 La-Module we analyzed *in vitro* binding assays by native gel electrophoresis. Because LARP1 associates with poly(A) RNA¹⁹⁹, and also with PABP *via* a PAM2 located within the La-Module^{167,170}, we hypothesized that the La-Module recognizes the poly(A) tails of TOP mRNAs. To test this, we determined the relative affinity of purified recombinant human LARP1 La-Module to homopolymeric RNA sequences using electrophoretic mobility shift assays (EMSAs) (Figure 2-1). EMSAs conducted with a 25-nucleotide poly(A) RNA in the presence of non-specific competitors demonstrated that the La-Module directly binds this sequence with an apparent K_d of 40 ± 1 nM (Figure 2-1A, B); the shifted poly(A) oligonucleotide largely remained in the well, possibly due to the formation of multimeric protein-RNA complexes (Figure 2-1A); as the recombinant protein was observed to be folded (Figure 2-2 A), and isothermal titration calorimetry (ITC) experiments confirmed the specific binding of the La-Module to poly(A) RNA (Figure 2-2 B), non-specific aggregation in the well is unlikely. Further, substitution of tRNA or salmon sperm DNA for poly(dI-dC) as a non-specific competitor allowed the La-Module-poly(A) complex to enter the gel (Figure 2-3).

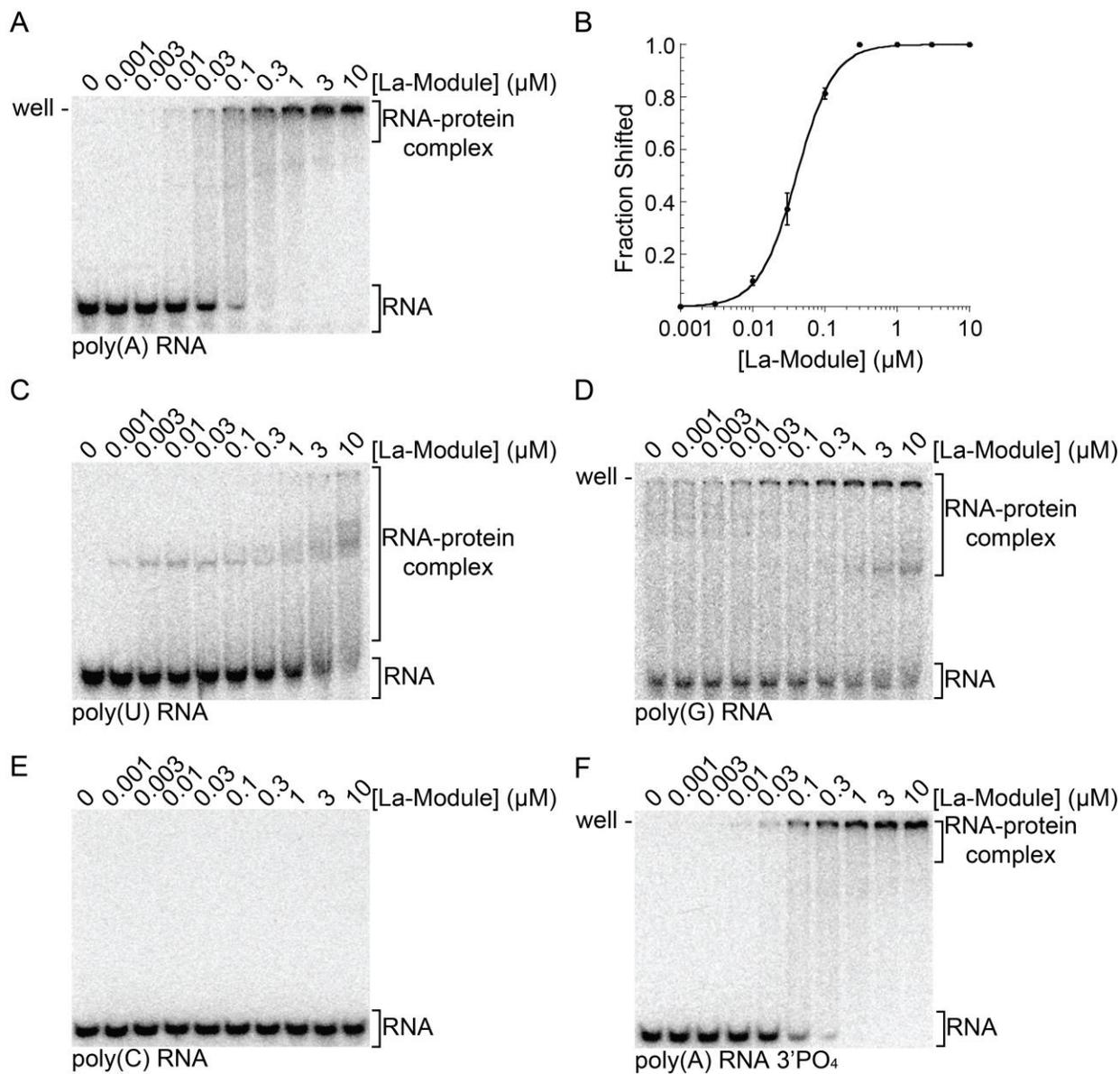
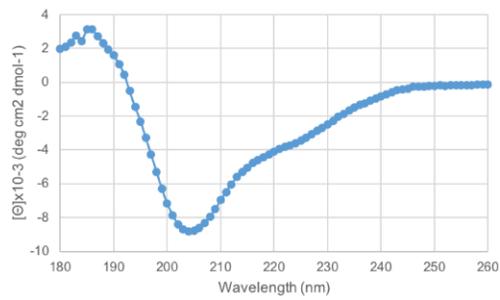


Figure 2-1 LARP1 La-Module binds poly(A) RNA. (A) EMSA analysis of binding assays of WT La-Module with poly(A) 25-mer. (B) Quantification of three independent EMSAs of WT La-Module with poly(A) RNA. Bars are standard deviation. (C-F) EMSAs analyzing WT La-Module with (C) poly(U) 20-mer, (D) poly(G) 19-mer (E) poly(C) 20-mer, and (F) poly(A) 25-mer 3'PO₄.

A



B

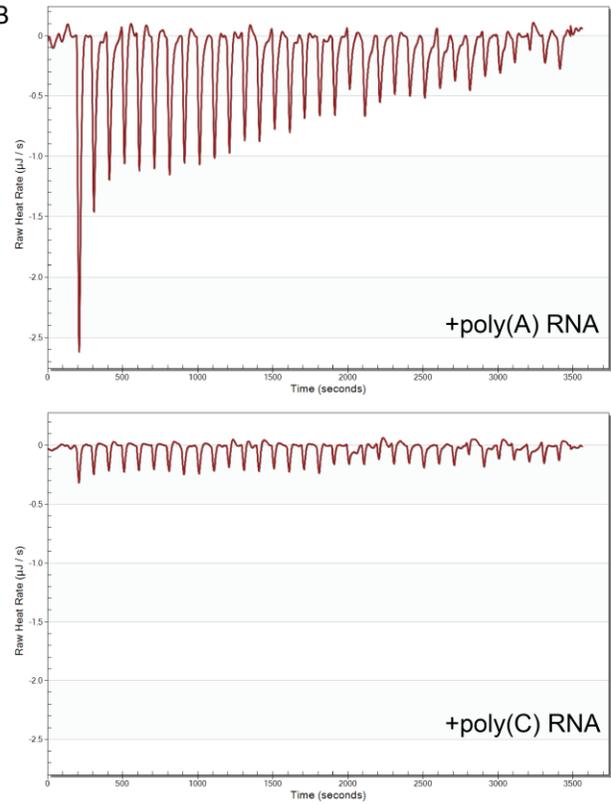


Figure 2-2 Biophysical characterization of the LARP1 La-Module demonstrates it is folded and interacts with poly(A), but not poly(C) RNA. (A) Far-UV CD spectrum of 20 μM LARP1 La-Module. (B) Thermograms of the heat released in isothermal titration calorimetry experiments analyzing the LARP1 La-Module with poly(A) RNA (top) and poly(C) RNA (bottom).

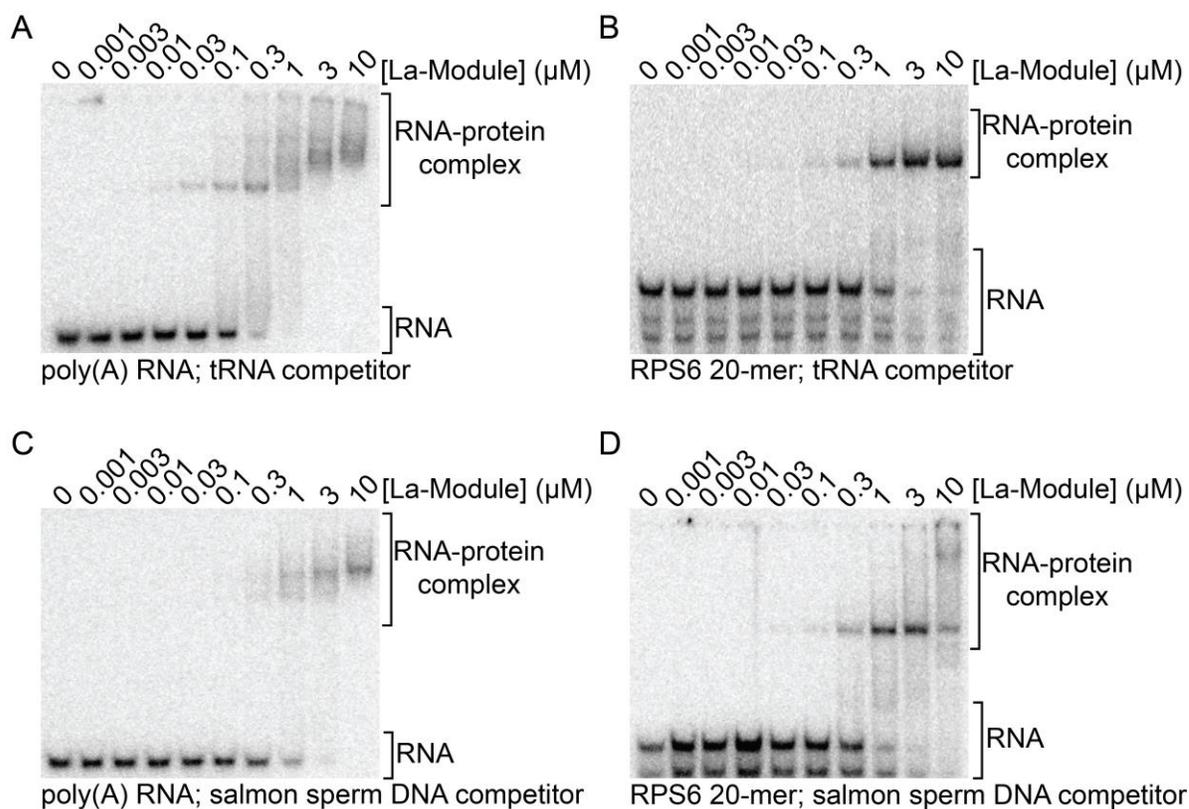


Figure 2-3 The La-Module poly(A) RNA complex is maintained the presence of different nucleic acid nonspecific competitors. EMSAs of binding assays of WT La-Module with: (A) poly(A) and (B) RPS6 20-mer RNA in the presence of tRNA, and (C) poly(A) and (D) RPS6 20-mer RNA in the presence of salmon sperm DNA.

We also determined the affinities of the La-Module for each of the other homopolymeric oligonucleotides. Importantly, the LARP1 La-Module bound poly(U) RNA, the cognate binding target of Genuine La and LARP7, with poor affinity (Figure 2-1 C). Similarly, the La-Module had low affinity for poly(G) RNA, and did not bind poly(C) RNA (Figure 2-1 D, E); ITC experiments confirmed that the La-Module does not recognize poly(C) RNA (Figure 2-2 B).

Because the chemical moiety at the 3' end of the RNA is required for binding by Genuine La and LARP7^{160,180,181}, we then tested the binding of the LARP1 La-Module with a poly(A) 25-mer modified with a 3'PO₄ group. While it did not abolish binding, assays conducted with this modified poly(A) RNA showed a two-fold decrease in apparent affinity of 85 ± 5 nM (Figure 2-1 F), suggesting a small role for the role of the 3' chemical moiety in LARP1 La-module recognition.

2.3.2 The LARP1 La-Module binds various TOP motifs

Given that the La-Module binds poly(A) RNA, we predicted that it would not bind the 5'UTRs of TOP transcripts. Surprisingly, the LARP1 La-Module shifts an oligonucleotide representing the complete 42-nucleotide 5'UTR of RPS6 mRNA (Figure 2-4 A). Furthermore, deletion of the first 10 nucleotides corresponding to the TOP motif abrogated binding (Figure 2-4 B). To validate the interaction between the La-Module and RPS6 TOP motif, we tested a 20-mer RNA representing the first 20 nucleotides of the RPS6 5'UTR (RPS6 20-mer), containing the 10-nucleotide pyrimidine tract followed by the GC-rich region. The La-Module bound the RPS6 20-mer with an apparent affinity of 31 ± 3 nM, comparable to the 40 ± 1 nM affinity observed for poly(A) RNA (Figure 2-4 C, E). Interestingly, unlike the DM15 region¹⁶⁸, the affinity of the La-Module for the RPS6 20-mer does not increase upon capping of the oligonucleotide, showing an apparent affinity of 32.8 ± 5 nM (Figure 2-4 D, E). The La-Module also shifted an oligonucleotide representing the RPL13A 5'UTR, and accordingly, deletion of the TOP motif pyrimidine tract also abrogated binding (Figure 2-4 F, G).

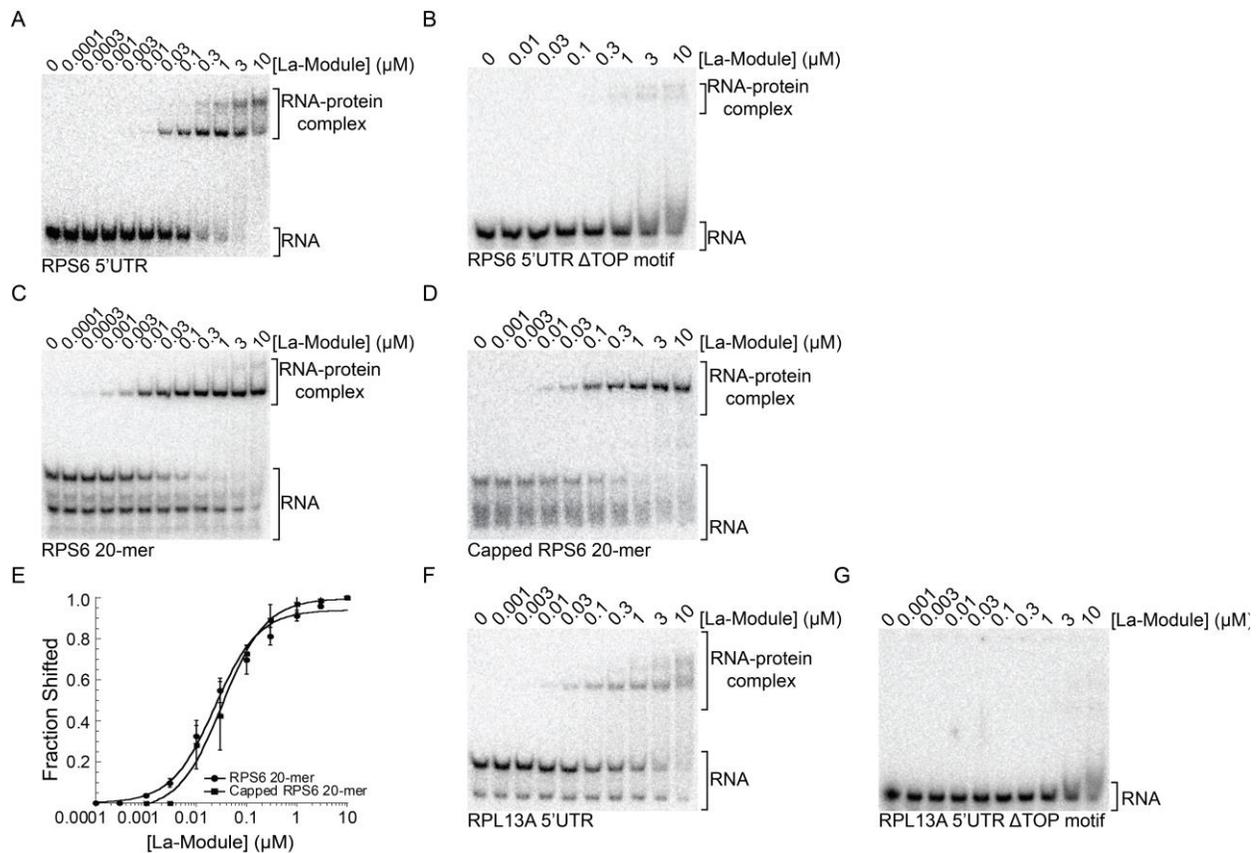


Figure 2-4 LARP1 La-Module binds TOP mRNA 5' UTRs in a TOP motif-dependent manner. EMSAs of binding assays of WT La-Module with (A) RPS6 5'UTR, (B) RPS6 5'UTR lacking the polypyrimidine region of the TOP motif, (C) RPS6 20-mer, and (D) Capped RPS6 20-mer. (E) Quantification of three independent EMSAs of WT La-Module with indicated RNAs. Bars are standard deviation. EMSAs of binding assays of the LARP1 La-Module with (F) RPL13A 5'UTR and (G) RPL13A lacking the polypyrimidine region of the TOP motif.

The La-Module did not bind oligonucleotides corresponding to the 5'UTR of RPS18 and PABPC1 mRNA (Figure 2-5 A, B), suggesting that, like the DM15 region¹⁷¹, the La-Module does not bind all TOP sequences. We next tested whether the LARP1 La-Module recognizes TOP mRNA 3'UTRs; it did not shift an oligonucleotide representing the 3'UTR of RPS6 mRNA (Figure 2-5 C).

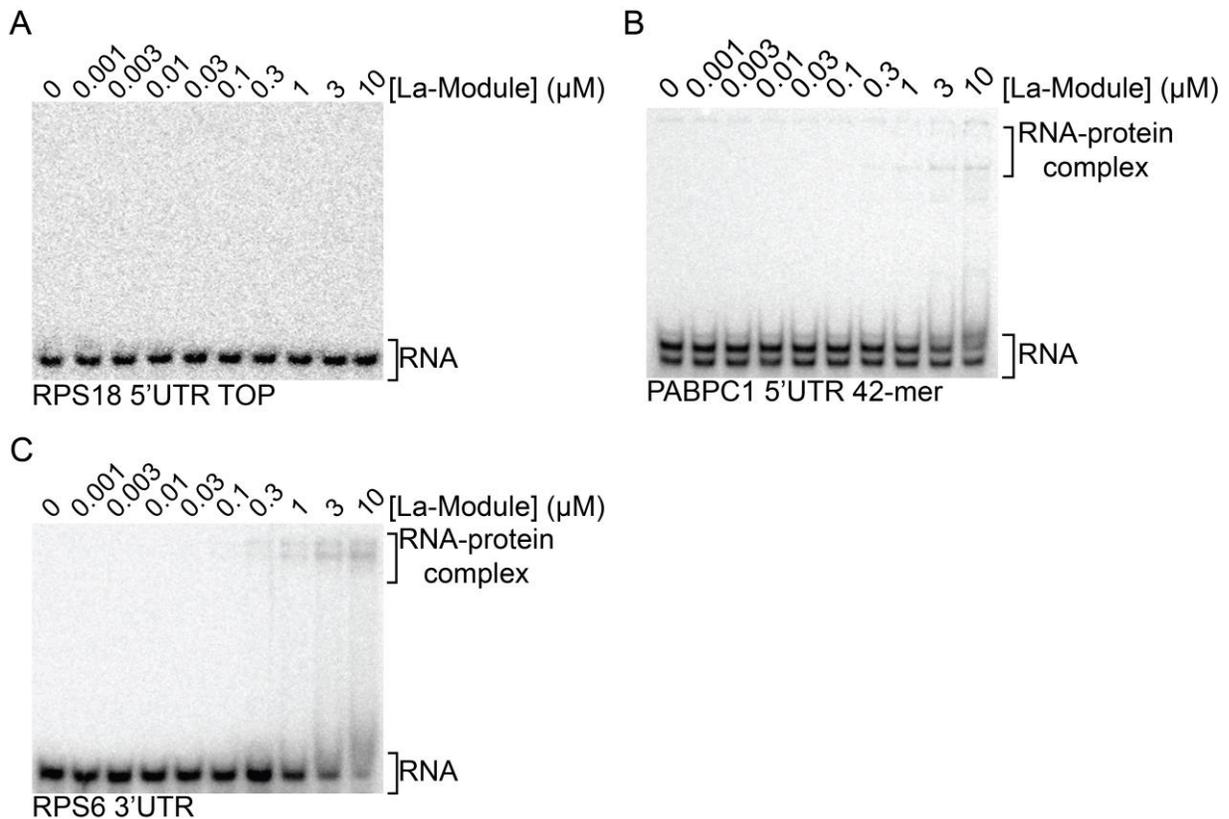


Figure 2-5 LARP1 La-Module does not bind all TOP motifs or the 3' UTR of RPS6 mRNA. EMSAs of binding assays of WT La-Module with (A) RPS18 5'UTR, (B) PABPC1 42-mer, and (C) RPS6 3'UTR.

2.3.3 LARP1 La-Module simultaneously engages poly(A) RNA and RPS6 TOP motif

Because the LARP1 La-Module bound the poly(A) and RPS6 20-mer RNAs with similar affinities, we utilized competition assays to delineate specificity (Figure 2-6). We first pre-bound the La-Module to radiolabeled RPS6 20-mer and competed with a titration of cold RPS6 20-mer or poly(A) RNA. Cold RPS6 20-mer at 50 nM was sufficient to displace the pre-bound RPS6 20-mer, whereas cold poly(A) RNA did not compete for the radiolabeled RNA even at 10 μM (Figure 2-6 A, B). We next pre-bound the La-Module to poly(A) RNA and competed with a titration of cold poly(A) or RPS6 20-mer RNA. Cold poly(A) RNA weakly displaced pre-bound poly(A) RNA, only achieving maximal displacement at 10 μM (Figure 2-6 C). However, when pre-bound poly(A) RNA was competed with cold RPS6 20-mer, beginning at 5 nM cold competitor, a complex forms that migrated further into the gel, between the shifted and free poly(A) RNA

(Figure 2-6 D; intermediate complex). Similarly, an intermediate complex formed when pre-bound poly(A) RNA was competed with cold RPL13A 5'UTR (Figure 2-6 E).

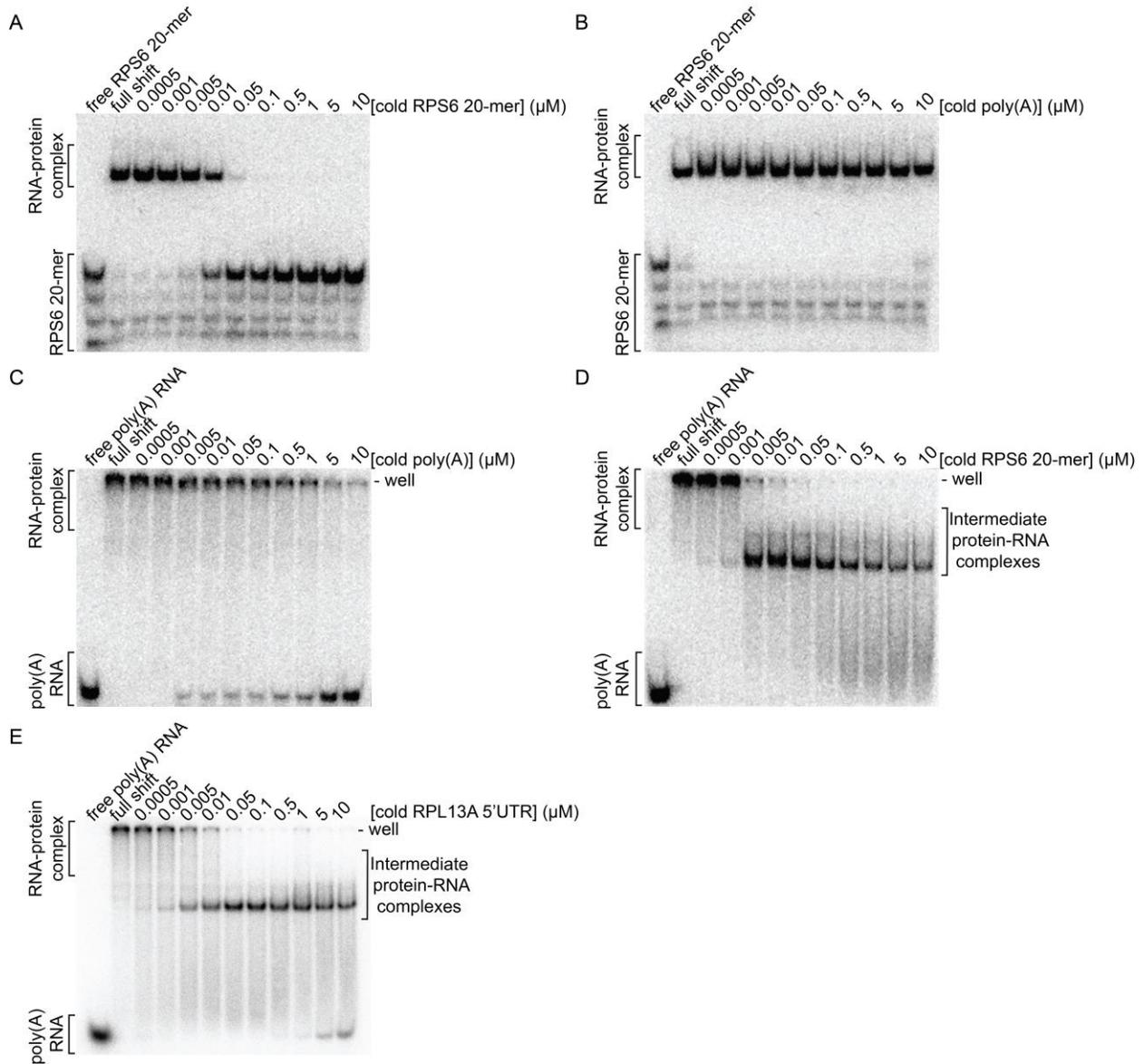


Figure 2-6 An intermediate complex forms upon the addition of cold RPS6 5'TOP motif or RPL13A 5' UTR to the La-Module poly(A) RNA complex. Competition assays conducted in the presence of poly(dI-dC) and analyzed by native gel of: the La-Module-RPS6 20-mer RNA complex with cold (A) RPS6 20-mer RNA, and (B) poly(A) 25-mer RNA. Competition assays conducted in the presence of poly(dI-dC) and analyzed by native gel of: the La-Module-poly(A) RNA complex with: cold (C) poly(A) 25-mer RNA, (D) RPS6 20-mer RNA, and (E) RPL13A 5'UTR.

These results led us to ask: in the absence of tRNA, why does the La-Module-poly(A) RNA complex get stuck in the well, but moves into the gel in the presence of cold RPS6 20-mer competitor? The simplest explanation is that the register of binding to poly(A) RNA is undefined because of its homopolymeric nature, thereby allowing for the assembly of multimeric complexes. We hypothesized that addition of a non-adenylate competitor could sequester molecules of La-Module away from the multimeric complex, thus allowing bound poly(A) RNA to enter the gel. To test this, we performed the competition assay with cold poly(C) RNA. However, the bound poly(A) remained within the well such that the intermediate complex was not recapitulated (Figure 2-7 A).

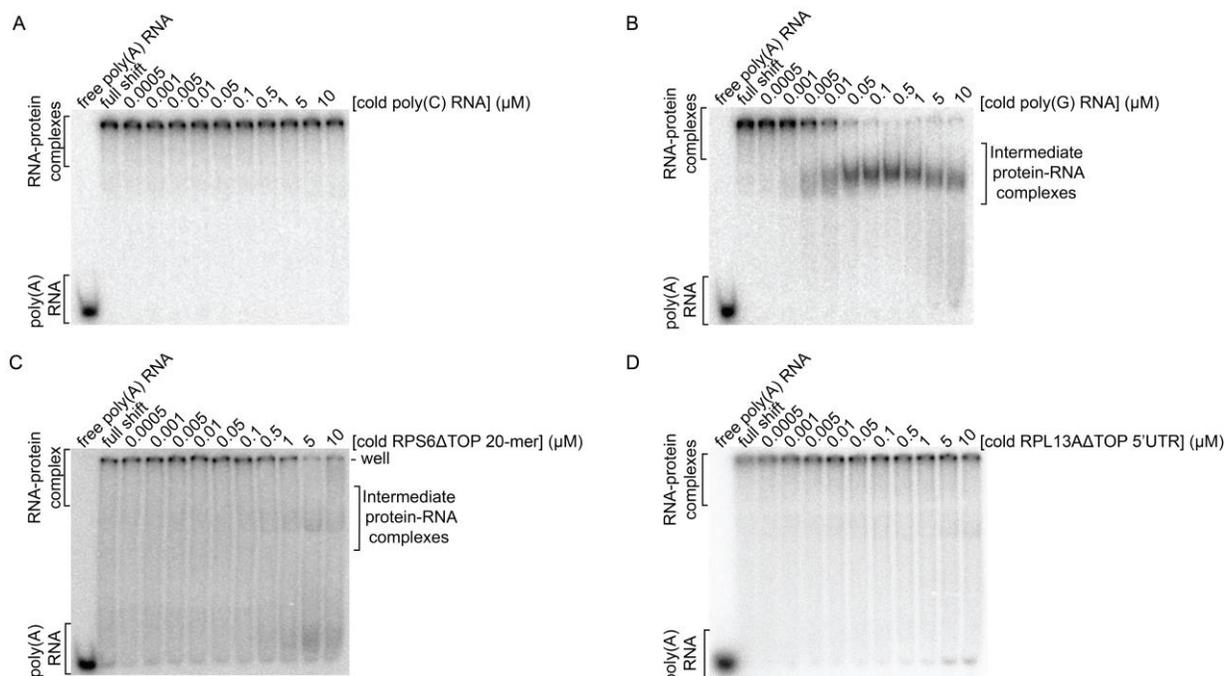


Figure 2-7 Formation of an intermediate complex is not an artifact of aggregation or RNA basepairing , and is dependent upon an intact TOP motif. Competition assays conducted in the presence of poly(dI-dC) and analyzed by native gel of the La-Module-poly(A) RNA complex with: cold (A) poly(C) RNA, (B) poly(G) RNA, (C) RPS6 5'UTR lacking the pyrimidine stretch of the TOP motif, and (D) RPL13A 5'UTR lacking the pyrimidine stretch of the TOP motif.

Given that the competitions were analyzed by native gel, we hypothesized that the intermediate complex has increased negative charge, and therefore migrated faster toward the positive electrode. This could occur if the La-Module simultaneously binds radiolabeled poly(A) and unlabeled RPS6 20-mer RNA. Alternatively, the intermediate complex could be an artifact of

uridines within the RPS6 20-mer base pairing with the poly(A) RNA. To exclude this possibility, we competed La-Module pre-bound to poly(A) RNA with cold poly(G) RNA (Figure 2-7 B); if the intermediate complex is an artifact of RNA duplex formation, we would not expect to reproduce the intermediate complex. However, we observed the formation of an intermediate complex that migrates between the shifted and free poly(A) RNA (Figure 2-7 B). Additionally, we competed the La-Module-poly(A) RNA complex with cold RPS6 20-mer and RPL13A 5'UTR with TOP motif deletions. Using the TOP motif deletion RNA also did not reproduce the intermediate complex (Figure 2-7 C, D). To further validate these results, we performed the competitions in the presence of tRNA as a non-specific competitor. La-Module pre-bound to RPS6 20-mer RNA was not competed by the addition of cold poly(A) RNA (Figure 2-8 B). However, once again, when we pre-bound the La-Module to poly(A) RNA and competed with cold RPS6 20-mer, a distinct RNP complex formed (Figure 2-8 A). Thus, we hypothesized that the intermediate complex might be a ternary complex of the La-Module, RPS6 20-mer, and poly(A) RNA.

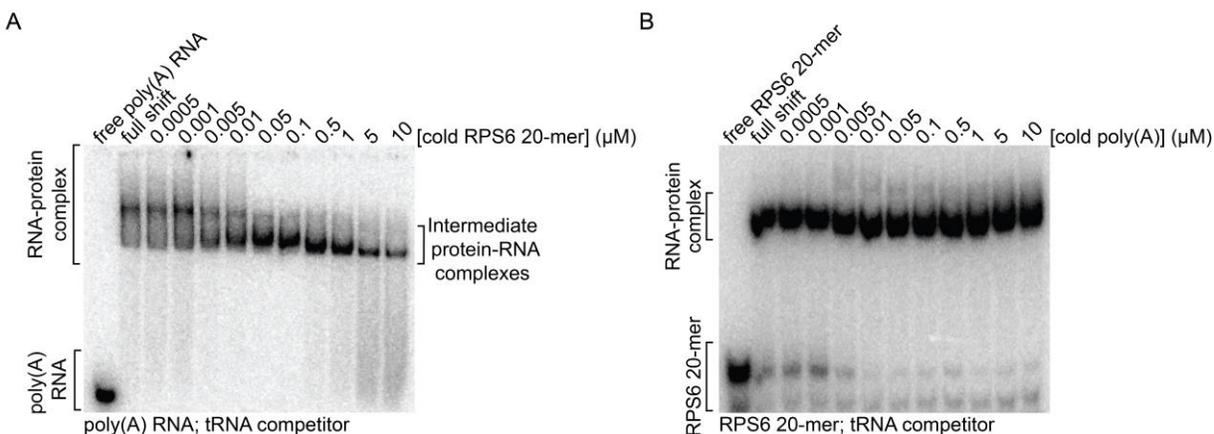


Figure 2-8 The intermediate complex forms in the presence of another non-specific competitor. Competition assays conducted in the presence of tRNA and analyzed by native gel of (A) the La-Module-poly(A) RNA complex with cold RPS6 20-mer RNA and (B) the La-Module-RPS6 20-mer RNA complex with poly(A) RNA.

