Deciphering the Binding of 5' Stem Loop RNA to the La Domain of Human LARP6

Blaine H. Gordon[a,b], Nolan Blackford[a,b], and Robert Silvers[a,b,*]

[a] Department of Chemistry & Biochemistry, Florida State University, Tallahassee, FL 32306, USA

[b] Institute of Molecular Biophysics, Florida State University, Tallahassee, FL 32306, USA

* To whom correspondence should be addressed: Robert Silvers, Department of Chemistry and Biochemistry, Institute of Molecular Biophysics, Florida State University, 91 Chieftan Way, Tallahassee, FL 32306, USA. Phone: +1 (850) 645-9649. Email: rsilvers@fsu.edu

ORCID:
Blaine Gordon 0000-0002-0758-5078
Nolan Blackford 0009-0008-5818-3870
Robert Silvers 0000-0003-0197-3878

Abstract

La-related protein 6 regulates the highly organized biosynthesis of type I procollagen polypeptides and affects proper assembly of procollagen peptides into heterotrimers of type I procollagen. LARP6-mediated regulation of collagen biosynthesis is mediated through interaction with the 5' stem loop motif found in type I and III collagen mRNA. Recent studies highlight the involvement of HsLARP6 in fibroproliferative diseases and its potential as a target for therapeutic intervention. The intrinsic instability of the La domain of HsLARP6 hampers studies probing the molecular basis of biologically- and disease-relevant structure-function relationship, particularly when high concentrations are required. This work provides detailed procedures to produce milligram amounts of RNase-free and functional La domain of HsLARP6. Furthermore, we investigated the effect of the construct length as well as RNA binding on protein stability. N- and C-terminal extensions greatly impact stability based on interactions with the core domain and modulation of the pI. When in complex with its cognate 5'SL RNA, the La domain shows unprecedented stability compared to the aggregation-prone unbound state. The protein-RNA complex remains stable for at least 50x longer than the unbound state, under identical conditions, likely due to a global change in conformational plasticity upon RNA binding. These results provide a foundation for further studies of the molecular recognition of 5'SL by HsLARP6 as well as a platform for refining potential antifibrotic therapeutics.

Keywords
RNA recognition, RNA-binding domain, La-related protein, mRNA, fibrosis

Introduction

Human La-related protein 6 (HsLARP6) belongs to a diverse, but highly conserved superfamily of RNA-binding proteins (RBPs) that is involved in gene regulation at the in post-transcriptional level.[1, 2] Due to their pivotal role in the regulation of the intricate balance of translation, breakdown, and storage of messenger RNAs (mRNAs), LARPs are closely linked to the onset and progression of various cancers, neurodegenerative disorders, and fibroproliferative diseases.[3-8]

The LARP6-mediated regulation of type I and III collagen biosynthesis plays a pivotal role in the onset and progression of fibroproliferative diseases.[7] Type I collagen is predominantly produced by fibroblasts and myofibroblasts and is one of the major constituents of the extracellular matrix of tendons, skin and bones, while it is only present in small amounts in healthy tissues and organs such as liver, heart, kidney, and lungs. In fibroblasts, the type I collagen polypeptides α1(I) and α2(I) are translated in a highly regulated and organized process and the folding of type I procollagen heterotrimers within the endoplasmic reticulum occurs in highly organized structures termed collagenosomes.[8] In fibroproliferative disease, the biosynthesis of type I procollagen peptides is significantly elevated in affected tissues, leading to scarring and impairment of tissue function.[9, 10] It was shown that LARP6 facilitates biosynthesis by recruiting accessory translational factors, which enhance the translational efficiency.[11-16] The interaction between LARP6 and the collagen mRNA is directly linked to this increase in collagen peptide biosynthesis.[17] Furthermore, it was recently shown that the interaction between LARP6 and collagen mRNAs is crucial for the proper formation of collagenosomes.[8, 16, 18, 19] Furthermore, in an animal model where the LARP6-dependent pathway of type I collagen biosynthesis is interrupted, it was shown that resulting animals are resistant to hepatic fibrosis, underlining the importance of LARP6 RNA interaction in fibrotic disease.[7] These findings

highlight the critical nature of the LARP6 interaction with collagen mRNA, that contributes to disease progression, providing potential targets for therapeutic intervention in fibrosis and other related conditions.

The defining structural feature of all members of the LARP superfamily is the presence of an RNA-binding domain (RBD), known as the La domain (Figure 1A). In the majority of LARPs, the La domain (LA) is paired with a downstream RNA recognition motif (RRM) connected by a short linker. Both domains form a functional unit, termed the La module, that is essential for mediating the interaction between the LARP and their target mRNAs.

Commonly, LARPs recognize single-stranded RNA motifs. LARP1 binds the poly(A) tail in mRNA as well as the 5' terminal oligopyrimidine (5'TOP) motif in TOP mRNAs,[20-22] LARPs 3 and 7 bind the UUU-3'OH motif of RNAPIII transcripts and 7SK snRNA, respectively,[23-25] while LARP4 interacts with AU-rich sequences.[26] LARP6 however recognizes a highly conserved, double-stranded RNA motif known as the 5' stem-loop (5'SL) found in the mRNAs coding for the type I collagen peptides α1(I) and α2(I) as well as type III collagen peptide α1(III) (Figure 1B).[16, 27, 28] The 5'SL motif is located at the interface between the 5'-UTR and coding region of type I and III collagen mRNAs and forms a stem loop structure containing an internal loop (bulge) that includes the start codon (Figure 1B & C). The 5'SL is between 46-48 nucleotides long and harbors nine highly conserved nucleotides that form the internal loop that is flanked by two stems with higher sequence variability. The 5'SL motif is highly conserved across all vertebrates and plays a critical role in regulating the translation of these collagen mRNAs.

The structure of the LAM:5'SL complex is currently unknown; hence the molecular basis of the interaction remains poorly understood. This is further complicated by conflicting studies regarding the individual roles of the La domain and RRM in 5'SL binding. Furthermore, previous

studies show that both La domain and La module are unstable and aggregation-prone, which complicates production, handling, and storage of protein preparations and hampers further studies into the interaction between LARP6 and 5'SL.

Here, we report an improved strategy for recombinant production of milligram quantities of high-purity La domain that is RNase-free and binding-competent, a prerequisite for biochemical and biophysical studies of 5'SL binding. We show that the choice of protein construct modulates the stability and binding competency of the La domain by probing N- and C-terminal extensions of different lengths. Furthermore, we show that the inherent instability of the La domain is suppressed entirely by binding of its cognate 5'SL RNA. These results provide insights into the binding of La domain to the 5'SL while also providing a significant platform for further studies into the molecular basis of binding.

