STRUCTURAL AND BIOCHEMICAL CHARACTERIZATION OF DOMAINS IN THE POSTTRANSCRIPTIONAL REGULATOR LARP6

By

Eliana L. Peña, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Biochemistry
May 2018

Committee Members:

Karen A. Lewis, Chair

Sean M. Kerwin

Steven T. Whitten

Lisa R. Warner

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

DEDICATION

For my family.

ACKNOWLEDGEMENTS

It's amazing to think that the same person, who, years ago, never thought she would ever accomplish anything is now this far along in her career. I know that this journey was not a lonely one. Throughout the years, there have been many people who have supported me in every aspect. I want to first and foremost thank my immediate family. Especially my mom, Susana, and my dad, Lawrence, for helping me both emotionally and financially throughout the years. They were always on the sidelines rooting for me, especially when the going got tough, and when it was too much to handle at times, they always were there to help pick me back up.

I like to thank my mentor Dr. Karen A. Lewis, who, throughout the years became one of the most influential person in my life. I will never forget the Spring semester of 2015 (during my Junior year of undergraduate) when I visited her consistently almost every few weeks to talk about the science she was doing. There was something about how well we communicated that made me realize that this was the person I wanted to work for. Finally, during the following summer a spot opened up for me to join her lab. I was a super grateful and eager student ready to learn everything I can. Her first-hand approach to teaching me was something I deeply admired about her. Throughout the years she has been not only a great principal investigator and mentor, but she has been an amazing inspiration for me to not sweat the small stuff, take care of myself, and be the

best young scientist I could be. She believed in me even when I didn't, and although she may not know it, it truly meant a lot to me Thank you.

The lab environment I grew up in (and yes, it feels like I grew up in this lab) was definitely accompanied by some really amazing people. I knew Samantha Zepeda as an undergraduate and was immediately blown away by how well she spoke. Little does she know, I aspire to have the same level of proficiency of the English language like she does. There is Melissa Carrizales, who is like a little sister to me. Her strong work ethic and outgoing personality reminds me so much of me and I really enjoyed our (sometimes lengthy) discussions about life, relationships, and what in the world to do with our lives. I know she is destined for greatness. Then there is Leticia (Lety) Gonzalez, who could always bring me to laughter. In the year that she has been in the lab, she has been an amazing friend, always there to listen to me when I need to vent or freak out about an upcoming talk. I will miss the laughing, joking around, and gossiping that accompanied our sushi lunch dates! Then there are the three previous graduate students from the lab: Francisco Betancourt, Eliseo Salas, and Jose M. Castro. These three were like big brothers to me, always being a point of resource, laughter, and sometimes some heated bickering. Most importantly, I want to thank Jose Castro, who was in the LewisKA lab when I first joined the lab and to whom I always looked up. He showed me almost everything I know now and inspired me to keep being the "sassy, smart, and caring person I am". I know he is always a phone call away to hear me freak out about whatever

big event is happening in my life. You all three have moved on now, but the impact you guys have left me will always remain.

My San Marcos friend group, I know my only connection with you all in the beginning was through my now fiancé, but we grew together through many meme-filled shenanigans. You guys really gave me a family away from home, and I couldn't be more grateful for that.

I would also like to thank a person very dear to me: my fiancé Rico Bendimez. He has been the biggest emotional supporter throughout my time in the master's program. He always encouraged me to get back up and keep pushing even when I was tired and weary. But, on the days where I had no more fight in me, he was always there to pick up the pieces, wrap me in a blanket, feed me pizza and snacks, hug me and put on a Harry Potter marathon. He showed me how to gain my assertive voice, push boundaries I never thought were possible, stand up for what I believe in and believe that I can do whatever I set my mind to.