To test this hypothesis, we conducted biotin pull-down assays. We pulled down the LARP1 La-Module using either biotinylated poly(A) or RPS6 20-mer RNA, prior to introducing radiolabeled RPS6 20-mer or poly(A) RNA (Figure 2-9). Enrichment of radiolabeled signal relative to the controls would suggest a ternary complex of the biotinylated bait, La-Module, and radiolabeled RNA. Using biotinylated poly(A) RNA as bait, we observed enrichment of the

radiolabeled RPS6 20-mer relative to controls when either bait or La-Module was omitted (Figure 2-9 A, B). Similarly, biotinylated RPS6 20-mer bait enriched radiolabeled poly(A) RNA relative to the controls (Figure 2-9 A, B). Thus, both biotinylated bait RNA and La-Module were required to capture the radiolabeled RNA.

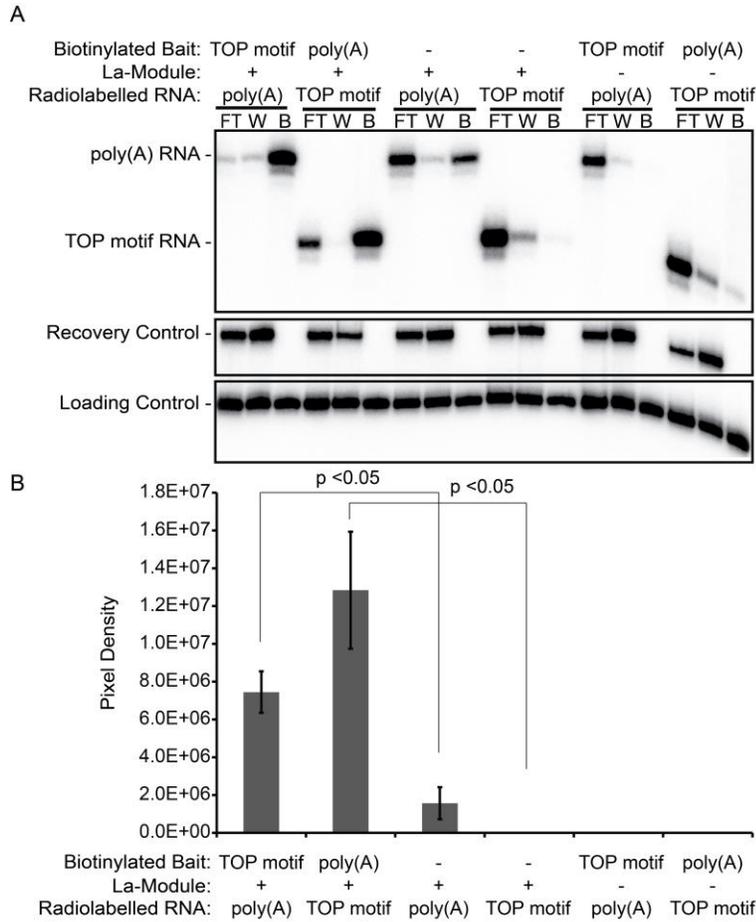


Figure 2-9 The La-Module simultaneously binds to poly(A) RNA and RPS6 TOP motif in biotin pull-down experiments. (A) Denaturing gel analyzing the flow through (FT), wash (W), and pull-down streptavidin beads (B) of the indicated binding experiments. Poly(A), poly(A) 25-mer; TOP motif, RPS6 20-mer. Ethanol precipitation recovery control, loading control, RPS6 20-mer and poly(A) RNA size markers are indicated. (B) Quantification of biotin pull-down assays. Bars indicate standard error.

2.3.4 Simultaneous binding of poly(A) and TOP motif RNA is faithful in the context of full-length LARP1

Note: this section was not a part of the article Al-Ashtal *et al.* RNA biology (2019).

We next wanted to examine the RNA-binding properties of the La-Module in the context of full-length LARP1 purified from human cells, as intramolecular interactions, posttranslational modifications, and conformational plasticity may influence RNA recognition. To do this, we compared a LARP1 R840E/Y883A (REYA) double mutant to wild-type (WT) LARP1. REYA LARP1 mutations reside within the DM15 and inhibit its ability to bind RNA (Figure 2-10)^{168,171}, allowing us to examine contributions outside of the DM15 to RNA binding.

We purified WT and REYA LARP1 from human cells and conducted EMSAs with poly(A) and RPS6 42-mer RNAs. As expected, WT LARP1 binds poly(A) and RPS6 42-mer (Figure 2-10 A, B). Consistent with our observations of recombinant purified La-Module, REYA LARP1 was also able to bind both poly(A) RNA and the RPS6 5'UTR with apparent affinities of 334 ± 5 nM and 274 ± 8 nM, respectively (Figure 2-10 C, D).

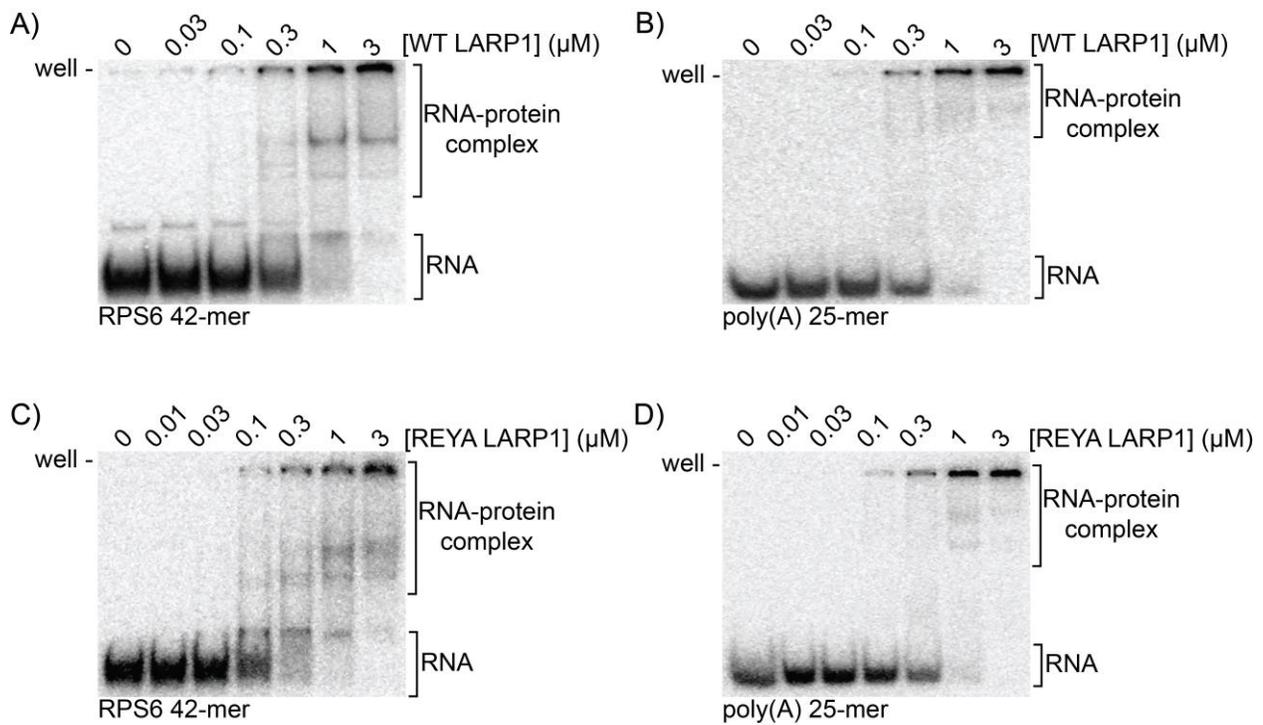


Figure 2-10 WT and REYA LARP1 bind poly(A) RNA and RPS6 5' UTR. EMSA analysis of: WT LARP1 with (A) RPS6 5' UTR; RPS6 42-mer, and (B) poly(A) RNA; poly(A) 25-mer, and REYA LARP1 with (C) RPS6 5' UTR; RPS6 42-mer, and (D) poly(A) RNA; poly(A) 25-mer.

Furthermore, REYA LARP1 also appears to simultaneously bind both poly(A) and the RPS6 20-mer in biotin pull-down experiments (Figure 2-11). In addition, REYA LARP1 did not bind ΔTOP RPS6 42-mer or the RPS6 3' UTR with appreciable affinity (Figure 2-11).

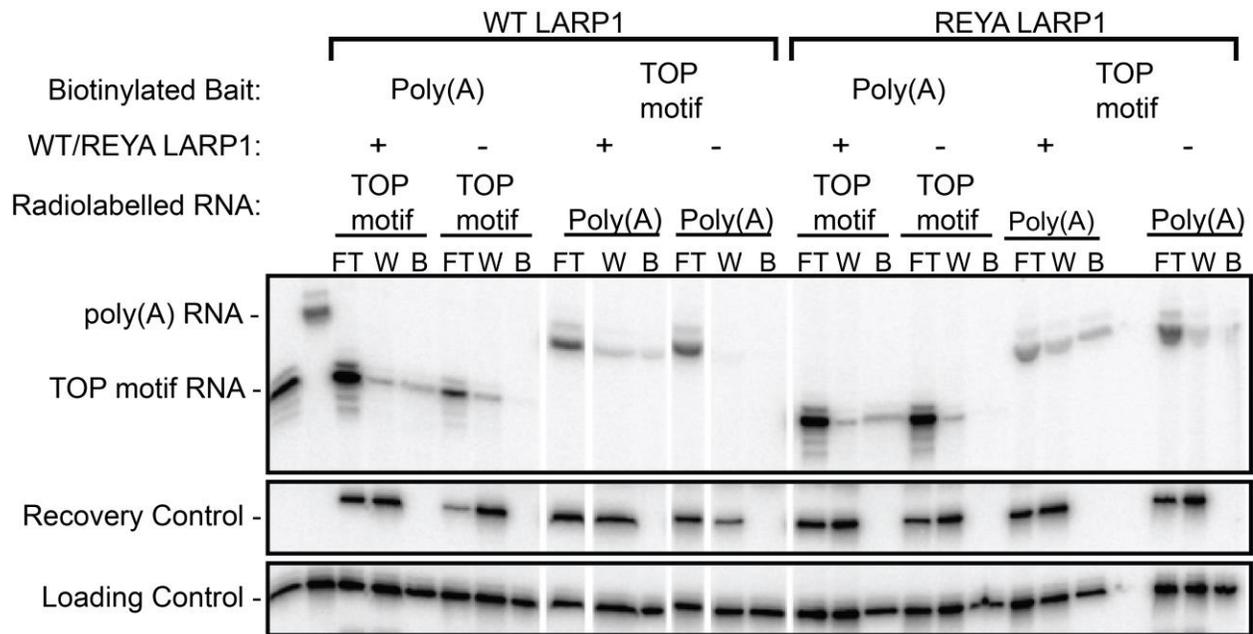


Figure 2-11 WT and REYA LARP1 simultaneously binds poly(A) RNA and RPS6 TOP motif in biotin pull-down experiments. Denaturing gel analyzing the flow through (FT), wash (W), and pull-down streptavidin beads (B) of the indicated binding experiments. Poly(A), poly(A) 25-mer; TOP motif, RPS6 20-mer. Ethanol precipitation recovery control, loading control, RPS6 20-mer and poly(A) RNA size markers are indicated.

2.4 Discussion

Here we identify the RNA binding partners of the LARP1 La-Module to elucidate its contribution to TOP mRNA recognition. We found that the LARP1 La-Module binds poly(A) RNA (Figure 2-1, 2-2, 2-3). Surprisingly, the LARP1 La-Module also binds the 5'UTRs of some TOP mRNAs (Figure 2-4, 2-5). Deletion of the TOP motif abrogated La-Module binding to TOP mRNA 5'UTRs, suggesting that the La-Module requires the polypyrimidine portion of TOP sequences (Figure 2-4). Indeed, the La-Module binds the RPS6 TOP motif with comparable affinity to poly(A) RNA, and does so in a cap-independent manner (Figure 2-1, 2-4).

The affinity of the LARP1 La-Module for poly(A) RNA may indicate a role at poly(A) tails. This is consistent with the association of LARP1 with PABP and poly(A) RNA^{167,170,199}. The PAM2 located within the interdomain linker between the LAM and RRM could bind PABP¹⁶⁷, while the LAM and RRM bind poly(A) tails of TOP transcripts or other mRNAs.

The coordinated translation of TOP mRNAs is mediated by a characteristic 5'TOP motif followed by a GC-rich region within the 5'UTRs of these transcripts¹⁴⁵. Our data also implicate the LARP1 La-Module at the 5'UTRs of TOP mRNAs *via* recognition of the TOP motif region. The La-Module binds the 5'UTRs of RPS6 and RPL13A mRNAs, but not those of RPS18 or PABPC1 (Figure 2-4). These data are consistent with differential levels of TOP mRNA repression by LARP1¹⁶⁷, and the specificity of the DM15 region for only certain TOP mRNAs¹⁷¹. While the DM15 region binds the 5' cap and first four nucleotides of TOP mRNAs¹⁶⁸, the La-Module binds in a cap-independent manner (Figure 2-4), suggesting that the La-Module recognizes sequences downstream of the DM15 binding-site. Consistent with this, the La-Module binds the RPS6 20-mer, but not an RPS6 15-mer that shortens the GC-rich region (data not shown), suggesting a contribution of that sequence La-Module binding; in addition, although it is weak, the La-Module binds to poly(G) RNA in the presence of non-specific competitor (Figure 2-1). These observations are consistent with data showing that, at least in some eukaryotic cell lines, both the 5'TOP motif and the GC-rich region immediately downstream are necessary for TOP mRNA translation regulation^{145,214}, and the mapping of LARP1 footprints to 3' ends of TOP 5'UTRs²¹⁵. Binding of the La-Module to TOP motifs may enhance TOP mRNA recognition by LARP1. As the C-terminal DM15 region binds the 5' cap and TOP motif to repress TOP mRNA translation^{168,169}, the N-terminal La-Module might bind cooperatively to the TOP motif to strengthen the interaction and promote stringent specificity for recognizing select TOP mRNAs.

More notably, the LARP1 La-Module can simultaneously bind both poly(A) and TOP motif RNA. We observed the formation of an intermediate complex upon competition of La-Module-poly(A) complex with TOP motif RNA (Figure 2-6). Biotin pull-down experiments indicate that this intermediate complex corresponds to the LARP1 La-Module simultaneously engaging both poly(A) RNA and TOP motif RNA (Figure 2-9). Furthermore, the RNA-binding properties of recombinant purified La-Module were faithful in the context of full-length REYA LARP1 purified from human cells (Figure 2-10, 2-11).

The unprecedented ability of the LARP1 La-Module to bind two distinct RNAs at the same time not only expands the known binding repertoire of LARPs, but also might explain the biological functions of LARP1. Importantly, the targets of the LARP1 La-Module correspond to the 5' and 3' ends of TOP mRNAs – TOP motifs and poly(A) tails (Fig. 5). The simultaneous

binding of poly(A) tails and the TOP motif would imply that the La-Module brings the 5' and 3' ends of bound transcripts into close proximity.

While the mRNA circularization model^{15,100} has recently been reassessed¹⁰⁴, other studies have provided evidence for its role in promoting oncogenic translation and tumorigenesis¹⁰⁸. Indeed, overexpression of LARP1 in epithelial malignancies, such as cervical and ovarian cancers is correlated with tumor progression and poor clinical outcome^{197,198}. Consistent with this, LARP1 increases the stability of several oncogenic transcripts^{197,198} and enhances protein synthesis^{170,195}.

The translational repression of transcripts in P-bodies and stress granules is also associated with 5' and 3' UTR proximity. Thus, binding of the LARP1 La-Module to both ends of a single transcript may also explain its localization to P-bodies and stress granules^{167,216} (Figure 2-12 A), where RNAs are translationally repressed in condensed conformations⁸⁴. This is consistent with the necessity of LARP1 for recruiting and anchoring TOP mRNAs to these granules²¹⁷.

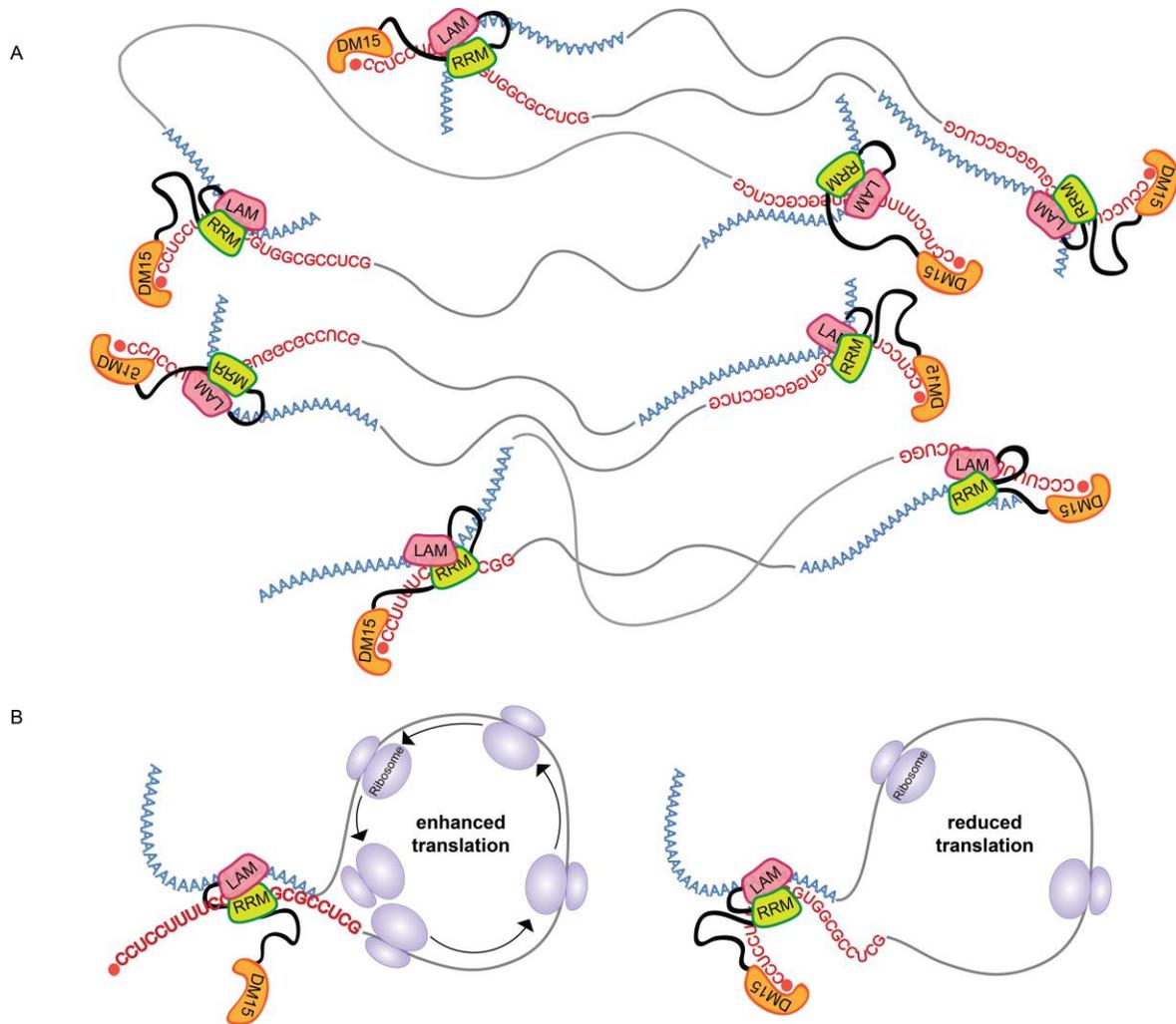


Figure 2-12 Model of potential roles for the LARP1 La-Module in TOP mRNA translation and localization to P-bodies and stress granules. (A) The LARP1 La-Module could bind the poly(A) tails and TOP motifs of different TOP transcripts to form multimeric complexes. This might contribute to the phase separation characteristic of P-bodies and stress granules, while repressing translation and minimizing transcript degradation. (B) The LARP1 La-Module could bind the poly(A) tail and TOP motif of the same TOP transcript to enhance translation (left); recognition of the 5' cap by the DM15 region of LARP1 could then repress the translation of bound transcripts (right).

The role of the LARP1 La-Module in TOP mRNA recognition may also shed light on seemingly incongruous data regarding the function of LARP1 in TOP mRNA translation regulation. If the La-Module binds to both ends of TOP mRNAs, it might increase TOP mRNA translation (Figure 2-12 B). However, upon the binding of the DM15 region to the 5' cap, the translation initiation complex would be occluded, resulting in translation inhibition (Figure 2-12

B). This idea perhaps explains the observations of LARP1 playing opposing stimulatory^{170,195} and repressive¹⁶⁷⁻¹⁶⁹ roles in TOP translation regulation. Since our data do not directly address translation levels, it is also possible that the simultaneous interaction of the LARP1 La-Module stabilizes TOP transcripts.

Other La-Modules have been shown to play a role in stimulating translation and stabilizing transcripts. Cytoplasmic La regulates both cap-dependent^{218,219} and cap-independent^{67,220-222} translation in a transcript-dependent manner. Cytoplasmic La also binds binding to poly(A) tails of messenger RNAs and facilitates entry into polysomes²²³. LARPs 4 and 5 associate with PABP, RACK1, polysomes, and increase mRNA translation^{164,166}. The LARP6 La-Module recognizes a stem-loop motif within the 5'UTR of collagen I- and III- encoding mRNAs to promote their translation^{161,162}. Thus, there is precedent for proposing a stimulatory role for the LARP1 La-Module.

Herein we demonstrate that the LARP1 La-Module binds both poly(A) RNA and TOP motifs with comparable affinities. While we observed that the LARP1 La-Module can bind poly(A) and TOP motif RNA at the same time, we do not know the stoichiometry of the ternary complex: does the La-Module dimerize, with each protein molecule binding one RNA? Or does one protein bind both RNAs using different binding surfaces? In most LARPs, the LAM and RRM bind synergistically to RNAs without using their canonical RNA-binding surfaces^{206,207}. It might be possible that the canonical binding surface of the LAM or RRM can bind a second RNA; indeed, Genuine La has recently been shown to bind poly(A) RNA *via* the canonical binding surface of the LAM²²³. Thus, much remains to be uncovered about the atypical binding modes of LARPs. Nonetheless, dual RNA recognition by the LARP1 La-Module offers a tantalizing explanation for the mechanisms by which LARP1 regulates the translation and stability of TOP transcripts, and might tie together the various roles of LARP1.

3.0 Identification of RNA-Binding Surfaces and Optimization of La-Module Construct

3.1 Introduction

After our discovery that the LARP1 La-Module simultaneously binds poly(A) and TOP motif RNA (Chapter 1), we wished to understand the molecular mechanism behind these interactions. We wanted to identify the RNA-binding surfaces of the La-Module and its stoichiometry in the La-Module-poly(A)-TOP motif ternary complex. More specifically: does one molecule of La-Module bind both poly(A) and TOP motif RNA using different RNA binding surfaces? Or does the La-Module form a dimer, with the RNAs binding to the same site within each monomer? We began by conducting a mutational analysis of the La-Module (amino acids 310-540). However, this approach was greatly hindered by difficulties generating sufficient quantities of La-Module mutants with sufficient yield and RNA-binding activity.

Indeed, despite our success at probing the RNA-binding properties of the LARP1 La-Module (amino acids 310-540) in Chapter 2, our efforts were slowed by the difficulties in reproducibly purifying high yields of active protein. For this reason, we speculated that the La-Module construct (amino acids 310-540) domain boundaries might be incorrect. To remedy this, we used sequence alignments, secondary structure predictions, and homology models (see Appendix 1) to guide construct design in order to identify an optimal construct. Analysis of these data led us to the hypothesis that the C-terminus of the RRM in our current La-Module construct (amino acids 310-540) was truncated. We hypothesized that this RRM truncation was the underlying cause for the poor expression, inconsistent purification, and low activity (<2%) of La-Module (amino acids 310-540). Based on this, we generated and tested various constructs that extend the C-terminus of the La-Module (summarized in Figure 3-3). We determined that La-Module (amino acids 310-647) provided the best expression, consistent purification and RNA-binding activity, as well as decreased aggregation.

We then identified the domain boundaries of the LAM and the RRM and worked towards understanding the molecular mechanism by which the La-Module binds RNAs. We used various biochemical and biophysical techniques to try to determine the RNA-binding surfaces of the La-

Module, as well as its stoichiometry in the La-Module-poly(A)-TOP motif ternary complex observed in Chapter 2. While we gained some insights, our results are largely inconclusive and future work is necessary to identify the stoichiometry and RNA-binding surfaces involved in the LARP1-poly(A) RNA- TOP motif ternary complex. Our current efforts are focused on testing combinatorial mutants for RNA binding and using crosslinking and mass-spectrometry to identify RNA-bound peptides.

3.2 Materials and methods

3.2.1 La-Module (310-540) mutants cloning, expression, and purification

La-Module (310-540) mutants were generating using site-directed mutagenesis. Mutants were then expressed and purified as done for wild-type La-Module (310-540) in Chapter 2.

3.2.2 UV cross-linking of La-Module (310-540)

Competition assays were conducted as previously described for La-Module (310-540) in Chapter 2. Reactions were then exposed to a 254 nm UV radiation light source 4 cm away for 40 min on ice. 10 μ L 2X SDS loading buffer was added to each reaction. Samples were resolved by running on a 12.5 % SDS-PAGE (29:1 polyacrylamide) at room temperature for 50 min at 125 V. Gels were dried, exposed overnight on phosphor screens (GE Healthcare Lifesciences), then imaged on a Typhoon FLA plate reader (GE Healthcare Lifesciences).

3.2.3 Size exclusion chromatography of La-Module (310-540)

Binding reactions were conducted as described in Chapter 2 with the following exceptions: 1 μ M cold A25 or RPS6 20-mer were used in lieu of radiolabelled RNA, and 10 μ M La-Module (310-540) was used accordingly. Binding reactions were loaded onto a Superdex 75 size exclusion

column (GE Healthcare Lifesciences) equilibrated in buffer [50 mM Tris-HCl pH 7.5, 100 mM NaCl, 5 % glycerol, 1 mM DTT] and eluted at 1 mL/min. Elution peaks were compared to elution of 500 μ L gel filtration standards (Bio-Rad, 1511901).

3.2.4 La-Module (amino acids 310-647), cloning, expression, and purification

The optimized LARP1 La-Module coding region (amino acids 310-647, isoform 2) (Integrated DNA Technologies) was PCR amplified and cloned into pET28a (Novagen Inc) using BamHI and SacI sites. Plasmids expressing La-Module point mutants and linker deletions were generated using site-directed mutagenesis. The resulting constructs, a 6XHis-SUMO-La-Module fusion protein, were expressed in *E. coli* BL21(DE3), cultured at 37°C for 2 hrs prior to shifting to 17.5 °C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

For purification, cells were resuspended and lysed by sonication in lysis buffer [50 mM Tris-HCl, pH 8.0, 750 mM NaCl, 10 mM imidazole, 10% v/v glycerol, 0.01% CHAPS, protease inhibitors (10 μ M leupeptin, aprotinin, 10 μ M bestatin, 1 μ M pepstatin, 10 μ M PMSF)]. The lysate was cleared *via* centrifugation. The 6XHis-La-Module was purified in batch using nickel agarose affinity chromatography (Thermo Fisher Scientific) and eluted [50 mM Tris-HCl, pH 8.0, 750 mM NaCl, 350 mM imidazole, 10 % v/v glycerol]. The 6XHis-SUMO tag was cleaved overnight, then the tag and 6XHis-ULP1 were separated as described for the RRM. The La-Module was collected from the flowthrough and buffer exchanged by dialysis for 2 hrs in 2 L [50 mM Tris-HCl, pH 8.0, 75 mM NaCl, 10 % v/v glycerol, 1 mM DTT], then further purified by HiTrap Heparin, followed by tandem HiTrap S and HiTrap QP (GE Healthcare Lifesciences) chromatography with a NaCl gradient (150 mM-1M). Fractions containing La-Module were collected, concentrated to 0.3-0.5 mL, then loaded onto Superdex 75 size exclusion column (GE Healthcare Lifesciences) equilibrated in buffer [50 mM Tris-HCl, pH 7.5, 850 mM NaCl, 5 % glycerol, 1 mM DTT] and eluted at 0.25 mL/min. Fractions containing La-Module were buffer exchanged and concentrated to 100 μ M for storage using a Vivaspin Turbo 30K MWCO Centrifugal Concentrator (Sartorius) [50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 % glycerol, 1 mM DTT] or 10-20 mg/mL for crystallization in buffer [50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 % glycerol, 1 mM DTT].

3.2.5 LAM cloning, expression, and purification

The region encoding the LARP1 LAM (amino acids 310-405 from 1019-amino acid isoform 2) (Integrated DNA Technologies) was PCR amplified and cloned into pET28a (Novagen Inc) using NdeI and BamHI sites. The 6XHis-LAM fusion protein was expressed in *E. coli* BL21(DE3), and cultured at 37°C for two hours prior to shifting to 17.5 °C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

For purification, cells were lysed, cleared, and was purified by nickel agarose affinity chromatography as described for the La-Module purification above. The 6XHis tag was also removed as described for the La-Module, and removed using a second nickel affinity chromatography step using a HiTrap Nickel FF (GE Healthcare Lifesciences). The LAM was collected from the flowthrough and buffer exchanged by dialysis for 2 hrs in 2L buffer [50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10 % v/v glycerol, 1 mM DTT]. The LAM was separated from protein and nucleic acid contaminants using tandem HiTrap S and HiTrap QP chromatography as described above for the La-Module, with LAM also eluting from the Q column. Fractions containing LAM were collected, concentrated to 0.5-1 mL, then loaded onto Superdex 75 size exclusion column (GE Healthcare Lifesciences) equilibrated in buffer [50 mM Tris-HCl, pH 7.5, 750 mM NaCl, 5 % glycerol, 1 mM DTT] and eluted at 0.25 mL/min. Fractions containing LAM were buffer exchanged and concentrated to 100 µM for storage [50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 % glycerol, 1 mM DTT] or 10-20 mg/mL for crystallization [50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 % glycerol, 1 mM DTT].

3.2.6 RRM cloning, expression, and purification

The LARP1 RRM coding region (amino acids 440-4647 from 1019-amino acid isoform 2) (Integrated DNA Technologies) was PCR amplified and cloned into pET28b-N-SUMO²²⁴ (Novagen Inc) using BamHI and SacI sites. The resulting construct, a 6XHis-SUMO-RRM fusion protein, was expressed in *E. coli* BL21(DE3), and cultured at 37°C for two hours prior to 17.5 °C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

For purification, cells were lysed, cleared, and protein was purified by nickel agarose affinity chromatography as described for the La-Module. The 6XHis-SUMO was removed using 0.5 mg 6XHis-ULP1 protease²²⁴ per 40 mL eluate overnight at 4 °C in 2 L dialysis buffer [50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 0.5 mM EDTA, 0.5 mM DTT]. The cleaved 6XHis-SUMO tag and 6XHis-ULP1 were separated *via* a second nickel affinity chromatography step using a HiTrap Nickel FF (GE Healthcare Lifesciences). The RRM was buffer exchanged and further purified using tandem HiTrap S and HiTrap QP chromatography as described above for LAM, with the RRM eluting from the Q column. Fractions containing RRM were collected, concentrated to 0.5-0.75 mL, then loaded onto Superdex 75 size exclusion column (GE Healthcare Lifesciences) equilibrated in buffer [50 mM Tris-HCl, pH 7.5, 850 mM NaCl, 5 % glycerol, 1 mM DTT] and eluted at 0.25 mL/min. Fractions containing RRM were buffer exchanged and concentrated to 100 µM for storage [50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 % glycerol, 1 mM DTT] or 10-20 mg/mL for crystallization [50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 5 % glycerol, 1 mM DTT].

3.2.7 La-Module (310-647) electrophoretic mobility shift assays

Electrophoretic mobility shift assays were conducted as described for La-Module (310-540) in Chapter 2.

3.2.8 La-Module (310-647) competition assays

Competition assays were conducted as described for La-Module (310-540) in Chapter 2.

3.2.9 RRM (440-647) electrophoretic mobility shift assays

Electrophoretic mobility shift assays were conducted as described for La-Module (310-540) in Chapter 1.

3.2.10 RRM (440-647) static light scattering (SLS)

5 μ M RRM was incubated on ice at a 1:1 molar ration with poly(A) 25-mer, RPS6 20-mer, or both. SLS was conducting using a DynaPro NanoStar Wyatt with 10 s acquisition times at 4°C.

3.2.11 La-Module (310-540) and RRM (440-647) native mass spectrometry

Native mass spectrometry experiments were conducted by the laboratory of Dr. Vicki Wysocki (Ohio State University). Briefly, La-Module (310-540) and RRM (440-647) were bound to poly(A) and RPS6 20-mer as described in Chapter 2, with the exception that 1X binding buffer consisted of 50 mM Tris-HCl, pH 7.5, 100 mM ammonium acetate, 7.5 % glycerol, 1mM DTT. Binding reactions were then subjected to direct infusion with nano-electrospray ionization mass spectrometry (nano-ESI-MS).

3.3 Results

3.3.1 Attempts to identify the RNA binding surfaces of the LARP1 La-Module (310-540)

We mutated conserved residues of the LARP1 La-Module (310-540) that are required for RNA binding in La and LARP7 (Figure 3-1 A, B), as well as mutations to basic residues in the linker that might bind RNA (Figure 3-1 C, D). Single mutations to these residues decreased binding to poly(A) 25-mer and RPS6 20-mer to a similar extent (Figure 3-1), which would suggest that these RNAs bind to the same site within the La-Module. However, the differences in RNA binding could not be corrected for protein activity, due to low yield and RNA-binding activity. Therefore, we could not directly compare the La-Module RNA-binding mutants. As an alternative, we attempted native mass spectrometry in collaboration with the laboratory of Dr. Vicki Wysocki (Ohio State University). However, our collaborators could not reconstitute RNA-binding activity of the La-Module.

