Structural and Functional Study of the La-related Protein 1 (LARP1)

Jianning Jiang

Department of Biochemistry

McGill University, Montreal

April 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master of Science.

TABLE OF CONTENTS

ABSTRACT

La-related proteins (LARPs) constitute a family of RNA-binding proteins involved in a variety of gene regulatory activities. LARPs are characterized by a unique tandem of two RNA-binding domains, La motif (LaM) and RNA recognition motif (RRM), collectively referred to as a La-module. In many LARPs, the LaM domain synergizes with downstream RRM to bind RNA. LARP1 is a protein involved in regulating synthesis of ribosomal proteins in response to mTOR signaling and mRNA stabilization. While its La-module has been studied functionally, its structure has not been characterized. Here, we characterized LARP1 La module. We showed that it does not contain a structured RRM. Instead, LARP1 LaM acts as a stand-alone RNA binding domain with submicromolar affinity and specificity for poly(A) RNA. We found that LARP1 LaM has enhanced binding for guanosine-substituted poly(A) at terminal, penultimate or internal positions. Through multiple high-resolution crystal structures of the LARP1 LaM domain in complex with poly(A), we elucidated the mechanism for its high specificity binding to the RNA 3′-end. Our structural studies provided insights into the molecular mechanism of LaM-poly(A) binding, and the role of LARP1 in poly(A) 3'-end protection.

RÉSUMÉ

Les protéines La-related (LARPs) constituent une famille de protéines de liaison à l'ARN impliquées dans une variété d'activités régulatrices des gènes. Les LARPs se caractérisent par un tandem unique de deux domaines de liaison à l'ARN, le motif La (LaM) et le motif de reconnaissance de l'ARN (RRM), collectivement appelé un module La. Dans de nombreux LARPs, le domaine LaM agit de manière synergique avec le RRM en aval pour lier l'ARN. LARP1 est une protéine impliquée dans la régulation de la synthèse des protéines ribosomiques en réponse au signal mTOR et dans la stabilisation de l'ARNm. Bien que son module La ait été étudié fonctionnellement, sa structure n'a pas été caractérisée. Dans cette étude, nous avons caractérisé le module La de LARP1. Nous avons montré qu'il ne contient pas de RRM structuré. Au lieu de cela, le domaine LaM de LARP1 agit comme un domaine de liaison à l'ARN autonome avec une affinité submicromolaire et une spécificité pour l'ARN poly(A). Nous avons constaté que LARP1 LaM présente une affinité accrue pour l'ARN poly(A) substitué par la guanosine aux positions terminales, pénultièmes ou internes. À l'aide de multiples structures cristallines haute résolution du domaine LaM de LARP1 en complexe avec l'ARN poly(A), nous avons élucidé le mécanisme de sa liaison hautement spécifique à l'extrémité 3'-ARN. Nos études structurelles ont fourni des informations sur le mécanisme moléculaire de la liaison LaM-poly(A) et le rôle de LARP1 dans la protection de l'extrémité 3'-ARN poly(A).

ACKNOWLEDGEMENTS

I am incredibly grateful for the support and guidance I have received throughout my studies, and I wish to extend my heartfelt thanks to all those who have contributed to making this experience exceptional. In particular, I want to express my deepest appreciation to Dr. Kalle Gehring for his unwavering faith in me, his exceptional mentorship, and his outstanding leadership. As an undergraduate student, Dr. Gehring welcomed me into his lab, and his support allowed me to flourish into the person I am today. I am grateful for his generosity and for the invaluable lessons I have learned from him.

I also want to thank Dr. Guennadi Kozlov for his daily guidance and encouragement. His patience and expertise have been indispensable to my research, and I am grateful for his unwavering support. I am particularly thankful for the technical training he provided me with, and for the valuable insights he shared with me that made my graduate project possible.

I would like to express my gratitude to our collaborators Dr. Sandy Mattijssen and Dr. Richard Maraia from the NICHD for their invaluable input from the cellular perspectives into this research projects. The meaningful discussions we had and their insights with respect to LARPs and RNA-binding proteins in general have been instrumental in shaping the direction and outcomes of this project.

I extend my sincere thanks to the former and current lab members of the Gehring lab: Luba, Ian, Samuel, Rayan, Ashkan, Ali, and Emma. I have cherished the time I have spent with them. Our discussions and collaborations have been instrumental to the success of my work.

To my Research Advisory Committee, Dr. Thomas Duchaine and Dr. Jerry Pelletier, I extend my deepest gratitude. I am grateful for their willingness to support my project, and for the

wise counsel they have provided me with. Their insights and advice will continue to guide me throughout my scientific career.

Finally, I want to express my profound appreciation to my parents. Their unwavering support, encouragement, and love have sustained me through the highs and lows of my studies. This accomplishment would not have been possible without them, and I will forever be grateful for their constant presence in my life.

AUTHOR CONTRIBUTION

Some of the results presented in this thesis were previously published in *Nucleic Acids Research*, Volume 50, Issue 16, 9 September 2022, Pages 9534–9547[1].

The plasmid constructs and truncations were designed by our research associate Dr. Guennadi Kozlov.

The protein constructs were expressed and purified by me with helpful feedback from Dr. Kozlov. The isothermal titration calorimetry (ITC) experiments were performed and analyzed by me and Dr. Guennadi Kozlov with feedback from Dr. Kalle Gehring.

Protein crystallization and optimization trials were conducted by Dr. Kozlov. Processing of the diffraction patterns was performed by Dr. Kozlov. Crystal structure refinement and molecular replacement was taught and performed by Kozlov.

In vitro cell assays were performed and analyzed by Dr. Sandy Mattijssen and colleagues from the group of Dr. Richard Maraia. Single-molecule poly(A) tail sequencing were also performed by the group of Dr. Richard Maraia.

LIST OF FIGURES

LIST OF ABBREVIATIONS

4E-BP	Eukaryotic translation initiation factor 4E-binding protein 1
5'TOP	5′ terminal oligopyrimidine
eIF4G	Eukaryotic initiation factor 4 G
EMSA	Electrophoretic mobility shift assay
GTP	Guanosine-5'-triphosphate
ITC	Isothermal calorimetry titration
LARP	La-related protein
mTORC1	Mammalian target of rapamycin complex 1
NMR	Nuclear magnetic resonance
p-TEFb	Positive transcription elongation factor-b
PABP	Poly(A) binding protein
Paip2	Poly(A) Binding Protein Interacting Protein 2
PAM2	Poly(A) binding protein interaction motif 2
RACK1	Receptor for activated protein kinase C 1
RAPTOR	Regulatory-associated protein of mTOR
RHA	RNA helicase A
RICTOR	Rapamycin-insensitive companion of mTOR
RPS6	Ribosomal protein S6
RRM	RNA recognition motif
SELEX	Systematic evolution of ligands by exponential enrichment
STRAP	Serine/Threonine kinase Receptor Associated Protein

CHAPTER 1. INTRODUCTION

1.1 The family of La-related proteins (LARPs)

Protein production is tightly regulated in all organisms. In eukaryotes, transcription and translation are spatially separated, allowing for post-transcriptional modifications that control pre-mRNA processing in the nucleus and the export of the mature mRNA to the cytosol. These modifications are regulated by a variety of mechanisms and factors, including RNA binding proteins (RBPs) that have diverse roles in controlling nuclear export rate, translocation and compartmentalization of mRNA, ribosome assembly and processivity during translation, and regulation of metabolism. La-related proteins (LARPs) are one such set of proteins, representing a superfamily of RNA-binding proteins that are conserved across eukaryotic organisms[2, 3], and play roles in gene expression regulation and RNA metabolism in both the nucleus and cytoplsm[4]. The LARPs have in common a highly conserved winged helix domain, termed the La motif (LaM), which is frequently accompanied by a downstream member-specific RNA-recognition motif (RRM) in a synergistic manner[5, 6]. The tandem arrangement of LaM and RRM, connected by a short linker, is commonly referred to as the La-module[2]. Evolutionary and structural analysis classified LARPs into five subfamilies (Figure 1), namely LARP1, genuine La protein (LARP3), LARP4, LARP6, and LARP7, each with unique family-specific motifs and domains that contribute to their differences in both structure and function. Each member in LARP protein family appears to associate with a particular class of RNAs, ranging from pre-tRNAs to mature mRNAs, and thereby carrying out a variety of post-transcriptional regulatory functions that are only beginning to be understood.

Figure 1

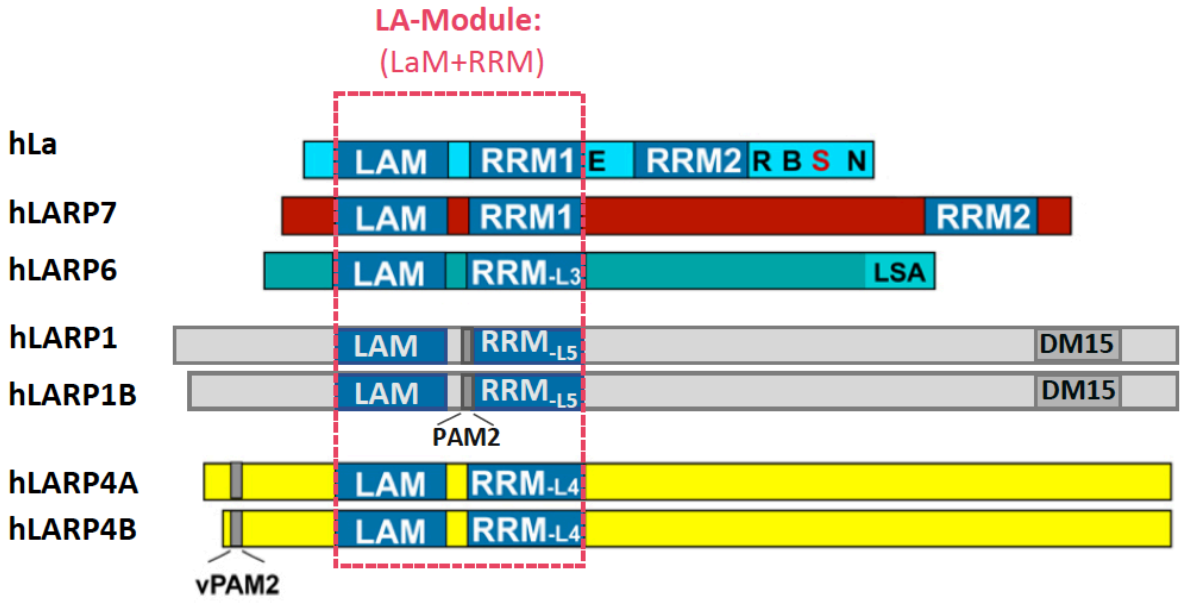


Figure 1. Schematic representation of architectures of La proteins and LARPs in human. The genuine La, LARP7, LARP6 and LARP4A/B contain the bipartite La Module domain, which consist of a La Motif (LaM) and RNA Recognition Motif (RRM) joined by a flexible linker. Various other domains can be seen such as a secondary RRM (RRM2), Poly(A)-binding protein interacting motif-2 (PAM2), DM15, and LaM and S1-associated motif (LSA). (*figure adapted from* [7].)

1.1.1 Nuclear LARPs: genuine La-protein and LARP7

Genuine La protein and LARP7 are highly related proteins characterized by the presence of a second RRM domain, designated RRM2α. The prototypic La-module in human La recognizes terminal UUU-3'OH in a sequence and length-dependent manner[8] and regulates RNA 3'end metabolism[3, 9, 10]mainly in the nucleus. This enables human La to function in controlling the biogenesis and maturation pathway for RNA polymerase III transcripts[11, 12], including precursors of small nuclear RNAs, nucleolar RNAs and tRNAs[13]. For tRNAs, La binding to the 3′UUU trailer releases the tRNA body to allow 3′-end maturation by endonucleolytic cleavage[9, 11]. La also exerts a chaperone-like activity[14] by prolonging the time window of protection for maturation events and assisting the pre-tRNAs in their local folding[8, 9].

LARP7 also recognizes a terminal UUU_{OH} stretch and exhibits similar activities in RNA metabolism control. In contrast to human La, LARP7 associates preferentially with the 7SK snRNA, a nuclear non-coding RNA[15, 16], and in turn regulates the activity level of positive transcription elongation factor-b (P-TEFb)[17].

1.1.2 Cytoplasmic LARPs: LARP1, LARP4, and LARP6

LARPs can be broadly defined as factors containing a high degree of homology with genuine La protein in the LaM-RRM arrangement of their La module, but have taken on independent and divergent functions from those associated with genuine La.

LARPs 1, 4 and 6 predominantly reside in the cytoplasm and are highly divergent. LARP1 and LARP4 appear to exhibit regulatory activities in RNA 3'- poly(A) metabolism through their peptide motif called poly(A) binding protein interaction motif-2 (**PAM2**) which is required for binding to poly(A)-binding protein (PABP). LARP 4 differs from LARP1 in the position of its La-module relative to the PAM2. LARP1 proteins carry a unique C-terminal DM15 domain[18], allowing specific recognition to the m^7GpppC cap of 5′ terminal oligopyrimidine (5′ TOP) motif[19, 20], a defining trait of mRNAs that encode for ribosomal proteins and other translation factors.

LARP4 modulates mRNA stability and enhance translation, through interactions with the 3′ UTR of mRNAs, the Poly(A) binding protein (PABP), RACK1 (receptor for activated protein kinase C 1) and polysomes[21, 22]. Vertebrate LARP6 binds stem-loop structure, comprising a large internal loop, found in the 5′ UTR of the collagen α1 and α2 of type I and α1 type III mRNAs. This association is involved in the collagen biosynthetic pathway[23]. It is however highly likely that LARP6 proteins perform additional biological roles, executed through recognition of other RNA and/or protein partners. In support of this, Human LARP6 has been identified as a

regulator of miR-141 and miR-145 biogenesis[23], and a recent high throughput RNA-SELEX analysis uncovered a complex binding specificity for Human LARP6, ranging from linear motifs, multiple dimeric motifs, and internal loops embedded in a double-stranded RNA stem[24]. Protein interactors of Human LARP6 include non-muscle myosin, RNA helicase A (RHA) and STRAP (Serine/Threonine kinase Receptor Associated Protein[23, 25]. The exact role of these associations is yet to be discovered, and the molecular details for the interactions are not known. Human LARP6 contains an LSA domain at its C-terminus. Neither the function nor the structure of this domain has been determined. It has been hypothesized to contribute to ligand sensitivity or protein-protein interaction[2, 26].

1.2 Conserved RNA recognition mechanism in LARPs

LARPs have adapted to bind a wide variety of RNA ligands to carry out their versatile functions in cells, while having La-module as the modular RNA-binding unit across most members of the family LARPs. The LaM was shown to adopt an elaborate winged helix-turn-helix fold featuring three α-helices are organized around a three-stranded β-sheet in α1-β1-α2-α3-β2-β3 topology, whereas the RRM1 adopts the typical β1α1β2β3α2β4 fold of canonical RRMs, which usually forms a β-sheet surface directly involved in RNA binding[2, 27]. Despite the fact that LaM and RRM domains fold independently, they work in full synergy to interact with the RNA[5, 28]. The best characterized interactions with structural analysis supports are for the Human La protein and LARP7.

The structures of the La-module of the human La protein in complex with short poly-uridine poly(U) stretches revealed a fixed ‘V-shaped’ conformational clamp that specifically recognizes the UUU-3’-OH target, forming a sequence-specific binding clef at the interdomain interface[8, 29]. Although the LaM side chains mediate most of the specific interactions with RNA,

cooperative protein-RNA and protein-protein interactions involving residues from both domains are required for the recognition of the penultimate uridylate (U-2) – the most sequence-specific determinant of its RNA target[8, 29]. This induced fit of the binding pocket around U-2 accounts for the high synergic nature of this interaction, and the co-crystal structure confirms the requirement for both motifs for RNA binding.

Additionally, this mode of binding was observed in a recent cryo-EM structure of LARP7 in complex with 7SK, where the La-module binds all 8 poly(U) nucleotides, rather than just the expected three terminal U_{330}–U_{332} that form a "V-shape" sandwiched between LaM and RRM1[30]. The 5 additional poly(U) nucleotides form an S-turn that interacts with RRM1, but not LaM, on the β2-β3 loop, thus providing a complete poly(U)-tail-La module interface[30]. These findings offer a comprehensive understanding of the interactions between LARPs and RNA and are considered applicable to the mechanism of other LARP members in their interaction with RNA.

Although the La-module is capable of binding RNA in some cases, it is often found to be in collaboration with the other regions to associate with target RNAs in some LARPs. Recent studies on LARP4 have further highlighted the diversity in RNA recognition, showing that the N-terminal region, rather than the La-module alone, is responsible for primary RNA recognition[31].