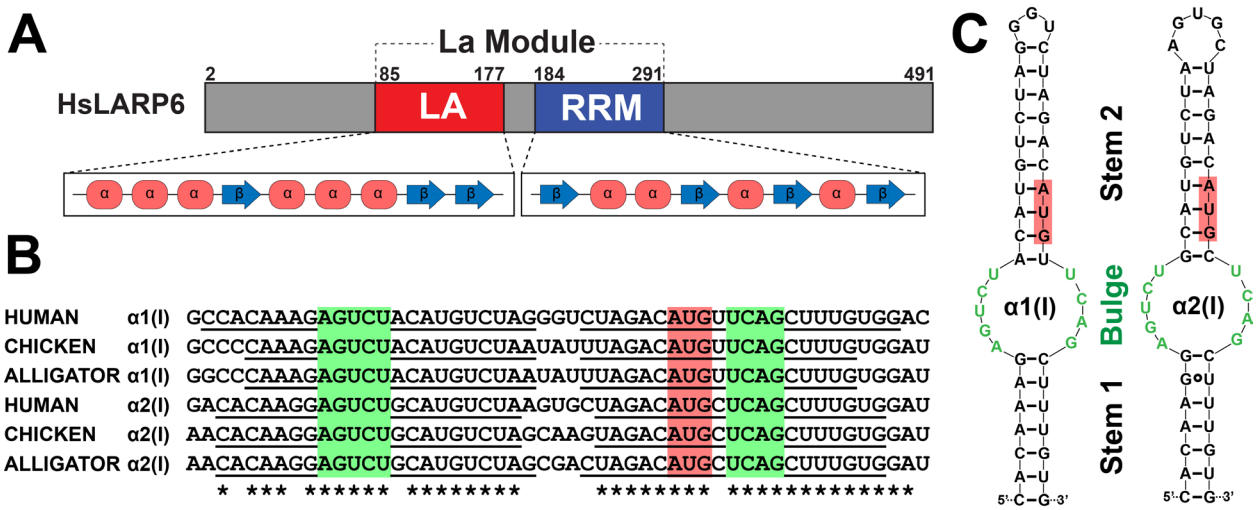

Figure 1: HsLARP6 and the 5' stem loop of type I collagen mRNAs. (A) Domain organization of HsLARP6 highlighting the La domain (LA) and RNA recognition motif (RRM) which compose the La module. Secondary structure of LA and RRM domains are as indicated. (B) Sequence alignment of mRNAs encoding the α1(I) and α2(I) strands of type I collagen from *H. sapiens*, *G. gallus*, and *A. mississippiensis*. Highly conserved stems (underlined) surround entirely conserved bulge regions (green). Conserved bases are indicated by an asterisk (*). The location of the start codon is shown in red. (C) Secondary structure prediction for α1(I) and α2(I) 5'SL. Bases belonging to the bulge and start codon are shown in green and red, respectively.

Results and Discussion

The post-translational regulation of collagen biosynthesis by HsLARP6 and its role in the onset and progression of fibroproliferative diseases is well documented. The interaction of the La module and 5'SL is directly involved in this process,[16, 18, 19] hence several previous studies have focused on studying the interaction employing biochemical and biophysical methods. Access to high-purity and binding competent protein is the foundation for success in further studies. Below, we delineate a detailed roadmap to obtaining milligram quantities of RNase-free, high-purity La domain, a prerequisite for further studies into the binding of the La domain to 5'SL RNA. Furthermore, we describe the dependency of La domain stability and binding competency on the length of N- and C-terminal extension. Binding of the 5'SL to the La domain suppresses the instability of the La domain entirely brought on by a global shift in dynamics of the La domain.

Preparation of RNase-free, high-purity HsLARP6(79-183)

The preparation of high-purity and RNase-free La domain of HsLARP6 is a prerequisite for investigations into the binding of 5'SL to the La domain and La module of HsLARP6, in particular where high concentrations are required. In our hands, following previously published protocols for protein expression and purification resulted in inconsistent results with protein preparations still containing host RNases and nucleic acid impurities that impeded biochemical and biophysical studies. Host ribonucleases (RNases) tend to co-purify with the La domain. Furthermore, previous studies claimed that improper handling and storage conditions result in loss of functionality.[29]

We recombinantly expressed in *E. coli* and purified constructs of the La domain with varying lengths of N- and C-terminal extensions, termed HsLARP6(73-183), HsLARP6(79-183), HsLARP6(84-183), HsLARP6(73-176), and HsLARP6(73-171). In the following, we detail the

expression and RNase-free purification of HsLARP6(79-183) to ensure the production of homogenous, monodisperse HsLARP6(79-183).

Using SDS-PAGE, the production and purity of HsLARP6(79-183) at various purification steps was analyzed (Figure 2A). Simultaneously, the presence of lingering host RNases during purification was determined by RNase activity assay (Figure 2B). After the cell lysate was cleared from insoluble debris, a strong overexpression band of His_6-HsLARP6(79-183) can be seen (Figure 2A, lane 1) and the cell lysate expectedly shows RNase activity (Figure 2B, lane 1). During the IMAC purification, two separate wash steps are involved. The first wash is done at low salt concentration, the second wash is performed using a high-salt buffer. RNase activity in the high-salt wash indicates that lingering host RNases can not be removed with the low-salt wash alone (Figure 2B, lane 2). After elution of His_6-HsLARP6(79-183) from the IMAC column, the protein is already quite pure (Figure 2A, lane 3), and more importantly, the elution does not contain any appreciable amounts of RNase (Figure 2B, lane 3). The high-salt wash is mandatory to remove any lingering host RNases. Interestingly, the intact mock RNA chosen for the RNase activity assay forms a large molecular complex that does not penetrate the gel. This behavior is not observed after His_6-tag removal by TEV digestion. High purity and RNase-free HsLARP6(79-183) is obtained after reverse IMAC to remove His_6-tagged TEV protease (Figures 2A & B, lane 4), as well as subsequent ion exchange chromatography (Figures 2A & B, lane 5) and size exclusion chromatography (Figures 2A & B, lane 6).

From the expression in 1 L LB medium, a 4.5 g wet cell pellet yielded a total of ~25 mg of high-purity, RNase-free HsLARP6(79-183). After size exclusion chromatography, an A260/A280 ratio of 0.54 was obtained, indicating a complete to near-complete removal of nucleic acids contaminants.[30]

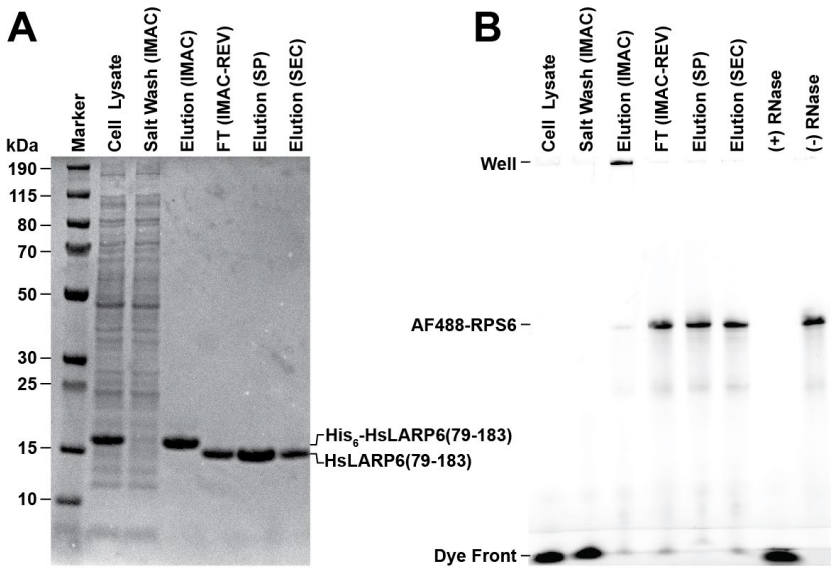

Figure 2: Expression and RNase-free purification of HsLARP6(79-183). (A) SDS-PAGE of several purification steps. Left to right: PageRuler™ Plus prestained protein ladder, cell culture lysate supernatant, pooled fractions from the high salt (1M NaCl) wash performed in the IMAC purification step, the first IMAC step, of the flowthrough and wash of the second IMAC step after TEV digestion, from the SP cation exchange step, and from SEC. Heights of each protein species before and after TEV digestion are marked. (B) RNase activity assay. Fluorescence image of 8M urea denaturing PAGE of the same samples from each purification step shown in A incubated at 37°C in the presence of a fluorescent RNA(AF488-RPS6 41mer) followed by a positive control containing 0.1mg/mL RNase A and a negative control. The heights of the well, AF488-RSP6 41mer, and dye front are marked.

The La domain exhibits high-affinity binding to 5'SL RNA

It was previously suggested that the La domain itself is unable to bind the 5'SL RNA and that the full-length La module is required for binding.[16, 27] ITC data of a HsLARP6(70-183) construct, presented alongside its solution NMR structure indicated no binding between the La domain and 5'SL.[27] We recently showed in collaboration with the Stefanovic laboratory that the La domain itself is sufficient to bind 5'SL RNA.[29] Indeed, HsLARP6(79-183) exhibits high-affinity binding to 5'SL in the absence of the downstream RRM domain (Figure 3).