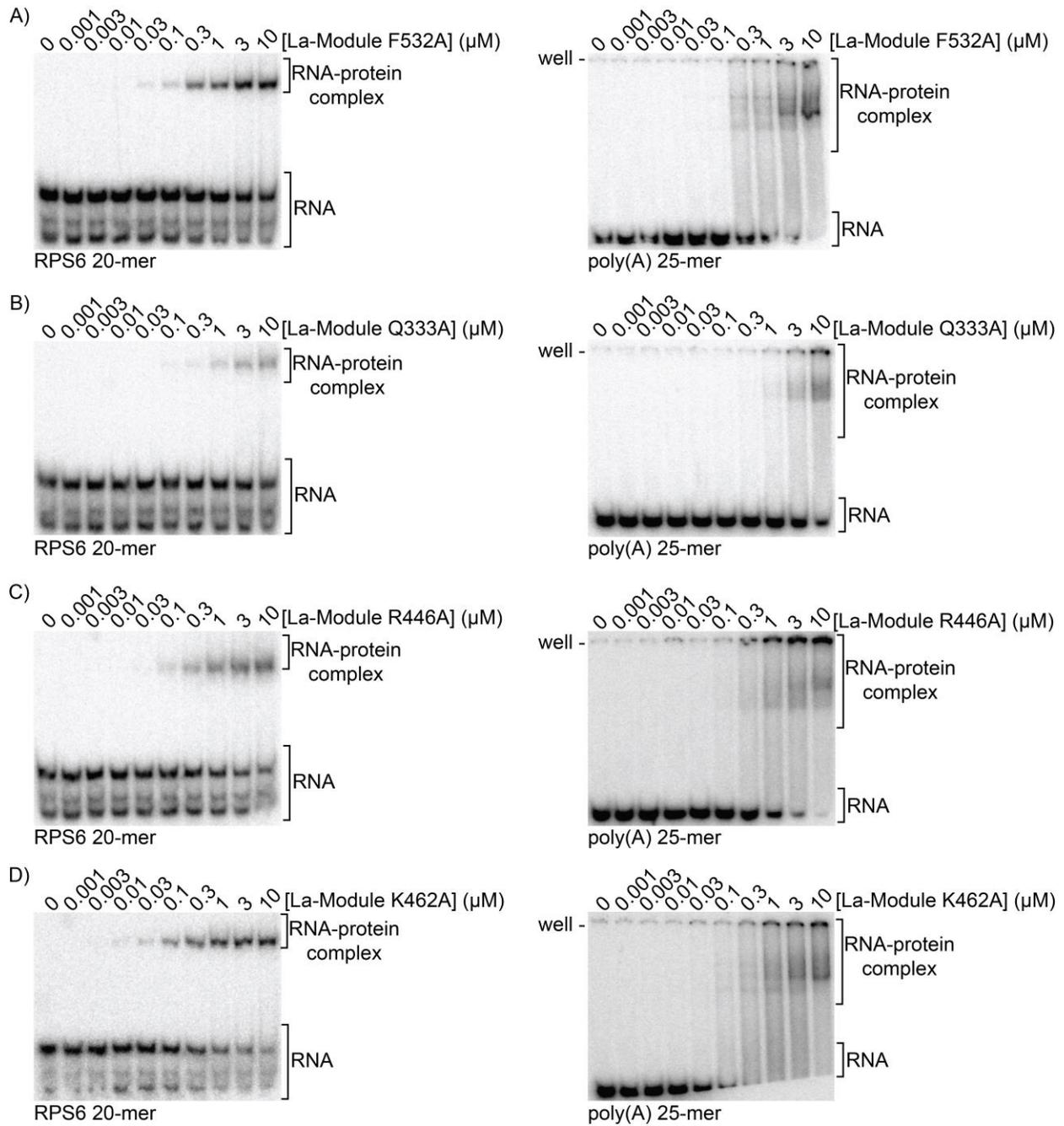


Figure 3-1 La-Module RNA-binding mutants show decreased affinity for RPS6 20-mer and poly(A) RNA.

EMSAs of La-Module RNA-binding mutants with RPS6 20-mer and poly(A) 25-mer.

3.3.2 Attempts to identify the stoichiometry of the LARP1 La-Module (310-540)

We next tried UV cross-linking experiments to determine the stoichiometry of the La-Module (310-540) when simultaneously bound to poly(A) and TOP motif RNA (Figure 3-2). We began by cross-linking La-Module bound to either radiolabelled poly(A) 25-mer or RPS6 20-mer RNA, and resolved the complexes *via* SDS-PAGE (Figure 3-2 A, B). A single La-Module (27 kDa) bound to either poly(A) RNA (8.3 kDa) or RPS6 20-mer RNA (6.3 kDa) would form a complex of 35.3 kDa or 33.3 kDa, respectively. While the majority of the RNA was not cross-linked, we observed complexes at the expected molecular weight (Figure 3-2 A, B). We also detected higher molecular weight complexes between 50 and 75 kDa that could be a result of multimerization of the La-Module or non-specific cross-linking (Figure 3-2 A, B). We also observed a lower molecular weight complex at ~20 kDa (Figure 3-2 A, B).

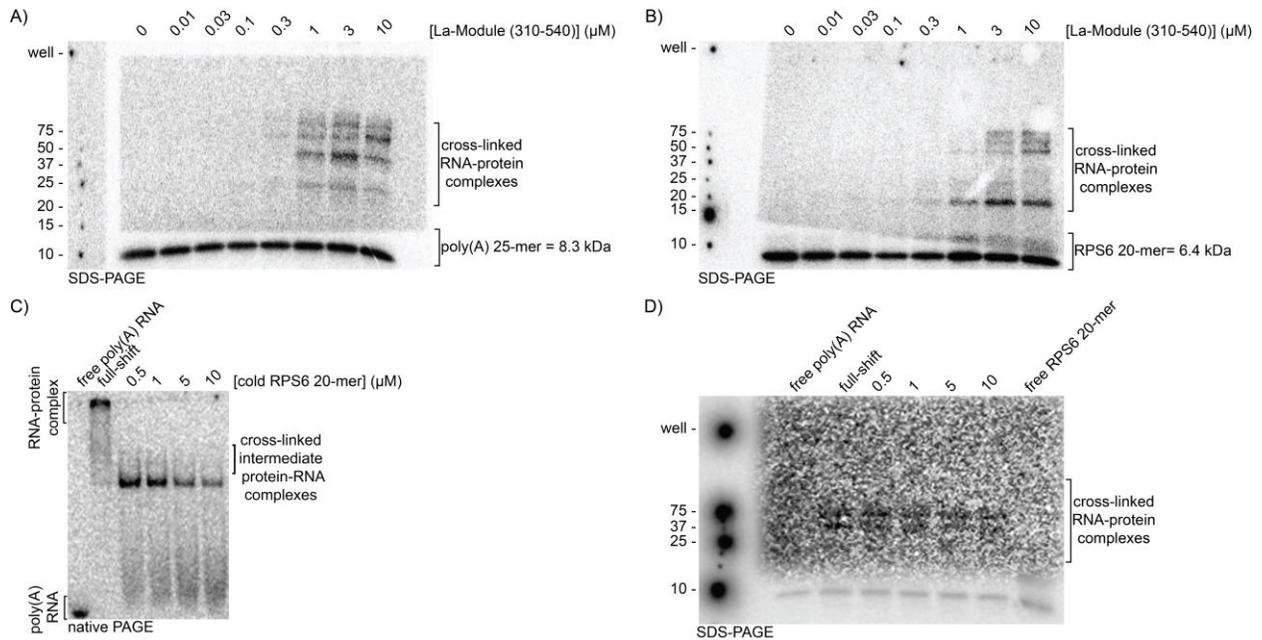


Figure 3-2 Cross-linking La-Module to poly(A) and RPS6 20-mer RNA was inefficient. UV cross-linking of La-Module EMSA with (A) poly(A) 25-mer and (B) RPS6 20-mer analyzed by SDS-PAGE. UV cross-linking of La-Module-poly(A) 25-mer-RPS620-mer ternary complex in competition assay resolved by (C) native PAGE and (D) SDS-PAGE. Contouring is consistent throughout gels in panels (A), (B), and (D); regions containing protein-RNA complexes were contoured for clarity.

Next, we performed a competition assay in which we pre-bound the La-Module to radiolabeled poly(A) RNA and competed with cold RPS6 20-mer, in order to reconstitute the La-Module-poly(A) RNA-RPS6 20-mer ternary complex. We UV cross-linked and resolved the complexes using native PAGE or SDS-PAGE (Figure 3-2 C, D). The native gel preserved the intermediate complex and showed the same migration patterns seen without cross-linking (Figure 3-2 C). However, analysis of these complexes by SDS-PAGE was not straightforward (Figure 3-2 D).

One La-Module bound to both poly(A) 25-mer and RPS6 20-mer would be 41.7 kDa, while a La-Module dimer with each monomer bound to one RNA would be 68.7 kDa. However, the detected complexes appeared as two bands: a lower band running above 35 kDa, as well as a higher band migrating below 75 kDa (Figure 3-2 D). Furthermore, there was no observable difference between the La-Module-poly(A) RNA control and the lanes in which a ternary complex would exist upon titration of RPS6 20-mer. It may be that the ternary complex was not effectively cross-linked due to the inefficiency of UV cross-linking (1-5% cross-link²²⁵) and the ternary complexes were not preserved within the denaturing SDS-PAGE (Figure 3-2).

We then attempted size exclusion chromatography (SEC) to identify the stoichiometry of the La-Module when simultaneously bound to poly(A) and TOP motif RNA. The SEC experiments showed polydisperse elution peaks and “tailing” (data not shown), which is indicative of conformational heterogeneity²²⁶. Importantly, the La-Module peak elutes in between that of a monomer and a dimer, even in the presence of poly(A) 25-mer or RPS6 20-mer (data not shown). Therefore, it was difficult to conclude whether the La-Module was a monomer or dimer when bound to both RNAs. These experiments were further complicated by the low yield RNA-binding activity of the La-Module.

3.3.3 Identification of La-Module (310-647) as the optimal construct

Due to the difficulty in working with the La-Module construct proposed in the literature¹⁵⁶ (amino acids 310-540), we sought to re-design the construct. Using sequence conservation, secondary structure predictions, and homology modeling (see Appendix B), we designed and tested various constructs (Figure 3-3).

Construct	
460-647	—————
440-647	—————
440-545	—————
543-647	—————
320-405	—————
320-540	—————
320-572	—————
320-647	—————
310-405	—————
310-540	—————
310-572	—————
310-647	—————

Figure 3-3 Constructs generated to identify the optimal La-Module construct. List of constructs generated with varying N- and C- termini to identify LAM and RRM domain boundaries. Constructs were expressed as 6XHis-SUMO fusion proteins.

Improved construct design was gauged by: 1) high expression yields, 2) reproducible purity, 3) lack of degradation and/or aggregation, and 4) reproducible RNA-binding activity. We determined that the LARP1 RRM ends at amino acid 647 rather than 540, because these constructs exhibited robust expression, consistent purification, and decreased degradation (Figure 3-4).

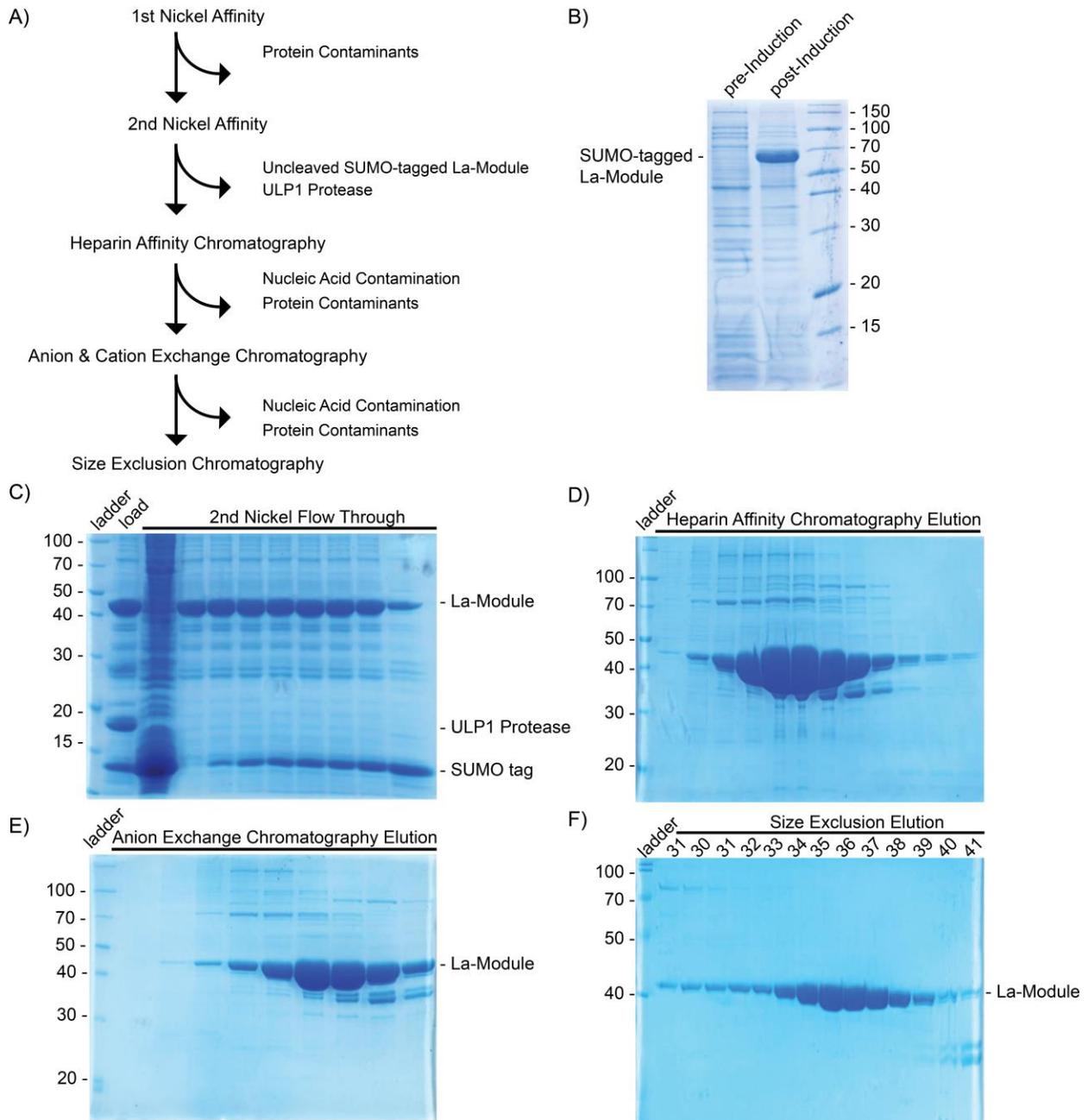


Figure 3-4 Purification of La-Module (amino acids 310-647). (A) Purification scheme, SDS-PAGE of: (B) Expression of 6XHis-SUMO-LaModule (310-647), (C) Second nickel affinity chromatography, (D) Heparin affinity chromatography elution, (E) Anion exchange chromatography elution, (F) Size exclusion elution.

Furthermore, LARP1 La-Module (310-647) does not exhibit well shifts when titrated into RNA during EMSAs, as observed with La-Module (310-540) in the absence of salmon sperm DNA or tRNA (Figure 3-5).

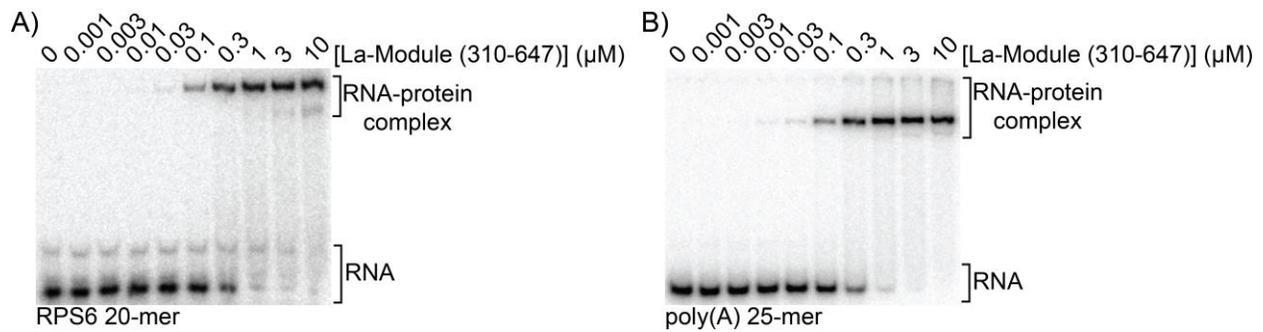


Figure 3-5 LARP1 La-Module (310-647) binds RPS20-mer and poly(A) 25-mer RNAs. EMSA analysis of La-Module(310-647) with (A) RPS6 20-mer, and (B) poly(A) 25-mer.

3.3.4 Contributions of the LAM and RRM to RNA binding

Having defined the C-terminus of the RRM, we next wanted to test the contribution of the LAM and RRM to RNA recognition by the La-Module. We purified recombinant LAM and RRM, and compared their affinities for poly(A) and TOP motif RNA using EMSAs (Figure 3-6). The LAM did not display RNA-binding activity for either poly(A) RNA or the RPS6 20-mer (Figure 3-6 A). By contrast, the RRM (amino acids 440-647) shifted oligonucleotides corresponding to the poly(A) 25-mer and RPS6 20-mer (Figure 3-6 B).

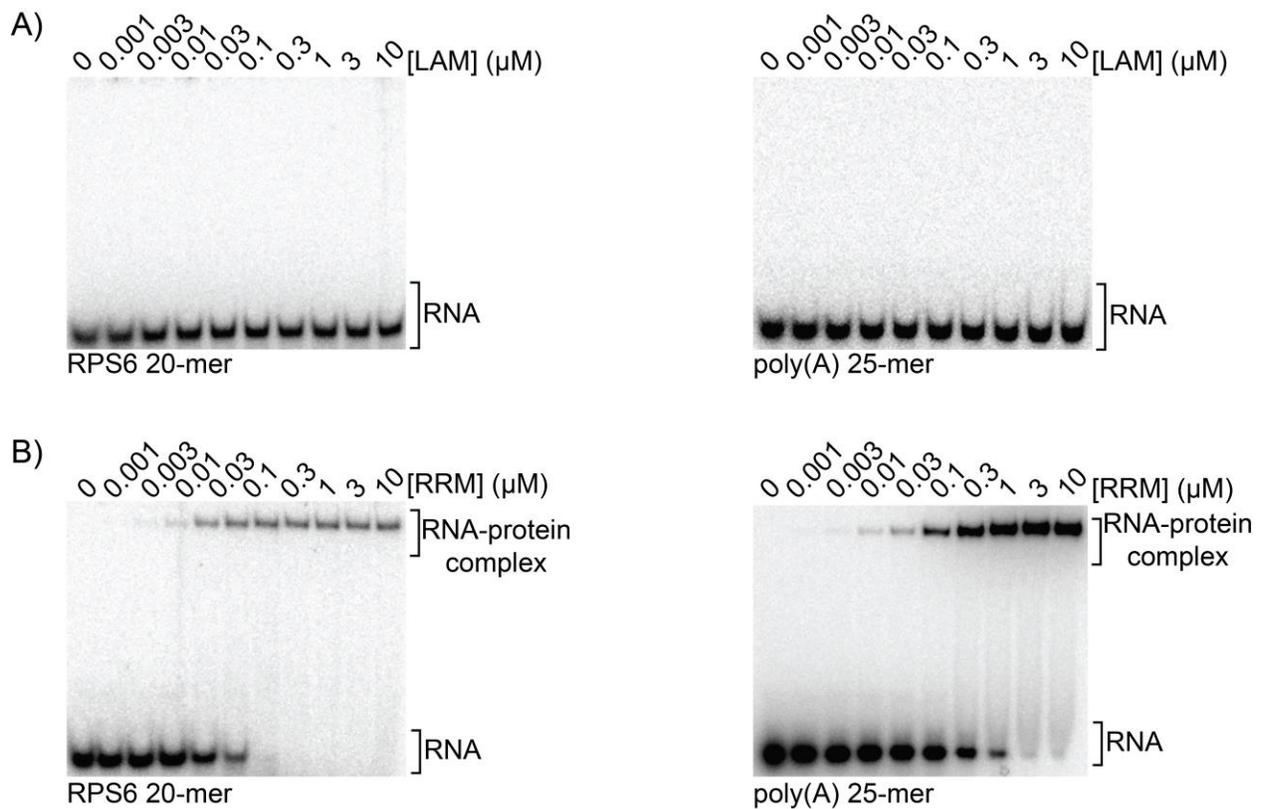


Figure 3-6 RNA-binding is driven by the linker and RRM region. EMSA analysis of: LAM (amino acids 310-405) with (A) RPS6 20-mer, and (B) poly(A) 25-mer; and linker-RRM region (amino acids 440-647) with (C) RPS6 20-mer, and (D) poly(A) 25-mer.

We next asked if tandem addition of the LAM or RRM would enhance RNA-binding of the RRM due to synergy of the domains (Figure 3-7). We first pre-bound the LAM and RRM, and titrated this complex into RPS6 20-mer (Figure 3-7 A). However, the observed shift corresponded to that of the RRM (Figure 3-7 A). We then pre-bound the RRM to RPS6 20-mer, then titrated in LAM (Figure 3-7 B). Again, the only observed shift corresponded to that of the RRM (Figure 3-7 B).

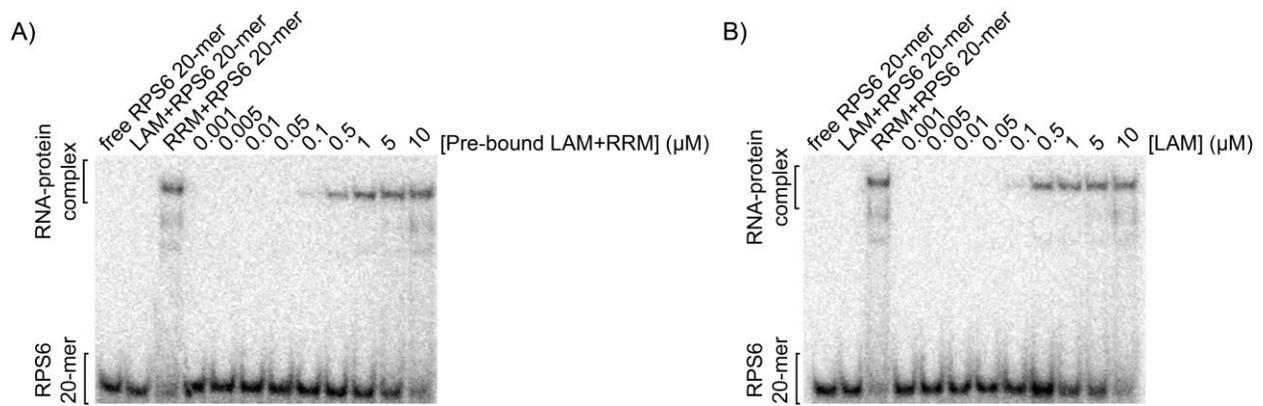


Figure 3-7 Tandem addition of the LAM and RRM suggest that the RRM region drives RNA-binding. EMSA analysis of (A) pre-bound LAM and RRM region titrated into RPS6 20-mer, and (B) LAM titrated into pre-bound RRM region and RPS6 20-mer.

3.3.5 La-Module (310-647) and RRM (440-647) can simultaneously bind poly(A) and TOP motif RNA

We next wished to confirm that the revised La-Module (amino acids 310-647) and RRM (amino acids 440-647) constructs can also bind poly(A) and TOP motif RNA simultaneously. Therefore, we conducted competition assays in which we pre-bound the La-Module or RRM to radiolabeled poly(A) RNA, then competed with cold RPS6 20-mer RNA (Figure 3-8 A, B). We detected an intermediate complex, albeit to a lesser extent than previously observed (Figure 3-8 A, B). Consistent with our previous findings, competition of RRM pre-bound to radiolabeled RPS6 20-mer with cold poly(A) RNA did not result in an intermediate complex (Figure 3-8 C).

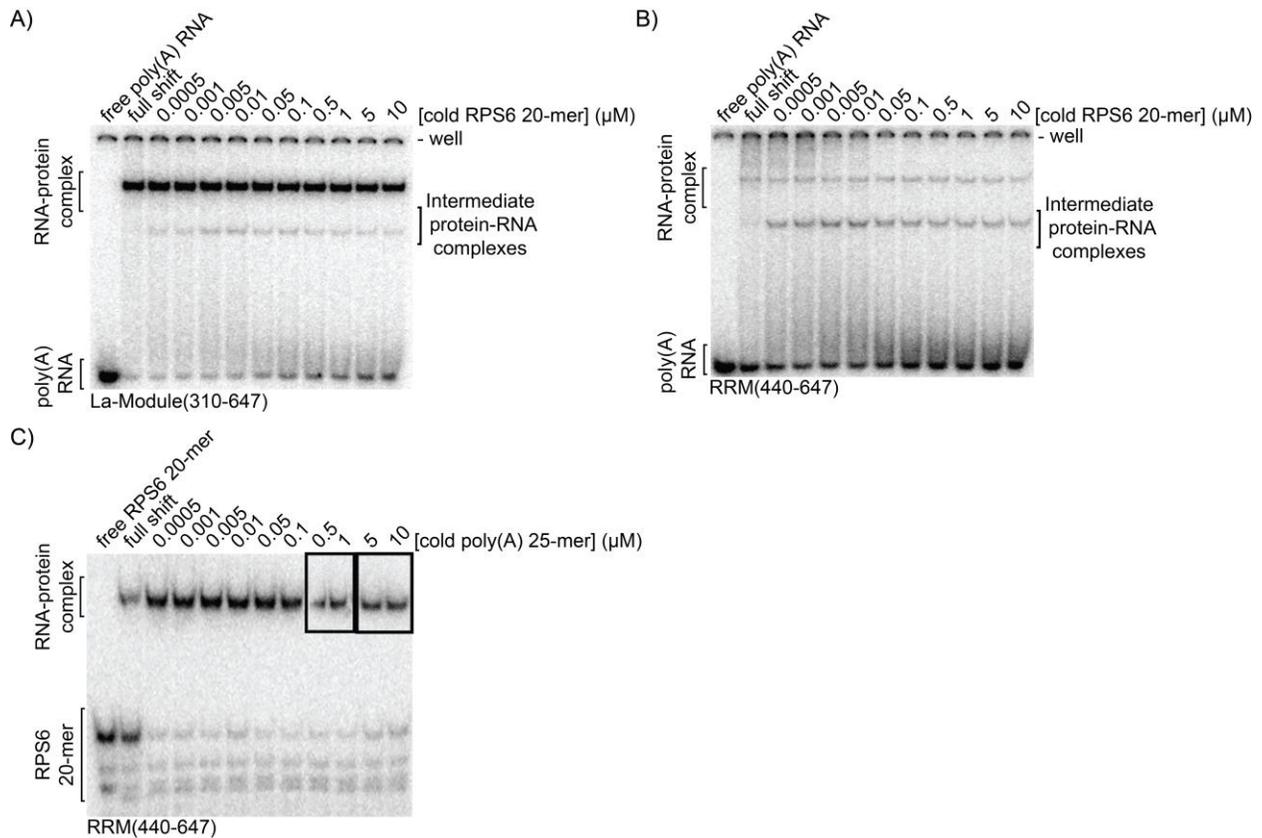


Figure 3-8 Revised La-Module and RRM form intermediate complex upon addition of RPS6 TOP motif to pre-bound poly(A) RNA. Competition assays conducted in the presence of poly(dI-dC) and analyzed by native gel of: (A) the La-Module-poly(A) RNA complex, (B) the RRM region-poly(A) RNA complex, and (C) the RRM region-RPS6 20-mer RNA, competed with cold RPS6 20-mer. Boxes in (C) indicate reconstructed region of the gel after a tear.

3.3.6 Attempts to identify the stoichiometry of the RRM (440-647) during simultaneous binding of poly(A) and TOP motif RNA

To determine the stoichiometry of the RRM (440-647) when simultaneously bound to poly(A) and TOP motif RNA, we first attempted static light scattering (SLS) experiments (Figure 3-9).

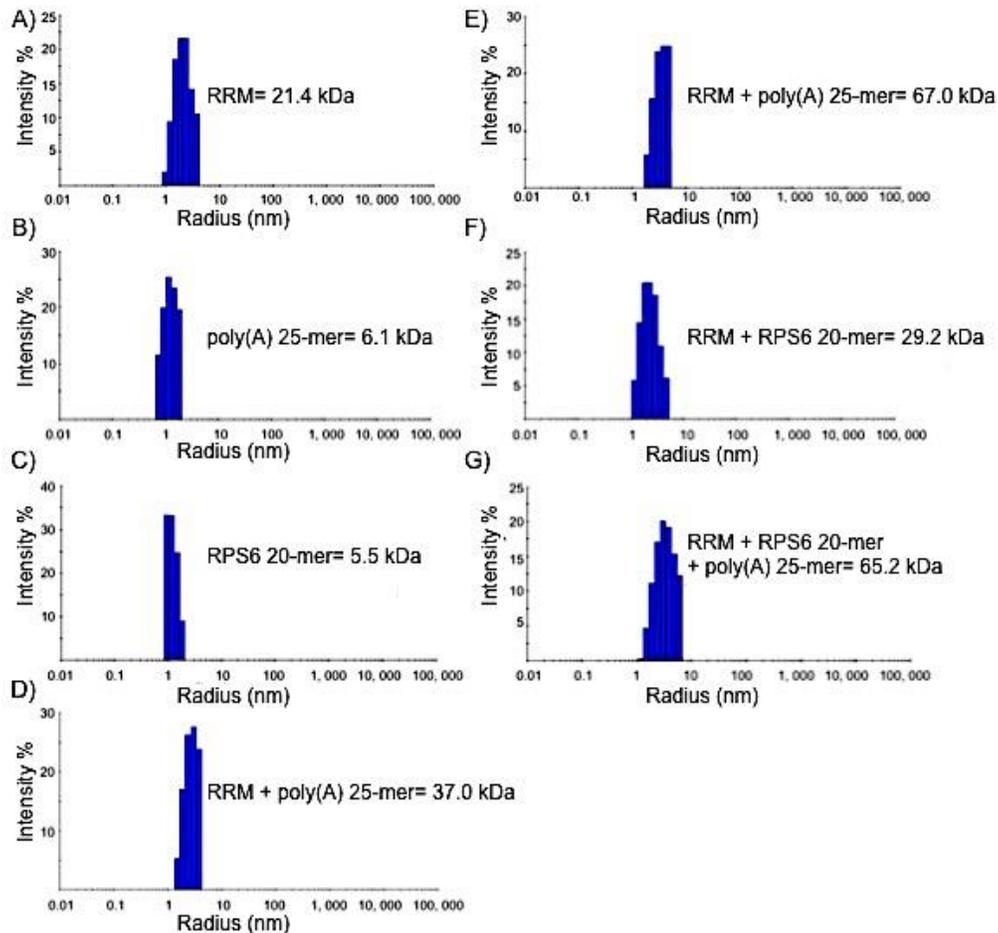


Figure 3-9 LARP1 RRM might bind poly(A) and TOP motif RNA as a monomer. SLS experiments measuring the molecular weight of: (A) RRM, (B) poly(A) 25-mer, (C) RPS6 20-mer, (D) RRM and poly(A) 25-mer, (E) RRM and poly(A) 25-mer, (F) RRM and RPS6 20-mer, and (G) RRM, poly(A) 25-mer, and RPS6 20-mer.

The SLS experiments suggested that native RRM has a molecular weight of ~21.4 kDa, close to its predicted 22 kDa molecular weight (Figure 3-9). The measured molecular weights of the poly(A) 25-mer and RPS6 20-mer were ~6.1 and 5.5 kDa, respectively (Figure 3-9). In addition, the RRM and poly(A) RNA complex was measured to be 67.0 kDa, which may suggest dimerization of the RRM during poly(A) binding (Figure 3-9). When we bound the RRM to either poly(A) 25-mer or RPS6 20-mer, the molecular weights of the complexes were 37.0 kDa and 29.2 kDa, respectively (Figure 3-9). SLS of the RRM, poly(A) 25-mer, and RPS6 20-mer, the resulted in a 65.2 kDa complex (Figure 3-9), which resembles the model of a single La-Module molecule binding to both RNAs using different RNA-binding surfaces. It is important to note that SLS **does not** provide accurate readings of molecular weight unless there is a significantly large difference

(see Discussion). Therefore, while these findings are interesting as they comment on the stoichiometry of the RRM when bound to either or both RNAs, they are not conclusive.

3.3.7 Attempts to identify the RNA binding surfaces of the LARP1 La-Module

To understand the molecular determinants of the La-Module-RNA interactions, we attempted to generate structural models using x-ray crystallography to help guide mutagenic analyses that would allow us to probe the functions of each domain with more specificity and rigor. We generated native crystals of the LAM (320-405) and RRM (440-647), as well as the La-Module (310-647) in the presence of either poly(A) 9-mer or RPS6 12-mer. Although we successfully generated and reproduced many of these crystals, they showed no diffraction (Figure 3-10).

Construct	No. Crystal Hits	Reproduced
460-647	Not attempted yet	N/A
440-647	7	Yes
440-545	0	N/A
543-647	Not attempted yet	N/A
320-405	3	Yes
320-540	0	N/A
320-572	Not attempted yet	N/A
320-647	Not attempted yet	N/A
310-405	1	Not attempted yet
310-540	0	N/A
310-572	Not attempted yet	N/A
310-647	2	Not attempted yet
	2	Yes
	Not attempted yet	N/A
	Not attempted yet	N/A

Figure 3-10 Crystallized LARP1 La-Module, LAM, and RRM constructs.

We hypothesized that the LARP1 La-Module contains a relatively long linker of ~100 amino acids (see Appendix 1) between the LAM and RRM, which could increase the intrinsic disorder and prohibit successful diffraction of the La-Module crystals. Therefore, we attempted glutaraldehyde cross-linking prior to cryopreservation of the RNA-bound La-Module crystals, in order to improve the diffraction quality. However, these did not diffract. We then generated and crystallized linker deletion mutants of the LARP1 La-Module (Figure 3-11). These crystals were

very fragile when manipulated for cryopreservation and the few crystals that were successfully cryoprotected did not diffract.