1.3 LARP1 on TOP mRNA translation

There are two paralogues of LARP1 genes in humans: LARP1a and LARP1b encoding 1096 and 914 amino acid proteins, respectively. Although there is some similarity between LARP1a and b, LARP1a is the more abundantly expressed and gives the strongest knockdown phenotype[2].

LARP1 was first identified by Natsume and colleagues to be selectively associated with TOP mRNAs in 2013[20], with their studies showing that depletion of LARP1 resulted in reduced level of 5'TOP mRNA. This connection between LARP1 and TOP mRNA was subsequently confirmed by various groups through biochemical and structural analyses.

TOP mRNAs are mRNAs characterized by a 5' terminal oligopyrimidine (5'TOP) motif, play a role in regulating mTOR signalling. 5' TOP motif consists of an invariant cytosine at the +1 position, followed by an uninterrupted series of 4 to 14 pyrimidine nucleotides, situated immediately downstream of the 7-methylguanosine triphosphate (m^7Gppp) cap at the very 5'-end. Currently, there are 97 well established classical TOP mRNAs, which encode nearly all ribosomal proteins, as well as all elongation and some translation initiation factors[32]. The translation of ribosomal protein mRNAs is hypersensitive to nutrient availability and a variety of growth signals including those transmitted by the mTORC1 pathway[32, 33]. Genome-wide ribosome profile studies[34, 35] confirm that mTORC1 play a central role in TOP mRNA translation, and that TOP mRNAs account for the largest class of mRNAs that are translationally-regulated by mTOR. Additionally, mRNA 5′ sequence analyses suggest that thousands more mRNAs may encode TOP sequences[33, 36], although many mTORC1-regulated mRNAs lack classical TOP sequences[33-35]. Whether these mRNAs are nevertheless controlled through the TOP mechanism or through other mTORC1-regulated mechanisms requires further investigation.

The observation that LARP1 is connected to TOP mRNAs through its C-terminal region[20] made by the group of Natsume led to further investigations by others[18, 37] looking for determined connection between LARP1, mTORC1, and ribosomal protein translation. Specifically, Philippe et al. confirmed in vitro that LARP1 functions as a repressor of TOP

mRNA translation and that its DM15 domain binds both the m^7Gppp cap and the adjacent 5' TOP motif of TOP mRNAs[34, 38]. Lahr et al. compared the binding affinity of DM15 domain binding to m^7Gppp cap and 5'TOP motif and observed that replacement of the first cytosine for a guanosine within the 5′TOP motif of ribosomal protein S6 (RPS6) or omission of the m^7Gppp-cap structure considerably reduces the affinity of the DM15 region for the RPS6 5′UTR[39], confirming that the DM15 interacts with both the 5′TOP sequence as well as the m^7Gppp-mRNA cap structure. Crystal structure of DM15 domain bound to the cap analogue 7-methylguanosine triphosphate (m^7GTP) provided direct evidence identifying LARP1 as a cap binding protein. The crystal structure showed that m^7GTP is stabilized by two tyrosines in cation-pi interactions, in an architecture similar to other cap-binding proteins[19, 39].

In addition, results from electrophoretic mobility shift assays (EMSAs) using a recombinant fragment of human LARP1 spanning residues 796–946 of the DM15 domain in combination with 7-methylguanosine triphosphate-capped RNA oligo spanning the entire 5'UTR (42-mer) of RPS6 mRNA revealed that DM15 directly binds m^7Gppp cap with high affinity in low nanomolar range. Interestingly, the DM15 domain of LARP1 demonstrated 90 fold greater affinity for a capped RNA substrate with a cytosine immediately following the cap, confirming that LARP1 preferentially interacts with capped mRNA that contain a cytosine at the +1 position[19, 40]. To further elucidate the molecular mechanism by which LARP1 interacts with TOP mRNA, Lahr et al. determined the crystal structure of DM15 of human LARP1 bound to m^7GpppC. The resulting structure supports the model of specific recognition of the invariant +1 cytosine that characterizes the 5′TOP motif of every TOP mRNA[37, 39]. This result is of particular importance because first cytosine is indispensable for repression of TOP mRNA translation[41].

The crystal structures provide a strong molecular basis by which LARP1 differentiates TOP mRNAs from all other cellular mRNAs.

With LARP1 recognition of mRNA 5' cap and TOP motif being confirmed, more experiments were carried out to understand the mechanism for translation regulation of LARP1 on TOP mRNAs. Cap-dependent translation initiates in the cell cytoplasm when eIF4E recognizes the m^7Gppp mRNA cap structure[42]. eIF4E assumes a cupped hand-shaped architecture whereby the ventral face binds the m^7Gppp moiety and the dorsal face mediates interactions with a number of proteins including the eukaryotic initiation factor 4G (eIF4G) and eIF4E-binding proteins (4E-BPs). Together, the binding of eIF4G to eIF4E stabilizes the association of eIF4E with the mRNA cap[19, 43, 44].

To better understand the mechanism by which LARP1 inhibits TOP mRNA translation, Lahr et al. compared the affinity of recombinant human eIF4E with that of the DM15 domain of human LARP1 for capped m^7Gppp-RPS6 5'UTR by EMSA and noted that LARP1 binds with higher affinity than eIF4E to this TOP mRNA. eIF4E exhibits lower affinity for a 5' cytosine than for 5' guanosine, consistent with LARP1 showing a preference for transcripts beginning with a cytosine and eIF4E a preference for transcripts beginning with a guanosine. In addition, DM15 efficiently displaces eIF4E from m^7Gppp-RPS6 5′UTR, even when added substoichiometrically, while eIF4E cannot dissociate the DM15-m^7Gppp-RPS6 complex, even when added in molar excess[19, 40]. To further test whether LARP1 competes for cap binding with eIF4E in TOP mRNA, Fonseca et al. carried out RNA IP experiments targeting endogenous eIF4G in extracts from human fibroblasts that had been stably depleted of LARP1 by RNAi. The results showed that knocking down LARP1 enhances the co-precipitation of TOP mRNAs with eIF4G[37], suggesting that LARP1 and eIF4G compete for association with TOP mRNAs.

Consistently, transient overexpression of LARP1 in human fibroblasts markedly reduces the co-precipitation of TOP mRNAs with eIF4G[37]. Notably, overexpression of a LARP1 mutant defective for m^7Gppp cap- and 5′TOP motif-binding does not impair the interaction between eIF4G1 and TOP mRNAs[19]. Together, the results suggested that LARP1 represses the translation of TOP mRNAs by competing binding with eIF4E on m^7Gppp cap, and thereby blocking the assembly of the eIF4F complex on the 5'end of TOP mRNA.

1.4 LARP1 regulates TOP translation through mTORC1 pathway

The exact role of LARP1 in mTORC1-mediated regulation of TOP mRNAs remains unclear, and several questions persist.

Fonseca et al. have demonstrated that LARP1 is a target of mTORC1, as evidenced by its association mTORC1-specific component RAPTOR (regulatory-associated protein of mTOR), but not with mTORC2 -specific component RICTOR (Rapamycin-insensitive companion of mTOR). Moreover, LARP1 acts as a repressor of TOP mRNA translation and its association with 5′TOP motif is essential for blocking TOP mRNA translation[37, 45].

mTORC1 elicits its effects on mRNA translation through the phosphorylation of multiple targets. LARP1 interacts with mTORC1[37] and phosphoproteomic studies suggest that it gets phosphorylated by mTOR kinase on multiple serine, threonine and tyrosine residues[40]. Notably, these phosphorylation sites are located just N-terminal to the LARP1-specific DM15 domain that mediates binding to the m^7Gppp cap and the 5′TOP motif. This observation suggests that mTORC1 may regulate LARP1 association with the 5′ end of TOP mRNAs through multisite phosphorylation[19, 39]. Work by Hong et al. has demonstrated that mTORC1-mediated phosphorylation of LARP1 dissociates it from 5'UTRs and relieves its inhibitory activity on ribosomal protein mRNA translation, indicating its role as a phospho-dependent molecular

switch that regulates TOP mRNAs in an mTORC1-dependent manner[46]. In this proposed model, eIF4F is destabilized via 4E-BPs and dephosphorylated LARP1 binds the 5′ ends of TOP mRNAs when mTORC1 is inactive, repressing their translation and promoting their stability by maintaining the length of poly(A) tails. Upon mTORC1 activation, LARP1 phosphorylation relieves its inhibitory activity, allowing eIF4F to bind TOP mRNAs and resume their translation[45]. The exact significance of these phosphorylation sites in the context of TOP mRNA remains unclear, and further research is required to elucidate the role of mTORC1-mediated phosphorylation of LARP1 in LARP1's binding to TOP mRNAs and its function in translational regulation of TOP mRNAs[40, 45].

1.5 Human LARP1 has oncogenic properties

Several independent studies reported that human LARP1 expression is mis-regulated in cancer cell lines and malignant tissues, revealing its oncogenic role.

Mura et al. observed that LARP1 is over-expressed in the majority of epithelial malignancies compared to their adjacent normal tissues in a study involving in-silico analysis of database[47, 48]. A similar study conducted by the same group on human cervical squamous carcinoma cells showed increased human LARP1 mRNA and protein levels as compared to non-cancerous cervical tissues[48].

A survey of 15 hepatocellular cancer cell lines showed that both human LARP1 mRNA and protein are significantly upregulated in malignant cells as compared to an immortalized normal liver epithelial cell line[49]. Moreover, when comparing protein levels by western blotting and immunohistochemistry between hepatocellular cancer cell and adjacent non-cancerous tissues, a marked increase in human LARP1 level was observed in hepatocellular cancer cell lesions[49]. In patients, high levels of LARP1 protein in tumour tissue correlate with approximately

35% increased risk of death by five years and with tumour size, survival time and Child-Pugh score[49]. Levels of LARP1 have also been noted to be high in prostate and breast cancer. In prostate cancer cells, tumour migration is attenuated upon LARP1 knockdown and LARP1 expression is negatively regulated by microRNAs mi-26a or b[50]. In breast cancer, LARP1 protein and mRNAs were found to be over-accumulate[47]. RNA-sequencing and high throughput software analyses of tissue samples revealed the presence of a novel LARP1 splice variation in 4/6 non-triple negative cases. This variant makes human LARP1 a putative useful biomarker for malignant tissues[47].

Work by S. Blagen's team suggested that human LARP1 also contributes to cancer progression[51, 52], including cell migration, invasion, EMT and tumorigenesis. They show that siRNA-mediated depletion of HsLARP1 significantly decreases cell migration and invasion abilities, while stable overexpression of HsLARP1 has opposite effects. Consistent with a positive role in cell migration, human LARP1 promotes lamellipodia formation and is more concentrated in lamellipodia at the leading edge of migrating cells[51]. Transwell Matrigel invasion assays showed that LARP1 overexpression significantly enhanced the invasive capabilities of HeLa cells[48]. A more pronounced effect was seen after LARP1 knockdown, with an 85% reduction in cell invasion. Similar effect was observed in Invasion assays performed in PC9 cells after LARP1 knockdown, showing a significant decrease in the number of invasive cells[48]. Additional in vivo analyses suggest that human LARP1 increases tumorigenicity: mice injected with HeLa cells overexpressing human LARP1 show an increased tumor progression as compared to those injected with non-transgenic HeLa cells[48]. This implies that LARP1 has a distinct interactome in cancer compared to non-malignant cells that could result from conformational changes to LARP1 driven by upstream signalling events, partner protein

interactions or associations with other RBPs, micro or non-coding RNAs[53]. Altogether these data support that the human LARP1 protein has an oncogenic potential and supports the physiological relevance and importance to study the mechanism and structure of human LARP1.

1.6 mRNA 3'-end processing

Poly(A) tails are present on nearly all eukaryotic mRNAs and are added during transcription. They are necessary for exporting mature mRNAs to the cytoplasm and function as key regulators of gene expression. The poly(A) tail impacts the stability and translational status of mRNAs and works synergistically with the 7-methylguanosine cap on the 5'-end of the mRNA to stimulate translation[54]. PABPC, which directly binds to poly(A) tails, has four N-terminal RRM domains that can bind poly(A) RNA. PABPC also has a proline-rich linker and a C-terminal MLLE domain, which recognizes PAM2, a domain found in many PABPC partner proteins that regulate poly(A) tail dynamics. These partner proteins contribute to regulating deadenylation. For example, the PAM2-containing protein TOB2 can be phosphorylated, which modifies its interaction with PABPC and promotes deadenylation by interacting with CCR4-NOT[55].

The length of poly(A) tails is regulated by exonucleases involved in deadenylation activities, such as the PAN2-PAN3 and CCR4-NOT complexes[56]. PABPC and deadenylases interact to control deadenylation, and many regulators have PAM2 motifs that allow direct interaction with the MLLE domain of PABPC. LARP1 binds directly to poly(A) RNA and PABPC1 through a PAM2 motif, protecting mRNAs from deadenylation and possibly stabilizing PABPC binding[4].

Recent sequencing data has provided a strong link between translation efficiency and deadenylation, reinforcing the correlation between mRNA stability and translation efficiency.

Highly translated RNAs have relatively short poly(A) tails of around 30 nucleotides, accommodating a single PABPC[57]. Longer poly(A) tails have been observed to be efficiently translated, challenging the correlation between longer tails and increased mRNA stability[58]. The insertion of ribonucleotides other than adenosine also impedes deadenylation and promotes mRNA stability, with guanosine being the least efficiently removed by deadenylation machinery[59].

1.7 LARP1-Poly(A)-binding protein (PABP) complex

The presence of a mademoiselle (MLLE) domain in PABP allows for docking of various PAM2-containing proteins[60]. The crystal structure of MLLE domain from PABP alone and in complex with PAM2 peptides from PABP-interacting protein 2 (Paip2)[61], and with GW182 has been determined[62, 63]. Sequence analysis identified a PAM2-like sequence in LARP1 that could also bind PABP. However, this sequence was very unusual, consisting of only eleven residues and lacking the invariant alanine at position -7[37]. Furthermore, a gap was required to align the critical Leucine at position 3 and phenylalanine at position 10[37]. Despite the differences, mutagenesis of the phenylalanine aligned with the PAM2 consensus position-10 resulted in decreased amount of PABP that coimmunoprecipitated with LARP1[64]. Nevertheless, this mutation does not rule out the possibility of atypical/asymmetrical PAM2 in LARP1. A recent study identified an N-terminal Phe missing in the original identification and characterized the PAM2 of LARP1. The study demonstrated the comparable affinity of a LARP1-derived peptide to other PAM2 peptides and emphasized the importance of key residues in the LARP1 PAM2 in cellular assays of mRNA 3′ poly(A) tail protection and stabilization by full length LARP1. PAM2 mutation that impaired co-immunoprecipitation of cellular PABP with the LARP1 La-

module, which also impaired mRNA poly(A) protection-stabilization, did not affect oligo(A) binding by the recombinant, purified La-module[64].

1.8 mRNA poly(A) tail protection and stabilization activities for LARP1

The mRNA mobility assays used to determine mRNA poly(A) length protection activity showed positive results for LARPs 1, 4 and 4B while LARPs, 6, 7 and Human La protein were negative[22]. Subsequent analysis identified a 304 amino acid La-module fragment from LARP1 that exhibited poly(A) protection and mRNA stabilization activity, and was demonstrated bind PABP through co-immunoprecipitation[64]. Another 230 amino acid La-module fragment with different terminal extensions was shown to bind poly(A) as a recombinant protein[65]. These two LARP1 La-module fragments exhibited different levels of poly(A) length protection and mRNA stabilization activities in HEK293 cells[64], and these activities were decreased upon point mutations to PAM2. In assays using stable GFP mRNA and on β-globin-ARE unstable reporter mRNA, LARP1 exhibits mRNA stabilization activity, with LARP4 serving as a positive control[64]. Recent findings from Park and colleagues demonstrated that LARP1 knockdown leads to a global reduction of messenger RNA abundance by interfering with the CCR4-NOT-mediated deadenylation in vitro, and this was shown to occur through the formation of a ternary complex with PABP and poly(A)[66]. Furthermore, depletion of LARP1 results in accelerated deadenylation specifically in the 30–60 nucleotides range[66].

1.9 LARP1 function in pathways other than for TOPs

LARP1 has been linked to mitochondrial DNA replication during oogenesis in drosophila[52]. It facilitates the translation of mRNA localized to the mitochondrial outer surface. PINK kinase phosphorylation of drosophila LARP1 inhibits the translation of these outer mitochondrial-located mRNAs[52]. In humans, LARP1 is also crucial for inner-mitochondrial

translation necessary for oxidative phosphorylation, thereby expanding the significance of drosophila link to energy production and mTOR1 metabolic control[4]. LARP1 homologs in yeast, Sro9p, Slf1p lack the DM15 domain. Sro9p and Slf1p associates with overlapping sets of mRNAs that include ribosomal proteins mRNAs, but slf1p has greater target mRNA accumulatio due to distinct stabilization activity[4, 36]. Overall these observations suggest conservation of LARP1 activities across organisms.