And finally, this opportunity to pursue my M.S. degree in Biochemistry would not have been possible without the support from the South Texas Doctoral Bridge Program and its principal investigators: Dr. Babatunde Oyajobi, Dr. Nicquet Blake, Dr. Rachel Booth, Dr. Ronald Walter and Dr. Rachel Salinas. These individuals have made been an amazing source of professional (and sometimes personal) help throughout my time here in the Master's program at Texas State University.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

ABSTRACT

The La-related protein 6 (LARP6) is an RNA binding protein that is involved in the regulation of Type-I collagen biosynthesis. Though the mechanism of binding remains to be determined, it is known this protein binds to a stem loop structure found in the 5' untranslated region of mRNAs that codes for Type-I collagen through a highly conserved binding domain known as the La Module. The LARP6 protein also contains an uncharacterized LSA domain which contains a well-conserved cysteine. The goal of this study was to characterize the structural significance and contribution to biochemical activity of the conserved domains found in the LARP6 protein. Isolated La Module domain constructs from human, zebrafish, and platyfish LARP6 proteins were used to determine the domain boundaries of each species. Limited proteolysis studies were used to determine the global stability of the isolated La Modules as well as suitability for future structural characterization. RNA binding activity of the three isolated La Modules from each species were measured against the stem loop structure present in the human COL1a1 mRNA using electrophoretic mobility shift assays. The conserved cysteines found in the RRM and LSA domains were targeted for serine mutagenesis in the full-length human protein construct, recombinantly expressed, and purified. Size exclusion chromatography of these serine mutant proteins implicate the cysteine in the LSA in an intramolecular disulfide bond.

I. INTRODUCTION

In all organisms, protein production is a highly-regulated process. In eukaryotes, the spatial separation of transcription and translation adds an additional layer of regulation, including the ability to control the processing of the pre-mRNA in the nucleus and the export of the mature mRNA transcripts to the cytosol.[1] Many regulatory mechanisms control this flow of information; however, none are as quickly responsive as post-transcriptional regulation.[2, 3] After the synthesis of the mRNA, a series of RNA binding proteins (RBPs) associate with the RNA transcript, and can control nuclear export rate into the cytosol, translocation and compartmentalization within the cytosol, assembly and processivity of the ribosome during translation, and protection from degradation.[3, 4]

One set of such proteins is the superfamily of La-related protein (LARPs), which bind to a variety of RNA ligands.[5] LARPs are characterized by a bipartite RNA-binding domain known as the "La module", comprised of a highly-conserved "La motif" and an RNA recognition motif (RRM).[6] The first LARP discovered was the LARP3 protein (formerly known as the "genuine La protein"), which binds to pre-tRNAs and pre-snoRNAs to chaperone their processing and folding (Figure 1).[6, 7] There are now six La-related protein subfamilies, and each appears to associate with a particular class of RNAs in the cell, ranging from pre-tRNAs to mature mRNAs. By binding to these RNA ligands (**Figure 1**). LARPs carry out a variety of post-transcriptional regulatory functions that are only just beginning to be understood.