Size exclusion chromatography of the unbound La domain compared to that of the La domain mixed with a slight molar excess of 5'SL RNA shows a shift in retention volume from 76.5 mL to 74.8 mL, indicating the formation of the larger ribonucleoprotein complex (Figure 3A). The formation of the complex is also evidenced by the inversion of A280/A260 ratio for the unbound

and bound forms. Microscale thermophoresis performed for fluoresceine-labeled LARP6(73-183)G74C demonstrated the ability of the La domain to bind 5'SL with an apparent K_D of 5.94 ± 0.99 nM (Figure 3B). K_D values previously reported for the complex formed between HsLARP6 and 5'SL RNA range from 0.5-40 nM.[19, 27, 29] [^{1}H, ^{15}N]-HSQC spectra of isotopically-labelled unbound and bound La domains show clear chemical shift perturbations (CSPs) which evidence the formation of the complex (Figure 3C). The absence of the resonances stemming from the unbound state indicate the complete conversion of the unbound state to the bound state in the 1:1 complex of La domain and 5'SL indicating high-affinity binding.

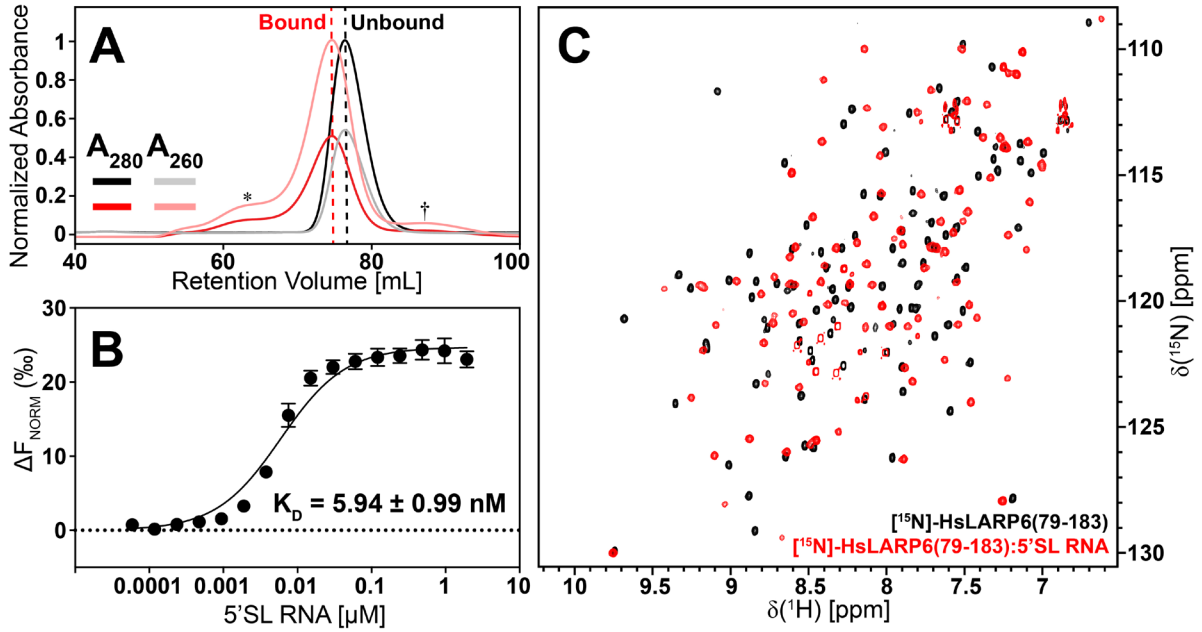

Figure 3: La domain binding to 5'SL RNA. (A) Size exclusion chromatography traces of normalized absorbance at 280 nm and 260 nm of ^{15}N-HsLARP6(79-183) bound to 5'SL RNA (red) and unbound (black). Minor amounts of duplexed 5'SL RNA and unbound 5'SL RNA are present as indicated by (*) and (†), respectively. Differences in retention volumes are emphasized with dashed lines. (B) Microscale thermophoresis of FL-HsLARP6(73-183) G74C against 5'SL RNA at 25°C. The dissociation constant K_D is as indicated. (C) 2D [^{1}H,^{15}N]-HSQC spectra of unbound and bound [^{15}N]-HsLARP6(79-183). Spectra were collected at 298K at 16.4T.

N- and C-terminal extensions modulate stability and 5'SL binding

Previous studies investigated the dependency of the length of the N- and C-terminus beyond the core domain on its ability to bind 5'SL. [16, 27, 29] Deletion of residues 1-88 of the full-length La module resulted in a loss of 5'SL binding. It is known from the solution NMR structure of La domain that the tryptophan residue located at position 85 (W85) anchors the flexible N-terminal sequence preceding helix 1 to an aliphatic patch located on helix 5. However, alanine point mutants of W85 do not result in a loss of binding. Deletion of residues from the C-terminal coil preceding position 183 also result in a loss of binding.[29] Based on these studies, it is not possible to ascribe the diminished binding capability of C-terminal truncations to either the diminished protein stability or a perturbation of intrinsic binding capability.

Here, we investigated the stability and binding of two N-terminal extensions and two C-terminal truncations. Compared to the constructs typically used spanning from residues 73 to 183 of HsLARP6, called HsLARP(73-183), we compared the sample preparation, stability, and 5'SL binding of HsLARP(79-183), HsLARP(84-183), HsLARP(73-176), and HsLARP(73-171). These five constructs were recombinantly expressed and purified identically to ensure side-by-side comparison. N-terminal truncations beyond residue E84 were not considered based on the previously stated role of W85. Total expression yield from 1L LB cultures of each construct expressed in *E. coli* Rosetta II (DE3) was evaluated by SDS-PAGE (Figure S1).

A similar overexpression level was observed for all constructs, with HsLARP6(79-183) having slightly increased levels. However, the behavior of these constructs during purification was very different. Particularly, the C-terminal truncations HsLARP6(73-176) and HsLARP6(73-171) did not remain soluble and aggregated quickly after elution from the IMAC column, removing them from further stability and binding studies.

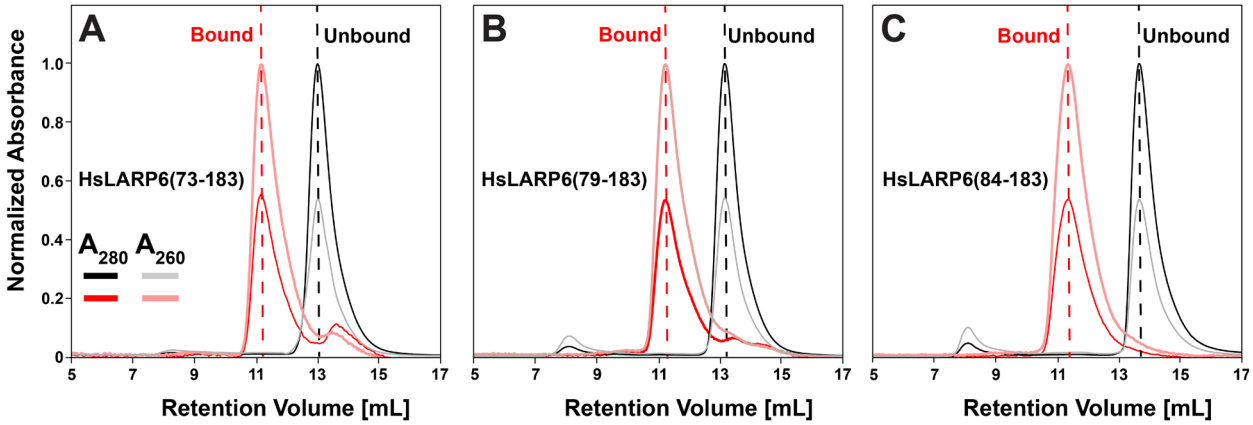

Figure 4: Size exclusion chromatography of 5'SL RNA-bound and unbound La domain N-terminal truncations. Overlay of the SEC profile of the normalized A_{280} (black/red) and A_{260} (gray/light red) traces of HsLARP6 residues(A) 73-183 (B) 79-183 (C) 84-183 before (black dash) and after (red dash) addition of an equimolar amount of 5'SL RNA.