1.10 Project goals

Although La-modules in other LARP family members have been characterized structurally and functionally, the LARP1 La-module has not received enough attention. While the DM15 domain of LARP1 has been structurally characterized, the structure and molecular mechanism of N-terminal half are not well understood. Although cell-based assays have suggested a role for LARP1 La-module in stabilizing mRNA poly(A) tail, the specific mechanism remained underlying this process remains unclear. The goal of this project is to closes this gap and study the detailed molecular basis of RNA binding by the LaM domain present in the LARP1 La-module. To do this, we expressed and purified LARP1 La-Module in E. coli. We used ITC to identify the affinity of different fragments of LARP1 La-module interacting with a variety of RNA targets. We also used NMR to look at the interaction in solution. Crystallization of the LARP1 LaM and in complex with RNAs was performed to reveal the molecular mechanism of their interaction. Altogether, the results of this project will provide structural and molecular evidence to support the important role of LARP1 in mRNA poly(A) tail processing and stabilization.

CHAPTER 2. MATERIALS AND METHODS

2.1 Construct design

Plasmids for bacterial expression of the LARP1 fragments 323–410, 323–417, 323–439, 323–509, 417–509 and 399–540 were obtained by mutagenesis of 310–540 fragment cloned into pET28a vector introducing deletions and stop codons at the appropriate positions using QuikChange Lightning Site-Directed Mutagenesis Kit (Thermo Fisher Scientific). LARP1 point mutants were obtained by site-directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit (Thermo Fisher Scientific). DNA sequencing was used to verify all sequence modifications.

2.3 RNA oligonucleotides

A_{25}, A_{11}, and A_6 were obtained from Sigma-Aldrich. U_6 was ordered at Integrated DNA Technologies and used without additional purification. A_3, A_4, A_5G, A_4GA, A_3GA_2, and additional samples of A_6 and A_{25} were synthesized on 2 × 2 μmol scale using an ABI 3400 synthesizer with standard β-cyanoethylphosphoramidite chemistry on long chain alkylamine controlled pore glass (LCAA-CPG, 500 Å) with standard synthesis protocols. Oligonucleotides were purified either by preparatory denaturing PAGE or ion-exchange HPLC and desalted using C-18 SEP PAK cartridges as previously published.

2.4 Expression of proteins

100 ml starter cultures were grown overnight at 37 ˚C and used to inoculate 1L cultures of LB medium containing 50 μg/mL kanamycin. The cells were grown at 37˚C in a shaking incubator rotating at 200 RPM until an OD600nm of ~0.8 was reached, at which point the temperature was reduced to 30˚C. For NMR experiments, the recombinant protein was isotopically labeled by

growth of E. coli BL21 in M9 minimal medium with 15N-ammonium sulfate as the sole source of nitrogen. Protein expression was induced by the addition of isopropyl β-d-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The cells were harvested after 4 hours incubation by centrifugation at 4500 rpm for 20 min at 4˚C, resuspended in lysis buffer (50 mM HEPES pH 7.6, 500 mM NaCl, 5% glycerol, 1 mM PMSF, 5 mM β-mercaptoethanol, 10 µg/mL DNase I, 0.1mg/L lysozyme, 5mM imidazole) and lysed by sonication in a Fisher Scientific Sonic Dismembrator 5000 (2-3 cycles of 1:10 min at 55% amplitude; ON:10s; OFF: 20s). The cell lysate was clarified by centrifugation at 20000 RPM for 45 minutes at 4˚C.The supernatant was then passed through Whatman filter paper with a pore size of 0.45 µm to remove residual cellular debris prior to binding to resin.

2.5 Purification of proteins

The filtered lysate containing the 6xHis-fusion proteins were incubated with Ni^{2+}-charged chelating sepharose affinity resin for 30 minutes at 4˚C. The suspension was then passed through a column and washed five times with 2 CV of wash buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 5 mM β-mercaptoethanol, 30 mM imidazole). The protein was eluted from the column with elution buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 5% glycerol, 5 mM β-ME, 500 mM imidazole). The eluate was concentrated to ≤ 5 mL and additionally purified using size-exclusion chromatogrpahgy using a HiLoad 16/600 Superdex 75 PG column (Cytiva) with HPLC buffer (10mM MES pH 6.3, 100mM NaCl, 1mM TCEP).

2.6 Isothermal titration calorimetry (ITC)

ITC experiments were performed on a MicroCal iTC200 and VP-ITC titration calorimeter (Malvern Instruments Ltd). The syringe contained 300 µM of proteins, while the sample cell

contained 30 µM RNA. All experiments were carried out at 293 K with 19 injections of 2 µl with stirring at 310 rpm on iTC200 or 29 injections of 10 µl on VP-ITC.

Results were analyzed using ORIGIN software (MicroCal) and fitted to a binding model with a single set of identical sites.

2.7 NMR spectroscopy

NMR samples were exchanged in 10 mM MES pH 6.3, 100 mM NaCl, 1 mM TCEP. For NMR titrations, A_2 and A_6 RNAs were added to ^{15}N-labeled LARP1 fragments to the final molar ratios of 1:2 or 1:1, respectively. Chemical shift perturbations were calculated as the weighted sum of proton and nitrogen shifts using the equation $(\Delta\delta H^2 + (\Delta\delta N/5)^2)^{1/2}$. All NMR experiments were performed at 25 °C using Bruker 600 MHz spectrometer. NMR spectra were processed using NMRPipe and analyzed with SPARKY.

2.8 Crystallization

Initial crystallization conditions were identified utilizing sitting drop vapor diffusion with the Classics II and Nucleix screens (QIAGEN). The best LARP1 LaM domain crystals were obtained by equilibrating a 0.6 µl drop of the protein (residues 323–410) at 20 mg/ml in HPLC buffer (10 mM MES pH 6.3, 100 mM NaCl, 1 mM TCEP), mixed with 0.6 µl of reservoir solution containing 0.2 M ammonium sulfate, 0.1 M Bis–Tris pH 6.5, 25% (w/v) PEG 3350. Crystals grew in 30–40 days at 20ºC. The best LARP1 LaM domain/RNA complex crystals were obtained by equilibrating a 0.6 µl drop of the LaM domain (residues 323–410) with oligonucleotide in a 1:1.1 molar ratio (10 mg/ml of protein) in buffer (10 mM MES pH 6.3, 100 mM NaCl, 1 mM TCEP), mixed with 0.6 µl of reservoir solution containing [0.056 M sodium phosphate, 1.344 M potassium phosphate] or [0.2 M sodium chloride, 0.1 M Bis–Tris pH 5.5, 25% (w/v) PEG3350] for A_3, [0.056 M sodium phosphate, 1.344 M potassium phosphate] for A_4,

[0.1 M HEPES pH 7.5, 2 M ammonium sulfate] or [0.1 M Bis–Tris pH 5.5, 25% (w/v) PEG3350] for A_6, [0.1 M BICINE pH 9.0, 1.6 M ammonium sulfate] for A_{11}, [0.1 M HEPES pH 7.5, 25% (w/v) PEG 3350] for A_3UA_2. Crystals grew in 3–14 days at 20ºC. For data collection, crystals were cryo-protected by soaking in the reservoir solution supplemented with 30% (v/v) ethylene glycol in conditions using PEG3350 or with 25% glycerol otherwise.

2.9 Structure solution and refinement

Diffraction data from single crystals of LARP1 LaM domain and its RNA complexes were collected at the Canadian Light Source (CLS), Cornell High-Energy Synchrotron Source (CHESS) and Advanced Photon Source (APS). Data processing and scaling were performed with HKL2000. The initial phases for the complex structure were determined by molecular replacement with Phaser[67], using the coordinates of the LARP3 LaM domain (PDB entry 1ZH5)[8]. The initial phases were improved by Autobuilder in PHENIX package[68]. The starting protein model was then completed and adjusted with the program Coot[69] and improved by several cycles of refinement, using the program phenix.refine[68] and model refitting. The resulting electron density maps revealed clear density for RNA oligonucleotide, which was manually built with the program Coot[69]. The final protein model was then used for phasing of data for RNA-free LARP1 LaM domain. At the latest stage of refinement for both structures, we also applied the translation-libration-screw (TLS) option[70]. The final models have all residues in the allowed regions of Ramachandran plot. The coordinates have been deposited with the Protein Data Bank (PDB).

CHAPTER 3. RESULTS

3.1 Sequence alignment and secondary structure prediction

The La motif (LaM) domain is highly conserved among La-related proteins (LARPs), as evidenced by the high sequence similarity between the LaM domains of LARP1 and LARP4 (Figure 2). However, the downstream RNA recognition motif (RRM) region shows much lower sequence conservation. In particular, LARP1 contains a PAM2 motif immediately following its LaM domain, which is absent in LARP4. Phylogenetic analyses involving 134 sequences from 29 eukaryotic species have revealed that LARP1 is an outlier in that 40% of LARP1 genes lacked a predicted RRM domain. Interestingly, LaM-containing proteins lacking an canonical RRM domain in yeast as well LARP1 homologs in *Arabidopsis thaliana, Caenorhabditis elegans, Danio rerio, Drosophila melanogaster, Escherichia coli, Homo sapiens, Mus musculus, Rattus norvegicus and Saccharomyces cerevisiae, Danio rerio, Drosophila melanogaster* and *Mus musculus, Schizosaccharomyces pombe elegans* and other metazoan[2]. Secondary structure prediction for the LaM and RRM regions of LARP4 using JPred4 is in agreement with the experimentally determined structure, whereas the RRM region of LARP1 does not have predicted secondary structure elements.

Figure 2

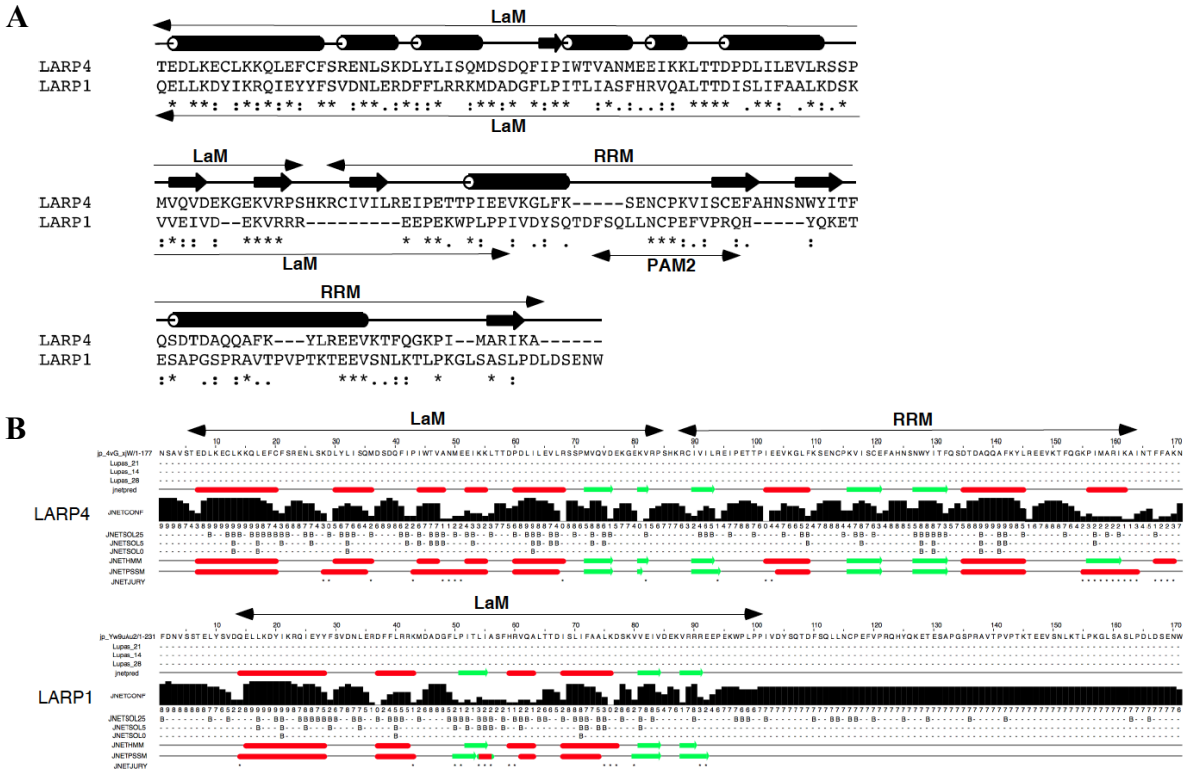


Figure 2. Comparison of La Module regions in LARP1 and LARP4.

- **A.** Sequence alignment of LARP4 La module with the corresponding region in LARP1
- **B.** JPred4 secondary structure prediction for LaM and RRM regions of LARP4 and LARP1. Red bars are alpha helices and green arrows are beta sheets.

3.2 LARP1 does not possess a classic tandem LaM-RRM structure

We analyzed NMR spectroscopy of LARP1 fragments of different lengths. The $^{1}H–^{15}N$ correlation spectrum of residues 323–509 which covers the LaM-RRM (numbered according to the 1019-residue long isoform) exhibited a mixture of weaker, dispersed signals and a central band of intense signals. This characteristic is consistent with the presence of a mix of ordered and disordered residues with the well-dispersed low-amplitude signals arising from a folded domain and the high-amplitude strong signals in the middle coming from unstructured residues. This fragment contains two tryptophans: Trp406 in the predicted LaM and Trp480 from the putative RRM. The $^{1}H–^{15}N$ correlation spectrum of residues 323–410 corresponding to only the

LaM domain showed good dispersion of signals characteristic of a well-folded domain. The spectrum is in agreement with the weaker signals in the (323–509) NMR spectrum including one Trp406 indole proton. In contrast, the $^{1}H–^{15}N$ correlation spectrum of residues–509 representing the putative RRM alone, produced a spectrum with poorly dispersed signals that matched the unfolded set of signals in the 323–509 fragment (Figure 3). The spectrum of constructs comprising LaM or putative RRM confirms the presence of only one folded structure of LaM domain and an absence of folded RRM in the predicted La-module. Larger constructs terminating at residue 540 retained the characteristics of an unfolded protein, which eliminated the possibility that the predicted RRM has been prematurely truncated and the residues required for proper folding were removed from the construct design. Furthermore, the addition of A_{25} does not alter the spectrum of the putative RRM (417-509) or (399-540) fragment, ruling out the possibility that the RRM folds or changes conformation upon accommodation of RNA (Figure 3).

Figure 3

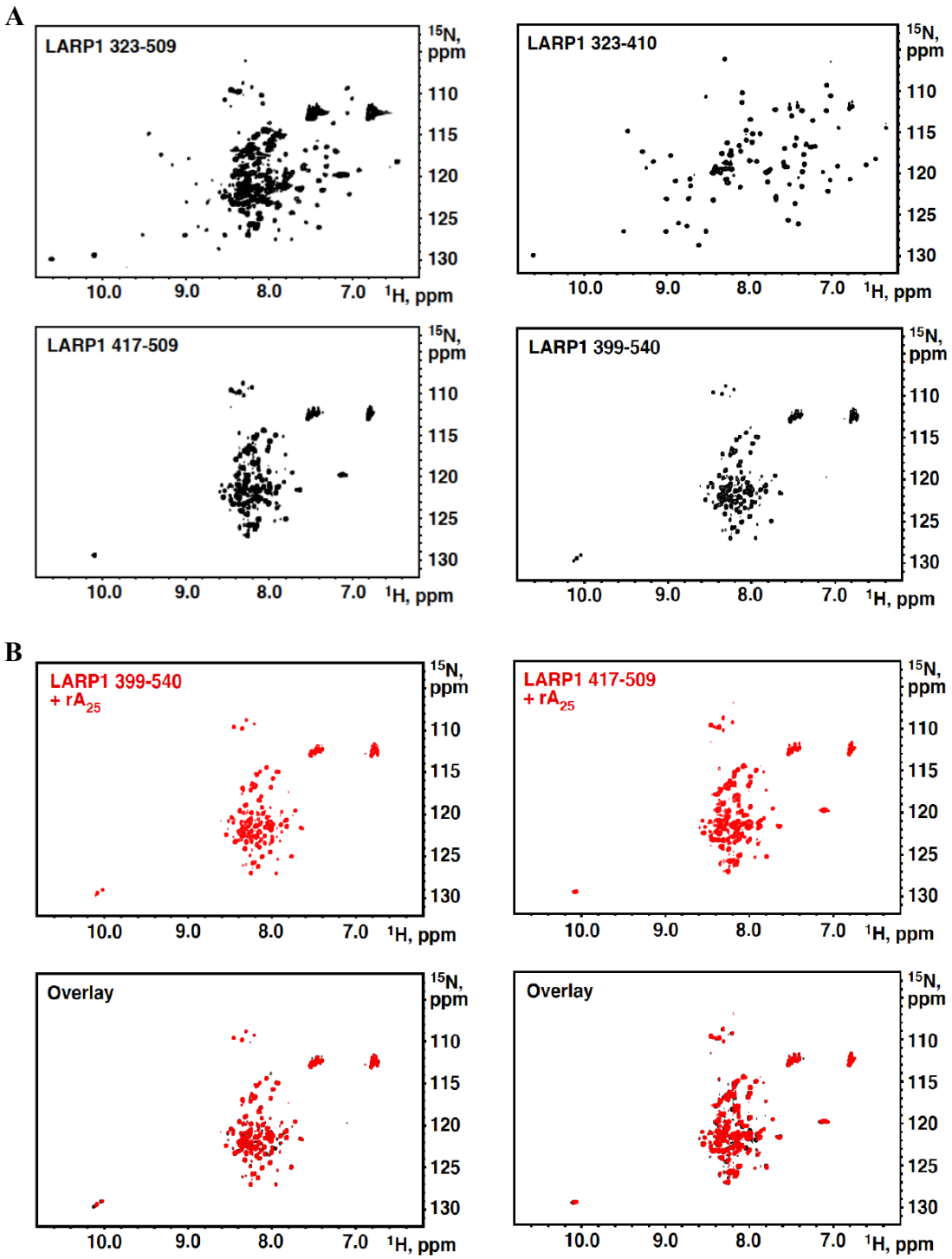


Figure 3. LARP1 does not contain a folded RRM domain.