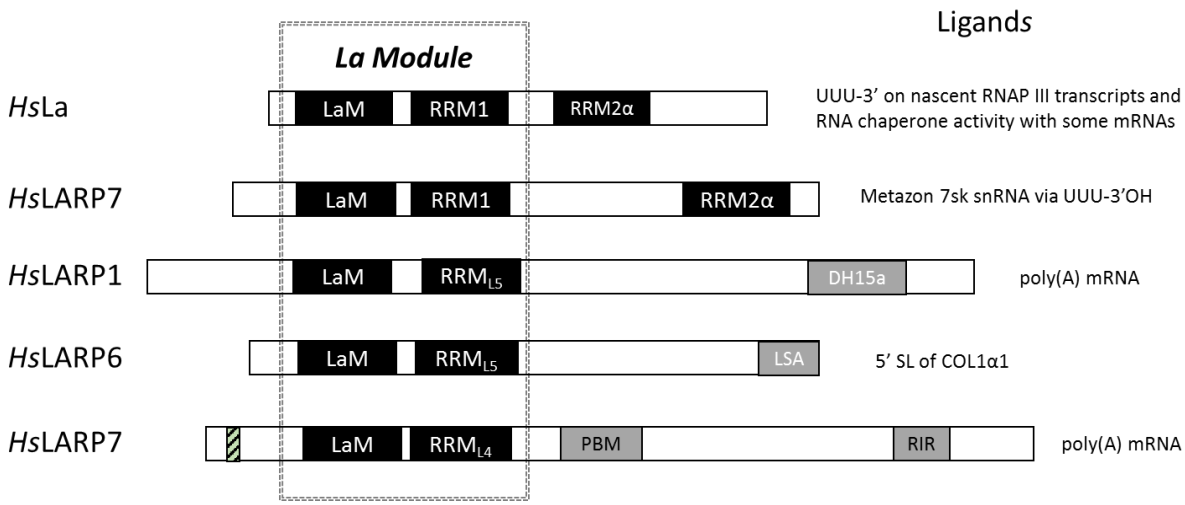

Figure 1. Topological representation of LARP proteins and their function. The genuine La, LARP7, LARP6 and LARP4A/B contain the bipartite La Module domain, which consist of a La Motif (LaM) and RNA Recognition Motif (RRM) joined by a flexible linker. Various other domains can be seen such as a secondary RRM (RRM2), Poly(A)-binding protein interacting motif-2 (PAM), DM15, and LaM and S1-associated motif (LSA). Adapted from **[8]**

Of the La-related proteins, the La-related protein 6 (LARP6) family is the least characterized. The only established function of LARP6 is to regulate collagen expression by binding to the mRNA that encodes collagen type I. [9-12]. When LARP6 expression increases, the expression of collagen type I also increases. A better understanding of the signals involved in upregulating LARP6 could lead to advances in treatment for fibrotic diseases, which are caused by excessive amounts of collagen type I.[11, 13]

La-related Protein 6 (LARP6)

In addition to the La module, the LARP6 subfamily is distinguished from other LARPs by a conserved sequence at the C-terminus of the protein called the "LSA motif" (for "La Module- and SS1-associated domain).[7] The molecular structures of the two La module subdomains from human LARP6 (*Hs*LARP6) have been solved via solution NMR (**Figure 2**).[9] The first domain is canonically known as the La Motif, comprised of

five α-helices arranged in a winged helix-turn-helix fold. The second domain, the RNA Recognition Motif (RRM), is connected to the La motif through a flexible linker. This RRM is α/β motif with an anti-parallel β pleated sheet at the core. In other RRM containing proteins, this domain is the sole binding domain.[14] However, in the case of LARP6 and the other La-related proteins, all three elements (La motif, RRM, and interdomain linker) are necessary for binding activity.[9]