A similar degree of stability was observed for the constructs of varying N-terminal lengths - HsLARP6(73-183), HsLARP6(79-183), and HsLARP6(84-183). To test their ability to bind 5'SL RNA, size exclusion chromatography was performed (Figure 4). Size exclusion chromatographs of the unbound-state and the 1:1 complex of the La domain and 5'SL were compared for all constructs. Complex formation was observed for all N-terminal variants tested here.

While all three N-terminal variants appear suitable for binding studies, each of these constructs varies significantly in isoelectric point (pI). The pI of HsLARP6(73-183), HsLARP6(79-183), and HsLARP6(84-183) are calculated to be 7.1, 8.2, and 9.4, respectively. The relative hydrophobicity[31, 32] (Figure 5A) and large number of glutamate and aspartate residues found between residues 70 and 90 in the N-terminus HsLARP6(73-183) (Figure 5B) provide the explanation for the significant difference in pI. Of these 20 residues, seven residues are glutamic acid, and one is aspartic acid. Physiologically relevant biochemical and biophysical studies operate within a specific pH range, while proteins near their pI are more likely to experience instability imposed by hydrophobic aggregation of neutral species.

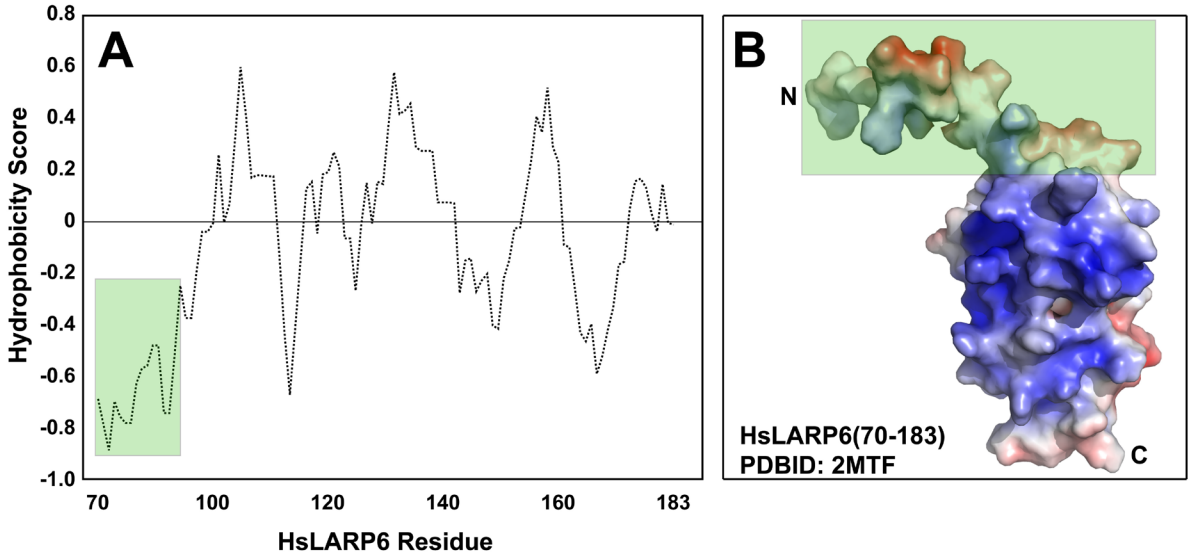

Figure 5: The negatively charged N-terminus of the La domain dictate the overall pI. (A) The hydrophobicity index of the HsLARP6 La domain residues 70-183 demonstrate the relatively high degree of polar residues in the flexible N-terminal tail. Residues 70-90 (green highlight) contain 8 Glu and Asp residues. (B) The electrostatic potential mapped onto the surface of the solution structure of the unbound La domain (PDB ID: 2MTF). Blue and red regions indicate positively and negatively charged surfaces, respectively. Gray regions are nonpolar. N- and C-termini are indicated.

5'SL binding stabilizes the La domain

The instability of the La domain and La module of HsLARP6 and tendency to aggregate and form precipitates is well documented.[29, 33] While observed over a large range of concentrations, the La domain of HsLARP6 is particularly unstable at higher concentrations above ~100 μM (~1.25 mg/mL), limiting further in vitro biochemical and biophysical studies.[27, 29, 33] Previous NMR spectroscopic studies of the unbound La domain and La module were conducted in the presence of 50 mM L-glutamate and 50 mM L-arginine to enhance solubility and slow down precipitation, with limited success.[27]

We investigated the intrinsic instability of the unbound state of HsLARP6(79-183) as well as bound state employing NMR spectroscopy. We acquired [^{1}H, ^{15}N]-HSQC spectra of HsLARP6(79-183) (unbound state) and HsLARP6(79-183):5'SL (1:1) (bound state) over the course of 3 hours and 7 days, respectively (Figure 6). The use of [^{1}H, ^{15}N]-HSQC spectra allows for the detection of backbone and side chain amide resonances, with shifts of resonances (chemical

shift perturbations, CSP) being indicative of localized changes in chemical environment and changes in signal intensity being indicative of loss of sample. Resonance positions and intensities of [^{1}H, ^{15}N]-HSQC spectra were compared at identical contour levels and scaling to detect any local or global structural changes in the La domain over time. Projections of a representative resonance were selected to show the change in intensity over time (Figure 6 insets).

The unbound La domain was poorly stable as evidenced by the dramatic loss in signal intensity within the first three hours of measurement at 25°C (Figure 6, left column), as expected. Compared to a [^{1}H, ^{15}N]-HSQC spectrum of a freshly prepared sample of 500 μM HsLARP6(79-183) (Figure 6A), peak intensities decreased by 70% within 1 hour (Figure 6C). Peak intensities decreased by 90% within 3 hours (Figure 6E) compared to the intensity of the original spectrum. The resonances observed over the course of these measurements do not show chemical shift perturbations and no other resonances are observed, indicating that the loss in signal can be attributed to the loss of the monomeric protein species without forming any appreciable amounts of soluble aggregates. Signal loss can therefore be attributed to the formation of large (solid) protein aggregates that do not present with resonances in solution NMR spectroscopic experiments. Furthermore, the previously clear NMR sample exhibited significant precipitation after being removed from the magnet at around 4 hours after starting the initial experiment (Figure S2). While the process of precipitation is naturally accelerated by temperature, sample concentration, and buffer conditions, these results clearly indicate an intrinsic instability of the La domain under most conditions.

Surprisingly, the binding of 5'SL RNA to the La domain of HsLARP6 counteracts this intrinsic instability and suspends precipitation entirely under otherwise identical conditions. Following the same experimental setup, the bound La domain was stable at 25°C over the course of 7 days

(Figure 6, right column). Compared to a [^{1}H, ^{15}N]-HSQC spectrum of a freshly prepared sample of 500 μM HsLARP6(79-183):5'SL (1:1) (Figure 6B), peak intensities in [^{1}H, ^{15}N]-HSQC spectra recorded after 1 day and 7 days remained stable (Figure 6D & F, respectively). The [^{1}H, ^{15}N]-HSQC spectra recorded over the course of 7 days show no change in peak intensity or position. In our experience, various NMR samples of the HsLARP6(79-183):5'SL (1:1) complex ranging from concentrations of 250 μM to 1.7 mM measured extensively under identical conditions at 25°C, but stored at 4°C between measurements, remained stable over at least 3 months. This demonstrates the tremendous shift in stability of the bound La domain compared to the unbound La domain.