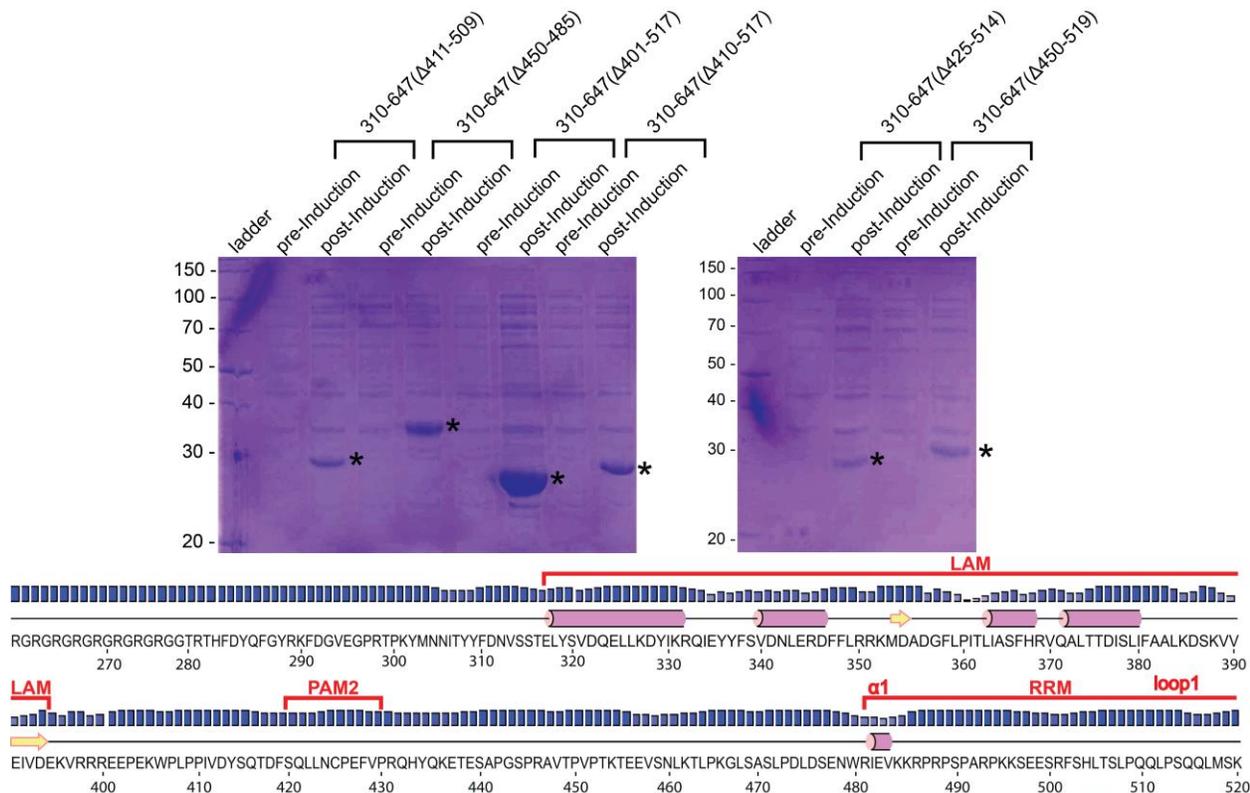


Figure 3-11 Expression of La-Module linker deletion constructs for crystallographic studies. Expression of LARP1 La-Module (amino acids 310-647) constructs with various linker deletions analyzed by SDS-PAGE. Corresponding secondary structure prediction shown below.

We speculated that loops within the RRM may contribute significant flexibility to the RRM. HSQC NMR in collaboration with Dr. Lisa Warner (Boise State University) showed that the LARP1 RRM region (amino acids 440-647) is intrinsically disordered (Figure 3-12 B). Furthermore, comparison of the HSQC NMR spectra of the LARP1 and LARP6 RRMs showed that the LARP1 RRM is much more disordered (Figure 3-12 C).

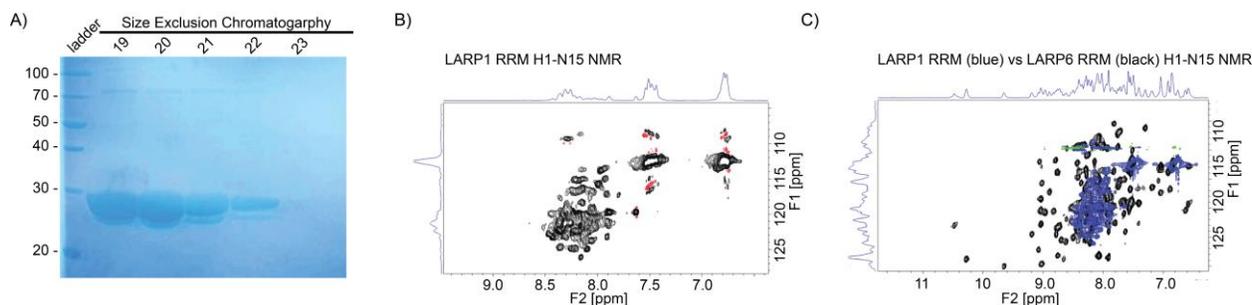


Figure 3-12 Recombinant purified LARP1 RRM is intrinsically disordered. (A) purified RRM region (amino acids 440-647), (B) HSCQC NMR of RRM region, (C) comparison of LARP1 RRM (blue) and LARP6 RRM (black) spectra.

3.4 Discussion

We attempted to identify the RNA-binding surfaces of the LARP1 La-Module construct (amino acids 310-540) by mutagenic analysis and native mass spectrometry. However, it presented great difficulties due to poor yield and RNA-binding activity. Due to the poor activity of the La-Module RNA-binding mutants (< 1%) we could not correct the affinities for RNA-binding activity, and thus could not directly compare affinities (Figure 3-1).

UV cross-linking to determine the stoichiometry of the La-Module when bound to poly(A) and TOP motif RNA was inconclusive as no difference in band migration was observed relative to the La-Module-poly(A) control (Figure 3-2 D). It may be that the La-Module-poly(A)-RPS6 20-mer ternary complex was not efficiently cross-linked due to the poor efficiency of UV cross-linking (typically less than 5% cross-link)²²⁵. Thus, the ternary complex could not be preserved in the SDS-PAGE as compared to the native gel (Figure 3-2 C, D). Indeed, cross-linked complex of the La-Module to either poly(A) RNA or RPS6 20-mer was only detected beginning at 0.3 μ M La-Module, with the majority of the RNAs uncross-linked even at 10 μ M (Figure 3-2 A, B).

We could not determine the stoichiometry of the La-Module in the ternary complex by SEC due peak broadening that could be caused by conformational heterogeneity²²⁶. The linker between the LAM and RRM in the LARP1 La-Module is predicted to be ~110 residues, and the RRM contains relatively long loops (see Appendix 1). This may allow for a significant amount of conformational plasticity, which is consistent with peak broadening. Furthermore, the low protein

yields and RNA-binding activity complicated SEC, as it was difficult to generate sufficient quantities of active protein that could be detected by UV absorbance.

To circumvent these issues, we revised the LARP1 La-Module (310-540) construct previously proposed in the literature¹⁵⁶ due the difficulties it presented. We generated constructs with varying C-termini in different vector backgrounds (Figure. 3-3) – guided by sequence conservation, secondary structure predictions, and homology models (see Appendix 1 for details) – to identify the optimal construct for future experiments. The revised construct, La-Module (310-647), exhibits robust expression, consistent purification and RNA-binding activity, as well as decreased aggregation and degradation (Figure 3-4, 3-5). Importantly, La-Module (310-647) binds both poly(A) and TOP motif RNAs with similar apparent affinities, consistent with our finds in La-Module (310-540) (Figure 3-5).

Using this optimized construct, we worked towards understanding the molecular mechanisms by the which the LARP1 La-Module recognizes poly(A) and TOP motif RNAs. We found that the linker and RRM region drive La-Module RNA binding (Figure 3-6). The LAM showed no RNA-binding, whereas the RRM region (440-647) shifted both poly(A) and RPS6 20-mer RNAs (Figure 3-5, 3-6). Tandem addition of the LAM and RRM region to RPS6 20-mer did not show formation of a supershifted complex relative to the RRM control (Figure 3-7). This may suggest that binding, at least to the RPS6 20-mer, is driven by the RRM and linker region (Figure 3-7). Taken together, these data suggest that RRM region (440-647) largely drives RNA-binding by the La-Module, while the LAM plays only a minor role, if any.

Importantly, the La-Module (310-647) and the RRM region (440-647) were able to simultaneously bind poly(A) and RPS6 20-mer RNAs, albeit to a lesser extent (Figure 3-8). Future work will be needed to determine the reason for this discrepancy. Consistent with our previous results (Chapter 2), the intermediate complex was only observed when RRM-poly(A) RNA complex was competed with cold RPS6 20-mer (Figure 3-8 B, C), suggesting ordered binding events that could reveal alternative conformations.

Next, we wanted to determine the stoichiometry and RNA-binding surfaces of the La-Module when simultaneously bound to poly(A) and RPS6 20-mer RNA. To determine the

stoichiometry of the RRM region when bound to both poly(A) and the RPS6 20-mer, we tried SLS. SLS suggested that the RRM region is a monomer when bound to both poly(A) RNA and RPS6 20-mer, and is also a monomer when bound to the RPS6 20-mer simultaneously (Figure 3-9). Interestingly, the RRM region bound was detected as both a monomer and a dimer when bound to poly(A) RNA, which may indicate that the RRM region has a propensity to dimerize (Figure 3-9). However, SLS cannot accurately compare changes in molecular weight, unless there is a significantly large difference, namely, the difference between a monomer and trimer (or larger)²²⁷. Small molecular weight differences, or even the difference between a monomer and dimer, cannot be concluded with confidence as it is out of the accuracy range of the equipment. Therefore, the results of the SLS experiments might be suggestive, but are not conclusive.

Our SLS experiments also showed polydispersity of the RRM region, which could indicate conformational heterogeneity due to disordered regions (Figure 3-9). Indeed, HSQC NMR spectra in collaboration with Dr. Lisa Warner (Boise State University) showed that the LARP1 RRM region (amino acids 440-647) is intrinsically disordered (Figure 3-12). The intrinsic disorder of the RRM and linker may allow for conformational flexibility, and may explain why the La-Module can bind poly(A) and TOP motif RNAs.

Currently, we are attempting to identify the RNA-binding surfaces of the La-Module (310-647) by two methods: 1) chemical cross-linking of the La-Module to RNAs, followed by protease digestion, and mass spectrometry, and 2) testing combinatorial RNA-binding mutants guided by sequence conservation and homology modeling for RNA-binding activity. In addition, we are exploring the use of cryoelectron microscopy to gain structural insight into the RNA-binding surfaces and conformation of the La-Module when bound to either poly(A) or RPS6 20-mer RNA.

4.0 Identification of LARP1-PABPC1 Interaction Interface

4.1 Introduction

PABP is a multifunctional protein with roles in nearly all stages of RNA metabolism. Cytoplasmic PABP (PABPC1), referred to as PABP from here on, functions in mRNA translation, stability, and deadenylation and decay. PABP plays a critical role in mRNA translation, as it binds mRNA poly(A) tails to protect from 3' → 5' exonucleolytic activity and interacts with eIF4G to enhance translation⁹⁴. The effect of PABP on mRNA translation can be both general and specific, depending on its interacting protein partner(s). Various effector proteins modulate mRNA translation through their interaction with PABP. Perhaps the most well characterized PABP-interacting proteins are PAIP1 and PAIP2, which compete to enhance and inhibit global mRNA translation, respectively^{48,228}.

Most proteins bind to PABP *via* its C-terminal MLLE domain, which consists of a five-helical bundle²²⁹ (Figure 4-1 A). The MLLE domain binds PABP-interacting proteins that harbor a PAM2 motif of ~15 amino acids²²⁹ (Figure 4-1 A). The MLLE-PAM2 interface is driven by hydrophobic interactions; conserved aromatic residues within the PAM2 bind hydrophobic pockets on either side of the MLLE²²⁹ (Figure 4-1 B). Some PABP-interacting proteins bind PABP through a PAM1, either independently or in concert with the PAM2²²⁸. While the sequence composition of the PAM2 is well defined, the PAM1 is defined more generally as an acidic region of ~ 25 amino acids. Known PAM1 sequences, such as that of PAIP1, bind RRM2, which is located toward the amino terminus of PABP²²⁸. However, the PAM2 is relatively more pervasive with numerous proteins utilizing this sequence to regulate the deadenylation, decay, stability, and translation of PABP-bound transcripts^{229,230}.

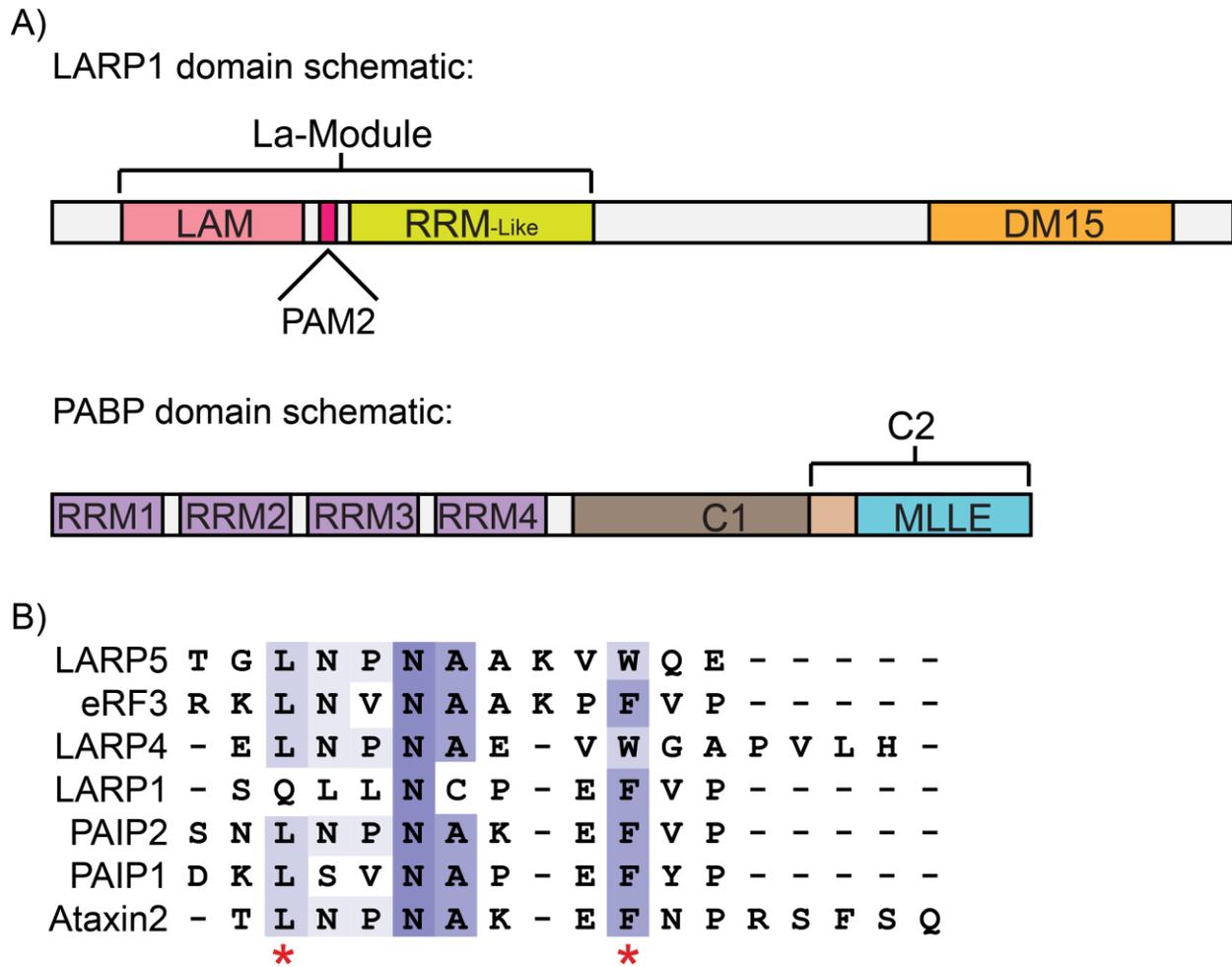


Figure 4-1 Schematic of LARP1 and PABP domain organization and PAM2 sequence conservation. (A) Schematic of LARP1 and PABP domain organization, (B) Alignment of various known PAM2 sequences. Key PAM2 residues (leucine and phenylalanine) denoted by red asterisks.

The interaction between the PABP MLLE domain and PAM2 sequences is critical to many processes of mRNA metabolism, including translation regulation²²⁹. PABP-interacting proteins, such as PAIP1 and PAIP2, regulate global protein translation levels through their interaction with PABP^{228,231}. Specific PABP-interacting proteins can be used to regulate the translation of certain subclasses of mRNAs^{232,233}. For example, ICP27 binds and recruits PABP to certain capped and polyadenylated viral and cellular mRNAs²³². More specifically, ICP27 competes with PAIP1 – a translation activator²²⁸ – for PABP binding²³². PABP bridges ICP27 and eIF4G, and this complex aids 40S recruitment during translation initiation²³². Similar to ICP27, Dazl promotes mRNA

translation in germ cells in an eIF4G-dependent manner downstream of cap-binding^{233,234}. Thus, translation regulation of a subclass of mRNAs can be mediated through RBPs with specificity for these transcripts and the ability to bind PABP.

LARP1 associates with PABP, seemingly through a putative PAM2 that is located in the linker between the LAM and RRM¹⁶⁷ (Figure 4-1 A, B). However, current studies disagree whether LARP1 associates with PABP in an RNA-dependent manner¹⁶⁷. LARP1 and PABP co-localize in the cell and co-sediment throughout polysome profiling experiments¹⁶⁷. In contrast to its interaction with TOP mRNAs, the LARP1-PABP association is unaffected by mTORC1 activity¹⁶⁷. This might suggest an importance for the LARP1-PABP complex that extends beyond TOP mRNA translation. For instance, the LARP1-PABP complex might be critical for targeting TOP mRNAs to SGs and P-bodies. Furthermore, PABP may aid in tethering the LARP1 La-Module to poly(A) tails or *vice versa*.

The LARP1 putative PAM2 is two amino acids shorter than canonical PAM2s¹⁶⁷ (Figure 4-1 B), which could lead to a novel binding modality. Numerous MLLE-PAM2 structural models have shown that, while the binding sites overlap, the PAM2 peptide conformations can vary. Typical PAM2 motifs stretch across the MLLE central helix ($\alpha 3$) – through a gap provided by the central glycine in the conserved KITGMLLE sequence – to dock the conserved leucine and phenylalanine residues into hydrophobic pockets on either side of the MLLE (Figure 4-1 B). GW182 contains an atypical PAM2; while the phenylalanine is conserved and binds to its corresponding pocket, the N-terminal leucine of the PAM2 is absent. The GW182 PAM2 flips back upon itself to form a hairpin-like structure that is facilitated through two β -turns that are stabilized by a network of hydrogen bonds, which includes additional contacts to the MLLE and several water molecules²³⁵. Thus, single amino acid changes to the PAM2 can lead to drastic changes to the PAM2 conformation. Importantly, variant PAM2 conformations allow PABP-interacting proteins to compete for PABP-binding through different binding affinities and modalities. The shorter LARP1 putative PAM2¹⁶⁷ might need to adopt a more “stretched” conformation in order to bind either side of the MLLE through the conserved leucine and phenylalanine residues (Figure 4-1 B).

We began our investigation by establishing that LARP1 and PABP directly interact with one another. We next asked which domain(s) of LARP1 and PABP mediate their interaction and the molecular determinants of binding. To do this, we mapped domain interactions using

recombinant purified proteins for GST pull-down assays. We found that the LARP1 La-Module binds the PABP MLLE domain in an RNA-independent manner. Unexpectedly, we also found preliminary evidence to suggest that the LARP1 La-Module could also contain a PAM1 region. Finally, we co-crystallized LARP1 putative PAM2 with the PABPC1 MLLE domain. We were able to collect diffraction data, but molecular replacement has been slowed by a pathology in the crystal in which two separate lattices seem to exist. Together, our data provide a framework for a mechanistic investigation as to the role of LARP1-PABP complex in TOP mRNA translation regulation.

4.2 Materials and methods

4.2.1 Cloning, expression, and purification of GST-tagged PABPC1 constructs

GST-tagged PABP constructs [RRM1+2 (1-190), RRM3+4 (191-368), C1C2 (369-633), C2 (495-636), MLLE (545-627)] were PCR amplified and cloned into pGEX6p1 expression vector using BamH1 and Sal1 sites. The resulting GST fusion proteins were expressed using *E. coli* BL21, and cultured at 37°C for two hours prior to IPTG induction and shifting to 17.5 °C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

For purification, 5 g frozen cells were resuspended in lysis buffer [25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10% v/v glycerol, 1 mM EDTA, 1 mM DTT, protease inhibitors (10 µM leupeptin, aprotinin, 10 µM bestatin, 1 µM pepstatin, 10 µM PMSF)]. Cells were lysed by homogenization and the lysate was cleared *via* centrifugation. GST-tagged constructs were purified in batch using 5 mL slurry of glutathione agarose resin equilibrated as per manufacturer's instructions (Goldbio, G-25-5). Cleared lysate was added to the resin and nutated for 3.5 hours at 4 °C. The resin was washed with three column volumes of lysis buffer. GST-tagged constructs were incubated in 1 cv elution buffer [10 mM reduced glutathione, pH 8] for 30 minutes to elute. The eluate was buffer exchanged overnight by dialysis at 4 °C in buffer [50mM Tris- HCl, pH 8, 150 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT]. The dialysate was concentrated to 0.75 mL by centrifugation and loaded onto a size exclusion column (Superdex 75, GE) equilibrated in

buffer [50 mM Tris-HCl, pH 8, 650 mM NaCl, 5% glycerol, 1 mM DTT]. Fractions containing GST-tagged protein of interest were buffer exchanged into storage buffer [50 mM Tris-HCl, pH 8, 250 mM NaCl, 10% glycerol] by 2L dialysis at 4 °C overnight, then concentrated to 50 μ M prior to storage.

4.2.2 LARP1 cloning, expression, and purification from *E. coli*

The sequence encoding LARP1 isoform 2 (Integrated DNA Technologies, codon optimized for bacterial expression) was PCR amplified and cloned into pET28a 6XHis-N-SUMO²²⁴ using BamH1 and Sac1 sites. The resulting 6XHis-SUMO-LARP1 construct was expressed in *E. coli* BL21(DE3) cells and cultured in autoinduction media at 37°C for two hours, then shifted to 17.5°C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

Cells were resuspended in buffer [25 mM HEPES-NaOH, pH 8.0, 750 mM NaCl, 10 mM imidazole, 10 % glycerol, 5 mM β -mercaptoethanol, and protease inhibitors (10 μ M leupeptin, aprotinin, 10 μ M bestatin, 1 μ M pepstatin, 10 μ M PMSF)]. Cells were lysed by sonication and the lysate was cleared *via* centrifugation. The cleared lysate was applied onto a HiTrap His FF (GE Healthcare Lifesciences) and 6XHis-SUMO-LARP1 was eluted with a 5-column volume imidazole gradient (20-350 mM). The 6XHis-SUMO was cleaved using 1 mg 6XHis-tagged ULP1 protease per 80 mL eluate for 2 hrs at 4°C in 2 L dialysis buffer [25 mM HEPES-NaOH, pH 8.0, 750 mM NaCl, 20 mM imidazole, 10 % glycerol, 5 mM β -mercaptoethanol, 5 μ M PMSF]. The cleaved 6XHis-SUMO tag and 6XHis-ULP1 were separated from LARP1 by re-applying onto HiTrap His FF and elution with an imidazole gradient (20-350 mM). LARP1 was buffer exchanged by dialysis for 2 hrs at 4°C in 2L buffer [25 mM Bis-Tris, pH 6.5, 200 mM NaCl, 10 % glycerol, 0.5 mM EDTA, 1 mM DTT].

LARP1 was further purified from protein and nucleic acid contaminants by tandem HiTrap S and HiTrap QP (GE Healthcare Lifesciences) chromatography with an NaCl gradient (200 mM-1M). LARP1 eluted from the S column, fractions were collected, concentrated to 0.5 mL on a Vivaspin 100K MWCO Centrifugal Concentrator (Sartorius), then loaded onto Superdex 200 size exclusion column (GE Healthcare Lifesciences) equilibrated in buffer [25 mM HEPES-NaOH, pH

7.5, 750 mM NaCl, 5 % glycerol, 0.5 mM β -mercaptoethanol] and eluted at 0.25 mL/min. Fractions containing LARP1 were collected, dialyzed into storage buffer (50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25% glycerol, 4 mM DTT) overnight at 4°C, concentrated to 50 μ M using a Vivaspin Turbo 50K MWCO Centrifugal Concentrator (Sartorius), frozen in liquid nitrogen, and stored at -80°C.

4.2.3 GST pull-downs

100 μ L glutathione agarose resin (Goldbio, G-250-5) slurry was washed as per manufacturer's instructions, then washed once more in 100 μ L pull-down buffer [20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 % glycerol, 0.02 % NP40]. Next, 50 μ M prey and bait proteins [stored in 50 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10 % glycerol] were combined in pull-down buffer for a final volume of 100 μ L. The pull-down reactions were added to the equilibrated resin, and a 7 μ L input sample was collected. Pull-down reactions were then nutated for 3.5 hours at 4 °C. Resin was collected by centrifugation at 500 x g for 5 min. 15 μ L flow-through samples were collected from each reaction, and the reactions were washed three times in 500 μ L pull-down buffer. The supernatant was dispensed and 15 μ L resin sample was collected. Input, flow-through, and pull-down samples were heat denatured in SDS loading buffer for 5 min at 95 °C then resolved on a 12.5 % SDS-PAGE.

4.2.4 MLLE cloning, expression, and purification

The PABP MLLE domain (545-627) was PCR amplified and cloned into a pHMG6 expression vector using Nhe1 and BamH1 sites. The MLLE domain contains no tryptophan residues for accurate determination of protein yields by spectrometry. Therefore, the 6XGly linker of pHMG6 was mutated to 2G-2W-2G using site-directed mutagenesis. The resulting fusion protein was expressed in *E. coli* BL21(DE3), and cultured in autoinduction media at 37°C for two hours prior to 17.5 °C for 18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80°C.

For purification, cells were resuspended in lysis buffer [50 mM Tris-HCl, pH 8.0, 400 mM NaCl, 10 mM imidazole, 10% v/v glycerol, protease inhibitors (10 μ M leupeptin, aprotinin, 10 μ M bestatin, 1 μ M pepstatin, 10 μ M PMSF)]. Cells were lysed by homogenization and the lysate was cleared *via* centrifugation. His-tagged MLLE was purified in batch using nickel agarose affinity chromatography (Thermo Fisher Scientific) and eluted [50 mM Tris-HCl, pH 8.0, 400 mM NaCl, 300 mM imidazole, 10% v/v glycerol]. The 6XHis-tag was removed by cleavage with 0.5 mg Tobacco Etch Virus protease per 10 mL eluate overnight at 4 °C in 2 L dialysis buffer [50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 0.5 mM EDTA, 0.5 mM DTT]. The cleaved 6XHis-tag was separated from the MLLE using a second nickel affinity chromatography step using a HiTrap Nickel FF (GE Healthcare Lifesciences). The MLLE was collected from the flowthrough and concentrated to 0.5-1 mL on a Vivaspin 5K MWCO Centrifugal Concentrator (Sartorius), then loaded onto Superdex 75 size exclusion column (GE Healthcare Lifesciences) equilibrated in buffer [50 mM Tris-HCl, pH 7.5, 750 mM NaCl, 5 % glycerol, 1 mM DTT] and eluted at 0.25 mL/min. Fractions containing MLLE were buffer exchanged *via* dialysis overnight at 4 °C and concentrated to 100 μ M for storage [50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 % glycerol, 1 mM DTT] or 10-15 mg/mL for crystallization [10 mM MES, pH 6.3, 100 mM NaCl]. MLLE and PAM2 co-crystallization

Co-crystallization conditions for MLLE and PAM2 peptide were identified using hanging drop vapor diffusion with the AM SO_4 crystallization suite (QIAGEN). The PAM2 peptide ($\text{H}_2\text{N-SQLLNCPEFVP-CONH}_2$) was synthesized and HPLC purified by the University of Pittsburgh Peptide Synthesis Facility. The best crystals were grown at room temperature using 1 μ L 1:1.2 ratio of MLLE:PAM2 pre-bound for 30 min on ice with 1-2 μ L reservoir solution [Condition 49: 0.1 M citric acid pH 4.0, 0.8 M ammonium sulfate]. Crystals appeared within three days and grew over the course of a week.

4.2.5 MLLE and PAM2 co-crystallization

Co-crystallization conditions for MLLE and PAM2 peptide were identified using hanging drop vapor diffusion with the AM SO_4 crystallization suite (QIAGEN). The PAM2 peptide ($\text{H}_2\text{N-SQLLNCPEFVP-CONH}_2$) was synthesized and HPLC purified by the University of Pittsburgh

Peptide Synthesis Facility. The best crystals were grown at room temperature using 1 μ L 1:1.2 ratio of MLLE:PAM2 pre-bound for 30 min on ice with 1-2 μ L reservoir solution [Condition 49: 0.1 M citric acid, pH 4.0, 0.8 M ammonium sulfate]. Crystals appeared within three days and grew to 250-500 μ m over the course of a week.

4.3 Results

4.3.1 LARP1 La-Module binds PABP MLLE

LARP1 associates with PABP *via* a putative PAM2 located between the LAM and RRM¹⁶⁷. Previous data in the lab suggested that the LARP1 La-Module binds the PABP C2 region through a PAM2 (Gabby Ciotti, data not shown), but not the MLLE domain. Based on these data, we wished to use GST pull-downs to identify the minimal region of the PABP C2 required for La-Module binding. First, we verified that the La-Module (amino acids 310-540) binds the PABP C2 region, which includes the MLLE domain (Figure 4-2 A). We then conducted pull-downs with N-terminal truncations of the C2 to narrow down the region that binds the La-Module (Figure 4-2 B, C, D). A construct containing the MLLE plus an additional 12 amino acids N-terminal binds the La-Module (Figure 4-2 D).

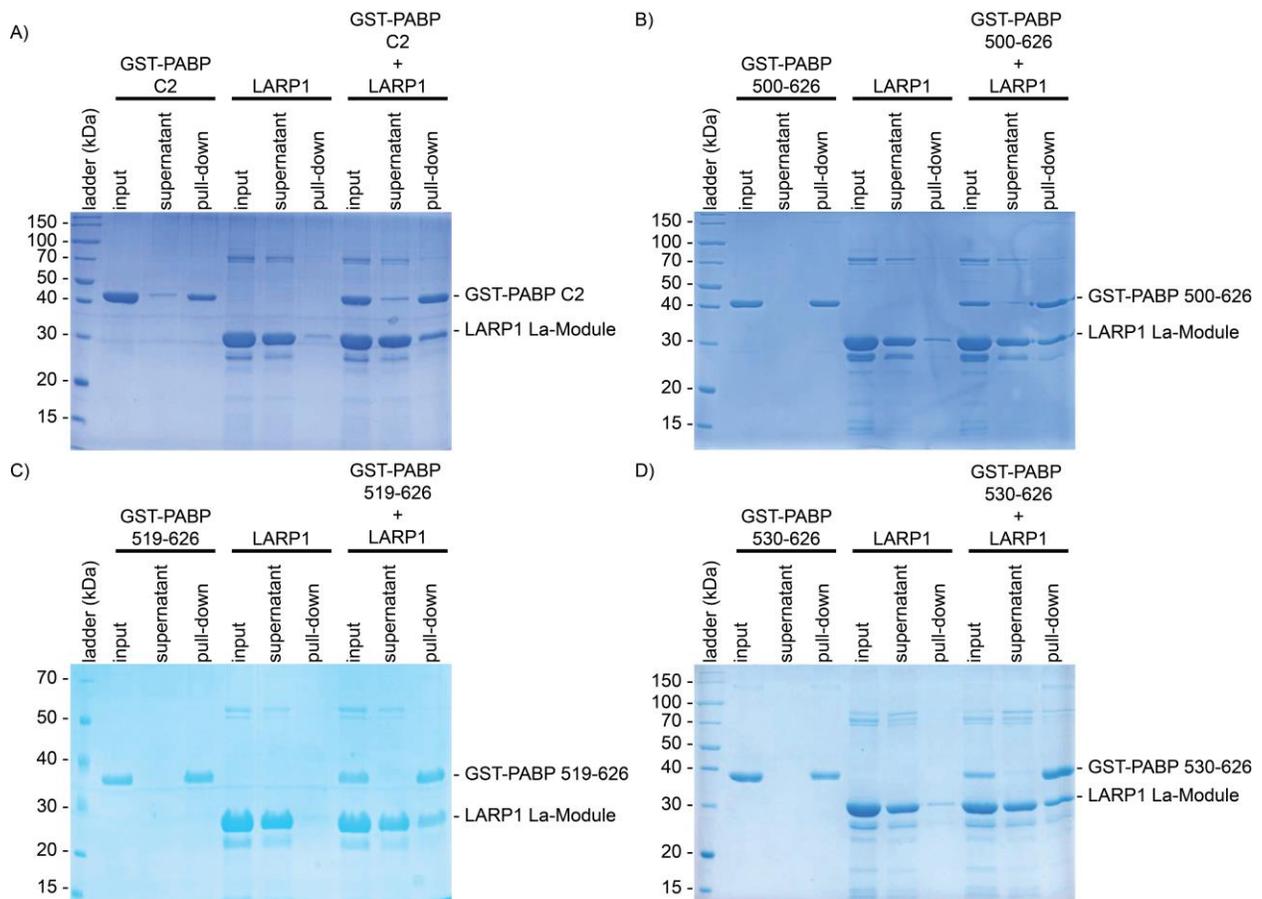


Figure 4-2 LARP1 La-Module binds the PABP MLE region. Pull-down of the La-Module with GST-tagged: (A) PABP C2, (B) PABP (amino acids 500-626), (C) PABP (amino acids 519-626), (D) PABP (amino acids 530-626). Proteins resolved by SDS-PAGE.