A. ^{15}N-^{1}H correlation spectrum of ^{15}N-labeled LARP1 (323-509) shows a presence of a folded domain characterized by well-dispersed low-amplitude signals. The ^{15}N-labeled LARP1 (323-410) fragment corresponds to the folded La motif (LaM) domain and the

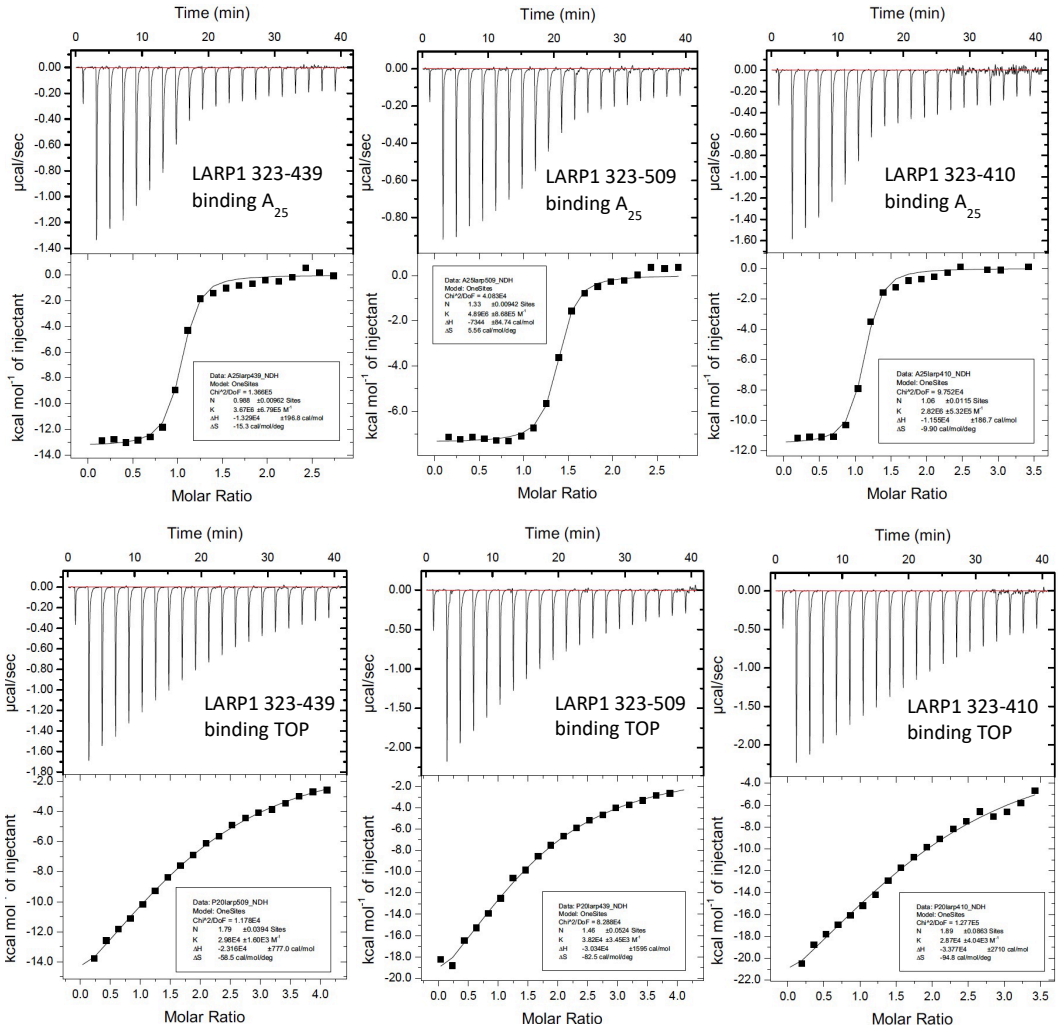
downstream region (residues 417-509) corresponds to the unfolded part of 323-509 fragment demonstrating an absence of structured RRM domain.

B. No change in the downstream region (residues 417-509) in addition to RNA indicated that the unstructured RRM does not fold upon binding to RNA.

3.3 The LaM alone is sufficient for RNA binding

To confirm that the unstructured RRM does not contribute to RNA binding in the context of intact LARP1 La-module, we used isothermal titration calorimetry (ITC) to determine the affinity of RNA binding for LARP1 constructs that were previously analyzed in NMR spectroscopy. ITC thermograms with LARP1 LaM-RRM (323–509) and A_{25} allowed us to determine K_d of ~0.2 μM, indicative of strong binding affinity with a stoichiometry close to 1:1. Comparable affinities were observed for the smaller fragment, LARP1 (323–439) that encompasses the LaM domain and the PAM2 motif, and LARP1 (residues 323–410) comprising only the LaM domain without PAM2. These results are in line with our NMR data, confirming that LARP1 LaM domain alone is sufficient for RNA 3'-end poly(A) binding with submicromolar affinity, thereby suggesting a distinct RNA recognition mechanism compared to the typical La protein and LARP7 where both LaM and RRM are required for recognition of UUU-3'OH[71]. Subsequently, we determined the binding affinity of LARP1 fragments for a previously characterized TOP RNA sequence. The intact LaM-RRM 323-509 fragment exhibited weaker binding affinity for the TOP RNA, with a K_d of 35 μM, which is approximately 150-fold reduced than the affinity observed for A_{25} binding. The binding affinity of the LaM-PAM2 (323-439) and LaM (323-410) fragments were similar to that of the intact LaM-RRM fragment. These data prompted us to further investigate the affinity and mechanism by which LARP1 LaM binds to poly(A) RNA. The ITC thermograms are presented in Figure 4.

Figure 4



RNA	Cell, µM	Syringe, µM	Chi-Square/DofF	N	Ka (M-1)	Error (M-1)	ΔH (cal/mol)	ΔS (cal/mol/deg)	Kd (nM)	error(nM)
					LARP1 (323-410)					
A_{25}	30 (RNA)	300 (protein)	97520	1.06±0.0115	2820000	532000	-11550±186.7	-9.9	355	57
TOP	30 (RNA)	300 (protein)	127700	1.89±00.0863	28700	4040	-33770±2710	-94.8	~35000	4000
					LARP1 (323-439)					
A_{25}	30 (RNA)	300 (protein)	136600	0.988±0.00962	3670000	679000	-13290±196.8	-15.3	272	42
TOP	30 (RNA)	300 (protein)	82880	1.46±0.0524	38200	3450	-30340±1595	-82.5	26178	2000
					LARP1 (323-509)					
A_{25}	30 (RNA)	300 (protein)	40830	1.33±0.00942	4890000	868000	-7344±84.74	5.56	204	40
TOP	30 (RNA)	300 (protein)	11780	1.79±0.0394	29800	1600	-23160±777	-58.8	35570	3750

Figure 4. Isothermal titration calorimetry of RNA binding to LARP1 LaM domain.
ITC thermograms of LARP1 constructs binding to A_{25}. RRM does not contribute to LARP1 binding to A_{25} or TOP RNA. The LaM domain binds A_{25} with high nanomolar affinity but micromolar affinity for a 20-mer TOP RNA.

3.4 LaM requires free 3’-end nucleotide for RNA recognition

The observation of a 1:1 stoichiometry observed with A_{25} suggested that there is a single binding site for RNA ligands on the LaM domain. To investigate whether the binding is mediated through 3’end of RNA oligonucleotides, we conducted some ITC experiments. Our results show that addition of a 3’ phosphate on A_6 led to a 200-fold loss of affinity, while 2′-O methylation of the 3′ nucleotide of A_6 completely blocked binding (Figure 5). These findings were validated by NMR titrations of single nucleotides into ^{15}N-labeled LARP1 LaM domain. The effects of 3’-AMP and 3’, 5'-ADP addition were smaller and restricted to just a few residues, in contrast to the significant and obvious peak shifts observed in the spectrum with 5'-AMP titration (Figure 6). The binding affinity of LARP1 LaM was found to be highly sensitive to the modification state and presence of 3’-end nucleotides, indicating a specific recognition mechanism for the 3’nucleotide of poly(A), which is similar to the mechanism employed by human La protein to 3’-end nucleotides of poly(U)[8, 29, 72].

Figure 5

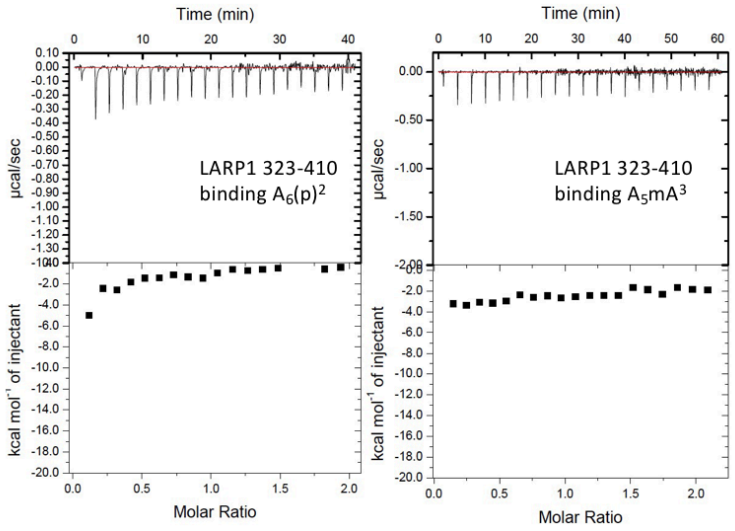


Figure 5. Isothermal titration calorimetry of RNA binding to LARP1 LaM domain. Affinities of 3’-end modified RNA confirms specificity of the LaM domain for the 3′ end of poly(A) RNA.

Figure 6

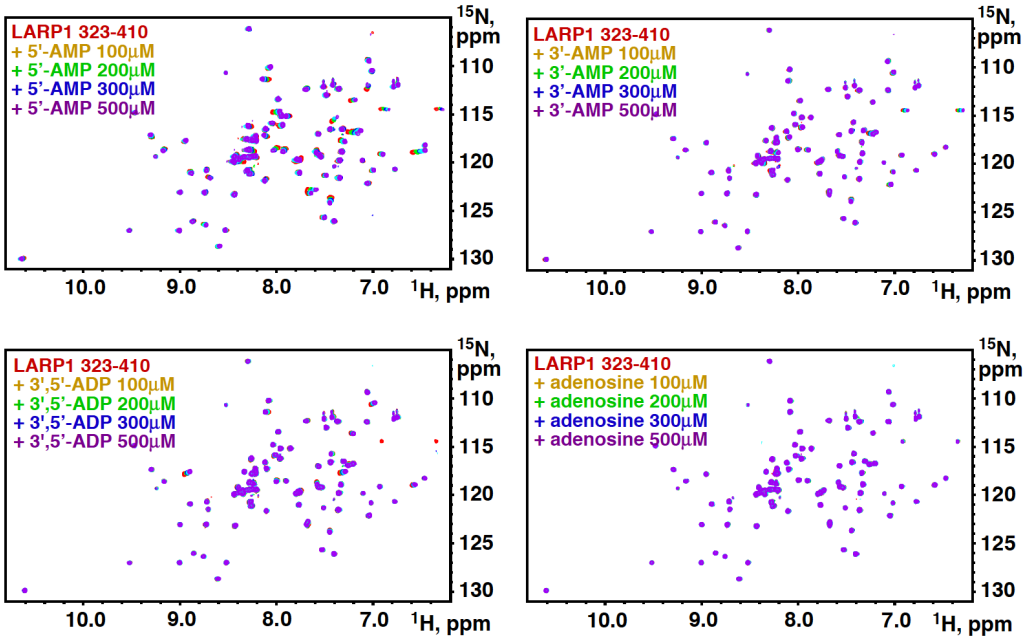
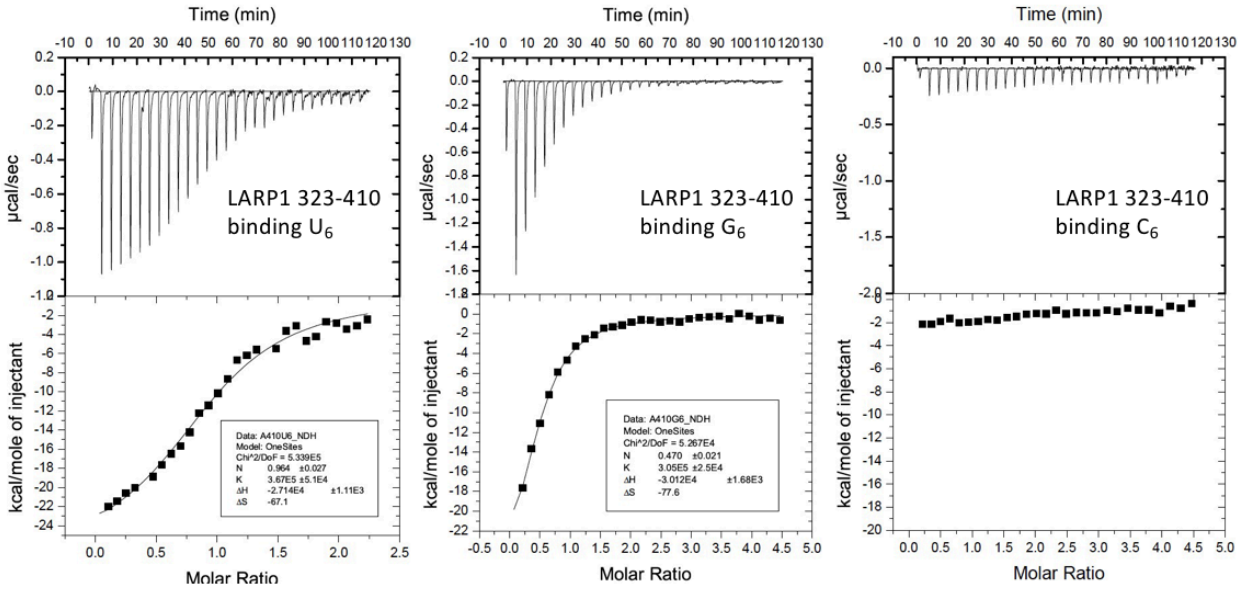


Figure 6. NMR of single nucleotide binding to LARP1 LaM domain.
Overlay of ^{15}N-^{1}H correlation spectra of ^{15}N-labeled LARP1 (323-410) upon addition of increasing nucleotide concentrations. Smaller and less peak shifts when LARP1 LaM was titrated with 3'-AMP and 3', 5'-ADP in comparison to 5'-AMP indicating weaker binding affinity for these RNA ligands with 3'-end modifications. LaM does not exhibit binding to adenosine.

3.5 Specificity for poly(A) RNA binding by the LARP1 LaM domain

We conducted additional ITC experiments to further investigate ligand binding specificity of the LARP1 LaM domain. Our results suggested that LARP1 displays nucleotide base discrimination with regard to RNA binding. Specifically, we found that LaM domain binds U_6 and C_6 RNA with a K_d of 1.9 μM and 4.7 μM, respectively. Notably, these affinities are approximate 10-fold lower than that observed for A_6. LaM binds C_6 RNA with an affinity weaker than 50 μM, which represents a more than 200-fold reduction in affinity relative to A_6 (Figure 7).