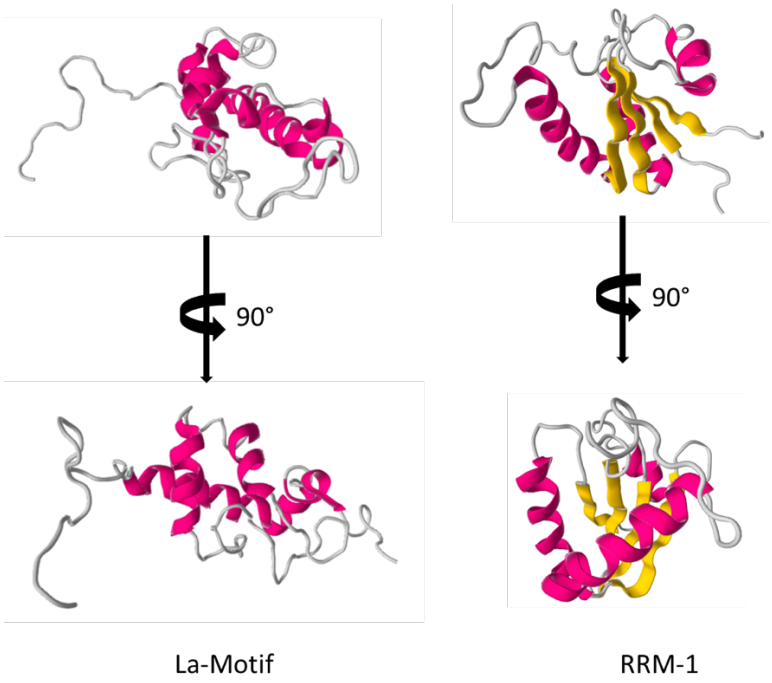

Figure 2. Structures of *Hs*LARP6 La Motif and RRM. (left) The *Hs*LARP6 La Motif is composed of five α-helices (PDB: 1MTF). (right) The *Hs*LARP6 RRM is mixed α/ß structure, with a core ß-sheet (PDB: 1MTG).[9] Images generated with Autodesk Molecule Viewer.

Currently, the only known endogenous ligand for LARP6 is a stem loop structure in the 5' untranslated region of collagen type I.[11] A general cellular function for a LARP6/RNA interaction remains to be determined.[5] LARP6 has been reported to

function in a variety of pathways, including the translocation of the mRNA from the nucleus to the cytosol, recruitment of other translation factors, and the correct compartmentalization of the collagen type I protein to the endoplasmic reticulum.[12, 15] Although there is a known RNA ligand, the mechanism of binding has yet to be fully elucidated. In particular, the RNA binding surface on the RRM and the orientation of each subdomain in relationship to each other during the binding event is unknown.

Previous research in our laboratory established that two orthologues of LARP6 from two fish species, zebrafish and platyfish, are suitable model proteins for structural and functional studies of LARP6.[16] The two fish LARP6 proteins were found to contain a protease-resistant 40 kDa region, in which was identified via mass spectrometry to be the uncharacterized N-terminal region and the RNA-binding La Module: residues 1-280 for *X. maculatus* (platyfish) and 1-290 in *D. rerio* (zebrafish). [16] When the human LARP6 protein was subjected to limited proteolysis, a similar protease-resistant region was observed, but the magnitude of resistance to proteolysis was noticeably less than that of the fish orthologues. To determine the whether the proteolytic resistance involves a stable interaction of the N-terminal domain and the La Module, each domain needs to be independently produced and characterized for comparison to the longer constructs.

The C-terminal domain of LARP6 is increasingly implicated in critical intermolecular interactions. Neither the function nor the structure of the LSA domain has been determined. It has been hypothesized by others that the LSA contributes to ligand selectivity and/or protein-protein interactions.[6, 13] Computational analysis of the C-

terminal domain of LARP6, including the LSA, predicts that this region is disordered, which is supported by limited proteolysis of the recombinant protein.[16] Disordered protein domains are well-established to be involved in protein-protein interactions which could allow for intermolecular interaction between LARP6 C-terminal regions (including the LSA) and other proteins, or intramolecular interactions between different domains of a single LARP6 molecule.[17-19]

A study using cultured mammalian cells found that the C-terminal domain of LARP6 protein interacts with the Serine-Threonine Receptor Associated Kinase (STRAP).[12, 20] A second study from the same research group showed that LARP6 is phosphorylated at multiple serines in the C-terminus, at positions 348, 396, 409, 421, 447, and 451. Two of these serines, Ser348 and Ser409, are directly phosphorylated by the mTOR kinase.[20] Phosphorylation of these two serines is important both for interaction with STRAP and for upregulating collagen type I biosynthesis.[21]

A multiple sequence alignment of vertebrate LARP6 proteins identified several conserved cysteines in the C-terminal domain (**Figure 3**). Most notably, a fully-conserved cysteine is located at position 258 in the RRM (**Figure 3**). Within individual classes of the Vertebrata subphylum (mammals, birds, amphibians, and bony fish), there are additional conserved cysteines within the C-terminal domain. However, the exact location of these cysteines varies across the classes. For example, mammals and birds share a highly-conserved cysteine, located at position 416 in human LARP6 (**Figure 4**). However, only mammals contain a cysteine at human position 378 (**Figure 4**). In

contrast, birds contain two additional cysteines within the RRM (**Figure 3**). Interestingly, all but two of the vertebrate sequences that were evaluated contain a cysteine within the LSA domain, located at position 490 in the human protein (**Figure 4**).