4.3.2 LARP1 La-Module binds the PABP RRM3s

Because typical PABP-interacting proteins bind PABP through the MLE domain, we hypothesized that LARP1 only binds PABP through the MLE domain. To test our hypothesis, we conducted pull-downs of recombinant purified LARP1 using GST-tagged PABP RRM1+2 and PABP RRM3+4 as bait (Figure 4-3 A, B). The GST-tagged RRM constructs captured LARP1 in a 1:1 ratio (Figure 4-3 A, B). We next asked which region(s) of LARP1 and the PABP RRM3s interact. We conducted pull-downs of the La-Module with GST-tagged PABP RRM1+2 or RRM3+4 as bait (Figure 4-3 C, D). We found that PABP RRM1+2 captured the La-Module in a 1:1 ratio (Figure 4-3 C, D).

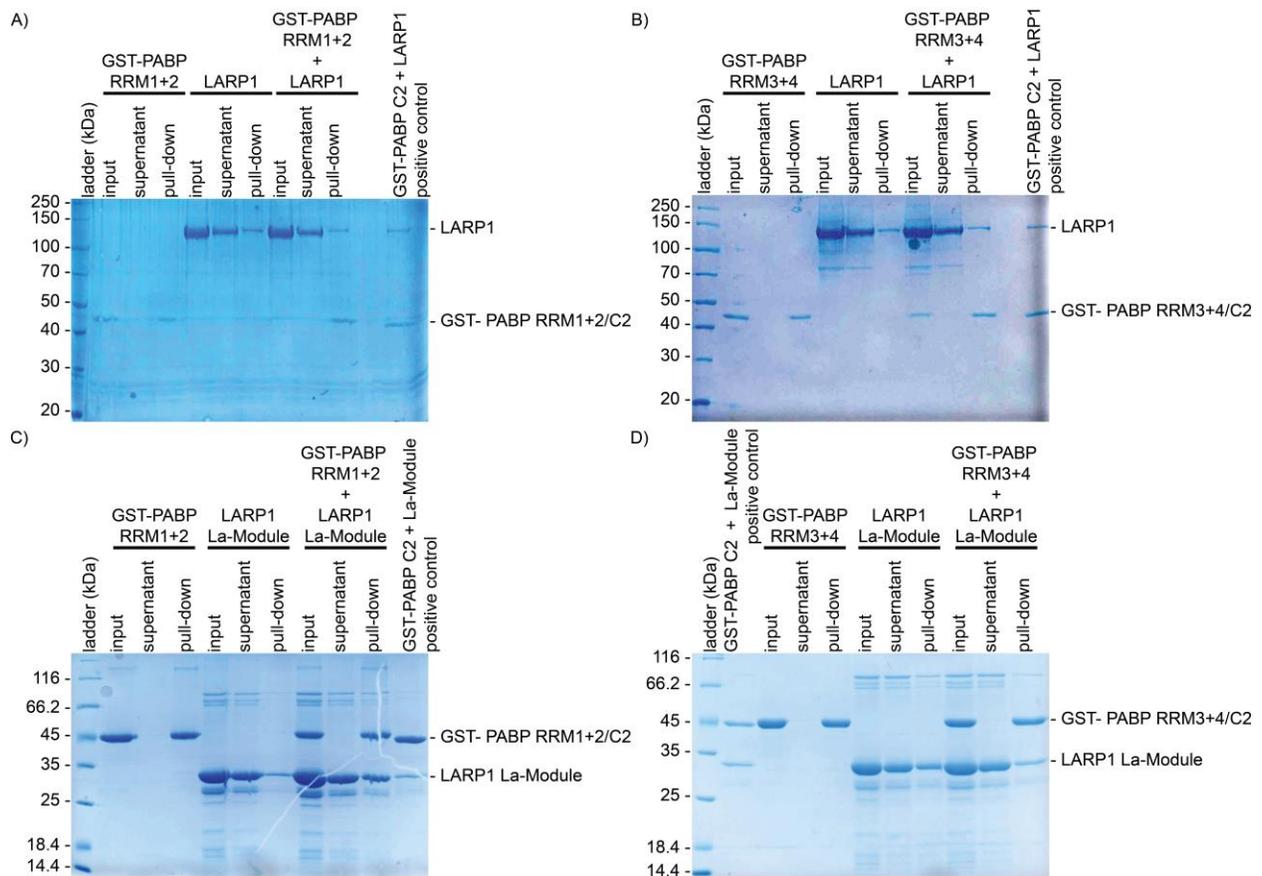


Figure 4-3 LARP1 La-Module binds PABP RRM1+2. Pull-down of recombinant purified WT LARP1 with GST-tagged (A) PABP RRM1+2, (B) PABP RRM3+4. Pull-down of recombinant purified La-Module with GST-tagged (A) PABP RRM1+2, (B) PABP RRM3+4. Proteins resolved by SDS-PAGE.

4.3.3 Crystallization of PABPC MLLE domain and LARP1 PAM2

The LARP1 PAM2 is two amino acids shorter than the consensus sequence (Figure 4-3). We hypothesized that this could lead to a novel conformation of the PAM2, in which it must adopt a more “stretched” conformation in order to bind the hydrophobic pockets on either side of the MLLE. For this reason, we sought to attain a structural model of the MLLE domain bound to the LARP1 PAM2. We co-crystallized the MLLE with LARP1 PAM2 and collected diffraction data. Thus far, molecular replacement has been complicated by a crystallographic pathology that is still being examined.

4.4 Discussion

LARP1 co-localizes with PABP to RNA granules and co-sediments with PABP throughout polysome profiling gradients¹⁶⁷. Furthermore, a putative PAM2 located between the LAM and RRM domains is necessary for the PABP association (Figure 4-1)¹⁶⁷. Interestingly, the LARP1-PABP association is insensitive to mTORC1 signaling¹⁶⁷, which could suggest a role for this interaction that extends even beyond mTORC1-dependent TOP mRNA translation regulation. From these data, we hypothesized that a direct LARP1-PABP interaction may promote TOP mRNA translation regulation and sequestration of these transcripts to SGs and P-bodies. We sought to determine how LARP1 and PABP interact in order to gain insight into the mechanism by which this complex could mediate TOP mRNA metabolism.

We found that the LARP1 La-Module binds the PABP MLLE region in an RNA-independent manner (Figure 4-2). Unexpectedly, we also found an interaction between the La-Module and PABP RRM1+2 (Figure 4-3), the drivers of poly(A) binding in PABP^{236,237}. This interaction could be mediated by a PAM1 sequence – as seen in PAIP which utilizes a PAM1 to bind PABP RRM1+2 – generally defined as a stretch of acidic residues²²⁸. The La-Module contains a helix rich in acidic residues (amino acids 521-540) within the RRM (Figure 5-4), which may be used as a PAM1. Though the putative PAM1 must be validated, it presents an interesting finding as it suggests that the La-Module may simultaneously bind the N-terminal RRMs and C-terminal MLLE domains of PABP (Figure 4-1). A ~255 linker separates the MLLE from the RRMs in PABP (Figure 4-1), which may provide sufficient flexibility for the C-terminal MLLE to localize near the N-terminal RRMs. This may allow simultaneous binding of the La-Module to RRM1+2 and the MLLE domains of PABP. Binding to PABP through two different motifs may enhance the LARP1-PABP interaction through cooperative binding events. In addition, PABP-interacting proteins that harbor either a PAM1 or PAM2 motif may displace LARP1 from one of its corresponding PABP binding sites. These proteins could mediate sequestration into RNA granules, as well as recruit or occlude mRNA translation or decay factors. Ultimately, the presence of two PABP binding motifs may serve as a switch for different biological processes.

The LARP1-PABP interaction could be used to specifically recruit PABP and PABP-associated proteins, to TOP mRNAs. Specificity of both the DM15 and the La-Module for TOP motifs would tether LARP1 to TOP mRNAs, while the La-Module PAM2 recruits PABP to the

poly(A) tails of these transcripts. The LARP1-PABP complex might have various purposes in TOP mRNA translation regulation. For instance, the LARP1-PABP complex may aid in 40S recruitment to TOP mRNAs. LARP1 depletion leads to a loss in TOP mRNAs associated with small non-polysomes, particularly the 40S²³⁸. However, the molecular constituents of 40S recruitment to TOP mRNAs are not defined²³⁸. Because PABP has been implicated in 40S recruitment, it's possible that the LARP1-PABP complex specifically recruits 40S to TOP transcripts. Anchoring of PABP to TOP mRNA poly(A) tails might also prevent deadenylation in translationally repressed TOP mRNAs within the cytoplasm and SGs. In P-bodies, LARP1 may be displaced from PABP to allow for TOP mRNA deadenylation and decay (Figure 4-4). P-body-associated protein GW182, which functions in microRNA-mediated deadenylation and decay, binds the PABP MLE domain through PAM2s²³⁵ (Figure 4-4). These PAM2s could displace the LARP1 PAM2 and recruit the mRNA decay machinery to TOP mRNAs (Figure 4-4).

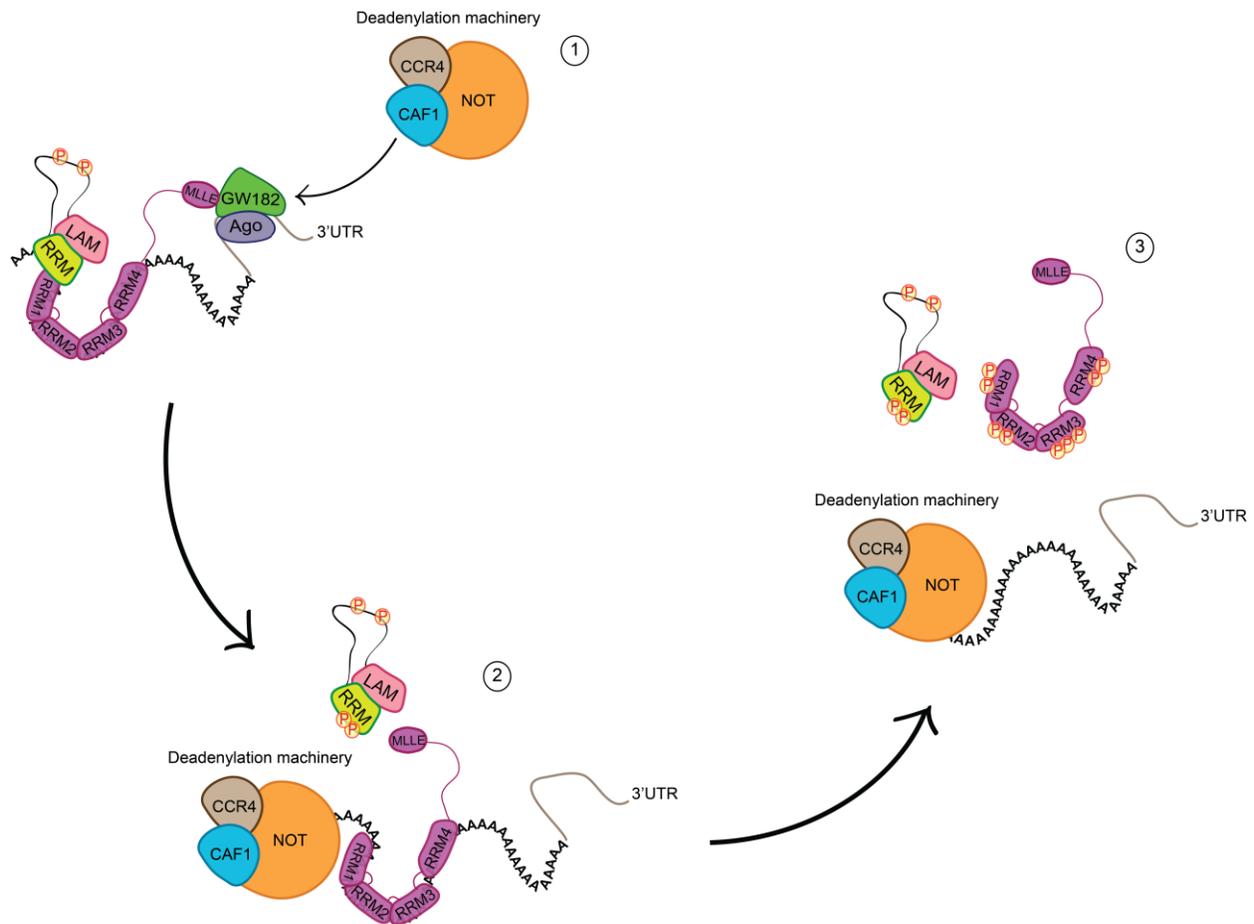


Figure 4-4 Model of LARP1 PAM2 displacement for deadenylation during mRNA decay in P-bodies. Step-wise phosphorylations to the La-Module and PABP can allow for recruitment of deadenylation machinery to TOP

mRNAs in P-bodies. (1) Phosphorylation near LARP1 PAM2 region allows GW182 PAM2 to displace LARP1 PAM2 and recruit deadenylation machinery, (2) phosphorylation to the LARP1 RRM releases La-Module from TOP mRNAs, (3) phosphorylation of PABP RRMs releases PABP from TOP mRNA poly(A) tails for deadenylation.

Because multiple PABPs can bind a single poly(A) tail, several opportunities exist for PABP-interacting proteins to bind and regulate post-transcriptional processes in the cytoplasm. What mechanisms coordinate the association and/or dissociation of the MLLE and PAM2-containing proteins? PAM2 sequences are located within IDRs and are usually proximal to potential serine and threonine phosphorylation sites²³⁹. Phosphorylation of residues near the PAM2 impedes MLLE binding, diminishes the interaction with PABP, and impairs biological functions²³⁹. For example, phosphorylation near the PAM2s of PAN3 and Tob2 decreases their ability to promote deadenylation in mammalian cells²³⁹.

In addition, PABP itself undergoes extensive post-translational modifications (PTMs), such as phosphorylation, methylation, and acetylation²⁴⁰. Modifications were identified at the MLLE, each of the RRMs, as well as the proline-rich linker between the RRMs and MLLE (Figure 4-1)²⁴⁰. The PTMs change throughout the cell cycle, and molecular modeling suggests that they can diminish interactions between PABP and PAM2-containing proteins as well as with RNA²⁴⁰. Thus, the interaction between PABP and PABP-interacting proteins can be regulated through reversible PTMs to either protein. Differential PTM of PABP and PABP-interacting proteins in response to upstream signaling may coordinate these interactions and the biological process that they mediate.

Various phosphorylation sites have been identified in LARP1, including phosphorylation near the PAM2²⁰⁴ (Figure 4-4, 4-5). Phosphorylation near the PAM2 may occlude binding to the MLLE, which was previously found to be necessary for the LARP1-PABP association¹⁶⁷ (Figure 4-4, 4-5). In P-bodies, this may release the LARP1 PAM2 and allow PAM2-containing deadenylase machinery to bind the MLLE for TOP mRNA degradation (Figure 4-4). Although not investigated yet, differential PTMs of PABP may also coordinate its localization, and therefore localization of associated mRNPs, to either SGs or P-bodies during translation repression.

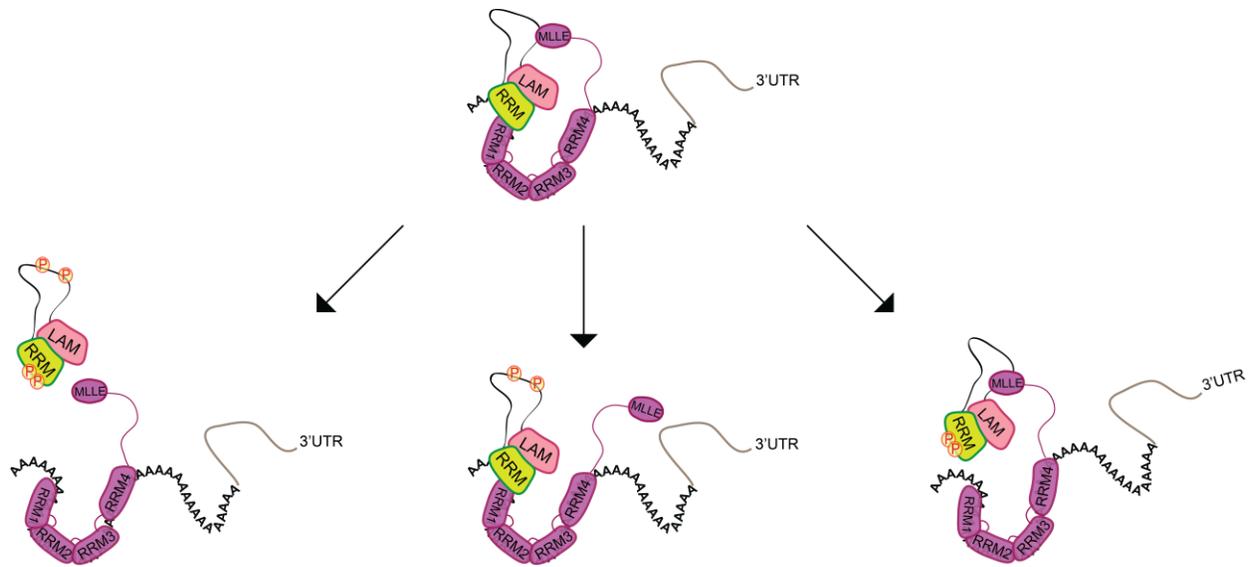


Figure 4-5 Phosphorylation of LARP1 PAM2 and PABP may affect interactions. Interactions between the LARP1 La-Module and PABP can be dynamically modified through reversible phosphorylations to the La-Module PAM2 and RRM.

5.0 Discussion

5.1 Role of the LARP1 La-Module in TOP mRNA recognition

Our investigation implicates the LARP1 La-Module in RNA recognition at both the 3' and 5' ends of TOP mRNAs. We found that the LARP1 La-Module directly binds poly(A) RNA and PABP, suggesting a role for the La-Module at poly(A) tails. This is consistent with previous observations that LARP1 associates with poly(A) RNA¹⁹⁹ and PABP, and that the PABP association is mediated by a putative PAM2 located between the LAM and RRM¹⁶⁷. Unexpectedly, we also found that the LARP1 La-Module binds the 5'UTRs of some TOP mRNAs. This observed binding to these 5'UTRs is cap-independent but requires both the pyrimidine and GC-rich region of the TOP motif. Furthermore, the LARP1 La-Module is able to simultaneously engage both poly(A) and TOP motif RNA.

We extended our study of the La-Module within the context of full-length LARP1. We found that the RNA-binding behavior of recombinant purified La-Module is faithful in the context of full-length LARP1 purified from human cells. Both WT and REYA (mutations in the DM15 region of LARP1 that abrogate its role in RNA-binding activity) LARP1 bind poly(A) and TOP motif RNA, and are also able to simultaneously bind these sequences. While these data do not exclude the possibility that regions outside of the La-Module and DM15 region contribute to RNA-binding, they do support the idea that the RNA-binding trend of recombinant purified La-Module is possible in the context of the 3 dimensional fold of full-length LARP1.

The ability of the LARP1 La-Module to simultaneously bind both poly(A) and TOP motif RNA could suggest that this RNA-binding unit exists at both ends of TOP mRNAs at the same time (Figure 5-1). The affinity of the La-Module for poly(A) RNA and PABP may not be sufficient to specifically recruit the La-Module to the poly(A) tails of only TOP mRNAs. Docking of the La-Module on TOP mRNA poly(A) tails could be enhanced through cooperative TOP motif recognition by the La-Module and the DM15 (Figure 5-1). The C-terminal LARP1 DM15 region binds the 5'cap and first four nucleotides of the TOP motif to inhibit formation of the eIF4F translation initiation complex during mTORC1 inhibition (Figure 5-1)^{167-169,171}. Accordingly, the

La-Module binds TOP motifs in a cap-independent manner. Furthermore, although the La-Module requires the TOP motif for binding, it does not bind sequences lacking the adjacent GC-rich region (Figure 2-2). These data suggest that the La-Module might bind sequences downstream of the DM15 binding site, perhaps at or near the junction between the TOP motif and GC-rich region (Figure 5-1).

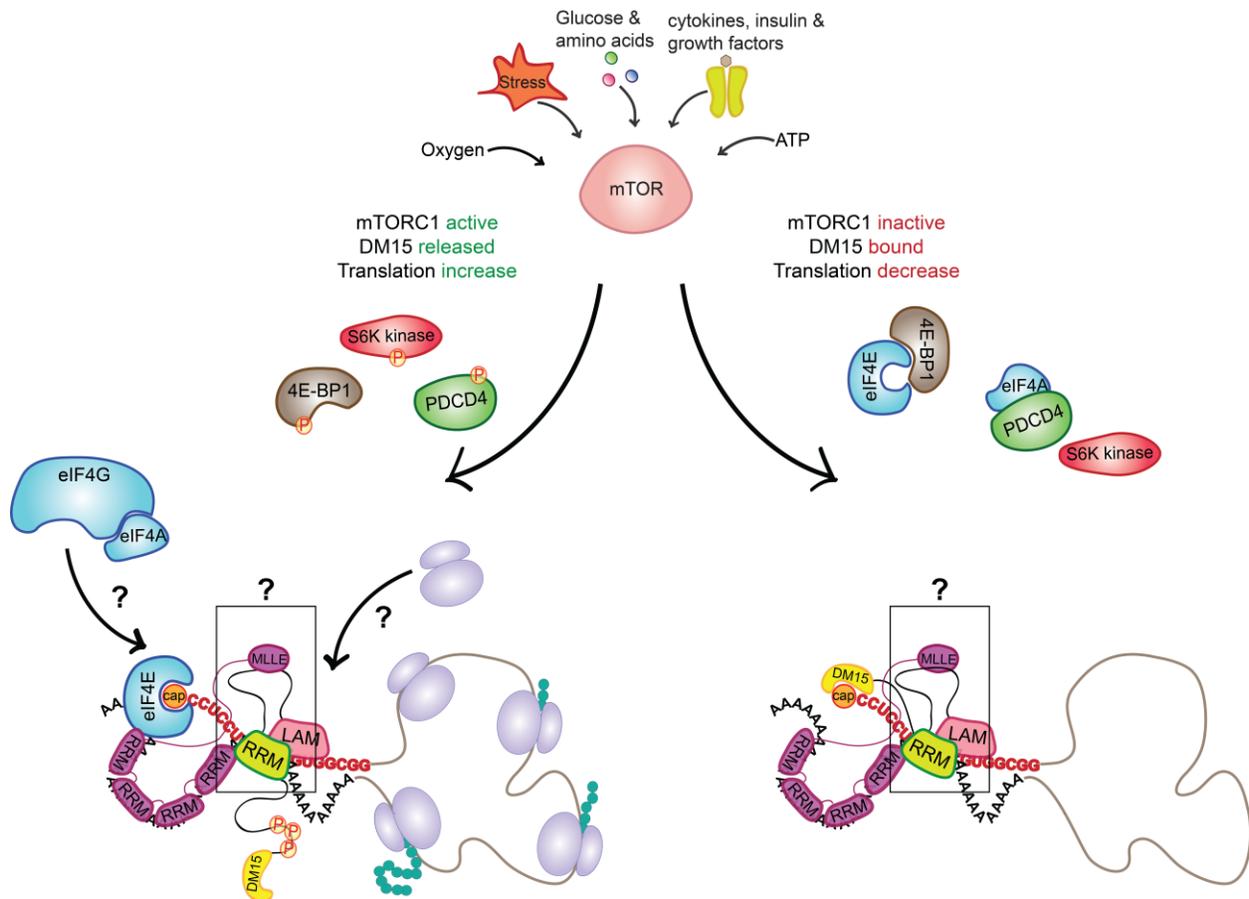


Figure 5-1 Model for TOP mRNA translation regulation through LARP1 and outstanding questions. Question marks denote questions: (1) How, if at all, does eIF4G bind while the La-Module is bound? (2) What rearrangements of the LAM, RRM, linker, MLE, poly(A) RNA, and TOP motif RNA permit translation activation versus repression? (3) How does ribosomal scanning occur if the La-Module is bound to the 5' UTR?

However, one important question is: why must the La-Module bind TOP motifs at all? Theoretically, binding of the DM15 to the 5' cap and TOP motif would be sufficient to anchor the La-Module to TOP mRNA poly(A) tails. One possibility is that the La-Module contributes to the translation state of TOP mRNAs by engaging both ends of the transcript. Several mTORC1

phosphorylation sites cluster at or near the DM15 region²⁰⁴. Depositing a negative charge in this region might cause the DM15 region to release the cap and TOP motif due to charge repulsion or induced conformational change (Figure 5-1). With the 5' cap freed, the eIF4F translation initiation complex can assemble (Figure 5-1). Simultaneous engagement of the TOP motif and poly(A) tail by the La-Module might facilitate TOP mRNA translation through circularization, which might enhance ribosome recycling from the 3' to the 5' end (Figure 5-1). In addition, the La-Module might also help to anchor PABP to TOP mRNA poly(A) tails, which can then aid in the recruitment of translation initiation factors (Figure 5-1). Precedence exists for the enhancement of translation by La-Modules in other LARPs, as La^{218,222,223,241}, LARP4¹⁶⁴, LARP5^{166,211}, and LARP6^{161,162} have been shown to increase mRNA translation. Assuming that the La-Module simultaneously engages TOP motifs and poly(A) tails in order to enhance translation, mechanistic details of TOP mRNA translation initiation must be characterized.

5.1.1 Future direction: How does the LARP1 La-Module affect TOP mRNA translation?

The biological function of the La-Module must be investigated to deepen our understanding of how LARP1 regulates TOP mRNA translation. Translation assays can be used to test whether the LARP1 La-Module enhances TOP mRNA translation, similar to the La-Modules of other LARPs^{162,167,220,222,241}. Specifically, WT La-Module can be compared to RNA-binding and PAM2 mutants to identify how binding to poly(A), TOP motifs, and PABP affect TOP mRNA translation (Figure 5-1). Furthermore, comparison of WT and phosphomimetic LARP1 can be used to test whether phosphorylation of LARP1 near the DM15 or PAM2 enhances TOP mRNA translation (Figure 5-1). Next, cryo-EM and/or single-molecule fluorescence studies can be used to investigate the possibility of mRNA circularization, as well as LARP1 conformational changes in response to phosphorylation, RNA, and PABP binding. This will allow us to understand how conformational changes instigated by mTORC1-mediated phosphorylation or binding partners could allow LARP1 to toggle between permitting and suppressing translation (Figure 5-1). These experiments can then be expanded upon to identify how the TOP mRNP is arranged during translation activation versus repression (Figure 5-1). More specifically: Does an intact eIF4F translation initiation complex assemble? If so, how do eIF4G and eIF4A arrange

themselves relative to the La-Module? Does the La-Module rearrange itself, and its interactions during translation regulation? How does the 43S scan to find the translation start codon if the La-Module is bound?

The La-Module might bind TOP motif sequences downstream of the eIF4G and eIF4A binding sites. This is supported by the potential requirement for the GC-rich region for TOP mRNA 5' UTR binding (discussed above). Furthermore, in yeast, eIF4G preferentially binds unstructured sequences containing poly(U) stretches²⁴², and uridine is enriched within the first few nucleotides of TOP mRNAs regulated by LARP1²⁴³. To test this hypothesis, the nucleotide specificity of the La-Module could first be determined by cross-linking immunoprecipitation sequencing (CLIP-seq) with REYA LARP1 and validated *in vitro* by RNase footprinting using recombinant purified La-Module. These data can then be compared to the eIF4G and eIF4A-binding sites determined using the same methods. Translation assays with reporters containing mutations to the RNA-binding sites of eIF4G, eIF4A, and La-Module can then be used to determine the affect of these proteins upon TOP mRNA translation.

The La-Module could change its interactions during TOP mRNA translation regulation. For example, the La-Module may simultaneously engage the TOP motif and poly(A) tails during translation repression. Phosphorylation of specific La-Module regions upon mTORC1 activation might cause it disengage from the TOP motif prior to translation. This could clear the TOP motif for eIF4A helicase activity and 43S scanning. To determine whether the La-Module remains bound to both ends of TOP mRNAs before and after mTORC1 activation, the RNA-binding surface(s) of the La-Module and their role in TOP mRNA translation must first be identified (discussed in 5.2.1). La-Module phosphomimetic mutants can then be compared to WT and RNA-binding mutants in translation assays to test if phosphorylation of certain regions of the La-Module releases it from TOP motifs to permit translation. Differences in binding affinity and kinetics of the La-Module for TOP motif and poly(A) RNA before and after phosphorylation by mTORC1 could mediate such rearrangements. This could be explored by comparing binding thermodynamics of WT and phosphomimetic La-Module for TOP motifs and poly(A) RNA using isothermal titration calorimetry (ITC).

5.2 Canonical faces of LAM and RRM for RNA-binding activity

In our study of the LARP1 La-Module we found that a region encompassing the linker and RRM (amino acids 440-647) drives RNA-binding. This RRM region was able to bind poly(A) and TOP motif RNAs with similar affinities, and initial experiments showed that it might also bind both of these sequences simultaneously. This observation is consistent with a recent study of the LARP4 La-Module, in which the RRM contributed to poly(A) binding while the LAM did not bind this sequence¹⁶⁵. Although structural studies in La and LARP7 suggested that the canonical faces of the LAM and RRM do not bind RNA, other studies of La indicate that these surfaces participate in RNA-binding and may even mediate discrete biological functions^{223,244}.

In Genuine La, synergistic binding of the LAM and RRM to the 3'UUU-OH pre-tRNAs is not sufficient to recognize and fold severely impaired pre-tRNAs *in vivo*²⁴⁴. The RRM β -sheet is required to fold these defective pre-tRNAs, suggesting functional modularity of the RRM and perhaps use in a distinct step of pre-tRNA folding and maturation²⁴⁴. Consistently, chemical shift mapping experiments detected some binding to the canonical surfaces of the LAM and RRM¹⁸¹. Recently, the canonical face of the LAM was shown to bind poly(A) RNA in a sequence and length dependent manner, and binding to poly(A) RNA allows La to enter polysomes²²³. These data agree with previous findings suggesting that La protein promotes the translation of both cap- dependent and independent translation of cellular and viral RNAs^{218,220,222,241}. Therefore, the canonical RNA-binding surfaces of the LAM and RRM in Genuine La not only engage different RNAs, but also facilitate distinct biological functions.

The use of different RNA-binding surfaces within the LAM and RRM could provide many advantages to LARPs. RNA recognition and specificity can be increased through binding of the same RBD to different regions of the same RNA molecule. This could aid biological functions, such as RNA chaperone activity, as multiple transient interactions can recognize and correct aberrant RNA folding. Such is the case with p53, which guides assembly of the telomerase ribonucleoprotein through step-wise structural rearrangements of telomerase RNA²⁴⁵. In addition, distinct surfaces within an RBD can recognize sequence and structural features within related transcripts to facilitate their coordinated metabolism. Unique binding sites within the same RBD might also permit simultaneous interactions with multiple RNAs in order to facilitate localization

and/or nucleation of RNA granule formation. Thus, distinct binding surfaces within the same RBD can expand the binding repertoire and function of LARPs.

In LARP1, distinct surfaces of the RRM could be used to bind poly(A) and TOP motif RNA. This would also allow both sequences to be bound simultaneously as they would not need to compete for binding to the same site. Alternatively, a distinct surface could mediate dimerization, which would allow each of the RNAs to occupy the same site. It is also possible that one La-Module can bind both RNAs using distinct binding sites in addition to the ability to form a dimer.

5.2.1 Future direction: How does the La-Module bind two RNAs simultaneously?

Identifying the interaction surfaces of the La-Module is crucial to a comprehensive understanding of the molecular mechanisms by which it contributes to TOP mRNA recognition and translation regulation. First, the stoichiometry of the La-Module in the La-Module-poly(A)-TOP motif ternary complex should be identified. This can be done using analytical ultracentrifugation or other methods that directly measure mass. NMR or cross-linking mass spectrometry can then determine the RNA-binding surfaces and/or the dimerization interface. These experiments can be verified by separation of function mutations and binding assays. In addition, identification of the RNA-binding and/or dimerization surfaces will guide more mutagenic analyses tested in translation assays. For instance, we can then ask whether simultaneous poly(A) and TOP motif binding are necessary for translation regulation. Thus, we can delineate binding events at the 5' and 3' ends of TOP mRNAs that are necessary and sufficient to TOP mRNA translation regulation.

5.3 Contribution of disordered regions to LARP1 RNA-binding activity

RBPs achieve RNA-binding through structured domains that form classical RBDs, such as the RRM, K-homology (KH), and zing-finger domains, as well as regions of low structural complexity or intrinsically disordered regions (IDRs)²⁴⁶. IDRs are classified into three categories

based on their lengths²⁴⁷: 1) short linear motifs (SLiMs) encompass 1-10 amino acids that make up consensus motifs for protein or nucleic acid binding, 2) molecular recognition features (MoRFs) are 10-70 amino acids that undergo a disorder-to-structure conformational change upon binding, and 3) low complexity sequences (LCs) that contain hundreds of repetitions of a few amino acids, which might also assume a structured conformation. Post-translational modifications can also lead to disorder-to-structured conformational changes, such is the case in 4E-BP2²⁴⁸. However, it is important to note that IDRs need not undergo a disorder-to-structure conformational change to execute their functions, and many IDRs remain in a disordered state throughout a protein's lifetime²⁴⁶. Disordered regions permit interactions with numerous proteins and nucleic acids by allowing a protein to assume various conformations.

While no studies have experimentally validated LARP IDRs yet, LARPs do contain disordered regions ranging from a few to over one hundred amino acids in length¹⁵⁶. Although the functions of disordered regions in LARPs have not been well characterized, some data suggest they play critical roles in regulating the binding modality, cellular distribution, and biological function of LARPs¹⁵⁷.

LARP1 contains many disordered regions that are potentially IDRs. Our secondary structure predictions suggest that the La-Module contains an unusually long (~110 amino acid) linker between the LAM and the RRM, as well as long loops within the RRM. Indeed, NMR spectra of the RRM (in collaboration with Dr. Lisa Warner, Boise State University) suggest that it is intrinsically disordered. Our size exclusion and SLS experiments also corroborate conformational plasticity of the RRM and linker region (amino acids 440-647) as evidenced by polydispersity. One interesting avenue for future work is to understand how disordered regions contribute to the binding repertoire of LARP1 through short binding motifs, conformational plasticity, and disorder-to-order transitions.