The exact source of this significant shift in stability remains elusive, however, the fact that almost all residues experience chemical shift perturbations upon RNA binding (Figure 3C) offers perspective. While it is known that the buried surface area of protein-RNA complexes is larger, on average, compared to protein-protein interactions [34], it is still most typical to only see CSPs at and around the binding interface [35, 36]. The presence of CSPs for nearly all residues of the bound-state protein is indicative of a global shift in conformational dynamics upon RNA binding. While the magnitude of CSPs does not indicate a completely different fold of the protein, binding of the RNA imposes an allosteric effect on residues far beyond the binding interface, likely stabilizing the protein fold and thus suppressing the formation of aggregation-prone species.

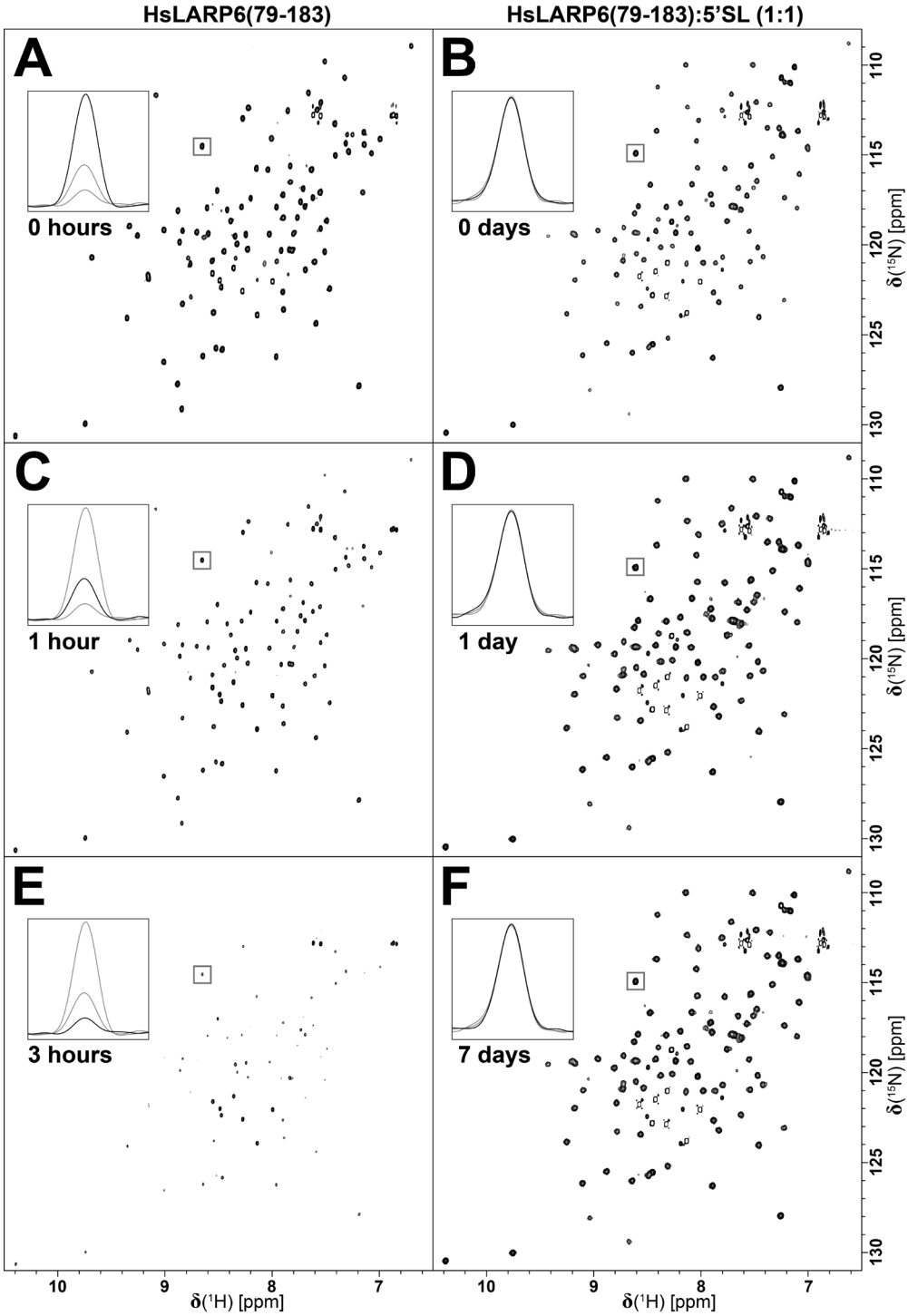

Figure 6: Stability of HsLAPR6(79-183) in the unbound and bound state. Signal position and intensity in 2D [^{1}H, ^{15}N]-HSQC, collected periodically, were employed to monitor physiochemical changes in HsLARP6 La domain. The HsLARP6(79-183), in the absence of 5'SL RNA, remains soluble at 500 μM after reaching 25°C (A) but rapidly precipitates losing nearly 70% of detectable protein in 1h (B). After 3 h, the remaining signal is only ~10% of the initial intensity (C). HsLARP6(79-183) bound to 5'SL RNA is far more stable. Signal intensity at 0 h (D) remains constant after 24 h (E) and even up to 7 (F). Insets show the 1D ^{1}H proton dimension of the first transient of the corresponding HSQC for both bound and unbound samples. All spectra are processed identically and scaled to the same contour level and contour increment.

Conclusions

The stability and binding of 5'SL RNA to the La domain of HsLARP6 was studied here. Varying the length of the N-terminus between residues 73 to 84 heavily modulates the pI of the La domain to the presence of a large number of acidic residues in the N-terminus of the La domain. The length of these N-terminal truncations does not affect 5'SL binding. However, removal of residues from the C-terminus before residue 183 greatly destabilizes the HsLARP6 La domain. Determining features that stabilize the HsLARP6 La domain will greatly improve the potential for the structural, and broader functional, characterization of RNA binding. Interestingly, the stability of the La domain is significantly increased when bound to 5'SL, even at high concentrations. The ribonucleoprotein complex formed is identical in NMR spectra for 7 days at room temperature while the unbound La domain in comparison shows ~90% loss in signal intensity within 3 hours under otherwise identical conditions. This work provides a platform for advances in RNA binding and studies into HsLARP6 as a potential target for antifibrotic therapeutics by simplifying the model of the interaction.

Materials and Methods

Constructs

Sequences for various HsLARP6 La domain constructs (UniProtKB accession number Q9BRS8, Table S1) named His_6-TEV-HsLARP6(73-183), His_6-TEV-HsLARP6(79-183), His_6-TEV-HsLARP6(84-183), His_6-TEV-HsLARP6(73-176), His_6-TEV-HsLARP6(73-171), and His_6-TEV-LARP(73-183) as well as a cysteine point mutation at position 74 named His_6-HsLARP(73-183)G74C were cloned into a pET28a vectors via *Nco*I and *Bam*HI restriction sites for expression in *E. coli* (Genscript USA Inc.). A tobacco etch virus (TEV) protease cleavage site (ENLYFQ▼G) was inserted between a six histidine fusion tag (His_6 tag) and the first residue of the native sequence.

Chemicals

All reagents and chemicals were purchased from commercial suppliers and used as obtained. For a complete list of chemicals and their abbreviations, see Table S2.