Figure 7

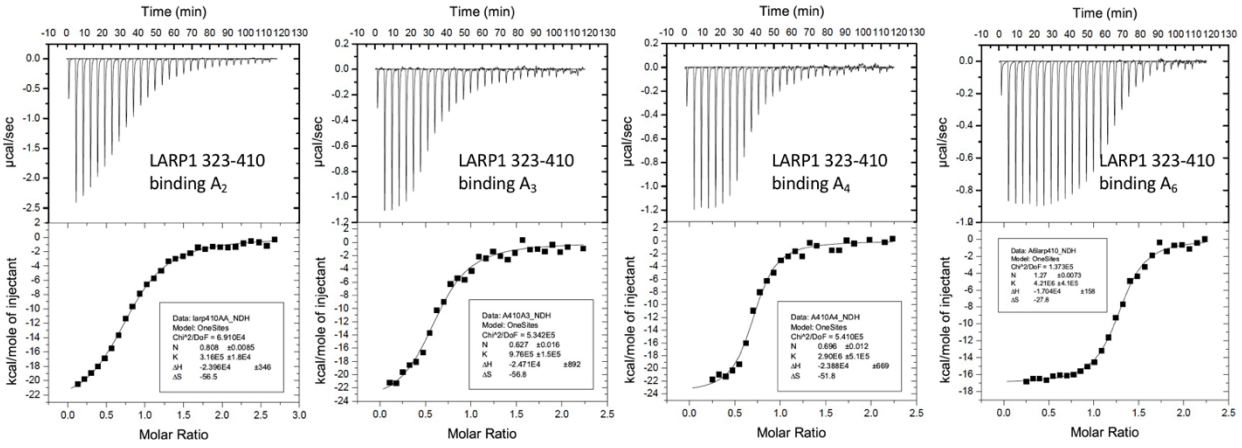


RNA	Cell, µM	Syringe, µM	Chi-Square/DofF	N	Ka (M-1)	Error (M-1)	ΔH (cal/mol)	ΔS (cal/mol/deg)	Kd (nM)	rel. affinity to A_6	error(nM)
	LARP1 (323-410)										
A_6	15 (RNA)	150 (protein)	137300	1.27±0.0073	4210000	410000	-17040±158	-27.8	240	1	24
U_6	15 (RNA)	150 (protein)	533900	0.946±0.027	367000	51000	-27140±1110	-67.1	2725	11.3	330
G_6	15 (RNA)	150 (protein)	52670	0.470±0.021	305000	25000	-30120±1680	-77.6	3279	13.8	249
C_6	15 (RNA)	300 (protein)	NB	NB	NB	NB	NB	NB	NB	NB	NB

Figure 7. Isothermal titration calorimetry of LaM binding to A_6, U_6, G_6 and C_6
LARP1 LaM had 10-fold lower affinity for U_6 and G_6 RNA binding relativeA_6. LaM does not prefer cytosine as C_6 RNA showed more than a 200-fold lower affinity relative to A_6

Our finding indicated that the length of oligonucleotides did not result in a significant change in binding affinity. While A_2 is too short for sufficient binding, three to four nucleotides were adequate and did not result in significant reduction in binding affinities compared to A_6, A_{11}, or A_{25} (Figure 8). The variations shown in their binding affinities are due to some precipitation of LARP1 LaM at high concentration during ITC experiments at 25°C. Taken together, LARP1-LaM show a preference for 3'-end RNA with A-rich sequences.

Figure 8



RNA	Cell, µM	Syringe, µM	Chi-Square/DofF	N	Ka (M-1)	Error (M-1)	ΔH (cal/mol)	ΔS (cal/mol/deg)	Kd (nM)	Rel. affinity to A_6	error(nM)
	LARP1 (323-410)										
A_2	30 (protein)	300 (RNA)	69100	0.808±0.0085	316000	18000.000	-23960±346	-56.5	3165	13.3	171
A_3	15 (RNA)	150 (protein)	534200	0.627±0.016	976000	150000	-24710±892	-56.8	1024	4.3	136
A_4	15 (RNA)	150 (protein)	541000	0.696±0.012	2900000	510000	-23880±669	-51.8	345	1.4	50
A_6	15 (RNA)	150 (protein)	137300	1.27±0.0073	4210000	410000	-17040±158	-27.8	240	1	24

Figure 8. Isothermal titration calorimetry of LaM binding to A_2, A_3, A_4 and A_6
Dinucleotide A_2 binds with10-fold weaker affinity than A_4 A_6, A_{11}, or A_{25}. LARP1-LaM does not require long stretch of oligonucleotide to exhibit high affinity bindings.

3.6 NMR analysis of RNA binding to LaM

We carried out two titrations acquiring NMR spectra of the LaM domain in the presence of increasing concentrations from of the dinucleotide A_2 and the hexamer A_6 to look at the dynamics in ligand binding (Figure 9). Both oligonucleotides caused peak shifts in the $^{1}H–^{15}N$ correlation spectra confirming binding observed previously in ITC. The shifts were generally larger for A_6, consistent with higher affinity binding; the most significant difference was in the increased dynamics of binding observed with A_2. Intermediate titration point with A_2 showed fast-exchange between the free and bound states with peaks sliding between the initial and final positions. Several peaks showed exchange broadening and disappeared during the titration. In contrast, the signals corresponding to the free and bound states in the A_6 titration with were in slow exchange as is typical for strong, sub-micromolar binding. The higher affinity of A_6 was

also apparent in the amount of RNA required to reach saturation. A_2 required over-titration to a 1:2 ratio of protein to RNA, while 1:1 was sufficient for A_6.

Figure 9

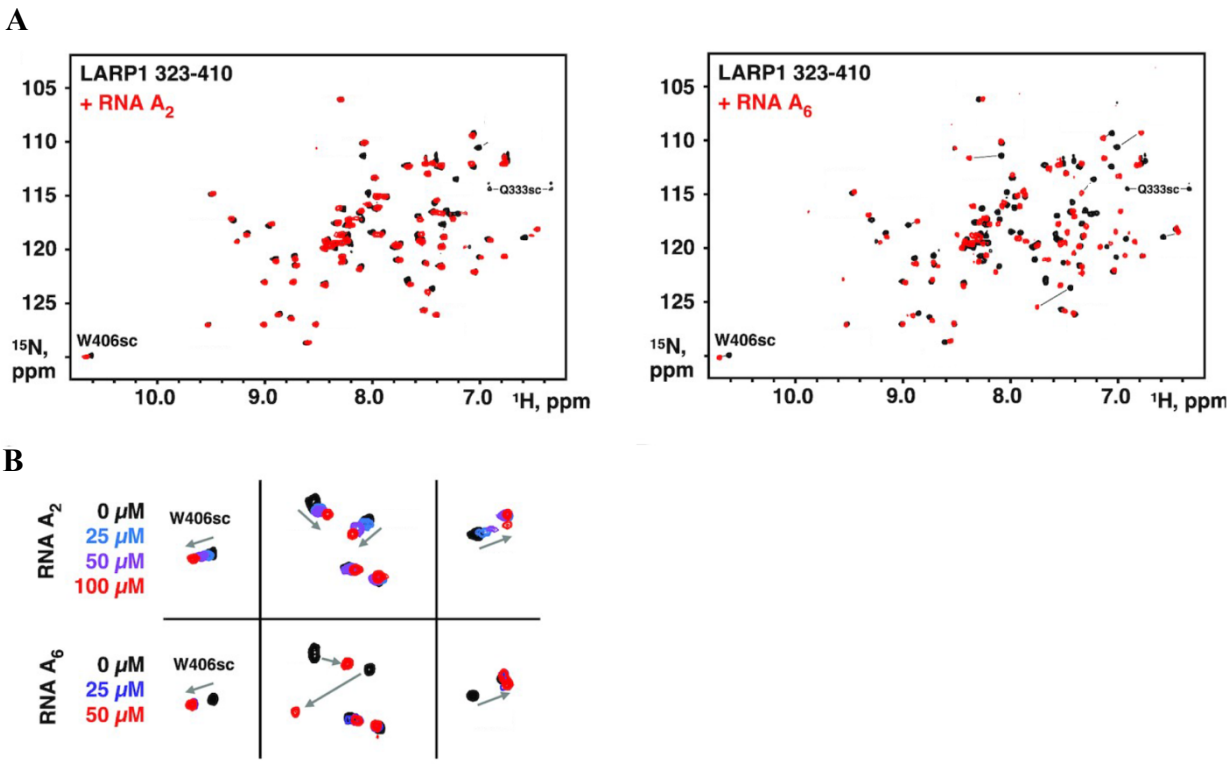


Figure 9. NMR of A_2 and A_6 binding.

- **A.** Spectra overlay of 50 µM ^{15}N-labeled LARP1 (323-410) alone (*black*) and in the presence of 100 µM A_2 or A_6 (*red*).
- **B.** Comparison of selected peak shifts showing fast-exchange in the A_2 spectra and slow-exchange in the A_6 spectra.

3.7 Plasticity in base recognition by LARP1 LaM domain

In order to explore the distinct base interactions of LARP1-LaM, NMR titrations were carried out using increasing concentration (100 µM to 500 µM) of 5'-GMP, 5'-UMP and 5'CMP nucleotides into ^{15}N-labeled LARP1 LaM, in addition to A_2 and A_6 (Figure 10). These nucleotides displayed spectral changes in fast exchange between the bound and free forms, with reduced amplitude to those observed with A_6, indicating their binding capacity to LARP1-LaM.

Interestingly, the titration with 5'-GMP produced the most significant change, followed by 5'-AMP, while 5'-UMP and 5'-CMP produced fewer and more subtle changes, suggesting a greater binding affinity for pyrimidine towards LARP1 LaM domain. These findings suggest that the inclusion of guanosine at the 3'-end of RNA may enhance the binding affinity of LARP1 LaM domain to 3'-end RNA.

Figure 10

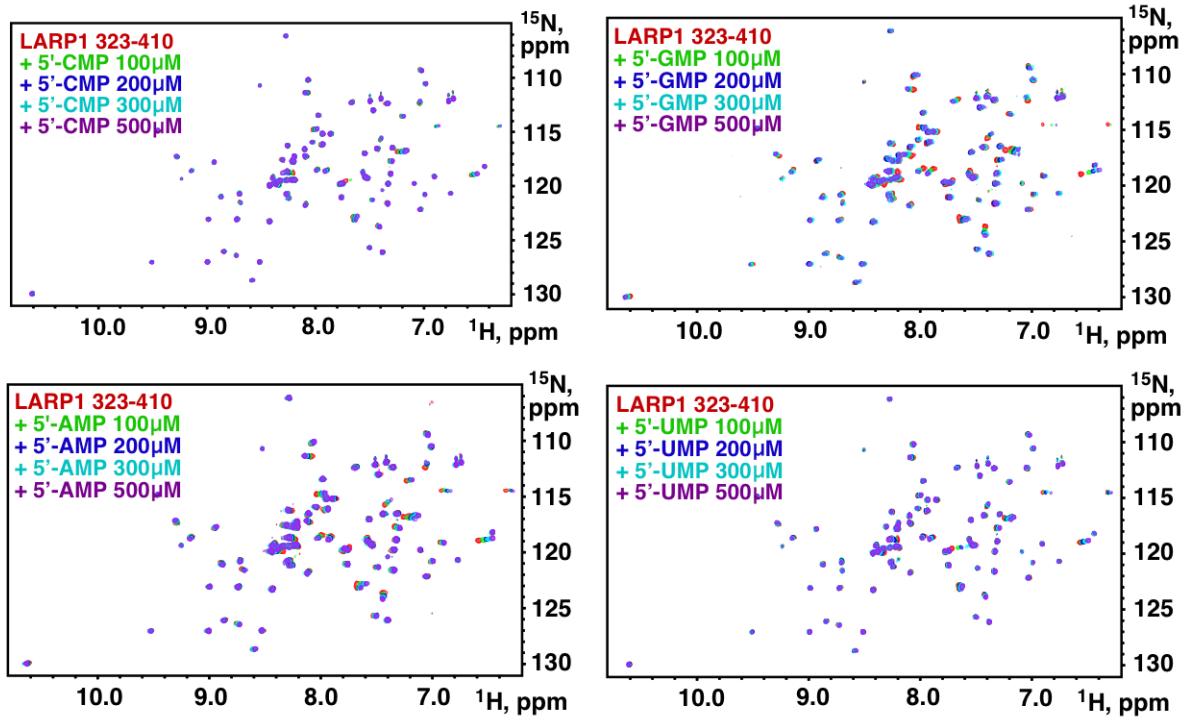
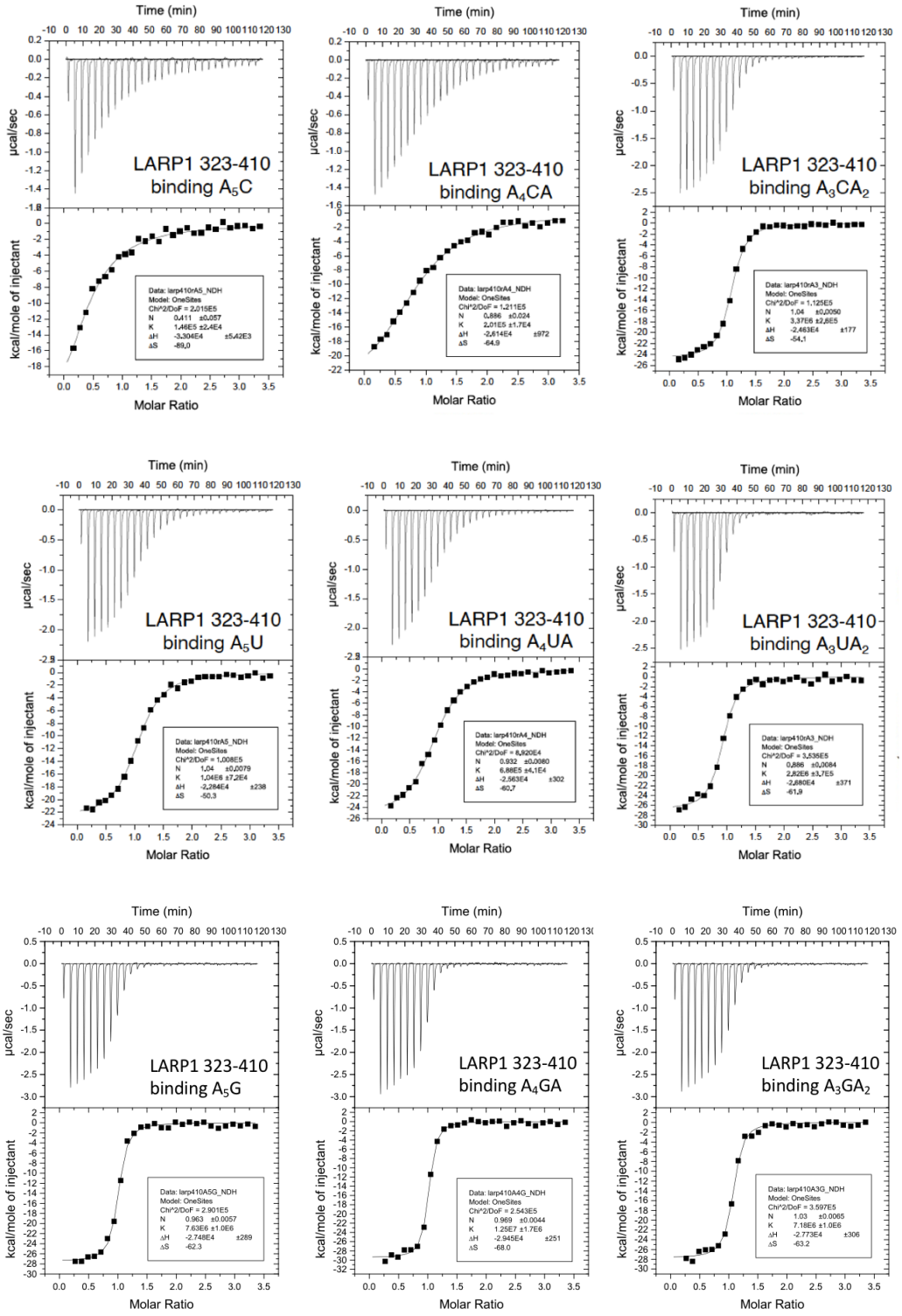


Figure 10. NMR of single nucleotides binding to LARP1 LaM.
Overlays of ^{15}N-^{1}H correlation spectra of ^{15}N-labeled LARP1 (323-410) upon addition of increasing nucleotide concentrations show peak shifts upon specific binding of nucleotides in fast exchange.