The linker between the La-Module and C-terminal binding domains or motifs of LARPs can allow for interdomain interactions. For instance, in Genuine La, the linker between RRM1 and RRM2 α permits a protein-protein interaction between these two domains²⁴⁹. Interestingly, the RRM1-RRM2 α interaction withstands poly(A) RNA-binding, but is disrupted upon binding to poly(U) RNA²⁴⁹. These data may suggest interplay between RNA-binding modes and conformational plasticity. In addition, the RRM1-RRM2 α interaction may also affect the

subcellular localization of Genuine La by gating the accessibility of the nuclear retention element and nuclear localization sequence within these motifs²⁴⁹.

LARPs also contain potential SLiMs. For instance, the disordered PAM2 consensus sequence found in LARPs 1, 4, and 5 that allow binding to the PABP MLLE domain could might serve as a potential SLiM. The *S. cerevisiae* homologue of Genuine La contains disordered GR/GK repeats that promotes post-transcriptional biogenesis of tRNAs and other cellular RNAs^{250,251}. Although its role has not been fully characterized, the disordered LARP6 LSA region recruits STRAP kinase to aid in the regulation of collagen mRNA translation²⁵².

The LARP4 PAM2 – located in the disordered region N-terminal (NTR) to La-Module – binds both PABP and poly(A) RNA¹⁶⁵. Transient secondary structures in the NTR are critical for poly(A) binding, but do not disrupt binding to the PABP MLLE domain¹⁶⁵. Therefore, poly(A) RNA and the MLLE might compete for binding to the LARP4 PAM2, with fluctuations to the conformational state of the NTR dictating which entity is bound¹⁶⁵. In addition, the NTR, makes transient intramolecular interactions between the La-Module¹⁶⁵. This may stabilize the secondary structure of the RNA-bound PAM2 and sterically occlude PABP binding.

The putative IDRs within LARP1 can serve as inherent molecular switches. By manipulating the overall conformation of LARP1, IDRs can dictate which RNA-binding surfaces are accessible. SLiMs containing positively charged and/or aromatic residues are commonly used for nucleic acid binding; a classic example is the RGG/RG motif²⁵³. N-terminal to the LARP1 La-Module is an RG repeat that is conserved amongst LARP1s that could have RNA-binding activity. The RG repeat could cooperatively bind RNAs with the La-Module or bind independently. Furthermore, its accessibility and purpose could be regulated by transient disorder-to-structure fluctuations. Conformational remodeling can also regulate which regions of LARP1 are exposed to post-translational modifications or can result as a consequence of these modifications. LARP1 IDRs could govern intramolecular interactions and the accessibility of RNA and protein binding motifs. Ultimately, this may not only affect its role in TOP mRNA translation repression, but also re-organization of TOP mRNPs for shuttling into SGs and P-bodies.

5.3.1 Future direction: How do unstructured regions facilitate LARP1 interactions with RNA and proteins? Do binding partners stimulate disorder-to-order changes?

NMR and/or circular dichroism experiments in the presence and absence of poly(A) and TOP motif RNAs can be used to determine whether RNA binding entails disorder-to-structure transitions in the LARP1 RG repeats and if different RNAs induce discrete conformational changes. Cryo-EM experiments of full-length LARP1 can be used to observe interdomain orientations of the LAM, RRM, and DM15 in the presence of poly(A) RNA and different TOP motif RNAs. Given that TOP mRNA 5'UTRs are predicted to have different secondary structures, it is possible that the LARP1 RBDs might alter their relative orientations in order to bind diverse sequence and structural elements of these UTRs. Although TOP mRNA translation is coordinated through the TOP motif and adjacent GC-rich region¹⁴⁵, the secondary structures of the sequences vary. Therefore, it is unlikely that LARP1 binds all 5'UTRs in the same orientation; conformational plasticity conferred by disordered regions will allow different placements of the RBDs for RNA recognition and, ultimately, coordinated translation regulation.

Finally, the La-Module interdomain likely provides significant conformational plasticity to this domain. Small-angle scattering of the La-Module in the presence and absence of RNAs and MLE can be used to identify conformational reorganization of the La-Module in response to different binding partners. For instance, binding of the MLE might encourage the interdomain linker to undergo a conformational change that brings the LAM and RRM closer together, whereas RNA binding may encourage a more extended state. Furthermore, the potential for ordered binding events in the LARP1-PABP interaction should be investigated, as one protein may bind TOP mRNA poly(A) tails first in order to recruit the second for specific metabolic events. For instance, the La-Module may specifically recruit PABP to TOP mRNA poly(A) tails. In P-bodies, PABP may then recruit components of the mRNA decay machinery for degradation of TOP mRNAs during the stress response.

5.4 Concluding remarks

This thesis presents the first investigation of the LARP1 La-Module. Here, we identify LARP1 La-Module RNA and protein binding partners. The La-Module engages features present at both ends of TOP mRNAs, namely, poly(A) RNA and the TOP motif. Unexpectedly, the La-Module binds both of these sequences simultaneously. This might enhance TOP mRNA translation upon release of the DM15 from the 5' cap in response to mTORC1-mediated phosphorylation. In addition, multivalent RNA interactions may mediate the selected sequestering and anchorage of TOP mRNAs into SGs and P-bodies during the stress response. Recognition of poly(A) and TOP motif RNAs is driven by a region encompassing the interdomain linker and the RRM. Future studies will be required to further define the minimal motifs and residues involved in RNA-binding.

We also detect a direct interaction between the La-Module and PABP MLE domain, driven by the PAM2 between the LAM and RRM. Additionally, we found an unexpected interaction between the La-Module and PABP RRMs that may be driven by a putative PAM1 region located the RRM. Binding of the La-Module to PABP may help anchor it to poly(A) tails. Conversely, the La-Module may help to dock PABP to TOP mRNA poly(A) tails. Future studies are required to discern the order of binding events between LARP1 and PABP, as well as the functional role of this complex in TOP mRNA translation regulation.

Appendix A Structural Predictions Used to Optimize La-Module Construct

Appendix A.1 Introduction

The expression and purification of the LARP1 La-Module (amino acids 310-540) presented tremendous difficulty due to low expression and purification yields, as well as substantial contamination by higher molecular weight proteins corresponding to protein chaperones (see Chapter 2). These troubles were compounded by aggregation, in addition to poor stability and RNA-binding activity, evident as degradation and 1-1.5 % RNA-binding activity, respectively (see Chapter 2). We hypothesized that these were symptoms of suboptimal construct design, and focused on revising the construct using a combination of sequence alignments, secondary structure predictions, and homology modeling to guide our efforts.

Appendix A.2 Materials and methods

Appendix A.2.1 Protein sequence alignment and secondary structure prediction

Protein sequence alignments were performed using Clustal Omega Multiple Sequence Alignment^{254,255}. The alignments were further analyzed using Jalview software²⁵⁶. The secondary structure prediction was generated using PSIPRED 4.0²⁵⁷ using the human LARP1 amino acid sequence (isoform 2).

Appendix A.2.2 Ab initio protein structure prediction

The *ab initio* homology model was generated *via* Phyre 2.0²⁵⁸ using the intensive mode and human LARP1 amino acids 310-647 (isoform 2) as the query sequence.

Appendix A.2.3 One-to-one threading

Homology models were generated by one-to-one threading using Phyre 2.0²⁵⁸ using both global and local alignment algorithms. Human LARP1 amino acids 310-647 (isoform 2) was used as the query sequence and threaded onto LARP6 RRM (PDBID: 2MTG), LARP4 (PDBID: 6I9B), and La (2VOO).

Appendix A.3 Results and discussion

Appendix A.3.1 LARP1 RRM region greatly diverges from defined RRMs

Secondary structure predictions suggested that the LARP1 La-Module region contains a significant proportion of disorder (Figure A-1). While the LARP1 LAM (amino acids 320-405) is conserved and predicted to adopt a winged helix-turn-helix motif, the RRM is more obscure (Figure A-1, A-2). As with other LARPs, the LARP1 RRM lacks the canonical RNP1 and RNP2 motifs^{163,165,175,176} (Figure A-1, A-3). However, unlike other LARPs, the LARP1 RRM region is much larger (~165 residues long as opposed to ~ 90) has large loops, and contains only two predicted β -strands; significant deviation from the canonical $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ structure of RRMs (Figure A-1, A-3). Additionally, the LARP1 RRM region has poor sequence conservation with other LARPs (as compared to the LAM conserved (Figure A-2, A-3). Importantly, the secondary structure prediction suggested that the LARP1 La-Module (amino acids 310-540) terminates within the RRM; D540 is preceded by two α -helices and followed by three α -helices and two β -strands (Figure A-1)

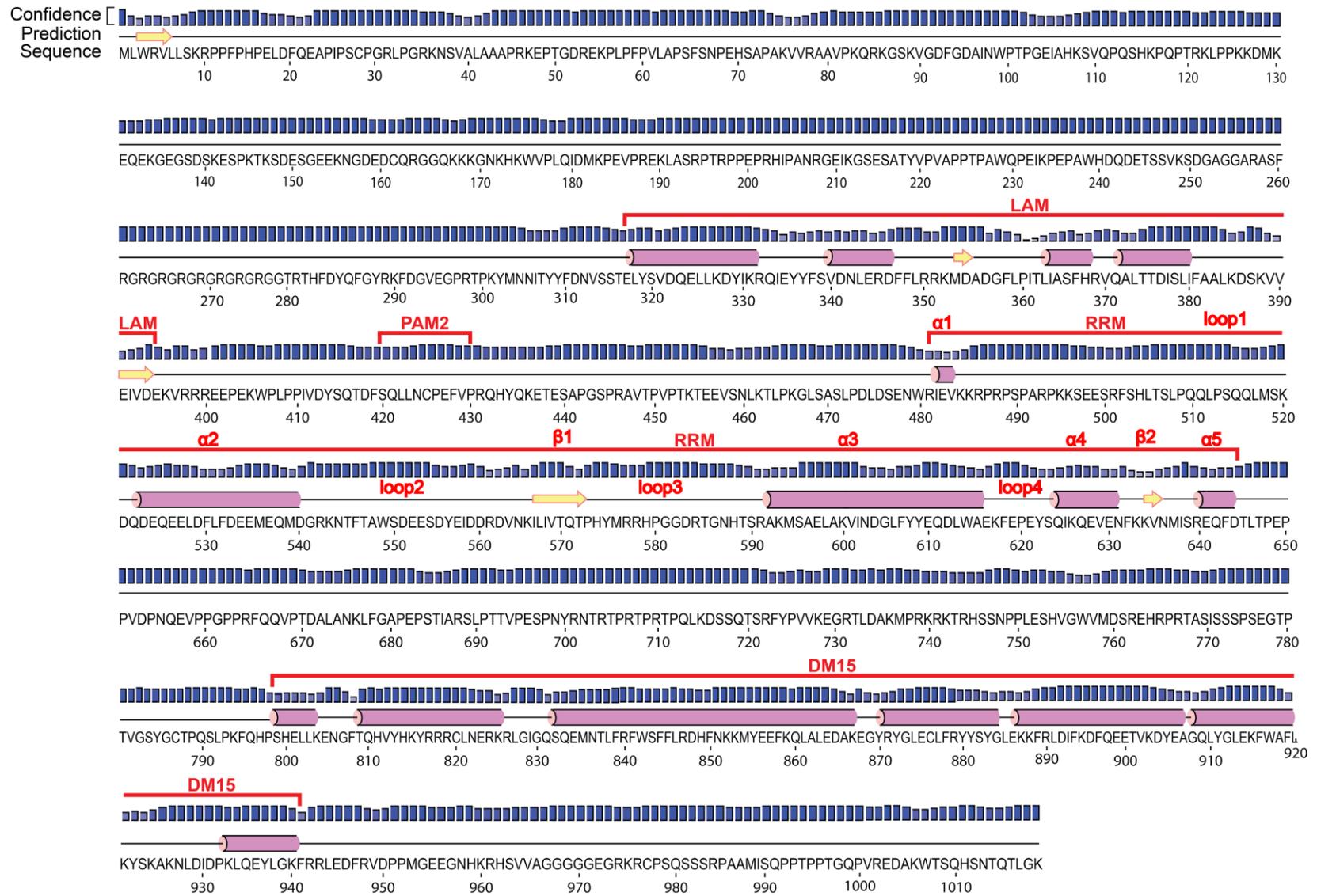


Figure A-1 Secondary structure prediction of LARP1.

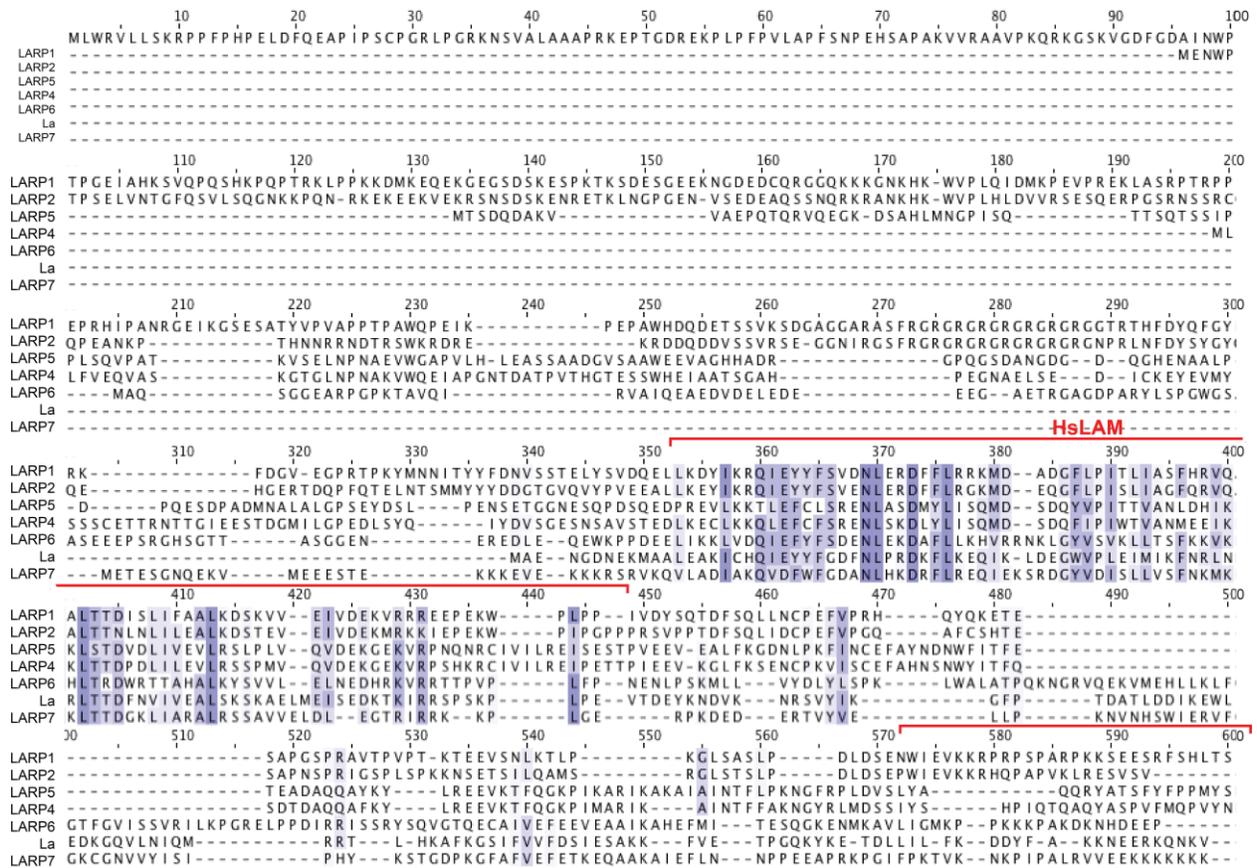


Figure A-2 Sequence conservation of LARP1 LAM region. Protein sequence alignment of the LAM region from human LARP1, LARP2, LARP4, LARP5, LARP6, LARP7, La.



Figure A-3 Sequence conservation of LARP1 RRM region. Protein sequence alignment of the RRM region from human LARP1, LARP2, LARP4, LARP5, LARP6, LARP7, La.

Based on these data, we hypothesized that the C-terminus of the LARP1 RRM occurs at ~ F643, which marks the end of the last structural motif preceding the DM15 (Figure A-1, A-3). However, because the predicted LARP1 RRM region contains relatively long loops – for example an 18 amino acid loop between β_1 and α_3 , as well as the low confidence of β_2 – we also considered that the RRM C-terminus might end at T572 or N631, followed by a small structural motif independent of the RRM (Figure A-1).

Prediction of the RRM N-terminus was more challenging. The interdomain linker between the LAM and RRM is at least ~85 amino acids in length (Figure A-1). After the LAM, RRM α_1 is predicted with ~ 50% confidence, perhaps because it might form a β -strand instead (Figure A-1). Although the residues encoding RRM α_1 have a propensity to form helices, some are β -branched

residues in the vicinity of large aromatics and other β -branched residues, both of which have a propensity to form β -strands²⁵⁹⁻²⁶¹ (Figure A-1). Alternatively, this region may be unstructured and the interdomain linker between the LAM and RRM would be ~ 125 amino acids long. RRM loop₁ is 38 amino acids long and followed by α_2 , predicted with relatively high confidence and conserved across LARP1 from various species (Figure A-1, A-4). Given the length of RRM loop₁, it is unclear whether α_1 , if it exists, resides within the RRM or corresponds to an independent motif between the LAM and RRM. To help identify the RRM C- and N- termini, we generated various homology models of the LARP1 La-Module and RRM.

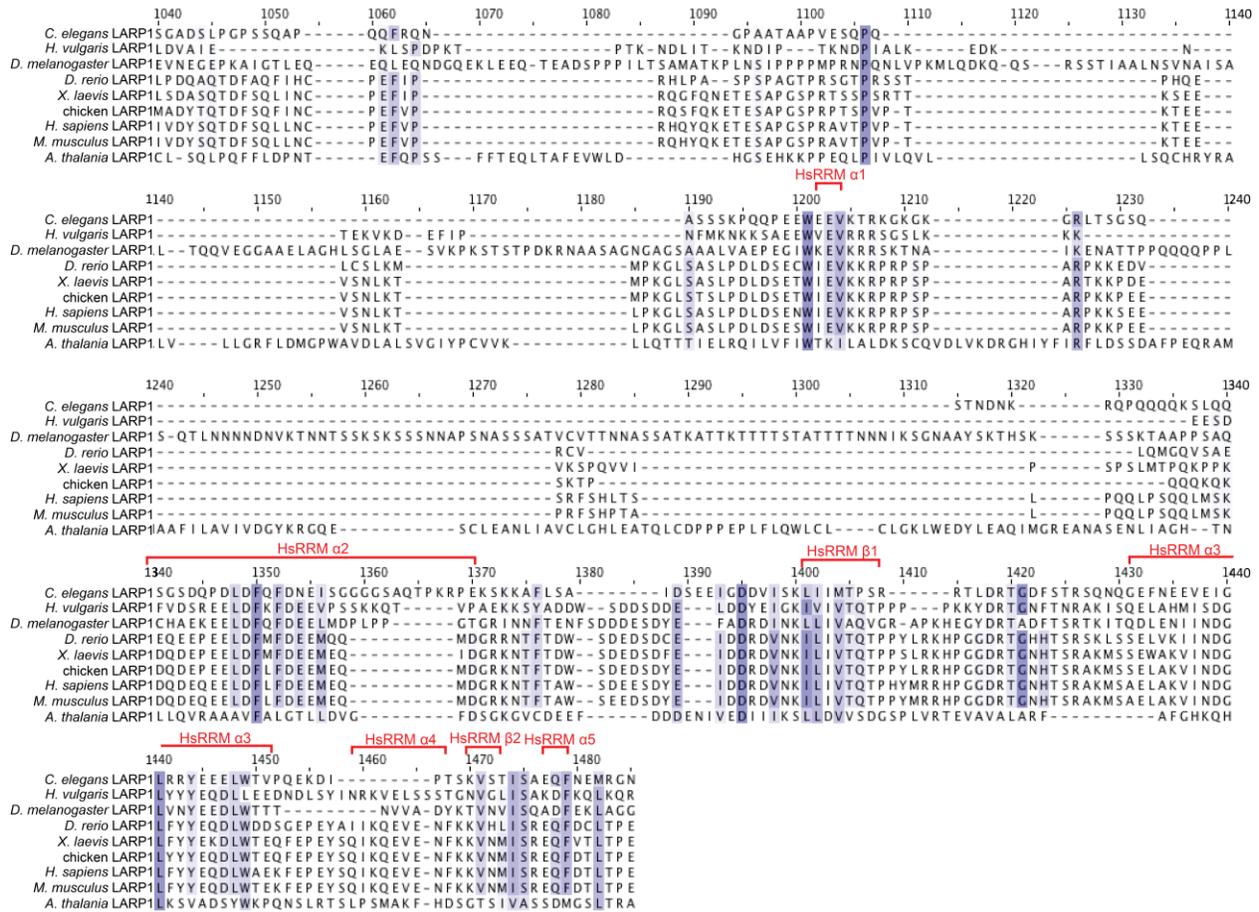


Figure A-4 Sequence conservation of the LARP1 RRM region amongst different species. Protein sequence alignment of the LARP1 RRM region from various species. Animals: *D. melanogaster*, *H. vulgaris*, *C. elegans*, *D. rerio*, *X. laevis*, *G. gallus*, *M. musculus*, *H. sapiens*. Plants: *A. thaliana*.

Appendix A.3.2 *Ab initio* homology model

First, we generated an *ab initio* homology model of the LARP1 La-Module region (amino acids 310-647) using Phyre2 (Figure A-5)²⁵⁸. *Ab initio* homology modeling produces an unbiased structural prediction of the query sequence by using physical principles rather than relying on previously solved structures²⁶². In agreement with conservation of the LAM and the secondary structure prediction, a winged helix-turn-helix was modeled with 90% confidence (Figure A-1, A-5). Also consistent with the secondary structure prediction, the PAM2 was predicted to be unstructured and resides within the interdomain linker between the LAM and RRM (Figure A-1, A-5). Some helices in the secondary structure prediction were also corroborated in the *ab initio* model. Specifically, RRM α_2 and α_5 in the *ab initio* model correspond to α_2 and α_5 in the secondary structure prediction (Figure A-1, A-5). In addition, RRM α_3 and α_4 , and their separating loop in the *ab initio* model correspond to α_3 in the secondary structure prediction (Figure A-1, A-5). The identity of the remaining two α helices differed (Figure A-1, A-5). Interestingly, the *ab initio* model suggested no β -strands exist within the LARP1 RRM and 64% of the query is modeled as disordered (Figure A-5). It is important to note that, orientations of the disordered regions are likely inaccurate due to the difficulty in modeling unstructured regions²⁵⁸ (Figure A-5).

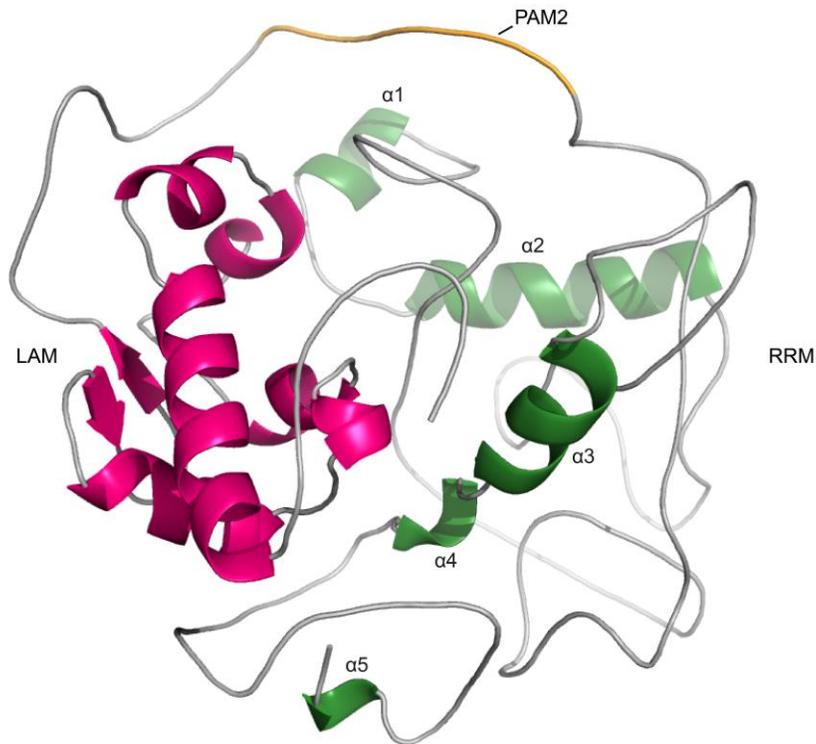


Figure A-5 *Ab initio* homology model of the LARP1 La-Module (310-647).

Appendix A.3.3 One-to-one threading

We next generated homology models by one-to-one threading of the LARP1 query sequence onto structural models of Genuine La, LARP6, and LARP4 using both local and global alignment algorithms using Phyre2²⁵⁸. Local alignments model only regions of high similarity within a query sequence that is substantially different²⁵⁸. Thus, local alignments are more stringent and are particularly useful for comparing disparate sequences²⁵⁸. For instance, in situations where the query is larger than the template, and has a divergent sequence with only few conserved structural motifs. Global alignments force the alignment to span the entire length of the query sequence, no matter how much larger than the template, and attempt to align every residue²⁵⁸. Therefore, this is best used when the query is similar in size and sequence to the template. Global alignments can be used in lieu of local alignments that yield no result due to high disparity between the query and template. However, the resulting homology model requires careful analysis because gaps are generated to accommodate the size difference as the alignment is forced end-to-end.

To start, we conducted a local and global thread of LARP1 amino acids 406-647 – from the end of the LAM to the hypothesized RRM C-terminus (see Appendix A.2.1 above) – using the LARP6 RRM as the template. The local homology model corroborated the secondary structure prediction of only two β -strands in the LARP1 RRM, as well as the identity of β_1 (Figure A-6 A). RRM β_1 is conserved among LARP1s from various species (Figure A-4) and contains basic residues that may participate in electrostatic interactions with the RNA (Figure A-6 B). However, the sequence encoding β_2 differed from the secondary structure prediction; the local thread suggested N587-R592, which also contains positively charged residues (Figure A-6 A). We next conducted a global homology model, which predicted four β -strands; β_2 and β_4 in the global homology model correspond to β_1 and β_2 in the secondary structure prediction, respectively (Figure A-6 C, D). The global homology model also identified N587- R592 as β_3 , which corresponds to β_2 in the local homology model (Figure A-6 C). RRM β_1 in the global homology model was encoded by R493-S497 (Figure A-6 C). However, the sequences in β_1 and β_3 in the global homology model are not conserved among LARP1s (Figure A-4). Furthermore, β_1 and β_3 are not supported by the secondary structure prediction (Figure A-1), likely due to the proline-rich nature of these sequences, which does not lend itself to the formation of α -helices²⁵⁹ or β -strands^{260,261}. Therefore, β -1 and β -3 in the global homology model are likely not real.

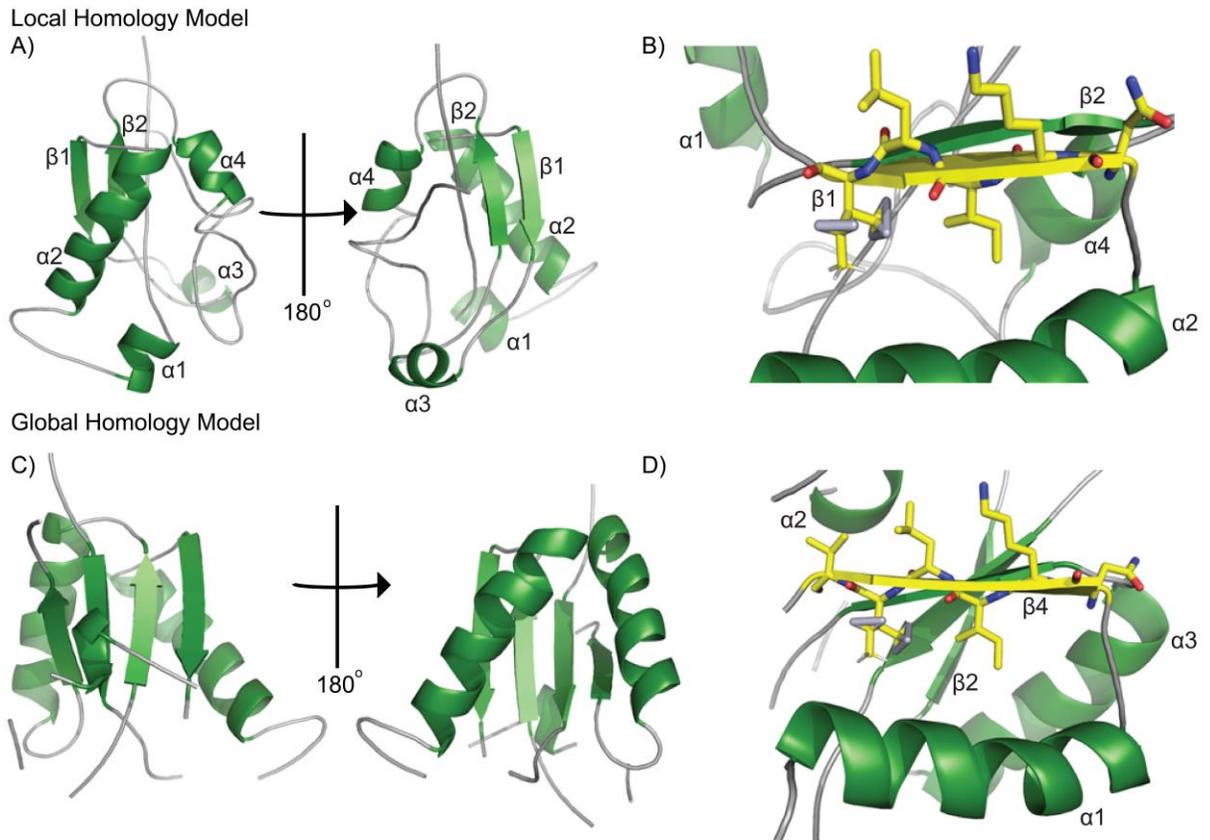
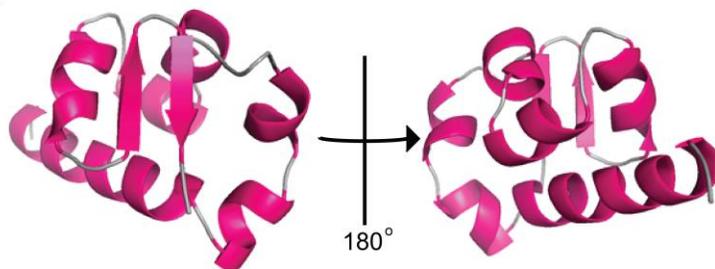


Figure A-6 Local and global homology models of the LARP1 RRM generated by threading on the LARP6 RRM.

We then generated homology models of the LARP1 La-Module (310-647) by threading onto the structure of the La-Module from Genuine La, the most distantly related LARP. The local alignment only predicted the LAM, suggesting significant dissimilarity between the LARP1 and Genuine La RRM (Figure A-7 A). The global homology model suggested the RRM contains three β -strands (Figure A-7 B). Here, RRM β_1 corresponds to the PAM2 motif (Figure A-7 B). This is unlikely because known PAM2s are unstructured regions that dock into hydrophobic pockets within the PABP MLLE domain²²⁹ (see Chapter 3). RRM β_2 corresponds to α_1 in the secondary structure prediction (Figure A-1, A-7 B), which may form a β -strand based on the encoding residues^{260,261}. RRM β_3 corresponds to loop₂ in the secondary structure prediction (Figure A-1, A-7 B). However, RRM β_3 was generated by omitting prolines in the query sequence to force the alignment. Furthermore, proline residues tend to break secondary structure motifs²⁵⁹⁻²⁶¹, therefore this region is likely unstructured.

Local Homology Model

A)



Global Homology Model

B)

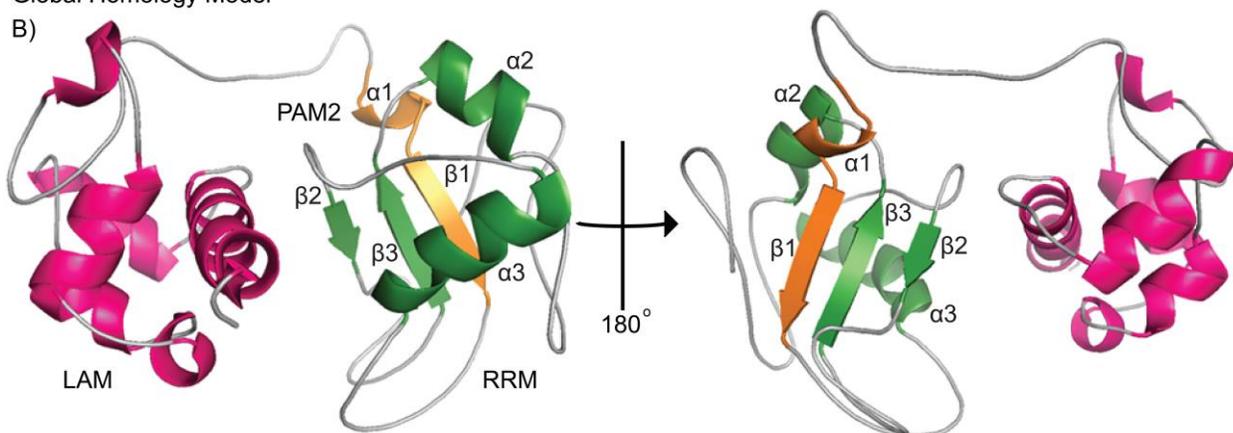


Figure A-7 Local and global homology models of the LARP1 La-Module (310-647) generated by threading on the La-Protein La-Module.