Expression of HsLARP6 La Domain

The plasmids coding for His_6-TEV-HsLARP6 La domain was transformed into chemically competent *E. coli* Rosetta II (DE3) (Sigma-Aldrich) for recombinant overexpression. Transformants were plated on LB-agar supplemented with 20 g/L D-glucose, 40 ug/mL kanamycin monosulfate, and 60 ug/mL chloramphenicol and grown at 37°C overnight. For precultures, 250 ml Erlenmeyer flasks were charged with 50 ml of lysogeny broth (LB) medium containing 10 g tryptone, 5 g yeast extract, and 10 g NaCl and autoclaved. Before use, precultures were supplemented with 10 g/L D-glucose, 40 ug/mL kanamycin monosulfate, and 60 ug/mL

chloramphenicol. Precultures were inoculated with colonies from freshly transformed plates and grown overnight at 37°C and 220 rpm in a MaxQ™ 8000 orbital shaker (Thermo Scientific, USA). For main cultures, 2.8 L baffled Fernbach flasks were charged with 1 L of LB medium containing 10 g tryptone, 5 g yeast extract, and 10 g NaCl and autoclaved. Before use, precultures were supplemented with 10 g/L D-glucose, 40 ug/mL kanamycin monosulfate, and 60 ug/mL chloramphenicol. Main cultures were inoculated from precultures to an OD_{600} of ~0.1. Cultures were incubated at 37°C and 220 rpm until an OD_{600} of 0.6-0.8 was reached. The cultures were subsequently cold-shocked by transferring them into an ice-water bath for 30 min. Recombinant expression was induced by addition of IPTG to a final concentration of 1 mM and was allowed to continue at 18°C and 220 rpm in an MaxQ™ 8000 orbital shaker for 16 h. Cells were collected by centrifugation at 4,000 × g and 4°C for 30 min using a Beckman Coulter™ Avanti™ J-20 XPI centrifuge equipped with a JLA-8.1000 rotor before discarding the supernatant. Cell pellets were subsequently resuspended in lysis buffer containing 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, 0.1% w/v CHAPS, 70 mM imidazole, and 5% v/v glycerol. A protease inhibitor mix was added to the cell suspension at final concentrations of 500 μM AEBSF, 1 μM E-64, 1 μM leupeptin, and 150 nM aprotinin before they were either flash frozen in LN_2 and stored at -20°C or immediately subjected to purification. All buffers and solutions used throughout all purifications were formulated in 0.1% v/v DEPC-treated ultrapure water. For His_6-TEV-HsLARP6(73-183)G74C, all buffers contained 5 mM 2-mercaptoethanol.

Purification of HsLARP6(79-183)

Cells were lysed by running the cell suspension through a Microfluidizer® M-110L (Microfluidics) in 3-5 cycles at a pressure of 50 PSI while continuously cooling the lysate to 4°C using a water-ice bath. The lysate was clarified by centrifugation at 37,500 × g and 4°C using a Beckman Coulter™ Avanti™ J-20 XPI centrifuge equipped with a JA-25.50 rotor. Subsequently, the supernatant was filtered through a 0.8 µm polyethersulfone (PES) membrane (Millipore-Sigma). Immobilized metal affinity chromatography (IMAC) was performed on a ÄKTA Start™ instrument (Cytiva, USA) monitoring absorbance at 280 nm. Clarified cell lysate was subsequently purified by immobilized metal affinity chromatography (IMAC); applied at a flow rate of 5 ml/min to a 5 mL HisTrap™ HP column (Cytiva, USA) previously equilibrated with HisTrap buffer A containing 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, 0.1% w/v CHAPS, and 70 mM imidazole. The column was then washed at a flow rate of 5 ml/min with HisTrap buffer A until the absorbance at 280 nm reached a stable baseline. Furthermore, the column rigorously washed at a flow rate of 5 ml/min with a high-salt wash buffer containing 20 mM Tris-HCl pH 7.4, 1 M NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, 0.1% w/v CHAPS, and 70 mM imidazole until the absorbance at 280 nm reached a stable baseline. The column was returned to the low-salt condition by washing the column with HisTrap A buffer until conductivity reached a stable baseline. Subsequently, His_6-TEV-HsLARP6(79-183) was eluted in a one-step gradient using a HisTrap B buffer containing 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, 0.1% w/v CHAPS, and 500 mM imidazole. His_6-tag removal by TEV digestion [37] was performed by adding 1 mL of 1 mg/mL in-house preparations of His_6-TEV protease [38] and placing it in a Spectra/Por® 3 dialysis membrane with a molecular weight cutoff (MWCO) of 3.5 kDa. The sample was dialyzed against

1 L of dialysis buffer containing 20 mM Tris-HCl pH 7.4, 200 mM NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, 0.1% w/v CHAPS, 70 mM imidazole, and 5 mM 2-mercaptoethanol for 16 hrs. The His_6-tag, His_6-tagged TEV protease and other impurities were removed by reverse HisTrap purification. The digested sample was applied at a flow rate of 5 ml/min to a 5 mL HisTrap™ HP column (Cytiva, USA) previously equilibrated with HisTrap A buffer. HsLARP6(79-183) was collected in the flow-through and wash fractions. The target protein was further purified by cation exchange chromatography using a 5 mL HiTrap™ SP Fast Flow column (Cytiva, USA) previously equilibrated with SP buffer A containing 20 mM MES pH 6.5, 50 mM monosodium L-glutamate, 50 mM of L-arginine, and 0.1% w/v CHAPS. The column was then washed at a flow rate of 5 ml/min with SP buffer A until the absorbance at 280 nm reached a stable baseline. Subsequently, HsLARP6(79-183) was eluted in a one-step gradient using a SP buffer B containing 20 mM MES pH 6.5, 1 M NaCl, 50 mM monosodium L-glutamate, 50 mM of L-arginine, and 0.1% w/v CHAPS. HsLARP6 La domain was further purified and buffer exchanged by size exclusion chromatography (SEC) using an ÄKTA pure™ 25 (Cytiva, USA) monitoring absorbance at 260 nm and 280 nm on a HiLoad 16/600 75pg column (Cytiva, USA). The detailed procedure depends on the intended application (*vide infra*). Protein concentration was determined by UV-VIS spectrophotometry using the extinction coefficients listed in Table S1 as determined using the ProtParam tool (ExPasy server). [32] Protein purity was confirmed by sodium-dodecyl-sulfate polyacrylamide electrophoresis (SDS-PAGE) on 4-15% gradient gels (Genscript, USA) stained with 0.0025% w/v each of Coomassie® Brilliant Blue G-250 and R-250 in 10% v/v ethanol, 5% v/v glacial acetic acid in water.

RNA Synthesis and Purification

The 48nt 5'stem-loop (5'SL) RNA (Table S3) from the mRNA of the α_2(I) polypeptide chain of human type 1 collagen was used for all binding assays. 5'SL RNA was synthesized by in vitro transcription (IVT) using in-house preparations of His_6-T7 RNA polymerase P266L. [39-41] IVT reactions were performed as described previously. [35] 500 nM T7 promoter sequence strand DNA oligonucleotides and the complementary template strand for the 5'SL RNA sequence containing 2'-methoxy modified bases at the last and penultimate residues (Integrated DNA Technologies, USA).[42] Reactions were incubated at 37°C for 16 h, then stopped by addition of EDTA to a final concentration of 50 mM. Protein components were removed from the solution by three rounds of extraction with an equal volume of 25:24:1 phenol:chloroform:2-methylbutanol. Excess phenol in solution was removed by three rounds of extraction using a 3:1 volume ratio of diethyl ether to IVT solution. Leftover diethyl ether was removed by heating to 50°C until it had boiled off. RNA solutions were diluted with an 8M urea solution to 5mL, ~4M urea to assist in the subsequent denaturing purification step. 5'SL RNA was further purified by denaturing SEC using a Hiload 16/600 75pg size exclusion column equilibrated with 20mM Tris-HCl pH 7.2, 200 mM KCl, and 7 M urea. Fractions containing purified 5'SL RNA were pooled and diluted with isopropyl alcohol to a 1:3 ratio and supplemented with ammonium acetate pH 5.2 to a final concentration of 300 mM. Precipitation reactions were incubated at 4°C for 16 h, RNA pellets were isolated by centrifugation for 15 min at 4000 × g at 4°C in a Beckman Coulter™ Allegra™ X-15R centrifuge with SX4750 swing bucket rotor. Pellets were washed three times with 80% v/v isopropanol to remove salt, then heated to 85°C for 1 h to remove excess isopropanol resulting in a glassy solid RNA pellet. Washed pellets were dissolved in 0.1% v/v DEPC-treated water to a final concentration of 20 μM RNA and 5 mM EDTA pH 7.4. 5'SL RNA was refolded at 95°C for 1 h then transferred to 4°C until the

solution was completely cooled. Refolded 5'SL RNA was exchanged into NMR or MST buffers by size exclusion chromatography (*vide infra*).