To investigate the effect of non-adenosine residues on LARP1-LaM binding at the 3'-end RNA, we employed nucleotide scanning binding experiments using ITC (Figure 11). Adenosine was replaced at positions -1, -2 and -3 with either uracil, cytosine or guanosine. Our results indicated that uracil substitutions were generally well tolerated, causing only a 3- to 4-fold

decrease in binding affinity at positions –1 and –2, with no effect observed at –3. Conversely, C-scanning showed a larger, 10-fold loss in affinity at positions –1 or –2, whereas cytosine substitution at position -3 had negligible effect on binding. Interestingly, guanosine substitution led to improved binding to the LaM domain. LARP1 LaM showed the highest affinity for guanosine at position -2 A_4GA with a K_d of 80 nM, representing a 3-fold improvement in affinity relative to A_6. A_5G and A_3GA_2 exhibited a smaller, approximately 2-fold affinity increase with K_d of 130 nM and 140 nM, respectively. While the binding improvement is not dramatic, it is reproducible and correlates well with the apparent steepness increase of the ITC experimental curves. Collectively, these results may suggest a functional role of guanosine presence at the 3'-end of RNA in LARP1 LaM domain binding.

Figure 11


Time (min)
µcal/sec
kcal/mole of injectant
Molar Ratio
LARP1 323-410 binding A5C
LARP1 323-410 binding A4CA
LARP1 323-410 binding A3CA2
LARP1 323-410 binding A5U
LARP1 323-410 binding A4UA
LARP1 323-410 binding A3UA2
LARP1 323-410 binding A5G
LARP1 323-410 binding A4GA
LARP1 323-410 binding A3GA2

RNA	Cell, μM	Syringe, μM	Chi-Square/DofF	N	Ka (M-1)	Error (M-1)	ΔH (cal/mol)	ΔS (cal/mol/deg)	Kd (nM)	error(nM)	Rel. affinity to A_6
	LARP1 (323-410)										
A_5U	20 (protein)	300 (RNA)	10080	1.04±0.0079	1040000	72000	-22840±238	-50.3	962	42	6.1
A_4UA	20 (protein)	300 (RNA)	89200	0.932±0.0080	688000	41000	-25630±302	-60.7	1453	82	1.5
A_3UA_2	20 (protein)	300 (RNA)	353500	0.886±0.0084	2820000	370000	-26800±371	-61.9	297	21	1
A_5C	20 (protein)	300 (RNA)	201500	0.411±0.057	146000	24000	-33040±5420	-89.0	6849	967	28.5
A_4CA	20 (protein)	300 (RNA)	121100	0.886±0.024	201000	17000	-26140±972	-64.9	4975	388	21
A_3CA_2	20 (protein)	300 (RNA)	112500	1.04±00.0050	3370000	260000	-24630±177	-54.1	297	21	1.2
A_5G	20 (protein)	300 (RNA)	290100	0.963±0.0057	763000	1000000	-27480±289	-62.3	130	20	0.52
A_4GA	20 (protein)	300 (RNA)	254300	0.969±0.0044	12570000	1700000	-29450±251	-68	80	10	0.32
A_3GA_2	20 (protein)	300 (RNA)	359700	1.03±0.0065	7180000	1000000	-27730±306	-63.2	140	20	0.56

Figure 11. ITC shows improved binding of guanosine containing RNAs.

ITC thermograms of LaM binding to different oligonucleotides (A_5U, A_4UA, A_3UA_2, A_5C, A_4CA, A_3CA_2, A_5G, A_4GA, A_3GA_2) and with non-adenosine substituted at different positions.

3.8 Crystal structure of LARP1 LaM domain binding to poly(A)

To investigate the molecular mechanism responsible for the binding specificity of LARP1. Dr. Guennadi Kozlov employed X-ray crystallography to obtain crystals LARP1 (323–410) diffracted to better than 2 Å. Dr. Kozlov then determined the structure of the unliganded protein using molecular replacement with the LaM domain of LARP3[8]. The DALI server[73] identified LARP7 (PDB 4WKR; 1.1 Å RMSD) and LARP3 (PDB 1S29; 1.2 Å RMSD) as the closest structural homologs[74, 75]. The structure is highly similar to previous LaM module structures but without the RRM domain (Figure 12). In the unliganded structure, a water molecule and two sulfate ions occupy the RNA binding site and are part of a conserved hydrogen bonding network formed by Asn342 and Asp346.

Figure 12

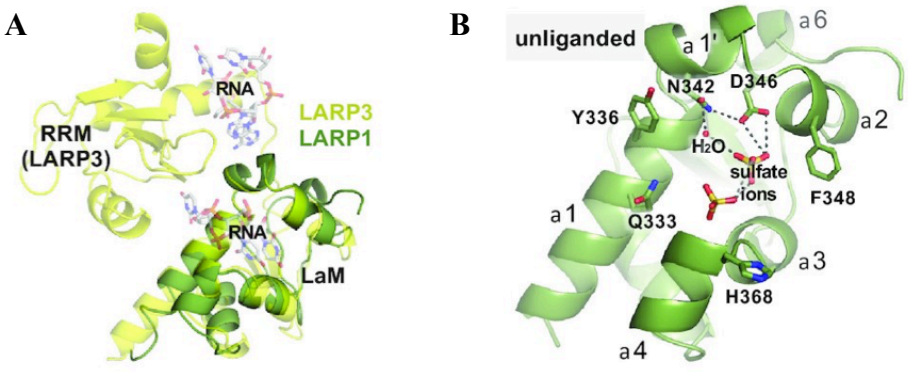


Figure 12. Structures of LARP1 LaM domain

- **A.** Overlay of LARP1 LaM (green) with La module of LARP3 (yellow; PDB 2VOD). RNA bound to the LARP3 LaM and RRM domains is shown semi-transparent.
- **B.** Structure of unliganded LaM.

Dr. Kozlov determined seven structures of LARP1 co-crystallized with RNAs ranging from three and eleven nucleotides in length. The crystals showed density only for the 3′-terminal nucleotides and were classified into two groups based on the stacking of the first and last base (Figure 13). In all the liganded structures, the A(–2) nucleotide ribose O2′ makes a hydrogen bond with the ordered water while the adenine base is flipped out and away from the 3′-nucleotide, and the A(–2) base stacks against the side chain of Tyr336, which may contribute to base selectivity. Additionally, the A(–2) adenine N3 forms a hydrogen bond with the side chain of Gln333.

Figure 13

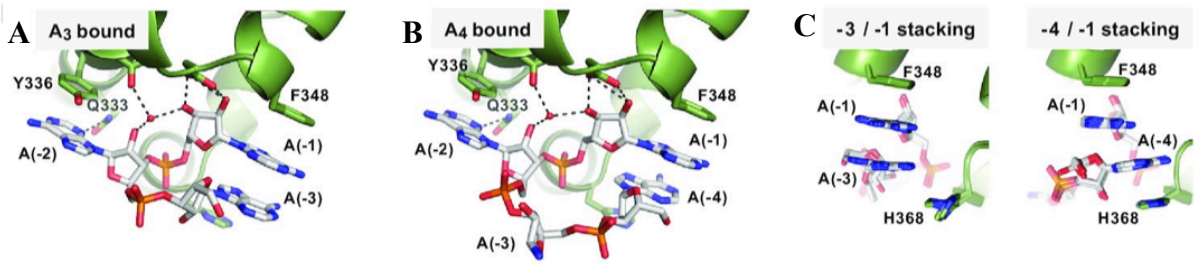


Figure 13. Structures of LARP1 LaM domain in complex with oligo-As[1].

- **A.** Structure of LaM with A_3 RNA bound (nucleotides numbered from the 3′-end).
- **B.** Structure of LaM with A_4 bound reveals an alternative stacking configuration.
- **C.** Comparison of the two different base stacking configurations.

The first group, observed in three of the seven liganded structures, has a stacking arrangement similar to the A_3 and A_6 bound structure and is characterized by a stacking arrangement of –3/–1. In this group, A(–3) is flipped back and stacked against A(–1), and the phosphate between the –2 and –3 nucleotides is stabilized by Arg345. Asn342 and Asp346 are responsible for binding the RNA 3′-end, and A(–1) and (–3) stack together adjacent to Phe348, while A(–2) stacks against

Tyr336. The second group, observed in four of the seven ligand structures, has the stacking arrangement is –4/–1, resembling the A_4 bound structure. In this group, A(–4) stacks on A(–1) with an additional contribution from His368, which stacks on the base of A(–4). The A(–3) nucleotide is partially disordered and does not make any contacts with the LaM. Moreover, the two groups can be distinguished by the shorter distance between His368's side chain and the adenylate base in the -4/-1 configuration. The altered stacking arrangements are reminiscent of poly(U) sequences interacting with LARP3, in which U(–1) stacks on U(–3) or alternatively U(–1) stacks on U(–4)[29].

3.9 Mutagenesis confirms the interactions observed in crystal structure

Following our identification of critical residues (Gln333, Tyr336, Arg345, Phe 348, His368) that interact with target RNA in the LARP1 LaM domain crystal structure, we performed ITC experiments to determine the effect of their mutation on RNA affinity. To confirm that the mutants were correctly folded, we isotopically labeled them and analyzed their $^{1}H–^{15}N$ correlation spectra. Our results showed that all mutants yielded NMR spectra similar to the wild-type protein, indicating that the mutations did not affect protein folding (Figure 14). Single point mutants in the binding site resulted in a decrease in A_4 binding affinity by one to two orders of magnitude, with loss of aromatic residues having the largest effects (Figure 15). Notably, loss of Gln333, which forms a hydrogen bond to the A(–2) adenine base, had a surprisingly large effect and decreased binding affinity by 20-fold. In contrast, the role of Arg345 in binding was found to be relatively minor, as its mutation to alanine only decreased the affinity by 2-fold. we generated and tested two double mutants, Q333A/F348A and Y336A/F348A, as the single-point mutants still displayed weak RNA binding. Our results showed that double mutations completely abolished binding to A_4. To further rule out a role of

the RRM region in poly(A) RNA binding, we tested the double mutants in the context of the longer LARP1 fragment that contains both the LaM and the supposed RRM. We found that the mutations always prevented RNA binding.

Figure 14

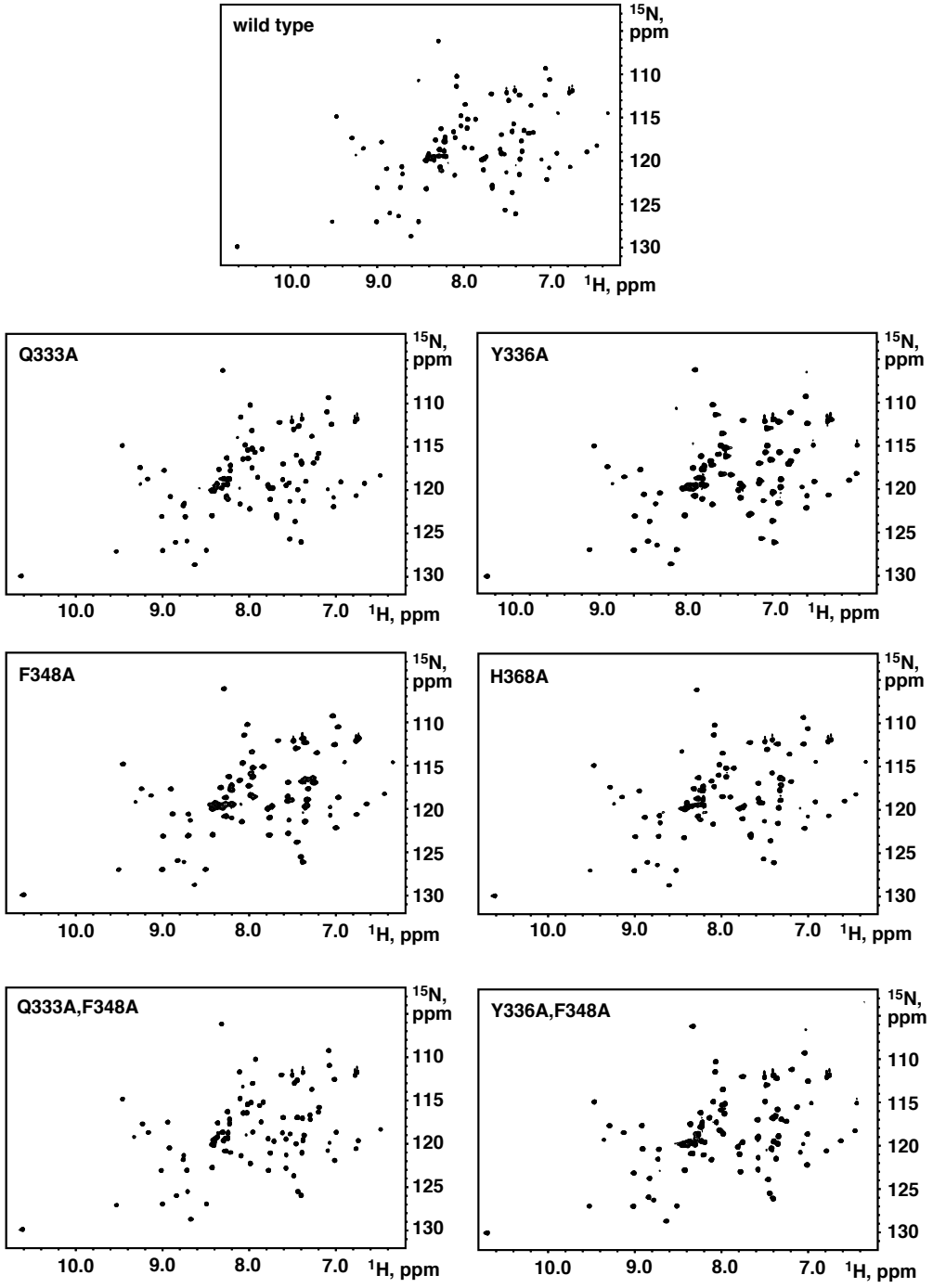
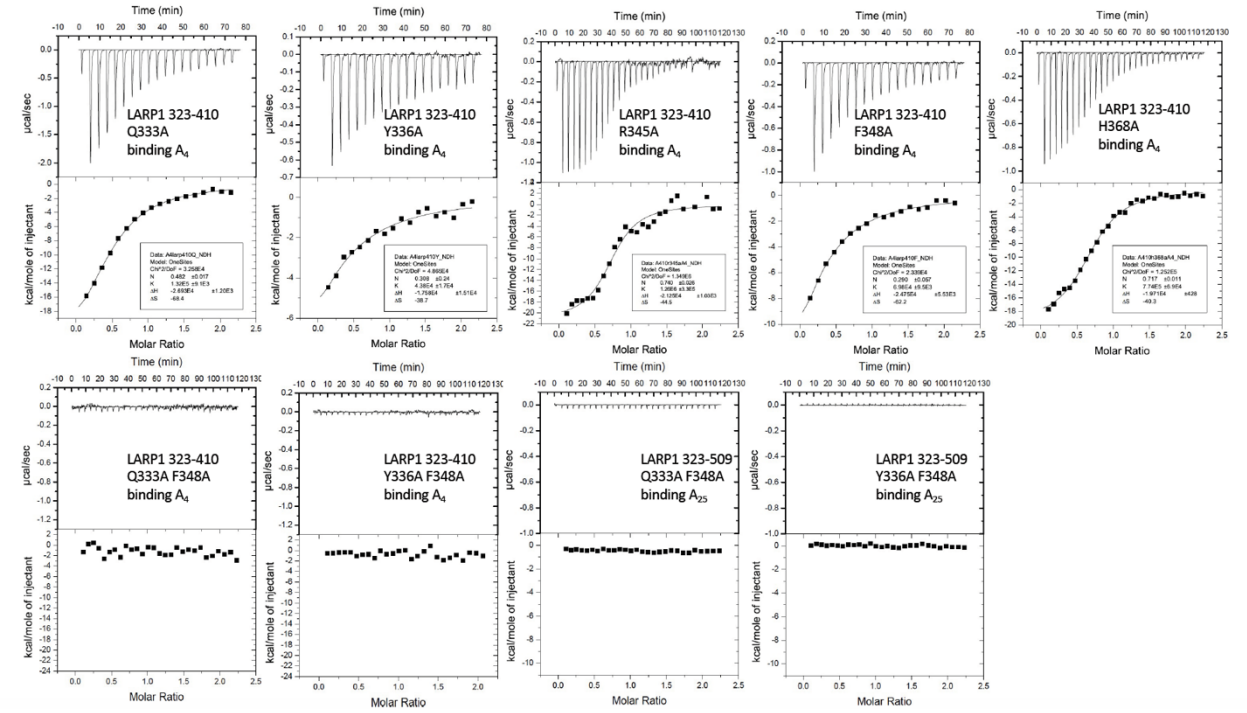


Figure 14. Verification of the structural integrity of the LaM domain mutants by NMR. ^{1}H–^{15}N correlation spectra of ^{15}N -labeled LARP1 (323-410) single and double mutants are very similar to the wild type spectrum, confirming that the mutations did not unfold the domain.