Finally, we modeled the LARP1 La-Module (amino acids 310-647) onto the LARP4 La-Module. Local homology modeling of the LARP1 La-Module (amino acids 310-647) produced only the LAM, again suggesting significant divergence of the LARP1 RRM (Figure A-8 A). The global homology model suggests the LARP1 RRM contains five β -strands, none of which agrees with sequence conservation, secondary structure predictions, or the LARP6- and La protein-based homology models (Figure A-8 B). This may be because the LARP4 RRM adopts a more canonical RRM fold¹⁶⁵, while the Genuine La and LARP6 RRMs contain α -helical insertions^{163,180}. This likely provides a better template for the LARP1 RRM modeling given that it likely has α -helix insertions.

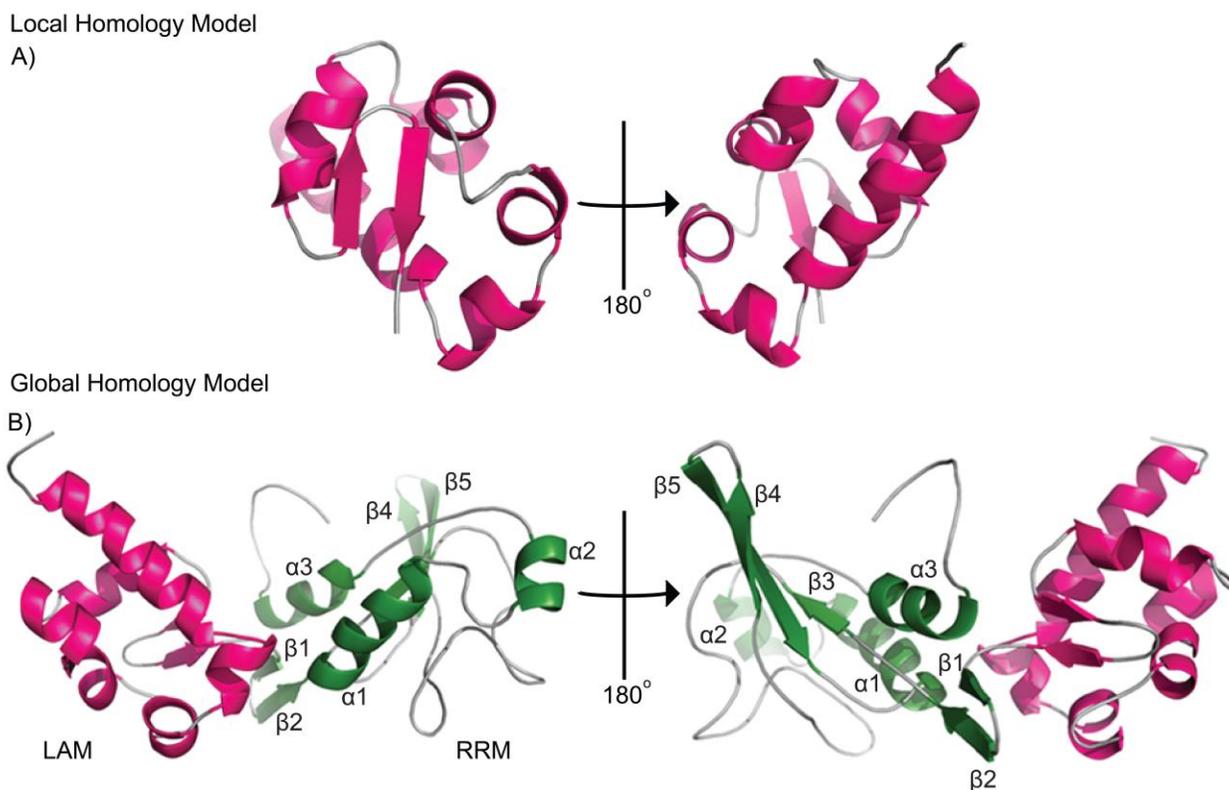


Figure A-8 Local and global homology model of the LARP1 La-Module (310-647) generated by threading on the LARP4 La-Module.

Appendix A.4 Conclusions

While the domain boundaries of the LARP1 LAM were straightforward to determine, the RRM presented a greater challenge. The LARP1 RRM substantially diverges from canonical RRMs, and even the RRMs of LARPs. Typically, RRMs are ~ 90 amino acids in length, with a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ fold, RNP1 and RNP2 motifs in the two central β -strands, and loops shorter than 10 amino acids^{175,176}. Although LARPs do not harbor RNP1 and RNP2 motifs, they are of the same length, adopt the $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ fold, and contain similarly sized loops as canonical RRMs^{160,163,165}. Strikingly, our secondary structure prediction suggests the LARP1 RRM is ~ 164 amino acids long, with relatively long loops – some greater than 20 amino acids – and α -helix insertions. Interestingly, the LARP1 RRM does not seem to adopt the $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ fold. Thus, the overall architecture of the LARP1 RRM is likely very different than RRMs found in other LARPs.

We predict that the LARP1 RRM contains at least two β -strands: β_1 and β_2 in the secondary structure prediction (Figure A-1). Consistent with this, the amino acids encoding β_1 and β_2 are β -branched residues that have a propensity to form β -strands. Furthermore, these residues are conserved amongst LARP1s from different species, as well as LARP2 (Figure A-1), suggesting functional importance. In addition, both local and global threading onto the LARP6 RRM corroborated the identity of β_1 in the secondary structure prediction (Figure A-6). It's possible that RRM α_1 in the secondary structure prediction forms a β -strand in reality, as suggested by the global homology model generated using the La protein La-Module (Figure A-7). This is due to the propensity for the contained residues to form a β -strand^{260,261}. Regardless, it is likely a secondary structure element due to its conservation amongst LARP1s (Figure A-4).

The secondary structure prediction suggested that the LARP1 RRM contains five α -helices (Figure A-1). The identities of RRM α_2 and α_3 in the secondary structure prediction are corroborated by all homology models generated. Furthermore, the residues within have a propensity to form helices and are conserved amongst LARP1s from different species²⁵⁹ (Figure A-1). In addition, RRM α_5 in the secondary structure prediction is conserved and corroborated by RRM α_5 in the *ab initio* model. The relatively longer RRM α_2 and α_3 may pack against one side of the β -strands, with shorter α -helices packing towards the other side. Indeed, the LARP6 RRM contains a helix insertion that orients itself on one side of the β -strands and is required for RNA binding¹⁶³.

The relatively long loops predicted within the LARP1 RRM may provide conformational plasticity to support binding to various RNAs and maybe even proteins. Our SEC, SLS, and HSQC-NMR experiments (see Chapter 3) corroborate conformational heterogeneity in the La-Module (310-647).

Based on the robust expression, ease and consistency of purification, decreased aggregation and degradation, as well as crystallization and RNA-binding activity (see Chapter 1), we concluded that the optimal LARP1 La-Module spans amino acids 310-647. Furthermore, we hypothesize that the RRM N-terminus begins between N460 and K520, as RRM construct 460-647 yielded crystals and displayed RNA-binding activity comparable to the La-Module (310-647) (see Chapter 1). However, future work is needed to define the RRM N-terminus more accurately.

Appendix B Microalgae LARP1

Appendix B.1 Introduction

Human LARP1 contains three RNA-binding domains separated by disordered regions (Figure B-1). These disordered regions may permit different conformations and interdomain interactions within LARP1 that dictate how it binds to protein and RNA. Structural plasticity – which may also be induced by post-translational modifications, such as mTORC1 phosphorylation^{167,170,201,204}, and binding partners, such as PABP¹⁶⁷ or RNA^{168,169,171} – may ultimately drive how and when LARP1 regulates transcript stability and translation. For instance, mTORC1 phosphorylation marks deposited throughout LARP1²⁰⁴ could lead to an extended conformation due to charge repulsion. This may inhibit synergistic binding by the La-Module and DM15, and ultimately release bound mRNAs.

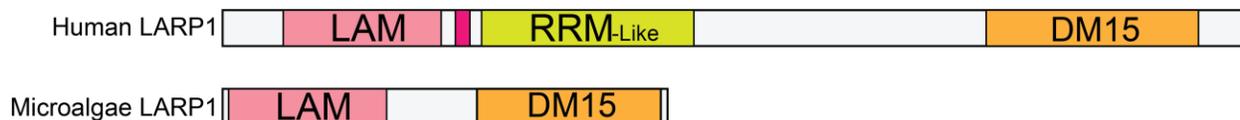


Figure B-1 Schematic of domain organization in human LARP1 versus microalgae LARP1. Microalgae LARP1 is smaller and fewer and shorter IDRs.

Structural studies of full-length LARP1 in its native and bound states would allow us to compare how the N-terminal La-Module and C-terminal DM15 orient themselves when bound to different targets, such as PABP, poly(A), or TOP motif RNA. This will allow us to understand how conformational plasticity within LARP1 governs its biological function. Additionally, we may uncover unique pockets that could be druggable for cancer therapeutics.

Unfortunately, full-length human LARP1 does not lend itself well to structural studies. Overexpression is not robust in *E. coli* or human cells. The unstructured regions are subject to protease activity, leading to an additional loss of protein yield. Furthermore, the long disordered regions decrease the chances of successful crystallization and collection of diffraction data. Therefore, we sought a homologue of LARP1 that is smaller and contains fewer unstructured regions to improve structure determination.

A LARP1 homologue exists in the microalgae *Auxenochlorella protothecoides*. The microalgae LARP1 (maLARP1) is 426 residues in length as opposed to the 1,019 amino acids of human LARP1a (Appendix Figure 9). Importantly, secondary structure predictions and homology modeling suggest that maLARP1 adopts a more globular conformation with fewer IDRs. In this study, we optimized the construct design, expression, and purification of maLARP1 in order to arrive at a structural model of LARP1 to identify potential interdomain interactions and druggable pockets.

Appendix B.2 Materials and methods

Appendix B.2.1 Protein cloning, expression, and purification

The sequence encoding the microalgae LARP1 (accession number 23616894) homologue (maLARP1) from *Auxenochlorella protothecoides* (Integrated DNA Technologies) was PCR amplified and inserted into a pET28b vector using BamH1 and SacI sites. The resulting construct expresses maLARP1 with an N-terminal 6XHis-SUMO tag. The 6XHis-SUMO-tagged maLARP1 was expressed in *E. coli* BL21(DE3) using autoinduction. Cultures were grown at 37 °C until an O.D.₆₀₀ of 0.4-0.5 then switched to 17.5 °C for 16-18 hours. Cells were harvested, frozen in liquid nitrogen, and stored at -80 °C.

Cells were resuspended in lysis buffer [50 mM Tris-HCl, pH 8.0, 650 mM NaCl, 10 mM imidazole, 10% glycerol]. Cells were lysed by homogenization in the presence of protease inhibitors [leupeptin, aprotinin, bestatin, PMSF] and lysate was cleared by ultracentrifugation. The cleared lysate was applied onto a HiTrap His FF (GE Healthcare Lifesciences) and 6XHis-SUMO-maLARP1 was eluted with a 5 column volume imidazole gradient (20-500 mM). The 6XHis-SUMO was cleaved using 0.5 mg 6XHis-tagged ULP1 protease per 80 mL eluate overnight at 4 °C in 2 L dialysis buffer [50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 1 mM DTT]. The cleaved 6XHis-SUMO tag and 6XHis-ULP1 were separated from maLARP1 by applying onto HiTrap His FF and eluting with an imidazole gradient (20-500 mM). The flow through was

collected and dialyzed for 2 hours at 4° C [50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% glycerol, 1 mM DTT] before applying onto tandem HiTrap Q and S columns (GE Healthcare Lifesciences) to separate nucleic acid and protein contaminants, respectively. The maLARP1 eluted from the Q column using a 5 column volume NaCl gradient [0.1-1M]. Fractions containing maLARP1 were concentrated in a 10K MWCO Vivaspin Turbo Concentrator (Sartorius) to 0.3 mL (45-50 mg/mL) and loaded onto an equilibrated Superdex 75 size exclusion column (GE Healthcare Lifesciences) [50 mM Tris-HCl, pH 7.5, 750 mM NaCl, 5 % glycerol, 1 mM DTT]. Fractions containing maLARP1 were buffer exchanged and concentrated to 100 µM for storage [50 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 % glycerol, 1 mM DTT] or 10-20 mg/mL (238 - 476 µM) for crystallization [50 mM HEPES, pH 7.5, 1 mM TCEP].

Appendix B.2.2 Secondary structure prediction and homology modeling

The secondary structure prediction was generated using PSIPRED 4.0 using the maLARP1 amino acid sequence²⁵⁷. The homology model was generated using Phyre 2.0 using the Intensive Modeling Mode. The software reported use of human DM15 (PDBID: 4CZ4) and various LAMs, with the top hit being human LAM from c-MPL protein (PDBID: 2CQK).

Appendix B.3 Results and discussion

Secondary structure predictions and homology modeling suggest that maLARP1 contains fewer disordered regions and adopts a more globular structure (Figure B-2, B-3). For instance, maLARP1 lacks the disordered region N-terminal to the La-Module in human LARP1 (Figure B-2, B-3). Unexpectedly, both the secondary structure prediction and homology model suggest that maLARP1 lacks an RRM (Figure B-2, B-3). This omits the ~110 amino acid linker predicted between the LAM and RRM in human LARP1, along with the relatively long loops within the human LARP1 RRM (Figure A-1, B-2).

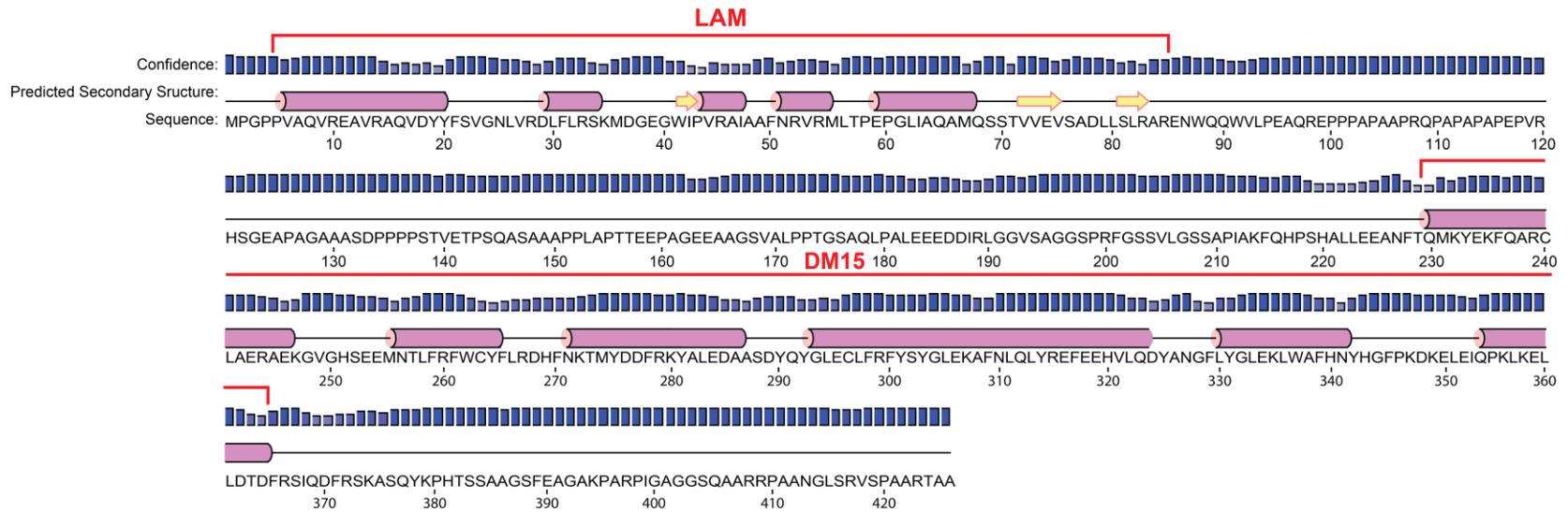


Figure B-2 Secondary structure prediction of maLARP1.

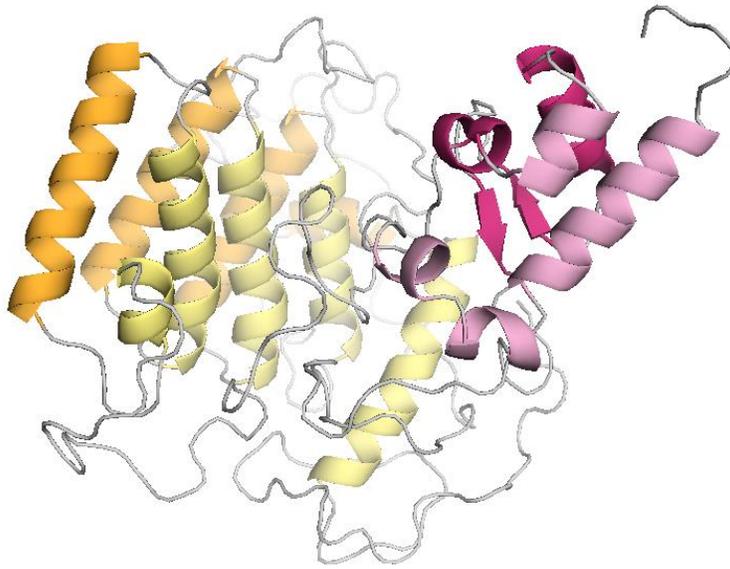


Figure B-3 Homology model of maLARP1. Pink: La-Motif, Orange and yellow: DM15 heat- like repeats.

maLARP1 is expressed robustly in *E. coli* and purification generates high yields of 25-30 mg protein with > 95 % purity (Figure B-4)

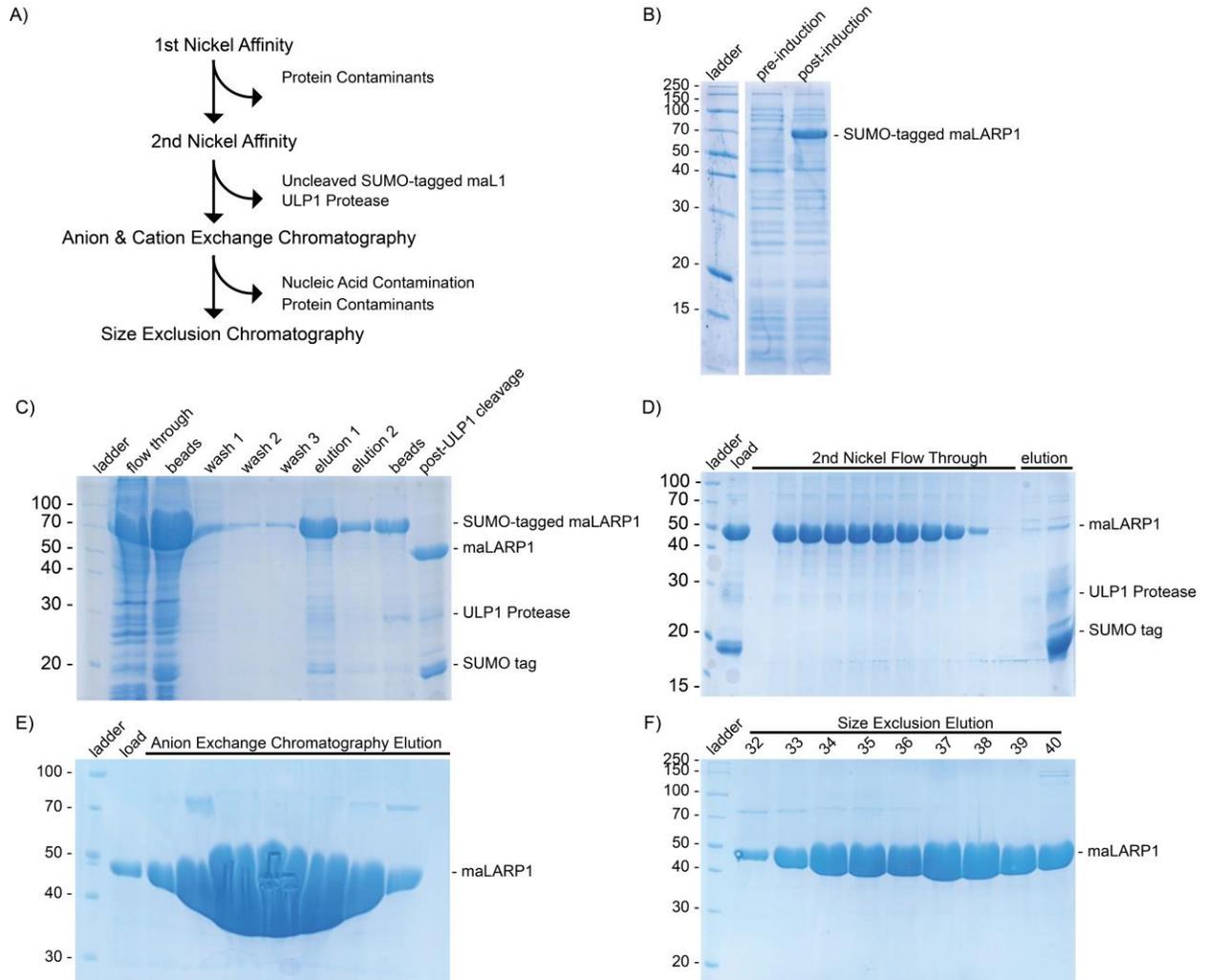


Figure B-4 maLARP1 exhibits robust expression and purification. (A) Purification scheme of maLARP1. SDS-PAGE showing (B) expression of 6XHis-SUMO-maLARP1, (C) first nickel affinity chromatography, (D) second nickel affinity chromatography, (E) anion exchange chromatography elution, (F) size exclusion chromatography elution.

Size exclusion chromatography showed a symmetrical peak suggesting protein stability and conformational homogeneity (Figure B-5).

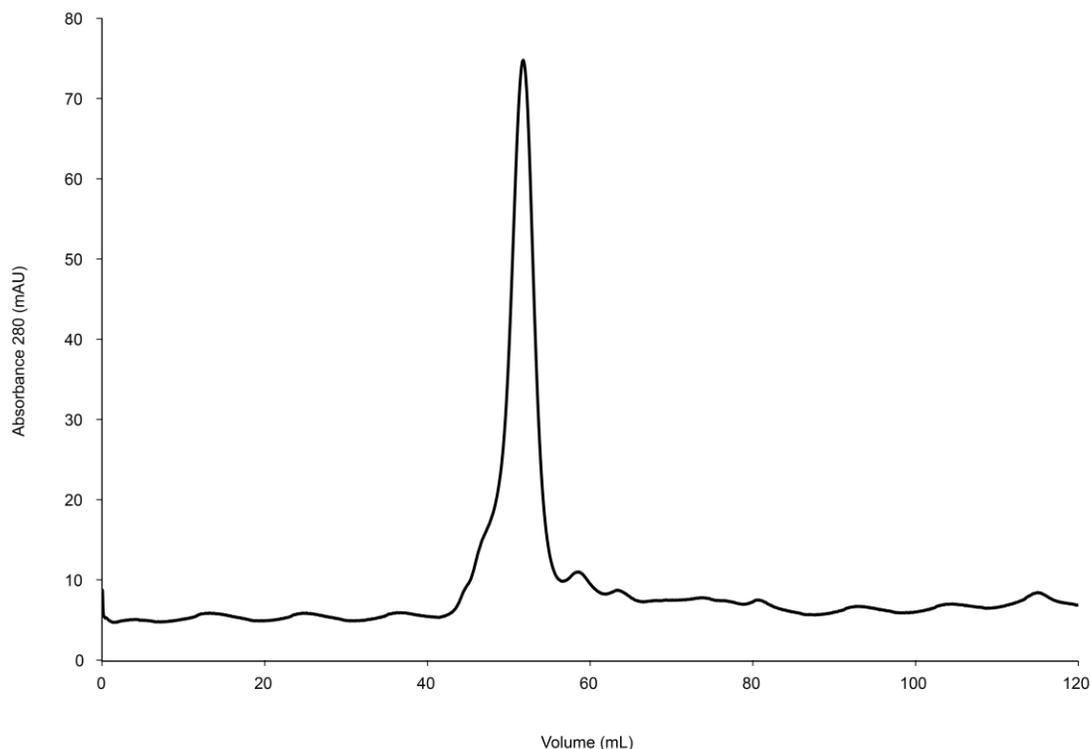


Figure B-5 Size exclusion chromatography suggests stability and homogeneity of purified maLARP1.

Relatively symmetrical elution peak of maLARP1 suggests homogeneity and stability.

Initial crystal screens were successful in producing a few crystal hits. However none of these hits diffracted at AMX (BNL) or LRL-CAT (APS). While maLARP1 contains fewer disordered regions than human LARP1, the linker between the LAM and DM15 is ~ 130 amino acids. Perhaps this linker adopts various conformations within the crystals that disturb crystal packing and prevent successful diffraction.

Appendix B.4 Conclusions

Secondary structure predications and homology modeling suggest that, while the LAM and D15 are present, maLARP1 lacks an RRM (Figure B-2, B-3). Residues within the interdomain linker between the LAM and DM15 may aid RNA binding in lieu of an RRM and may even adopt a structured motif upon binding in a disorder-to-structure transition. Although size exclusion

suggests sample homogeneity and stability (Figure B-5), collection of diffraction data was not successful. Future crystallization endeavors will try additive screens or screening in the presence of a ligand, such as RNA. Binding to a small molecule may induce the maLARP1 interdomain linker to form a structured motif and improve diffraction. Alternatively, glutaraldehyde crosslinking or annealing post-crystallization may produce diffracting crystals. A structural model of full-length LARP1 might highlight interdomain interactions, perhaps between the LAM and RRM, which might be of importance in TOP mRNA recognition and translation regulation.

Bibliography

1. Müller-McNicoll, M.a.N., Karla M. . How cells get the message: dynamic assembly and function of mRNA–protein complexes. *Nature Reviews Genetics* **14**, 275–287 (2013).
2. Yang, E. et al. Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. *Gene Research* **13**, 1863-1872 (2003).
3. Raghavan, A. et al. Patterns of coordinate down-regulation of ARE-containing transcripts following immune cell activation. *Genomics* **84**, 1002-1013 (2004).
4. Vlasova, I.A. et al. Coordinate stabilization of growth-regulatory transcripts in T cell malignancies. *Genomics* **86**, 159-171 (2005).
5. Thoreen, C.C. et al. A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109-13 (2012).
6. Schott, J. et al. Translational Regulation of Specific mRNAs Controls Feedback Inhibition and Survival during Macrophage Activation. *PLoS Genetics* **10**, e1004368 - e1004368 (2014).
7. Mathews, M.B., Sonenberg, N. & Hershey, J.W.B. *Origins and Principles of Translation Control*. (2000).
8. Kruh, J. & Borsook, H. Hemoglobin Synthesis in Rabbit Reticulocytes In vitro. *Journal of Biological Chemistry* **220**, 905-915 (1956).
9. Datta, A., de Haro, C., Sierra, J.M. & Ochoa, S. Mechanism of translational control by hemin in reticulocyte lysates. *Proceedings of the National Academy of Sciences* **74**, 3326-3329 (1977).
10. Laskey, R.A., Mills, A.D., Gurdon, J.B. & Partington, G.A. Protein synthesis in oocytes of *xenopus laevis* is not regulated by the supply of messenger RNA. *Cell* **11**, 345-351 (1977).
11. Humphreys, T. Measurements of messenger RNA entering polysomes upon fertilization of sea urchin eggs. *Developmental Biology* **26**, 201-208 (1971).
12. Hinnebusch, A.G. & Lorsch, J.R. The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. *Cold Spring Harbor Laboratory Perspectives in Biology* **4**(2012).
13. Dever, N.S.a.T.E. Eukaryotic translation initiation factors and regulators *Current Opinion in Structural Biology* **13**, 56-63 (2003).

14. Jackson, R.J., Hellen, C.U.T. & Pestova, T.V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews Molecular Cell Biology* **10**, 113-127 (2010).
15. Wells, S.E., Hillner, P.E., Vale, R.D. & Sachs, A.B. Circularization of mRNA by Eukaryotic Translation Initiation Factors. *Molecular Cell* **2**, 135-140 (1998).
16. Dever, T.E. & Green, R. The Elongation, Termination, and Recycling Phases of Translation in Eukaryotes. *Cold Spring Harbor Laboratory Perspectives in Biology* (2012).
17. Hellen, C.U.T. Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harbor Laboratory Perspectives in Biology* (2018).
18. Shatkin, A.J. Capping of Eucaryotic mRNAs. *Cll* **9**, 645-653 (1976).
19. Ramanathan, A., Robb, G.B. & Chan, S.-H. mRNA capping: biological functions and applications. *Nucleic Acids Research* **44**, 7511–7526 (2016).
20. Monteki, S. & Price, D. Functional coupling of capping and transcription of mRNA. *Molecular and Cellular Biology* **10**, 599-609 (2000).
21. Shimotohno, K., Kodama, Y., Hashimoto, J. & Miura, K.-I. Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc. Natl. Acad. Sci.* **74**, 2734–2738 (1977).
22. Konarska, M.M., Padgett, R.A. & Sharp, P.A. Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38**, 731-736 (1984).
23. Izaurralde, E. et al. A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* **78**, 657-668 (1994).
24. Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B. & Mattaj, I.W. A nuclear cap-binding complex binds Balbiani ring pre-mRNA cotranscriptionally and accompanies the ribonucleoprotein particle during nuclear export. *Journal of Cell Biology* **133**, 5-14 (1996).
25. Sonenberg, N. & Gingrast, A.-C. The mRNA 5' cap-binding protein eIF4E and control of cell growth. *Current Opinion in Cell Biology* **10**, 268-275 (1998).
26. Tamarkin-Ben-Harush, A., Vasseur, J.-J., Debart, F., Ulitsky, I. & Dikstein, R. Cap-proximal nucleotides via differential eIF4E binding and alternative promoter usage mediate translational response to energy stress. *eLife* **6**, e21907 (2017).
27. Marcotrigiano, J., Gingras, A.-C., Sonenberg, N. & Burley, S.K. Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* **89**, 951-961 (1997).
28. Daffis, S. et al. 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* **468**, 452-456 (2010).

29. Werner, M. et al. 2'-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Research* **39**, 4756-4768 (2011).
30. Adesnik, M., Salditt, M., Thomas, W. & Darnell, J.E. Evidence that all messenger RNA molecules (except histone messenger RNA) contain Poly (A) sequences and that the Poly(A) has a nuclear function. *Journal of Molecular Biology* **71**, 21-30 (1972).
31. Kates, J. Detection and Utilization of Poly(A) Sequences in Messenger RNA. *Methods in Cell Biology* **7**, 53-65 (1974).
32. Keller, W., Bienroth, S., Lang, K.M. & Christofori, G. Cleavage and polyadenylation factor CPF specifically interacts with the pre-mRNA 3' processing signal AAUAAA. *The EMBO Journal* **10**, 4241-4249 (1991).
33. Kerwitz, Y. et al. Stimulation of poly(A) polymerase through a direct interaction with the nuclear poly(A) binding protein allosterically regulated by RNA. *The EMBO Journal* **22**, 3705-3714 (2003).
34. Kuhn, U. et al. Poly(A) tail length is controlled by the nuclear poly(A)-binding protein regulating the interaction between poly(A) polymerase and the cleavage and polyadenylation specificity factor. *Journal of Biological Chemistry* **284**, 22803-22814 (2009).
35. Lim, J., Lee, M., Son, A., Chang, H. & Kim, V.N. mTAIL-seq reveals dynamic poly(A) tail regulation in oocyte-to-embryo development. *Genes & Development* **30**, 1671-1682 (2016).
36. Park, J.-E., Yi, H., Kim, Y., Chang, H. & Kim, V.N. Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. *Molecular Cell* **26**, 462-471 (2016).
37. Shepard, P.J. et al. Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. *RNA* **17**, 761-772 (2011).
38. Proudfoot, N.J. Ending the message: poly(A) signals then and now. *Genes & Development* **25**, 1770-1782 (2011).
39. Mangus, D.A., Evans, M.C. & Jacobson, A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* **4**, 223 (2003).
40. Blobel, G. A Protein of Molecular Weight 78,000 Bound to the Polyadenylate Region of Eukaryotic Messenger RNAs. *Proc. Natl. Acad. Sci.* **70**, 924-928 (1973).
41. Görlach, M., Burd, C.G. & Dreyfuss, G. The mRNA Poly(A)-Binding Protein: Localization, Abundance, and RNA-Binding Specificity. *Experimental Cell Research* **211**, 400-407 (1994).

42. Yongsheng Shi et al. Molecular Architecture of the Human Pre-mRNA 3' Processing Complex. *Molecular Cell* **33**, 365-376 (2009).
43. Gray, N.K., Hrabálkova, L., Scanlon, J.P. & W.P. Smith, R. Poly(A)-binding proteins and mRNA localization: who rules the roost? *Biochemical Society Transactions* **43**, 1277–1284 (2015).
44. Hosoda, N. et al. Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. *J Biol Chem* **278**, 38287-91 (2003).
45. Kahvejian, A., Roy, G. & Sonenberg, N. The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harbor symposia on quantitative biology* **66**, 293-300 (2001).
46. Burgess, H.M. & Gray, N.K. mRNA-specific regulation of translation by poly(A)-binding proteins. *Biochem Soc Trans* **38**, 1517-22 (2010).
47. Craig, A.W.B., Haghighat, A., Yu, A.T.K. & Sonenberg, N. Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392**, 520-523 (1998).
48. Khaleghpour, K. et al. Translational repression by a novel partner of human poly(A) binding protein, Paip2. *Mol Cell* **7**, 205-16 (2001).
49. Wu, J. & Bag, J. Negative control of the poly(A)-binding protein mRNA translation is mediated by the adenine-rich region of its 5'-untranslated region. *Journal of Biological Chemistry* **273**, 34535-34542 (1998).
50. de Melo Neto, O.P., Standart, N. & de Sa, C.M. Autoregulation of poly(A)-binding protein synthesis in vitro. *Nucleic Acids Research* **23**, 2198–2205 (1995).
51. al., J.C.V.e. The Sequence of the Human Genome. *Science* **291**, 1304-1351 (2001).
52. Lander, E.S., et al.; Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
53. Pesole, G. et al. Structural and functional features of eukaryotic mRNA untranslated regions. *Gene* **276**, 73-81 (2001).
54. Kozak, M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283-292 (1986).
55. Kathrin, L., Rhiju, D. & Maria, B. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews Molecular Cell Biology* **19**, 158–174 (2018).