RNase Activity Assay

Nuclease activity was performed at each purification step by exposing fluorescently labeled RNA to the fraction of interest. Pooled fractions from the cell lysate supernatant, HisTrap salt wash, HisTrap eluate, SP eluate, and SEC eluate were incubated with a final concentration of 1 μM AF488-RPS6 41mer (Table S3) for 1 h at 37°C. AF488-RPS6 41mer was synthesized following the 5' sequence of ribosomal protein s6 mRNA (AF488-RPS6 41mer) (Integrated DNA Technologies, USA). Positive and negative controls were composed of 1 μM AF488-RPS6 41mer solutions supplemented with 0.1 mg/mL RNase A or 0.1% v/v DEPC treated water, respectively. Samples were quenched by the addition of equal volume of sample loading buffer containing 90% v/v formamide in water, 10mM Tris-HCl pH 8.0, 10mM boric acid, 0.2 mM EDTA and heating to 90°C and subsequently analyzed by denaturing PAGE on 12% w/v (19:1 Acryl:Bis-acryl) 8M urea PAGE gel run at 17.5 V/cm for 1 h. Gels were imaged on a ChemiDoc™ MP Imaging System (Bio-Rad Laboratories, Inc.) using the excitation and emission filter settings for AF488.

Microscale Thermophoresis

Microscale thermophoresis was performed using freshly prepared HsLARP6(73-183)G74C, expressed and purified as described above. Purified protein was exchanged into 20 mM MES pH 6.6, 100 mM NaCl, 2 mM TCEP-HCl (Gold Bio) using a Superdex Increase 75pg analytical scale column. 1mL of 50-80 μM (0.66-1.06mg/mL) HsLARP6(73-183) G74C was incubated overnight with a 20x molar excess of fluorescein-5-maleimide. Fluorescently tagged protein (FL-LARP6(73-

183) G74C) was separated from free label by SEC using a Superdex Increase 75pg analytical scale column equilibrated with 20 mM Tris pH 7.25, 100 mM KCl, 0.05% w/v CHAPS, and 2 mM TCEP-HCl (MST buffer). MST buffer was supplemented with 0.1 mg/mL yeast tRNA as a nonspecific competitor and 0.4mg/mL BSA to prevent protein aggregation. The 48nt 5' stem-loop (5'SL) RNA sequence used was from the α2(I) strand of type 1 collagen peptide mRNA. Microscale thermophoresis was performed with a constant concentration of 20 nM FL-LARP6(73-183) G74C and a 16 1:1 serial dilutions of 5' SL RNA ranging from 1.93μM-58.9pM. Thermophoresis of the 16 capillaries were recorded with a 3 s delay period prior to 25 s of IR irradiation, set to medium intensity, and a 3 s recovery period. Thermophoresis profiles were collected for 3 replicate titrations. Averaged fluorescence intensity from before IR laser on-time and 0.5 s to 1.5 s after the IR laser on-time was used to determine the change in normalized florescence intensity (ΔF_{NORM}). [43-45] ΔF_{norm} values were averaged across 3 replicate measurement, then fitted to extract the dissociation constant using an asymmetric fit of a one-site binding model [46] (equation 1) in Prism 9.5.0 (GraphPad Software).

$$Fraction\ of\ bound = Unbound + \frac{(Bound - Unbound)\cdot[5'SL\ RNA]}{K_D + [5'SL\ RNA]} \quad (1)$$

where Bound and Unbound are those respective fractions of RNA populations in solution. Fraction of bound is the dose-response value for each data point. [5'SL RNA] is the total concentration of 5'SL RNA for each data point. K_D is the dissociation constant, and its error was estimated using a 95% confidence interval.

NMR Spectroscopy

[^{15}N]-HsLARP6(79-183) free state protein and a 1:1 [^{15}N]-HsLARP(79-183):5'SL RNA complex samples were prepared at a final concentration of 10 mM MES, 50 mM KCl after addition of 10% v/v D_2O, and 0.01% w/v DSS for internal ^{1}H referencing. ^{13}C and ^{15}N resonances were referenced indirectly. [47] NMR samples were prepared in 5 mm thin-walled Shigemi BMS-005B NMR tubes (Shigemi Co., Japan). Unbound and bound [^{15}N]-HsLARP6 La domain samples were concentrated to 500 μM using Vivaspin® centrifugal concentrators with MWCO of 3.5kDa. All NMR spectra were acquired at 298 K at 16.4 T on a Bruker Avance III NMR spectrometer equipped with a 3 channel ^{1}H-^{13}C/^{15}N-^{2}H, triple resonance TCI cryoprobe. 2D [^{1}H, ^{15}N]-HSQC experiments were recorded with 2048 by 256 complex points, 16 scans per transient.

HsLARP6 NMR Stability Assay

Stability in a typical NMR analysis condition for a ribonucleoprotein complex was evaluated for both the unbound and 5'SL RNA-bound HsLARP6 La domain using 2D [^{1}H-^{15}N]-HSQC experiments.[48] Both species were evaluated by periodic NMR analysis at 25°C. 5'SL-bound La domain was also evaluated by analysis at 25°C with long term storage for 7 days at 4°C between subsequent measurements. Spectra for unbound HsLARP6(79-183) at 25°C were recorded at 0 h, 1 h, and 3 h bound-state samples at 25°C were recorded at 0 h, 24 h, and 168 h. Spectra of the bound-state sample stored at 4°C between measurements were collected at 0 h and 168 h.

Declaration of competing interest

The authors declare that they have no competing conflict of interest.

Acknowledgements

Research reported in this publication was supported by NIGMS of the National Institutes of Health under award number R35GM142912. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would like to thank Branko Stefanovic for fruitful discussion.

Accession Codes

Sequence of Human LARP6: Uniprot Q9BRS8-1

References

[1] Bousquet-Antonelli C, Deragon JM. A comprehensive analysis of the La-motif protein superfamily. RNA. 2009;15:750-64.
[2] Deragon JM. Distribution, organization an evolutionary history of La and LARPs in eukaryotes. RNA Biol. 2021;18:159-67.
[3] Sommer G, Heise T. Role of the RNA-binding protein La in cancer pathobiology. RNA Biol. 2021;18:218-36.
[4] Stavraka C, Blagden S. The La-Related Proteins, a Family with Connections to Cancer. Biomolecules. 2015;5:2701-22.
[5] Xie C, Huang L, Xie S, Xie D, Zhang G, Wang P, et al. LARP1 predict the prognosis for early-stage and AFP-normal hepatocellular carcinoma. J Transl Med. 2013;11:272.
[6] Zou T, Wang PL, Gao Y, Liang WT. Circular RNA_LARP4 is lower expressed and serves as a potential biomarker of ovarian cancer prognosis. Eur Rev Med Pharmacol Sci. 2018;22:7178-82.
[7] Parsons CJ, Stefanovic B, Seki E, Aoyama T, Latour AM, Marzluff WF, et al. Mutation of the 5' untranslated region stem-loop structure inhibits {alpha}1(i) collagen expression in vivo. J Biol Chem. 2011;286:8609-19.
[8] Stefanovic B, Stefanovic L, Manojlovic Z. Imaging of type I procollagen biosynthesis in cells reveals biogenesis in highly organized bodies; Collagenosomes. Matrix Biol Plus. 2021;12:100076.