Figure 15



RNA	Cell, µM	Syringe, µM	Chi-Square/DofF	N	Ka (M-1)	Error (M-1)	ΔH (cal/mol)	ΔS (cal/mol/deg)	Kd (nM)	error(nM)	relative affinity
					LARP1 (323-410) Q333A						
A_4	30 (RNA)	300 (protein)	32580	0.482±0.017	132000	9100	-26930±1200	-68.4	7575	500	20
					LARP1 (323-410) Y336A						
A_4	30 (RNA)	300 (protein)	48650	0.308±0.24	43800	17000	-17580±15100	-38.7	22831	6335	70
					LARP1 (323-410) R345A						
A_4	30 (RNA)	300 (protein)	1349000	0.740±0.026	1260000	330000	-21250±1030	-44.5	793	5490	2.3
					LARP1 (323-410) F348A						
A_4	30 (RNA)	300 (protein)	23390	0.290±0.057	69800	9500	-24750±5530	-62.2	14327	1717	40
					LARP1 (323-410) H368A						
A_4	30 (RNA)	300 (protein)	125200	0.717±0.011	774000	69000	-19710±428	-40.3	1290	108	4
					LARP1 (323-410) Q333A, F348A						
A_4	30 (RNA)	300 (protein)	NB	NB	NB	NB	NB	NB	NB	NB	NB
					LARP1 (323-410) Y336A, F348A						
A_4	30 (RNA)	300 (protein)	NB	NB	NB	NB	NB	NB	NB	NB	NB
					LARP1 (323-509) Q333A, F348A						
A_{25}	15 (RNA)	150 (protein)	NB	NB	NB	NB	NB	NB	NB	NB	NB
					LARP1 (323-509) Y336A, F348A						
A_{25}	15 (RNA)	150 (protein)	NB	NB	NB	NB	NB	NB	NB	NB	NB

Figure 15. ITC thermograms of binding between mutated LARP1 fragments and RNAs. ITC results reveal reduced affinity of the Q333A, Y336A and F348A LaM mutants, and complete loss of binding for Q333A/F348A and Y336A/F348A double mutants. The mutations also prevented binding of A_{25} RNA in the context of the larger LARP1 fragment (323-509).

CHAPTER 4. DISCUSSION

The protein LARP1 is a multifunctional protein with two RNA recognition domains that are independently folded. LARP1 contains a unique RNA-binding, DM15 domain, at the C-terminus that is not present in other La-related proteins. The DM15 domain is known to bind the mRNA 5'-cap and TOP sequences, leading to negative regulation of TOP mRNA translation in response to mTORC1 inhibition[19]. In contrast, the N-terminal half of the LARP1 contains a highly conserved La motif, immediately followed by a conserved PAM2 motif and a putative RRM. While La-modules in other LARP family members have been studies structurally and functionally, the La-module in LARP1 remained characterized. Previous research has suggested that the N-terminal half, which contains the La-module, is implicated in the protection and stabilization of poly(A)[64, 65]. However, the structure and molecular mechanism underlying these functions are yet to be elucidated.

The present project employed ITC and NMR spectroscopy to investigate the RNA-binding properties of LARP1. Our results revealed that the putative RRM in LARP1 does not have a structured fold. Instead, the La motif serves as a stand-alone RNA-binding domain with high submicromolar binding affinity for poly(A) RNA. The results were complemented with our ITC experiments with poly(U), poly(G) and poly(G) titrated, confirming the preference of LaM for A-rich sequences. Our ITC experiments with oligonucleotides modified at the 3'-ribose indicated that LaM specifically recognizes 3'-end of poly(A). These results together provided compelling evidence that the LaM of LARP1 functions as a stand-alone 3′ poly(A) binding domain.

In order to validate our in-vitro affinity measurements, Dr. Kozlov obtained high-resolution crystal structures of the LARP1 LaM in complex with seven different

oligonucleotides. The structural analysis confirmed that RNA binding is driven by recognition of the 3'-end RNA. Interestingly, the position and conformation of the two 3'-adenylates were observed in all the structures whereas the 5′-nucleotides displayed considerable structural plasticity, including two different stacking configurations. We identified His368, a residue unique to LARP1, which contributes to this flexibility by providing a stacking interaction with the nucleotide at position – 4. Although the affinity difference was limited, the –4/–1 stacking conformation was found to be dominant in solution. We also identified Gln333 as an invariant residue whose side chain hydrogen bonds with the adenine at position –2 and found that its loss resulted in approximately 20-fold decrease in RNA binding affinity without affecting the protein folding. The crystal structures further allowed us to design point mutants that specifically block RNA binding. Overall, these results provide insights into the RNA binding mechanism of LARP1-LaM and its specific interaction with poly(A) RNA.

4.1 LARP1 in mRNA poly(A) tail stabilization and protection

In complement to our in vitro binding experiments, we collaborated with Dr. Sandy Mattijssen from the group of Dr. Richard Maraia to conduct in vivo cell assays to provide further evidence of LARP1's ability to stabilize mRNA and protect poly(A) tail using the β-glo-ARE mRNA expression level reporter, a widely used indicator of mRNA stability[1] (Figure 16). In these cellular experiments, the HEK293T LARP1-KO cells were transfected with empty vector, Flag-tagged LARP1 wild-type, or Flag-tagged LARP1-Q333A. These cells were co-transfected with aliquots of a mixture of expression plasmids for β-glo-ARE, GFP, and the VA1 small noncoding RNA transcribed by RNA polymerase III, the latter as a transfection control. After forty-eight hours, the cells were harvested, and RNA was examined to assess mRNA levels and poly(A) tail length via northern blot. The results showed that cells transfected with LARP1 wide-

type accumulated higher levels of mRNAs for β-glo-ARE mRNA than empty vector and LARP1-Gln333A. Additionally, there was an upward mobility shift, indicative of increased poly(A) tail length in the LARP1-wide type cells, consistent with 3′ end protection. To directly measure of the effect of LARP1 in poly(A) protection, single-molecule poly(A) tail sequencing (SM-PAT-seq) was also employed. Comparison of poly(A) tail length distributions obtained by poly(A) tail sequencing for β-glo-ARE and GFP confirmed the poly(A) tail length increase observed on northern blots in LARP1-wide type cells relative to Gln333A and the empty vector. These results were consistent with very recent publication that identified LARP1 as a deadenylation inhibitor through the formation of a stable ternary complex with PABPC1–poly(A), decelerating the CNOT complex in vitro, thereby, interferes with the CCR4-NOT-mediated deadenylation and shields the poly(A) tail from deadenylation[66].

Figure 16

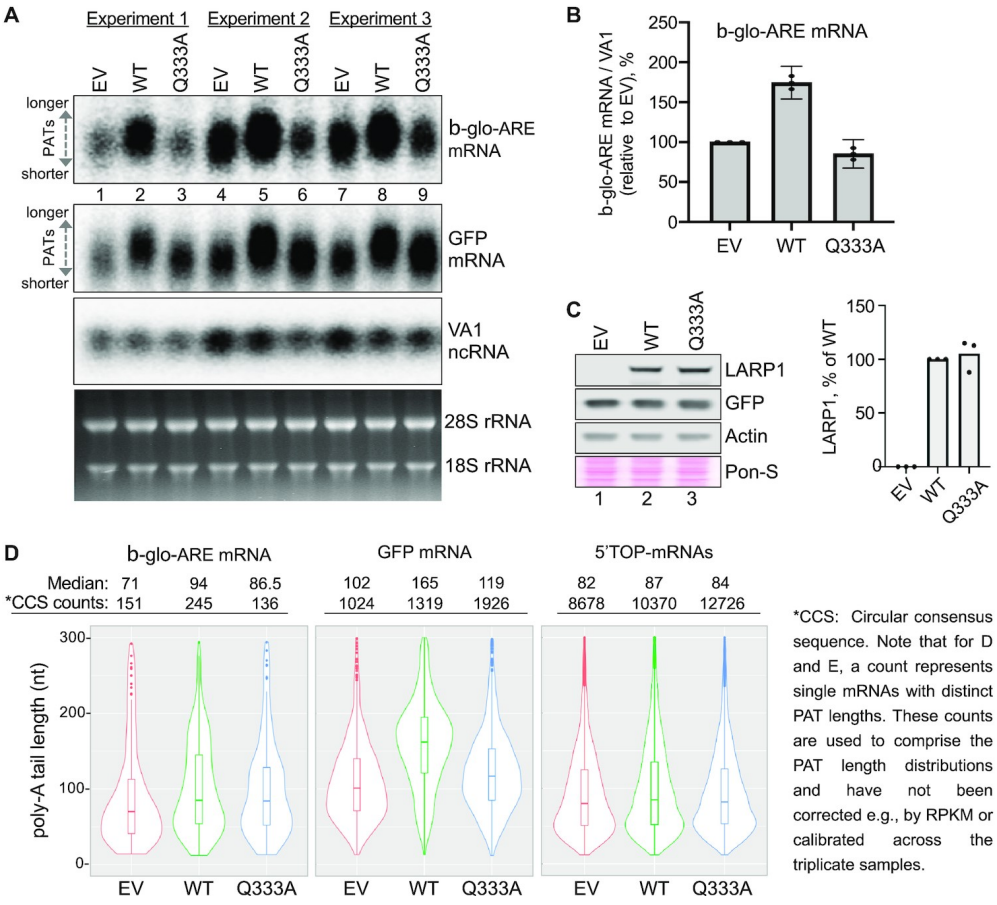


Figure 16. RNA-binding by the LARP1 LaM is required for PAT protection and mRNA stabilization.

A. Northern blot analysis of total RNA isolated from HEK293T cells 48 hours post transfection with the constructs indicated above the lanes

B. Quantitation of the β-globin-ARE mRNA signals from the northern blot in (A) normalized by the VA1 signals for replicate biological experiments.

C. Western blot (left) and quantification of protein levels (right) of the three experiments in panel A. The blot was probed with antibodies against LARP1, GFP and actin. Ponceau S (Pon-S) was used to stain total protein. LARP1 levels were normalized by actin.

D. Results of SM-PAT-seq analysis combined from three independent transfection experiments. The poly(A) tail lengths obtained by SM-PAT-seq are represented by violin plots in which the rectangles show the 95% confidence interval. The median PAT lengths above the plots are the circular consensus sequence (CCS) read counts (each CCS count represents an mRNA molecule with a specific PAT length).

Figure adapted from *Nucleic Acids Research*, Volume 50, Issue 16, 9 September 2022, Pages 9534–9547

4.2 Plasticity in base recognition by the LaM of LARP1

The plasticity in base recognition by the LaM of LARP1 is a notable finding, providing insight into another pathway that LARP1 LaM adopted is through stabilizing and protecting the guanylated-poly(A) tail. Our single nucleotide titration results showed that 5'-GMP bound better than 5'-AMP, 5'-UMP or 5'-CMP, indicating that LARP1 LaM has a preference for binding to 5'-GMP, compared to other nucleotides. Moreover, LARP1 LaM exhibits 2-3-fold increase in binding affinity to oligonucleotides accommodated with guanine bases at the position -1, -2 or -3. Guanylation of the poly(A) 3'-end results in a tighter binding to the LARP1 LaM domain, which is expected to enhance the LARP1's stabilizing effect on mRNAs. This finding is consistent with an earlier report indicating a positive correlation between 3'-end guanylation, mRNA stability and poly(A) tail length[76]. Subsequent studies have suggested that even a single guanosine residue included in the poly(A) is sufficient to impede the deadenylase CCR4-NOT complex, which shields mRNA from rapid deadenylation[59]. While our new data provide a plausible mechanism, further studies are necessary to experimentally verify that LARP1 and, in particular, its LaM domain are responsible for this guanylation-mediated poly(A) tail protection.

4.3 Full-length LARP1 RNA binding

The full-length LARP1 protein's RNA binding behavior is not well-understood, particularly with regards to the LaM domain. Previous research on a fragment of LARP1 (310–540) showed RNA binding with a pyrimidine-rich sequence with 40 nM affinity[65] using EMSA assays, with pull-down assays suggesting the potential for simultaneous binding of two RNA molecules, possibly through protein oligomerization[65]. Our ITC experiments with the shorter LARP1 fragment (residues 323–509) measured one-to-one binding stoichiometry of the same RNA with a thousand-fold weaker affinity. These discrepancies suggest that regions outside of

the LaM domain may contribute to RNA binding, despite lacking a well-folded, static structure. When we were attempting to express larger LARP1 fragments for structural studies, we observed a tendency of longer constructs to aggregate, potentially leading to the binding of multiple RNAs and increasing the affinity through avidity effects. The presence of long stretches of disordered regions in LARP1 raises intriguing questions about their function. The N-terminal half of LARP1 contains only one short, folded RNA-recognition motif and has a tendency to aggregate. One tempting speculation is that the disordered regions may play a role in the condensation of P-bodies and stress granules, thereby affecting protein-RNA binding. Consistent with this hypothesis, Jeffrey Chao's lab has demonstrated a TOP-anchoring role for LARP1 in stress granules and P-bodies through single-molecule imaging[77].

It is worth noting that LARP1 is not the only member in the LARP protein family containing intrinsically disorder regions (IDRs). Emerging studies on these IDRs from different members in LARPs have reported similar behaviors, where regions appearing unfolded in NMR spectra and without known RNA-binding motifs bind poly(A) RNA with low micromolar affinity. Recent work on human LARP4A[31] suggests that its La-module is not the main locus of interaction for RNA. Instead, its association with poly(A) is dominated by the N-terminal regions (NTRs) preceding the La-module, an intrinsically disordered region lacking recognizable RNA-binding sequences or motifs, in a novel mechanism of RNA recognition that is not fully characterized. Additional insights into the potential role of disorder region in RNA recognition have been provided by the studies of LARP6, where although La-module of human and zebrafish LARP6 is necessary and sufficient for high-affinity interaction with the conserved stem-loop of the 5′UTR collagen mRNA[26, 78, 79], the entire N-terminal domain of LARP6 comprising NTR and La-module displays different binding than the La-module alone, suggesting that the NTR may

modulate RNA-binding activity in zebrafish. Surprisingly, NTR sequences are fairly conserved, with 61% identity between human and zebrafish NTR, suggesting that the NTR may employ functional machinery and contribute to mammalian LARP6 binding activity as well. The molecular mechanism and function of the N-terminus is an ongoing topic to be studied.

One other potentially plausible question to be investigated with full-length LARP1 centers around its two independently functional RNA-recognition domains. As previously discussed, LARP1 contains a second RNA-binding DM15 domain at its C-terminus, which binds to the mRNA 5′-cap and TOP sequences to negatively regulate TOP mRNA translation in response to mTORC1 inhibition. The presence of the two RNA binding domains suggests that LARP1 could circularize TOP mRNAs, with the LaM domain binding to the 3′-end of poly(A) while the DM15 domain binds the 5′-end. The presence of the PABPC1-binding PAM2 motif adjacent to the LaM domain likely provides additional specificity and affinity for poly(A) RNA via recruitment of PABPC1.

4.4 Concluding remarks

In conclusion, our studies add new information and insight into the structural basis of how LARP1 stabilizes mRNA. Despite the large size of full-length LARP1, a single mutation in the LaM domain can significantly affect its ability to stabilize mRNA. Our previous observation that the LaM domain was sufficient for mRNA poly(A) tail protection and mRNA stabilization highlights the central role of the LaM domain in N-terminus La-module. The identification of point mutations that specifically block RNA binding will be useful in future studies of LARP1 activities and functions beyond mRNA stabilization.

REFERENCES

1. Kozlov, G.; Mattijssen, S.; Jiang, J.; Nyandwi, S.; Sprules, T.; Iben, James R.; Coon, Steven L.; Gaidamakov, S.; Noronha, A. M.; Wilds, Christopher J.; Maraia, Richard J.; Gehring, K., Structural basis of 3′-end poly(A) RNA recognition by LARP1. *Nucleic Acids Research* **2022,** *50* (16), 9534-9547.

2. Bousquet-Antonelli, C.; Deragon, J. M., A comprehensive analysis of the La-motif protein superfamily. *RNA* **2009,** *15* (5), 750-764.

3. 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* **2010,** *1799* (5-6), 365-78.

4. Mattijssen, S.; Kozlov, G.; Fonseca, B. D.; Gehring, K.; Maraia, R. J., LARP1 and LARP4: up close with PABP for mRNA 3' poly(A) protection and stabilization. *RNA Biology* **2021,** *18* (2), 259-274.

5. 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 Interdisciplinary Reviews-RNA* **2017,** *8* (6).

6. Merret, R.; Martino, L.; Bousquet-Antonelli, C.; Fneich, S.; Descombin, J.; Billey, E.; Conte, M. R.; Deragon, J. M., The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of La-related proteins. *RNA* **2013,** *19* (1), 36-50.

7. Dock-Bregeon, A. C.; Lewis, K. A.; Conte, M. R., The La-related proteins: structures and interactions of a versatile superfamily of RNA-binding proteins. *RNA Biology* **2021,** *18* (2), 178-193.