56. Somers, J., Pöyry, T. & Willis, A.E. A perspective on mammalian upstream open reading frame function. *The International Journal of Biochemistry & Cell Biology* **45**, 1690-1700 (2013).
57. Pelletier, J. & Sonenberg, N. Insertion mutagenesis to increase secondary structure within the 5' noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**, 15–526 (1985).
58. Pantopoulos, K., Hentze, M.W. & Rapid responses to oxidative stress mediated by iron regulatory protein. *EMBO* **14**, 2917-2924 (1995).
59. Gray, N.K. & Henze, M.W. Iron regulatory protein prevents binding of the 43S translation pre-initiation complex to ferritin and eALAS mRNAs. *EMBO* **13**, 3882-3891 (1994).
60. Babendure, J.R., Babendure, J.L., Ding, J.-H. & Tsien, R.Y. Control of mammalian translation by mRNA structure near caps. *RNA* **12**, 851-861 (2006).
61. Leppek, K., Das, R. & Barna, M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nature Reviews Molecular Cell Biology* **19**, 158–174 (2018).
62. Peselis, A. & Serganov, A. Structure and function of pseudoknots involved in gene expression control. *Wiley Interdisciplinary Reviews: RNA* **5**, 803-822 (2015).
63. Dai, W. et al. The 5'-UTR of DDB2 harbors an IRES element and upregulates translation during stress conditions. *Gene* **573**, 57-63 (2015).
64. Pyronnet, S., Pradayrol, L. & Sonenberg, N. A Cell Cycle–Dependent Internal Ribosome Entry Site. *Molecular Cell* **5**, 607-616 (2000).
65. Holcik, M. & Sonenberg, N. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* **6**, 318-27 (2005).
66. Wang, G., Guo, X., Silveyra, P., Kimball, S.R. & Floros, J. Cap-independent translation of human SP-A 5'-UTR variants: a double-loop structure and cis-element contribution. *Am J Physiol Lung Cell Mol Physiol* **296**, L635-47 (2009).
67. Ali, N., Pruijn, G.J.M., Kenanl, D.J., Keene, J.D. & Siddiqui, A. Human La antigen is required for the hepatitis C virus internal ribosome entry site-mediated translation. *Journal of Biological Chemistry* **275**, 27531–27540 (2000).
68. Bhattacharyya, S. & Das, S. An apical GAGA loop within 5' UTR of the coxsackievirus B3 RNA maintains structural organization of the IRES element required for efficient ribosome entry. *RNA Biol* **3**, 60-8 (2006).
69. Liu, G., Yanguéz, E., Chen, Z. & Li, C. The duck hepatitis virus 5'-UTR possesses HCV-like IRES activity that is independent of eIF4F complex and modulated by downstream coding sequences. *Virol J* **8**, 147 (2011).

70. Carrieri, C. et al. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. . *Nature* **491**, 454–457 (2012).
71. Mayr, C. Evolution and Biological Roles of Alternative 3'UTRs. *Trends in Cell Biology* **26**, 227–237 (2016).
72. Barreau, C., Paillard, L. & Osborne, H.B. AU-rich elements and associated factors: are there unifying principles? *Nucleic Acids Research* **33**, 7138-7150 (2006).
73. Bartel, D.P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215-233 (2009).
74. Sharova, L.V. et al. Database for mRNA half-life of 19,977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Research* **16**, 45-58 (2009).
75. Martin, K., C. & Ephrussi, A. mRNA Localization: Gene Expression in the Spatial Dimension. *Cell* **136**, 719-730 (2009).
76. Johnstone, O. & Lasko, P. Translational Regulation and RNA Localization in Drosophila Oocytes and Embryos. *Annual Review of Genetics* **35**, 365-406 (2001).
77. King, M.L., Messitt, T.J. & Mowry, K.L. Putting RNAs in the right place at the right time: RNA localization in the frog oocyte. *Biology of the cell* **97**, 19-33 (2005).
78. Pandey, N.B. & Marzluff, W.F. The Stem-Loop Structure at the 3' End of Histone mRNA Is Necessary and Sufficient for Regulation of Histone mRNA Stability. *Molecular and Cellular Biology* **7**, 4557-4559 (1987).
79. Whitfield, M.L. et al. Stem-Loop Binding Protein, the Protein That Binds the 3' End of Histone mRNA, Is Cell Cycle Regulated by Both Translational and Posttranslational Mechanisms. *Molecular and Cellular Biology* **20**, 4188-4198 (2000).
80. Tian, B., Hu, J., Zhang, H. & Lutz, C.S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Research* **33**, 201-212 (2005).
81. Mayr, C. & Bartel, D.P. Widespread Shortening of 3'UTRs by Alternative Cleavage and Polyadenylation Activates Oncogenes in Cancer Cells. *Cell* **138**, 673-684 (2009).
82. Dibble, C.C. & Manning, B.D. Signal integration by mTORC1 coordinates nutrient input with biosynthetic output. *Nature Cell Biology* **15**(2013).
83. Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nature Reviews Molecular Cell Biology* **8**, 113-126 (2007).
84. Decker, C.J. & Parker, R. P-Bodies and Stress Granules: Possible Roles in the Control of Translation and mRNA Degradation. *Cold Spring Harbor Perspectives in Biology* **4**(2012).

85. Munroe, D. & Jacobson, A. mRNA poly(A) tail, a 3' enhancer of translational initiation. *Molecular and Cellular Biology* **10**, 3441-3455 (1990).
86. Palatnik, C.M., Wilkins, C. & Jacobson, A. Translational control during early dictyostelium development: Possible involvement of poly(A) sequences. *Cell* **36**, 1011-1025 (1984).
87. Jacobson, A. & Favreau, M. Possible Involvement of poly(A) in protein syntheses. *Nucleic Acids Research* **11**, 6353-6368 (1983).
88. Gallie, D.R. The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development* **5**, 2108-2116 (1991).
89. Iizuka, N., Najita, L., Franzusoff, A. & Sarnow, P. Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **14**, 7322-7330 (1994).
90. Tarun Jr, S.Z. & Sachs, A.B. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO* **15**, 7168-7177, (1996).
91. Le, H. et al. Translation Initiation Factors eIF-iso4G and eIF-4B Interact with the Poly(A)-binding Protein and Increase Its RNA Binding Activity. *Journal of Biological Chemistry* **272**, 16247-16255 (1997).
92. Imataka, H., Gradi, A. & Sonenberg, N. A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *17* **24**, 7480-7489, (1998).
93. Piron, M., Vende, P., Cohen, J. & Poncet, D. Rotavirus RNA-binding protein NSP3 interacts with eIF4GI and evicts the poly(A) binding protein from eIF4F. *EMBO* **17**, 5811-5821 (1998).
94. Safaee, N. et al. Interdomain Allosterity Promotes Assembly of the Poly(A) mRNA Complex with PABP and eIF4G. *Molecular Cell* **48**, 375-386 (2012).
95. Borman, A.M., Michel, Y.M. & Kean, K.M. Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5'-end. *Nucleic Acids Research* **28**, 4068-4075 (2000).
96. Kahvejian, A., Svitkin, Y.V., Sukarieh, R., M'Boutchou, M.N. & Sonenberg, N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev* **19**, 104-113 (2005).
97. Tarun Jr, S.Z., Wells, S.E., Deardorff, J.A. & Sachs, A.B. Translation initiation factor eIF4G mediates in vitro poly(A) tail-dependent translation. *Proceedings of the National Academy of Sciences* **94**, 9046-9051 (1997).

98. Wakiyama, M., Imataka, H. & Sonenberg, N. Interaction of eIF4G with poly(A)-binding protein stimulates translation and is critical for *Xenopus* oocyte maturation. *Current Biology* **10**, 1147-1150 (2000).
99. Christensen, A.K., Kahn, L.E. & Bourne, C.M. Circular polysomes predominate on the rough endoplasmic reticulum of somatotropes and mammatropes in the rat anterior pituitary. *American Journal of Anatomy* **178**, 1-10 (1987).
100. Afonina, Z.A., Myasnikov, A.G., Shirokov, V.A., Klaholz, B.P. & Spirin, A.S. Formation of circular polyribosomes on eukaryotic mRNA without cap-structure and poly(A)-tail: a cryo electron tomography study *Nucleic Acids Research* **42**(2014).
101. Chou, T. Ribosome Recycling, Diffusion, and mRNA Loop Formation in Translational Regulation. *Biophysical Journal* **85**, 755-773 (2003).
102. Uchida, N., Hoshino, S.-i., Imataka, H., Sonenberg, N. & Katada, T. A Novel Role of the Mammalian GSPT/eRF3 Associating with Poly(A)-binding Protein in Cap/Poly(A)-dependent Translation. *Journal of Biological Chemistry* **277**, 50286–50292 (2002).
103. Tarun Jr, S.Z. & Sachs, A.B. A common function for mRNA 5' and 3' ends in translation initiation in yeast *Genes & Development* **9**, 2997-3007 (1995).
104. Vicens, Q., Kieft, J.S. & Rissland, O.S. Revisiting the Closed-Loop Model and the Nature of mRNA 5'–3' Communication. *Molecular Cell* **72**, 805-812 (2018).
105. Adivarahan, S. et al. Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway. *Molecular Cell* **72**, 727-738.e5 (2018).
106. Van Treeck, B. et al. RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome. *Proc. Natl. Acad. Sci.* **115**, 2734–2739 (2018).
107. Protter, D.S.W. & Parker, P. Principles and Properties of Stress Granules. *Trends in Cell Biology* **26**, 668-679 (2016).
108. Choe, J. et al. mRNA circularization by METTL3–eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556–560 (2018).
109. Rissland, O.S. et al. The influence of microRNAs and poly(A) tail length on endogenous mRNA–protein complexes. *Genome Biology* **18**(2017).
110. Archer, S.K., Shirokikh, N.E., Hallwirth, C.V., Beilharz, T.H. & Preiss, T. Probing the closed-loop model of mRNA translation in living cells. *RNA Biology* **12**, 248-254 (2015).
111. Yan, X., Hoek, T.A., Vale, R.D. & Tanenbaum, M.E. Dynamics of Translation of Single mRNA Molecules In Vivo. *Cell* **165**, 976-989 (2016).
112. Wu, B., Eliscovich, C., Yoon, Y.J. & Singer, R.H. Translation dynamics of single mRNAs in live cells and neurons. *Science* **352**, 1430-1435 (2016).

113. Pichon, X. et al. Visualization of single endogenous polysomes reveals the dynamics of translation in live human cells. *Journal of Cell Biology* **214**, 769 (2016).
114. Khong, A. & Parker, R. mRNP architecture in translating and stress conditions reveals an ordered pathway of mRNP compaction. *Journal of Biological Chemistry* **217**, 4124–4140 (2018).
115. Ross, J. mRNA Stability in Mammalian Cells. *Microbiological Reviews* **59**, 423-450 (1995).
116. Hambraeus, G., Wachenfeldt, C.v. & Hederstedt, L. Genome-wide survey of mRNA half-lives in *Bacillus subtilis* identifies extremely stable mRNAs. *Molecular Genetics and Genomics* **269**, 706–714 (2003).
117. Berger, S.L. & Cooper, H.L. Very short-lived and stable mRNAs from resting human lymphocytes. *Proc. Natl. Acad. Sci.* **72**, 3873-3877, (1975).
118. Herrick, D., Parker, R. & Jacobson, A. Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **10**, 2269-2284 (1990).
119. Russell, J.E., Morales, J. & Liebhaber, S.A. The role of mRNA stability in the control of globin gene expression. *Prog. Nucleic Acid Res. Mol. Biol.* **57**, 249-287 (1997).
120. Theodorakis, N.G. & Morimoto, R.I. Posttranscriptional regulation of hsp70 expression in human cells: effects of heat shock, inhibition of protein synthesis, and adenovirus infection on translation and mRNA stability. *Mol Cell Biol* **7**, 4357-68 (1987).
121. Schiavi, S.C., Belasco, J.G. & Greenberg, M.E. Regulation of proto-oncogene mRNA stability. *Biochimica et Biophysica Acta* **1114**, 95-106 (1992).
122. Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nature Reviews Molecular Cell Biology* **8**, 113–126 (2007).
123. Guhaniyogi, J. & Brewer, G. Regulation of mRNA stability in mammalian cells. *Gene* **265**, 11-23 (2001).
124. Jacobson, A. & Peltz, S.W. Interrelationships of the Pathways of mRNA Decay and Translation in Eukaryotic Cells. *Annual Review of Biochemistry* **65**, 693-739 (1996).
125. Spector, D.L. SnapShot: cellular bodies. *Cell* **1070.e1-1070.e2**, 1070 (2006).
126. Ivanov, P., Kedersha, N. & Anderson, P. Stress Granules and Processing Bodies in Translational Control. *Cold Spring Harbor Laboratory Perspectives in Biology* (2019).
127. Kedersha, N. et al. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *Journal of Cell Biology* **169**, 871- 884 (2005).

128. Anderson, P. & Kedersha, N. Stressful initiations. *Journal of Cell Science* **115**, 3227-3234 (2002).
129. Wek, R.C. Role of eIF2 α Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harbor Laboratory Perspectives in Biology* (2018).
130. Jain, S. et al. ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**, 487-498 (2016).
131. Kedersha, N. et al. Dynamic Shuttling of Tia-1 Accompanies the Recruitment of mRNA to Mammalian Stress Granules. *Journal of Cell Biology* **151**, 1257-1268 (2000).
132. Khong, A. et al. The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Molecular Cell* **68**, 808-820 (2017).
133. Luo, Y., Na, Z. & Slavoff, S.A. P-bodies: Composition, Properties, and Functions. *Biochemistry* **57**, 2424–2431 (2018).
134. Sheth, U. & Parker, R. Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies. *Science* **300**, 805-808 (2003).
135. Hubstenberger, A. et al. P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. *Molecular Cell* **68**, 144-157.e5 (2017).
136. Brengues, M., Teixeira, D. & Parker, R. Movement of Eukaryotic mRNAs Between Polysomes and Cytoplasmic Processing Bodies. *Science* **310**, 486-489 (2005).
137. Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I. & Filipowicz, W. Stress-induced Reversal of MicroRNA Repression and mRNA P-body Localization in Human Cells. *Cold Spring Harbor Symposia on Quantitative Biology* **71**, 513-521 (2006).
138. Tourrière, H. et al. The RasGAP-associated endoribonuclease G3BP assembles stress granules. *The Journal of Cell Biology* **160**, 823–831 (2003).
139. Molliex, A. et al. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **163**, 123-133 (2015).
140. Lin, Y., Protter, D.S.W., Rosen, M.K. & Parker, R. Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Molecular Cell* **60**, 208-219 (2015).
141. Kroschwald, S. et al. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *eLife* **4**, e06807 (2015).
142. Nott, T.J. et al. Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles. *Molecular Cell* **57**, 936–947 (2015).
143. Kato, M. et al. Cell-free Formation of RNA Granules: Low Complexity Sequence Domains Form Dynamic Fibers within Hydrogels. *Cell* **149**, 753–767 (2012).

144. Ma, X.M. & Blenis, J. Molecular mechanisms of mTOR-mediated translational control. *Nature Reviews Molecular Cell Biology* **10**, 307-318 (2009).
145. Meyuhas, O. & Kahan, T. The race to decipher the top secrets of TOP mRNAs. *Biochim Biophys Acta* (2014).
146. Schibler, U., Kelley, D.E. & Perry, R.P. Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. *Journal of Molecular Biology* **115**, 695-714 (1977).
147. Yoshihama, M. et al. The Human Ribosomal Protein Genes: Sequencing and Comparative Analysis of 73 Genes. *Genome Research* **12**, 379-390 (2002).
148. Levy, S., Avni, D., Hariharan, N., Perry, R.P. & Meyuhas, O. Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. *Proc. Natl. Acad. Sci.* **88**, 3319-3323 (1991).
149. Valentina Iadevaia, S.C., Elisa Tino, Francesco Amaldi, and Fabrizio Loreni. All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. *RNA* **14**, 1730–1736 (2008).
150. Meyuhas, O. & Drazan, A. Chapter 3 Ribosomal Protein S6 Kinase: From TOP mRNAs to Cell Size. *Progress in Molecular Biology and Translational Science* **90**, 109-153 (2009).
151. Hornstein, E., Tang, H. & Meyuhas, O. Mitogenic and Nutritional Signals Are Transduced into Translational Efficiency of TOP mRNAs. *Cold Spring Harbor Symposium Quantitative Biology* **66**(2001).
152. Biberman, Y. & Meyuhas, O. TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. *FEBS Letters* **456**, 357-360 (1999).
153. Fonseca, B.D., Lahr, R.M., Damgaard, C.K., Alain, T., Berman, A.J. LARP1 on TOP of ribosome production. *WIREs RNA* **9**(2018).
154. Mattioli, M. & Reichlin, M. Heterogeneity of RNA protein antigens reactive with sera of patients with systemic lupus erythematosus. Description of a cytoplasmic nonribosomal antigen. *Arthritis & Rheumatology* **17**, 421-429 (1974).
155. Alspaugh, M.A. & Tan, E.M. Antibodies to cellular antigens in Sjögren's syndrome. *The Journal of Clinical Investigation* **55**, 1067-1073 (1974).
156. Bousquet-Antonelli, C. & Deragon, J.M. A comprehensive analysis of the La-motif protein superfamily. *RNA* **15**, 750-64 (2009).
157. Maraia, R.J., Mattijssen, S., Cruz-Gallardo, I. & Conte, M.R. The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. *WIREs RNA* **8**, e1430 (2017).

158. Andreas Markert, M.G., Javier Martinez, Julia Wiesner, Andreas Meyerhans, Oded Meyuhas, Albert Sickmann, Utz Fischer. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *Scientific Report* **9**, 569-575 (2008).
159. Muniz, L., Egloff, S. & Kiss, T. RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Research* **41**, 4686–4698 (2013).
160. Uchikawa, E. et al. Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Research* **43**, 3373–3388 (2015).
161. Cai, L., Fritz, D., Stefanovic, L. & Stefanovic, B. Binding of LARP6 to the conserved 5' stem-loop regulates translation of mRNAs encoding type I collagen. *J Mol Biol* **395**, 309-26 (2010).
162. Stefanovic, L., Longo, L., Zhang, Y. & Stefanovic, B. Characterization of binding of LARP6 to the 5' stem-loop of collagen mRNAs: implications for synthesis of type I collagen. *RNA Biol* **11**, 1386-401 (2014).
163. Martino, L. et al. Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module. *Nucleic Acids Research* **43**, 645-660 (2015).
164. Yang, R. et al. La-related protein 4 binds poly(A), interacts with the poly(A)-binding protein MLE domain via a variant PAM2w motif, and can promote mRNA stability. *Mol Cell Biol* **31**, 542-56 (2011).
165. Isabel Cruz-Gallardo, L.M., Geoff Kelly, R. Andrew Atkinson, Roberta Trotta., Stefano De Tito, P.C., Zainab Ahdash, Yifei Gu, Tam T.T. Bui, and Maria & Conte, R. LARP4A recognizes polyA RNA via a novel binding mechanism mediated by disordered regions and involving the PAM2w motif, revealing. *Nucleic Acids Research* (2019).
166. Küspert, M. et al. LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. *RNA* **21**, 1294-1305 (2015).
167. Fonseca, B.D. et al. La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). *J Biol Chem* **290**, 15996-6020 (2015).
168. Lahr, R.M. et al. La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *eLife* **6**, e24146 (2017).
169. Philippe, L., Vasseur, J., Debart, F. & Thoreen, C.C. La-related protein 1 (LARP1) repression of TOP mRNA translation is mediated through its cap-binding domain and controlled by an adjacent regulatory region. *Nucleic Acids Research* **46**, 1457–1469 (2018).

170. Tcherkezian, J. et al. Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. *Genes Dev* **28**, 357-71 (2014).
171. Lahr, R.M. et al. The La-related protein 1-specific domain repurposes HEAT-like repeats to directly bind a 5'TOP sequence. *Nucleic Acids Res* **43**, 8077-88 (2015).
172. Dong, G., Chakshusmathi, G., Wolin, S.L. & Reinisch, K.M. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J* **23**, 1000-7 (2004).
173. Teichmann, M., Dumay-Odelot, H. & Fribourg, S. Structural and functional aspects of winged-helix domains at the core of transcription initiation complexes. *Transcription* **3**, 2-7 (2012).
174. Dong, G., Chakshusmathi, G., Wolin, S.L. & Reinisch, K.M. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *The EMBO Journal* **23**, 1000-1007 (2004).
175. Clery, A., Blatter, M. & Allain, F.H. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290-8 (2008).
176. Maris, C., Dominguez, C. & Allain, F.H.T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *The FEBS Journal* **272**, 2118–2131 (2005).
177. Kuwasako, K. et al. Solution structure of the second RNA recognition motif (RRM) domain of murine T cell intracellular antigen-1 (TIA-1) and its RNA recognition mode. *Biochemistry* **47**, 7437-6450 (2008).
178. Singh, M., Choi, C.P. & Feigon, J. xRRM. *RNA Biology* **10**, 353–359 (2013).
179. Jacks, A. et al. Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* **11**, 833-43 (2003).
180. Teplova, M. et al. Structural basis for recognition and sequestration of UUU(OH) 3' termini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* **21**, 75-85 (2006).
181. Alfano, C. et al. Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* **11**, 323-9 (2004).
182. Kotik-Kogan, O., Valentine, E.R., Sanfelice, D., Conte, M.R. & Curry, S. Structural analysis reveals conformational plasticity in the recognition of RNA 3' ends by the human La protein. *Structure* **16**, 852-62 (2008).
183. Singh, M., Choi, C.P. & Feigon, J. xRRM: A new class of RRM found in the telomerase La family protein p65. *RNA Biology* **10**, 353-359 (2013).

184. Jacks, A. et al. Structure of the C-Terminal Domain of Human La Protein Reveals a Novel RNA Recognition Motif Coupled to a Helical Nuclear Retention Element. *Structure* **11**, 833-843 (2003).
185. Martino, L. et al. Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein. *Nucleic Acids Research* **40**, 1381-1394 (2012).
186. Singh, M. et al. Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* **47**, 16-26 (2012).
187. Holohan, B. et al. Impaired telomere maintenance in Alzami syndrome patients with LARP7 deficiency. *BMC Genomics*, 17(Suppl 9):749 (2016).
188. Stavra, C. & Blagden, S. The La-Related Proteins, a Family with Connections to Cancer. *Biomolecules* **5**, 2701-2722 (2015).
189. Shao, R. et al. The novel lupus antigen related protein acheron enhances the development of human breast cancer. *International Journal of cancer* **130**, 544–554 (2012).
190. Bai, S.W. et al. Identification and characterization of a set of conserved and new regulators of cytoskeletal organization, cell morphology and migration. *BMC Biology* **9**, 54 (2011).
191. Seetharaman, S., Flemmyng, E., Shen, J., Conte, M.R. & Ridley, A.J. The RNA-binding protein LARP4 regulates cancer cell migration and invasion. *Cytoskeleton* **73**, 680-690 (2016).
192. Koso, H. et al. Identification of RNA-Binding Protein LARP4B as a Tumor Suppressor in Glioma. *Cancer Research* **76**, 2254-2264 (2016).
193. Li, Y., Jiao, Y., Yang, L. & Liu, Y. Expression of La Ribonucleoprotein Domain Family Member 4B (LARP4B) in Liver Cancer and Their Clinical and Prognostic Significance. *Disease Markers*, eCollection 2019 (2019).
194. Zhang, Y. et al. La-related protein 4B maintains murine MLL-AF9 leukemia stem cell self-renewal by regulating cell cycle progression. *Experimental Hematology* **43**, 309-318 (2015).
195. Burrows, C. et al. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Res* **38**, 5542-53 (2010).
196. Xie, C. et al. LARP1 predict the prognosis for early-stage and AFP-normal hepatocellular carcinoma. *J Transl Med* **11**, 272 (2013).
197. Mura, M. et al. LARP1 post-transcriptionally regulates mTOR and contributes to cancer progression. *Oncogene* (2014).

198. Hopkins, T.G. et al. The RNA-binding protein LARP1 regulates survival and tumorigenesis in ovarian cancer. (Imperial College London, London, UK, 2015).
199. Aoki, K. et al. LARP1 specifically recognizes the 3' terminus of poly(A) mRNA. *FEBS Lett* **587**, 2173-8 (2013).
200. Fonseca, B.D. et al. The ever-evolving role of mTOR in translation. *Semin Cell Dev Biol* (2014).
201. Hong, S. et al. LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *eLife* **2017**;6:e25237 (2017).
202. Hsu, P.P. et al. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* **332**, 1317-1322 (2011).
203. Kang, S.A. et al. mTORC1 Phosphorylation Sites Encode Their Sensitivity to Starvation and Rapamycin. *Science* **341**, 364-374 (2013).
204. Fonseca, B.D. et al. LARP1 is a major phosphorylation substrate of mTORC1. *bioRxiv* (2018).
205. Al-Ashtal, H.A., Rubottom, C.M., Leeper, T.C. & Berman, A.J. The LARP1 La-Module recognizes both ends of TOP mRNAs. *RNA Biology*, 1-11 (2019).
206. Maraia, R.J., Mattijssen, S., Cruz-Gallardo, I. & Conte, M.R. The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives. *Wiley Interdiscip Rev RNA* **8**(2017).
207. Bayfield, M.A., Yang, R. & Maraia, R.J. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* **1799**, 365-78 (2010).
208. Alfano, C., Babon, J., Kelly, G., Curry, S. & Conte, M.R. Resonance assignment and secondary structure of an N-terminal fragment of the human La protein. *J Biomol NMR* **27**, 93-4 (2003).
209. Chakshusmathi, G., Kim, S.D., Rubinson, D.A. & S.L., W. A La protein requirement for efficient pre-tRNA folding. *EMBO* **22**, 6562-6572 (2003).
210. Krueger, B.J. et al. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Research* **36**, 2219–2229 (2008).
211. Schäffler, K. et al. A stimulatory role for the La-related protein 4B in translation. *RNA* **16**, 1488–1499 (2010).
212. Küspert, M. et al. LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. *RNA* **21**, 1294-305 (2015).

213. Savitzky, A. & Golay, M.J.E. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry* **36**, 1627-& (1964).
214. Dror Avni, Silvian Shama, Gabrizio Loreni & Meyuhas, O. Vertebrate mRNAs with a 5'-Terminal Pyrimidine Tract Are Candidates for Translational Repression in Quiescent Cells: Characterization of the Translational cis-Regulatory Element. *Molecular and Cellular Biology* **14**, 3822-3833 (1994).
215. Hong, S. et al. LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *Elife* **6**(2017).
216. Nykamp, K., Lee, M.H. & Kimble, J. C. elegans La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. *RNA* **14**, 1378-89 (2008).
217. Wilbertz, J.H. et al. Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Molecular Cell* **73**, 946-958 (2019).
218. Cardinali, B., Carissimi, C., Gravina, P. & Pierandrei-Amaldi, P. La protein is associated with terminal oligopyrimidine mRNAs in actively translating polysomes. *J Biol Chem* **278**, 35145-51 (2003).
219. Crosio, C., Boyl, P.P., Loreni, F., Pierandrei-Amaldi, P. & Amaldi, F. La protein has a positive effect on the translation of TOP mRNAs in vivo. *Nucleic Acids Research* **28**, 2927-2934 (2000).
220. Meerovitch, K. et al. La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *Journal of Virology* **67**, 3798-3807 (1993).
221. Izumi, R.E., Das, S., Barat, B., Raychaudhuri, S. & Dasgupta, A. A Peptide from Autoantigen La Blocks Poliovirus and Hepatitis C Virus Cap-Independent Translation and Reveals a Single Tyrosine Critical for La RNA Binding and Translation Stimulation. *Journal of Virology* **78**, 3763-3776 (2004).
222. Kim, Y.K., Back, S.H., Rho, J., Lee, S.H. & Jang, S.K. La autoantigen enhances translation of BiP mRNA. *Nucleic Acids Research* **29**, 5009-5016 (2001).
223. Vinayak, J. et al. Human La binds mRNAs through contacts to the poly(A) tail. *Nucleic Acids Research* **4**, 4228-4240 (2018).
224. Mossessova, E. & Lima, C.D. Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. *Molecular Cell* **5**, 865-876 (2000).

225. Li, X., Song, J. & Yi, C. Genome-wide Mapping of Cellular Protein–RNA Interactions Enabled by Chemical Crosslinking. *Genomics Proteomics Bioinformatics* **12**, 72-78 (2014).
226. Popovici, S.-T., Kok, W.T. & Schoemakers, P.J. Band broadening in size-exclusion chromatography of polydisperse samples. *Journal of Chromatography A* **1060**, 237-252 (2004).
227. Corporation, W.T. *DynaPro NanoStar User's Guide*, (2010).
228. Roy, G. et al. Paip1 Interacts with Poly(A) Binding Protein through Two Independent Binding Motifs. *Molecular and Cellular Biology* **22**, 3769–3782 (2002).
229. Xie, J., Kozlov, G. & Gehring, K. The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* **1839**, 1062-8 (2014).
230. Albrecht, M. & Lengauer, T. Survey on the PABC recognition motif PAM2. *Biochem Biophys Res Commun* **316**, 129-38 (2004).
231. Khaleghpour, K. et al. Translational repression by a novel partner of human poly(A) binding protein, Paip2. *Molecular Cell* **7**, 205-216 (2001).
232. Smith, R.W.P. et al. Viral and cellular mRNA-specific activators harness PABP and eIF4G to promote translation initiation downstream of cap binding. *Proceedings of the National Academy of Sciences* **114**, 6310-6315 (2017).
233. Collier, B., Gorgoni, B., Loveridge, C., Cooke, H.J. & Gray, N.K. The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. *The EMBO Journal* **24**, 2656-26666 (2005).
234. Kleene, K.C. Connecting cis-elements and trans-factors with mechanisms of developmental regulation of mRNA translation in meiotic and haploid mammalian spermatogenic cells. *Reproduction* **146**, R1-19 (2013).
235. Kozlov, G., Safaee, N., Rosenauer, A. & Gehring, K. Structural Basis of Binding of P-body-associated Proteins GW182 and Ataxin-2 by the Mlle Domain of Poly(A)-binding Protein. *The Journal of Biological Chemistry* **285**, 13599-13606 (2010).
236. Deardorff, J.A. & Sachs, A.B. Differential effects of aromatic and charged residue substitutions in the RNA binding domains of the yeast Poly(A)-binding protein. *Journal of Molecular Biology* **269**, 67-81 (1997).
237. Kühn, U. & Pieler, T. Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *Journal of Molecular Biology* **256**, 20-30 (1996).
238. Gentilella, A. et al. Autogenous Control of 5'TOP mRNA Stability by 40S Ribosomes. *Molecular Cell* **67**, 55-70 (2017).

239. Huang, K.L., Chadee, A.B., Chen, C.Y., Zhang, Y. & Shyu, A.B. Phosphorylation at intrinsically disordered regions of PAM2 motif-containing proteins modulates their interactions with PABPC1 and influences mRNA fate. *RNA* **19**, 295-305 (2013).
240. Brook, M. et al. The multifunctional poly(A)-binding protein (PABP) 1 is subject to extensive dynamic post-translational modification, which molecular modelling suggests plays an important role in co-ordinating its activities. *Biochemical Journal* **441**, 803–816 (2012).
241. Craig, A.W.B., Svitkin, Y.V., Lee, H.S., Belsham, G.J. & Sonenberg, N. The La autoantigen contains a dimerization domain that is essential for enhancing translation. *Molecular and Cellular Biology* **17**, 163-169 (1997).
242. Zinshteyn, B., Rojas-Duran, M.F. & Gilbert, W.V. Translation initiation factor eIF4G1 preferentially binds yeast transcript leaders containing conserved oligo-uridine motifs. *RNA* **23**, 1365-1375 (2017).
243. Philippe, L., van den Elzen, A.M.G., Watson, M.J. & Thoreen, C.C. Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5' TOP motifs. *Proc. Natl. Acad. Sci.* (2020).
244. Huang, Y., Bayfield, M.A., Intine, R.V. & Maraia, R.J. Separate RNA-binding surfaces on the multifunctional La protein mediate distinguishable activities in tRNA maturation. *Nature Structural & Molecular Biology* **13**, 611-618 (2006).
245. Stone, M.D. et al. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. **446**, 458–461 (2006).
246. Calabretta, S. & Richard, S. Emerging Roles of Disordered Sequences in RNA-Binding Proteins. *Trends in Biochemical Sciences* **40**, 662-672 (2015).
247. van der Lee, R. et al. Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews* **114**, 6589-6631 (2014).
248. Bah, A. et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106-109 (2015).
249. Marella, S.A. et al. An interdomain bridge influences RNA binding of the human La protein. *Journal of Biological Chemistry* **294**, 1529-1540 (2018).
250. Kucera, N.J., Hodsdon, M.E. & Wolin, S.L. An intrinsically disordered C terminus allows the La protein to assist the biogenesis of diverse noncoding RNA precursors. *Proceedings of the National Academy of Sciences* **108**, 1308-1313 (2011).
251. Kuehnert, J. et al. Novel RNA chaperone domain of RNA-binding protein La is regulated by AKT phosphorylation. *Nucleic Acids Research* **43**, 581-594 (2015).

252. Vukmirovic, M., Manojlovic, Z. & Stefanovic, B. Serine threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. *Molecular and Cellular Biology* **33**, 3893-3906 (2013).
253. Thandapani, P., O'Connor, T.R., Bailey, T.L. & Richard, S. Defining the RGG/RG motif. *Molecular Cell Review* **50**, 612-623 (2013).
254. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
255. Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
256. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. & Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91 (2009).
257. Jones, D.T. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol* **292**, 195-202 (1999).
258. Kelley, L.A. & Sternberg, M.J. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**, 363-71 (2009).
259. Pace, C.N. & Scholtz, J.M. A helix propensity scale based on experimental studies of peptides and proteins. *Biophysical Journal* **75**, 422-427 (1998).
260. Bhattacharjee, N. & Biswas, P. Position-specific propensities of amino acids in the β -strand. *BMC Structural Biology* **10**, 29 (2010).
261. Minor, D.L. & Kim, P.S. Measurement of the β -sheet-forming propensities of amino acids. *Nature* **367**, 660-663 (1994).
262. Lee, J., Freddolino, P.L. & Zhang, Y. *From Protein Structure to Function with Bioinformatics*, 3-25 (Springer, Dordrecht, 2009).