[9] Makarev E, Izumchenko E, Aihara F, Wysocki PT, Zhu Q, Buzdin A, et al. Common pathway signature in lung and liver fibrosis. Cell Cycle. 2016;15:1667-73.
[10] Karsdal MA, Manon-Jensen T, Genovese F, Kristensen JH, Nielsen MJ, Sand JM, et al. Novel insights into the function and dynamics of extracellular matrix in liver fibrosis. Am J Physiol Gastrointest Liver Physiol. 2015;308:G807-30.
[11] Zhang Y, Stefanovic B. Akt mediated phosphorylation of LARP6; critical step in biosynthesis of type I collagen. Sci Rep. 2016;6:22597.
[12] Wang H, Stefanovic B. Role of LARP6 and Nonmuscle Myosin in Partitioning of Collagen mRNAs to the ER Membrane. PLoS One. 2014;9:e108870.
[13] Blackstock CD, Higashi Y, Sukhanov S, Shai SY, Stefanovic B, Tabony AM, et al. Insulin-like growth factor-1 increases synthesis of collagen type I via induction of the mRNA-binding protein LARP6 expression and binding to the 5' stem-loop of COL1a1 and COL1a2 mRNA. J Biol Chem. 2014;289:7264-74.
[14] Vukmirovic M, Manojlovic Z, Stefanovic B. Serine-threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. Mol Cell Biol. 2013;33:3893-906.
[15] Manojlovic Z, Stefanovic B. A novel role of RNA helicase A in regulation of translation of type I collagen mRNAs. RNA. 2012;18:321-34.
[16] Cai L, Fritz D, Stefanovic L, Stefanovic B. Binding of LARP6 to the conserved 5' stem-loop regulates translation of mRNAs encoding type I collagen. J Mol Biol. 2010;395:309-26.
[17] Zhang Y, Stefanovic B. LARP6 Meets Collagen mRNA: Specific Regulation of Type I Collagen Expression. Int J Mol Sci. 2016;17:419.
[18] Stefanovic B, Manojlovic Z, Vied C, Badger CD, Stefanovic L. Discovery and evaluation of inhibitor of LARP6 as specific antifibrotic compound. Sci Rep. 2019;9:326.
[19] Stefanovic L, Longo L, Zhang Y, Stefanovic B. Characterization of binding of LARP6 to the 5' stem-loop of collagen mRNAs: implications for synthesis of type I collagen. RNA Biol. 2014;11:1386-401.
[20] Al-Ashtal HA, Rubottom CM, Leeper TC, Berman AJ. The LARP1 La-Module recognizes both ends of TOP mRNAs. RNA Biol. 2021;18:248-58.
[21] Aoki K, Adachi S, Homoto M, Kusano H, Koike K, Natsume T. LARP1 specifically recognizes the 3′ terminus of poly(A) mRNA. FEBS Letters. 2013;587:2173-8.
[22] Kozlov G, Mattijssen S, Jiang J, Nyandwi S, Sprules T, Iben James R, et al. Structural basis of 3′-end poly(A) RNA recognition by LARP1. Nucleic Acids Research. 2022;50:9534-47.
[23] Teplova M, Yuan Y-R, Phan AT, Malinina L, Ilin S, Teplov A, et al. Structural Basis for Recognition and Sequestration of UUUOH 3′ Temini of Nascent RNA Polymerase III Transcripts by La, a Rheumatic Disease Autoantigen. Molecular Cell. 2006;21:75-85.
[24] Uchikawa E, Natchiar KS, Han X, Proux F, Roblin P, Zhang E, et al. Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. Nucleic Acids Research. 2015;43:3373-88.
[25] Wang Y, He Y, Wang Y, Yang Y, Singh M, Eichhorn CD, et al. Structure of LARP7 Protein p65–telomerase RNA Complex in Telomerase Revealed by Cryo-EM and NMR. Journal of Molecular Biology. 2023;435:168044.
[26] Kuspert M, Murakawa Y, Schaffler K, Vanselow JT, Wolf E, Juranek S, et al. LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. Rna. 2015;21:1294-305.

[27] Martino L, Pennell S, Kelly G, Busi B, Brown P, Atkinson RA, et al. Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module. Nucleic Acids Res. 2015;43:645-60.
[28] Yamada Y, Mudryj M, de Crombrugghe B. A uniquely conserved regulatory signal is found around the translation initiation site in three different collagen genes. J Biol Chem. 1983;258:14914-9.
[29] Stefanovic L, Gordon BH, Silvers R, Stefanovic B. Characterization of Sequence-Specific Binding of LARP6 to the 5' Stem-Loop of Type I Collagen mRNAs and Implications for Rational Design of Antifibrotic Drugs. J Mol Biol. 2022;434:167394.
[30] Manchester KL. Use of UV methods for measurement of protein and nucleic acid concentrations. Biotechniques. 1996;20:968-70.
[31] Sweet RM, Eisenberg D. Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. J Mol Biol. 1983;171:479-88.
[32] Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005. p. 571-607.
[33] Lizarrondo J, Dock-Bregeon AC, Martino L, Conte MR. Structural dynamics in the La-module of La-related proteins. RNA Biol. 2021;18:194-206.
[34] Kruger DM, Neubacher S, Grossmann TN. Protein-RNA interactions: structural characteristics and hotspot amino acids. RNA. 2018;24:1457-65.
[35] Schnieders R, Knezic B, Zetzsche H, Sudakov A, Matzel T, Richter C, et al. NMR Spectroscopy of Large Functional RNAs: From Sample Preparation to Low-Gamma Detection. Curr Protoc Nucleic Acid Chem. 2020;82:e116.
[36] Furtig B, Richter C, Wohnert J, Schwalbe H. NMR spectroscopy of RNA. Chembiochem. 2003;4:936-62.
[37] Raran-Kurussi S, Cherry S, Zhang D, Waugh DS. Removal of Affinity Tags with TEV Protease. Methods Mol Biol. 2017;1586:221-30.
[38] Tropea JE, Cherry S, Waugh DS. Expression and purification of soluble His(6)-tagged TEV protease. Methods Mol Biol. 2009;498:297-307.
[39] Guillerez J, Lopez PJ, Proux F, Launay H, Dreyfus M. A mutation in T7 RNA polymerase that facilitates promoter clearance. Proc Natl Acad Sci U S A. 2005;102:5958-63.
[40] Ramirez-Tapia LE, Martin CT. New insights into the mechanism of initial transcription: the T7 RNA polymerase mutant P266L transitions to elongation at longer RNA lengths than wild type. J Biol Chem. 2012;287:37352-61.
[41] Ellinger T, Ehricht R. Single-step purification of T7 RNA polymerase with a 6-histidine tag. Biotechniques. 1998;24:718-20.
[42] Kao C, Zheng M, Rudisser S. A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. RNA. 1999;5:1268-72.
[43] Wienken CJ, Baaske P, Rothbauer U, Braun D, Duhr S. Protein-binding assays in biological liquids using microscale thermophoresis. Nat Commun. 2010;1:100.
[44] Jerabek-Willemsen M, Wienken CJ, Braun D, Baaske P, Duhr S. Molecular interaction studies using microscale thermophoresis. Assay Drug Dev Technol. 2011;9:342-53.
[45] Silvers R, Saxena K, Kudlinzki D, Schwalbe H. Recombinant expression and purification of human TATA binding protein using a chimeric fusion. Protein Expression and Purification. 2012;85:142-7.

[46] Hulme EC, Trevethick MA. Ligand binding assays at equilibrium: validation and interpretation. Br J Pharmacol. 2010;161:1219-37.
[47] Harris RK, Becker ED, Cabral De Menezes SM, Granger P, Hoffman RE, Zilm KW, et al. Further conventions for NMR shielding and chemical shifts IUPAC recommendations 2008. Solid State Nucl Magn Reson. 2008;33:41-56.
[48] Lehner F, Kudlinzki D, Richter C, Müller-Werkmeister HM, Eberl KB, Bredenbeck J, et al. Impact of Azidohomoalanine Incorporation on Protein Structure and Ligand Binding. ChemBioChem. 2017;18:2340-50.