8. Teplova, M.; Yuan, Y. R.; Phan, A. T.; Malinina, L.; Ilin, S.; Teplov, A.; Patel, D. J., Structural basis for recognition and sequestration of UUU(OH) 3' temini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* **2006,** *21* (1), 75-85.

9. Yoo, C. J.; Wolin, S. L., The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell* **1997,** *89* (3), 393-402.

10. Maraia, R. J.; Kenan, D. J.; Keene, J. D., Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. *Mol Cell Biol* **1994,** *14* (3), 2147-58.

11. Bayfield, M. A.; Maraia, R. J., Precursor-product discrimination by La protein during tRNA metabolism. *Nature Structural & Molecular Biology* **2009,** *16* (4), 430-7.

12. Stefano, J. E., Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts. *Cell* **1984,** *36* (1), 145-54.

13. Eichhorn, K.; Jackson, S. P., A role for TAF3B2 in the repression of human RNA polymerase III transcription in nonproliferating cells. *Journal of Biological Chemistry* **2001,** *276* (24), 21158-65.

14. Pannone, B. K.; Xue, D.; Wolin, S. L., A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *Embo J* **1998,** *17* (24), 7442-53.

15. Krueger, B. J.; Jeronimo, C.; Roy, B. B.; Bouchard, A.; Barrandon, C.; Byers, S. A.; Searcey, C. E.; Cooper, J. J.; Bensaude, O.; Cohen, E. A.; Coulombe, B.; Price, D. H., LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Research* **2008,** *36* (7), 2219-29.

16. He, N.; Jahchan, N. S.; Hong, E.; Li, Q.; Bayfield, M. A.; Maraia, R. J.; Luo, K.; Zhou, Q., A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* **2008,** *29* (5), 588-99.

17. Markert, A.; Grimm, M.; Martinez, J.; Wiesner, J.; Meyerhans, A.; Meyuhas, O.; Sickmann, A.; Fischer, U., The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep* **2008,** *9* (6), 569-75.

18. Tcherkezian, J.; Cargnello, M.; Romeo, Y.; Huttlin, E. L.; Lavoie, G.; Gygi, S. P.; Roux, P. P., Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. *Genes Dev* **2014,** *28* (4), 357-71.

19. Lahr, R. M.; Fonseca, B. D.; Ciotti, G. E.; Al-Ashtal, H. A.; Jia, J.-J.; Niklaus, M. R.; Blagden, S. P.; Alain, T.; Berman, A. J., La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *eLife* **2017,** *6*, e24146.

20. Aoki, K.; Adachi, S.; Homoto, M.; Kusano, H.; Koike, K.; Natsume, T., LARP1 specifically recognizes the 3' terminus of poly(A) mRNA. *FEBS Lett* **2013,** *587* (14), 2173-8.

21. Yang, R.; Gaidamakov, S. A.; Xie, J.; Lee, J.; Martino, L.; Kozlov, G.; Crawford, A. K.; Russo, A. N.; Conte, M. R.; Gehring, K.; Maraia, R. J., La-related protein 4 binds poly(A), interacts with the poly(A)-binding protein MLLE domain via a variant PAM2w motif, and can promote mRNA stability. *Mol Cell Biol* **2011,** *31* (3), 542-56.

22. Mattijssen, S.; Arimbasseri, A. G.; Iben, J. R.; Gaidamakov, S.; Lee, J.; Hafner, M.; Maraia, R. J., LARP4 mRNA codon-tRNA match contributes to LARP4 activity for ribosomal protein mRNA poly(A) tail length protection. *Elife* **2017,** *6*. e28889.

23. Stefanovic, B., RNA protein interactions governing expression of the most abundant protein in human body, type I collagen. *Wiley Interdiscip Rev RNA* **2013,** *4* (5), 535-45.

24. Jolma, A.; Zhang, J.; Mondragón, E.; Morgunova, E.; Kivioja, T.; Laverty, K. U.; Yin, Y.; Zhu, F.; Bourenkov, G.; Morris, Q.; Hughes, T. R.; Maher, L. J., 3rd; Taipale, J., Binding specificities of human RNA-binding proteins toward structured and linear RNA sequences. *Genome Res* **2020,** *30* (7), 962-973.

25. Vukmirovic, M.; Manojlovic, Z.; Stefanovic, B., Serine-threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. *Mol Cell Biol* **2013,** *33* (19), 3893-906.

26. 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. *Journal of Molecular Biology* **2010,** *395* (2), 309-26.

27. Clark, K. L.; Halay, E. D.; Lai, E.; Burley, S. K., Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **1993,** *364* (6436), 412-20.

28. Alfano, C.; Sanfelice, D.; Babon, J.; Kelly, G.; Jacks, A.; Curry, S.; Conte, M. R., Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nature Structural & Molecular Biology* **2004,** *11* (4), 323-9.

29. 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* **2008,** *16* (6), 852-862.

30. Yang, Y.; Liu, S.; Egloff, S.; Eichhorn, C. D.; Hadjian, T.; Zhen, J.; Kiss, T.; Zhou, Z. H.; Feigon, J., Structural basis of RNA conformational switching in the transcriptional regulator 7SK RNP. *Mol Cell* **2022,** *82* (9), 1724-1736.e7.

31. Cruz-Gallardo, I.; Martino, L.; Kelly, G.; Atkinson, R. A.; Trotta, R.; De Tito, S.; Coleman, P.; Ahdash, Z.; Gu, Y.; Bui, T. T. T.; Conte, M. R., LARP4A recognizes polyA RNA via a novel binding mechanism mediated by disordered regions and involving the PAM2w motif, revealing interplay between PABP, LARP4A and mRNA. *Nucleic Acids Research* **2019,** *47* (8), 4272-4291.

32. Meyuhas, O.; Kahan, T., The race to decipher the top secrets of TOP mRNAs. *Biochim Biophys Acta* **2015,** *1849* (7), 801-11.

33. 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 U S A* **2020,** *117* (10), 5319-5328.

34. Thoreen, C. C.; Chantranupong, L.; Keys, H. R.; Wang, T.; Gray, N. S.; Sabatini, D. M., A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **2012,** *485* (7396), 109-13.

35. Hsieh, A. C.; Liu, Y.; Edlind, M. P.; Ingolia, N. T.; Janes, M. R.; Sher, A.; Shi, E. Y.; Stumpf, C. R.; Christensen, C.; Bonham, M. J.; Wang, S.; Ren, P.; Martin, M.; Jessen, K.; Feldman, M. E.; Weissman, J. S.; Shokat, K. M.; Rommel, C.; Ruggero, D., The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **2012,** *485* (7396), 55-61.

36. Yamashita, R.; Suzuki, Y.; Takeuchi, N.; Wakaguri, H.; Ueda, T.; Sugano, S.; Nakai, K., Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. *Nucleic Acids Research* **2008,** *36* (11), 3707-15.

37. Fonseca, B. D.; Zakaria, C.; Jia, J. J.; Graber, T. E.; Svitkin, Y.; Tahmasebi, S.; Healy, D.; Hoang, H. D.; Jensen, J. M.; Diao, I. T.; Lussier, A.; Dajadian, C.; Padmanabhan,

N.; Wang, W.; Matta-Camacho, E.; Hearnden, J.; Smith, E. M.; Tsukumo, Y.; Yanagiya, A.; Morita, M.; Petroulakis, E.; González, J. L.; Hernández, G.; Alain, T.; Damgaard, C. K., La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). *Journal of Biological Chemistry* **2015,** *290* (26), 15996-6020.

38. Philippe, L.; Vasseur, J. 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* **2018,** *46* (3), 1457-1469.

39. Lahr, R. M.; Mack, S. M.; Héroux, A.; Blagden, S. P.; Bousquet-Antonelli, C.; Deragon, J. M.; Berman, A. J., The La-related protein 1-specific domain repurposes HEAT-like repeats to directly bind a 5'TOP sequence. *Nucleic Acids Research* **2015,** *43* (16), 8077-88.

40. Fonseca, B. D.; Lahr, R. M.; Damgaard, C. K.; Alain, T.; Berman, A. J., LARP1 on TOP of ribosome production. *Wiley Interdisciplinary Reviews: RNA* **2018,** *9* (5), e1480.

41. Biberman, Y.; Meyuhas, O., TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. *FEBS Lett* **1999,** *456* (3), 357-60.

42. Topisirovic, I.; Svitkin, Y. V.; Sonenberg, N.; Shatkin, A. J., Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA* **2011,** *2* (2), 277-98.

43. von Der Haar, T.; Ball, P. D.; McCarthy, J. E., Stabilization of eukaryotic initiation factor 4E binding to the mRNA 5'-Cap by domains of eIF4G. *Journal of Biological Chemistry* **2000,** *275* (39), 30551-5.

44. Haghighat, A.; Sonenberg, N., eIF4G dramatically enhances the binding of eIF4E to the mRNA 5-cap structure. *Journal of Biological Chemistry* **1997,** *272* (35), 21677-80.

45. Berman, A. J.; Thoreen, C. C.; Dedeic, Z.; Chettle, J.; Roux, P. P.; Blagden, S. P., Controversies around the function of LARP1. *RNA Biology* **2021,** *18* (2), 207-217.

46. Hong, S.; Freeberg, M. A.; Han, T.; Kamath, A.; Yao, Y.; Fukuda, T.; Suzuki, T.; Kim, J. K.; Inoki, K., LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *Elife* **2017,** *6*. e25237.

47. Eswaran, J.; Horvath, A.; Godbole, S.; Reddy, S. D.; Mudvari, P.; Ohshiro, K.; Cyanam, D.; Nair, S.; Fuqua, S. A.; Polyak, K.; Florea, L. D.; Kumar, R., RNA sequencing of cancer reveals novel splicing alterations. *Sci Rep* **2013,** *3*, 1689.

48. Mura, M.; Hopkins, T. G.; Michael, T.; Abd-Latip, N.; Weir, J.; Aboagye, E.; Mauri, F.; Jameson, C.; Sturge, J.; Gabra, H.; Bushell, M.; Willis, A. E.; Curry, E.; Blagden, S. P., LARP1 post-transcriptionally regulates mTOR and contributes to cancer progression. *Oncogene* **2015,** *34* (39), 5025-36.

49. Xie, C.; Huang, L.; Xie, S.; Xie, D.; Zhang, G.; Wang, P.; Peng, L.; Gao, Z., LARP1 predict the prognosis for early-stage and AFP-normal hepatocellular carcinoma. *J Transl Med* **2013,** *11*, 272.

50. Kato, M.; Goto, Y.; Matsushita, R.; Kurozumi, A.; Fukumoto, I.; Nishikawa, R.; Sakamoto, S.; Enokida, H.; Nakagawa, M.; Ichikawa, T.; Seki, N., MicroRNA-26a/b directly regulate La-related protein 1 and inhibit cancer cell invasion in prostate cancer. *Int J Oncol* **2015,** *47* (2), 710-8.

51. Burrows, C.; Abd Latip, N.; Lam, S. J.; Carpenter, L.; Sawicka, K.; Tzolovsky, G.; Gabra, H.; Bushell, M.; Glover, D. M.; Willis, A. E.; Blagden, S. P., The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Research* **2010,** *38* (16), 5542-53.

52. Deragon, J. M.; Bousquet-Antonelli, C., The role of LARP1 in translation and beyond. *Wiley Interdiscip Rev RNA* **2015,** *6* (4), 399-417.

53. Stavraka, C.; Blagden, S., The La-Related Proteins, a Family with Connections to Cancer. *Biomolecules* **2015,** *5* (4), 2701-22.

54. Gallie, D. R., The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes Dev* **1991,** *5* (11), 2108-16.

55. Chen, C. A.; Strouz, K.; Huang, K. L.; Shyu, A. B., Tob2 phosphorylation regulates global mRNA turnover to reshape transcriptome and impact cell proliferation. *RNA* **2020,** *26* (9), 1143-1159.

56. Tang, T. T. L.; Stowell, J. A. W.; Hill, C. H.; Passmore, L. A., The intrinsic structure of poly(A) RNA determines the specificity of Pan2 and Caf1 deadenylases. *Nature Structural & Molecular Biology* **2019,** *26* (6), 433-442.

57. Baer, B. W.; Kornberg, R. D., Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc Natl Acad Sci U S A* **1980,** *77* (4), 1890-2.

58. Passmore, L. A.; Coller, J., Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol* **2022,** *23* (2), 93-106.

59. Lim, J.; Kim, D.; Lee, Y. S.; Ha, M.; Lee, M.; Yeo, J.; Chang, H.; Song, J.; Ahn, K.; Kim, V. N., Mixed tailing by TENT4A and TENT4B shields mRNA from rapid deadenylation. *Science* **2018,** *361* (6403), 701-704.

60. Xie, J.; Kozlov, G.; Gehring, K., The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* **2014,** *1839* (11), 1062-8.

61. Kozlov, G.; Ménade, M.; Rosenauer, A.; Nguyen, L.; Gehring, K., Molecular Determinants of PAM2 Recognition by the MLLE Domain of Poly(A)-Binding Protein. *Journal of Molecular Biology* **2010,** *397* (2), 397-407.

62. 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. *Journal of Biological Chemistry* **2010,** *285* (18), 13599-606.

63. Jinek, M.; Fabian, M. R.; Coyle, S. M.; Sonenberg, N.; Doudna, J. A., Structural insights into the human GW182-PABC interaction in microRNA-mediated deadenylation. *Nature Structural & Molecular Biology* **2010,** *17* (2), 238-240.

64. Mattijssen, S.; Kozlov, G.; Gaidamakov, S.; Ranjan, A.; Fonseca, B. D.; Gehring, K.; Maraia, R. J., The isolated La-module of LARP1 mediates 3'poly (A) protection and mRNA stabilization, dependent on its intrinsic PAM2 binding to PABPC1. *RNA biology* **2021,** *18* (2), 275-289.

65. 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* **2021,** *18* (2), 248-258.

66. Park, J.; Kim, M.; Yi, H.; Baeg, K.; Choi, Y.; Lee, Y.-s.; Lim, J.; Kim, V. N., Short poly(A) tails are protected from deadenylation by the LARP1–PABP complex. *Nature Structural & Molecular Biology* **2023**, *30* (3), 330–338.

67. McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J., Phaser crystallographic software. J Appl Crystallogr 2007, 40 (Pt 4), 658-674.

68. Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Echols, N.; Headd, J. J.; Hung, L. W.; Jain, S.; Kapral, G. J.; Grosse Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R. D.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P.

H., The Phenix software for automated determination of macromolecular structures. *Methods* **2011,** *55* (1), 94-106.

69. Emsley, P.; Cowtan, K., Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **2004,** *60* (Pt 12 Pt 1), 2126-32.

70. Winn, M. D.; Murshudov, G. N.; Papiz, M. Z., Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* **2003,** *374*, 300-21.

71. Horke, S.; Reumann, K.; Rang, A.; Heise, T., Molecular characterization of the human La protein.hepatitis B virus RNA.B interaction in vitro. *Journal of Biological Chemistry* **2002,** *277* (38), 34949-58.

72. Curry, S.; Kotik-Kogan, O.; Conte, M. R.; Brick, P., Getting to the end of RNA: Structural analysis of protein recognition of 5′ and 3′ termini. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **2009,** *1789* (9), 653-666.

73. Holm, L.; Sander, C., Dali: a network tool for protein structure comparison. *Trends Biochem Sci* **1995,** *20* (11), 478-80.

74. Uchikawa, E.; Natchiar, K. S.; Han, X.; Proux, F.; Roblin, P.; Zhang, E.; Durand, A.; Klaholz, B. P.; Dock-Bregeon, A. C., Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Research* **2015,** *43* (6), 3373-88.

75. 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* **2004,** *23* (5), 1000-7.

76. Chang, H.; Lim, J.; Ha, M.; Kim, V. N., TAIL-seq: genome-wide determination of poly(A) tail length and 3' end modifications. *Mol cell* **2014,** *53* (6), 1044-52.

77. Wilbertz, J. H.; Voigt, F.; Horvathova, I.; Roth, G.; Zhan, Y.; Chao, J. A., Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Mol Cell* **2019,** *73* (5), 946-958.e7.

78. Martino, L.; Pennell, S.; Kelly, G.; Busi, B.; Brown, P.; Atkinson, R. A.; Salisbury, N. J.; Ooi, Z. H.; See, K. W.; Smerdon, S. J.; Alfano, C.; Bui, T. T.; Conte, M. R., 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* **2015,** *43* (1), 645-60.

79. Castro, J. M.; Horn, D. A.; Pu, X.; Lewis, K. A., Recombinant expression and purification of the RNA-binding LARP6 proteins from fish genetic model organisms. *Protein Expr Purif* **2017,** *134*, 147-153.