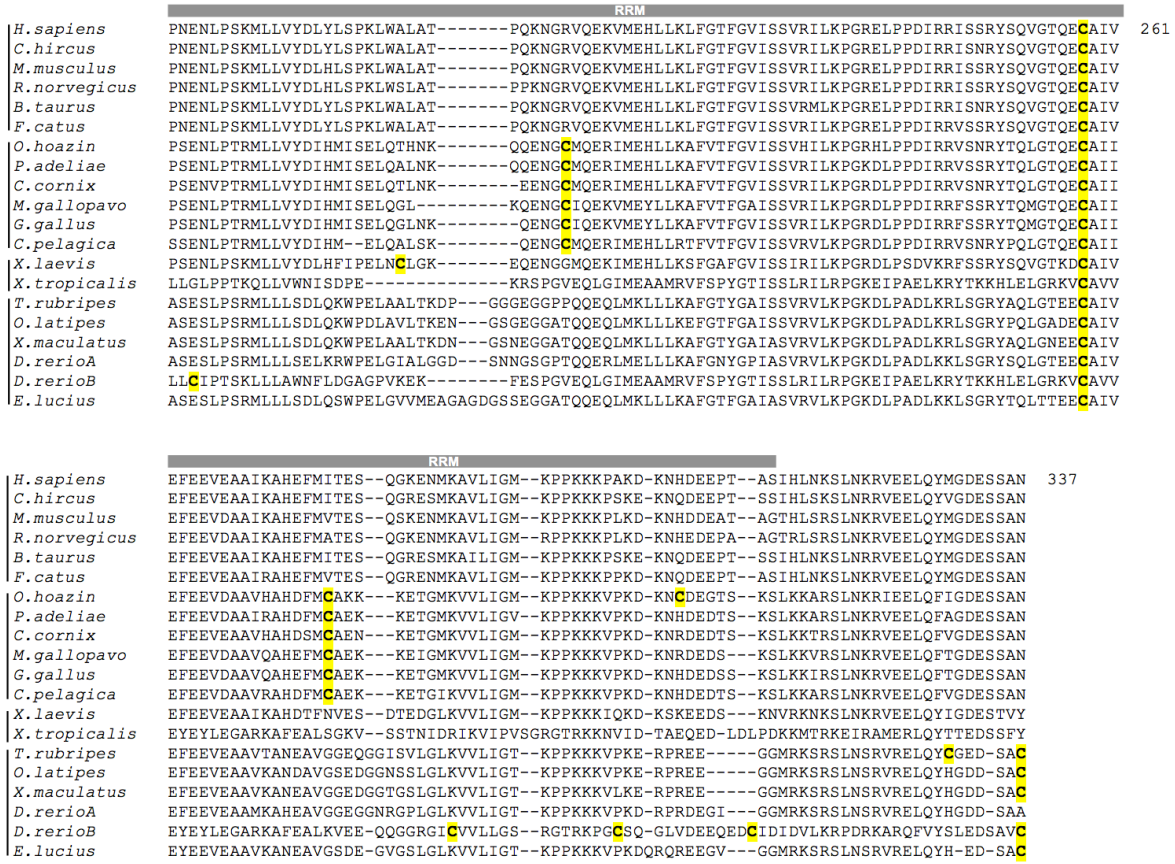

Figure 3. Multiple alignment of vertebrate LARP6 RRM Sequences. Vertebrate sequences were aligned with PROMALS3D[22], and vertically organize by class as indicated by vertical bars on left; in descending order are mammals, birds, amphibians, and bony fish. The domain topology is marked above the sequences, identifying the structurally characterized RRM. Cysteine residues are highlighted in bold text and yellow background. The numbering on the right is for the human sequence.

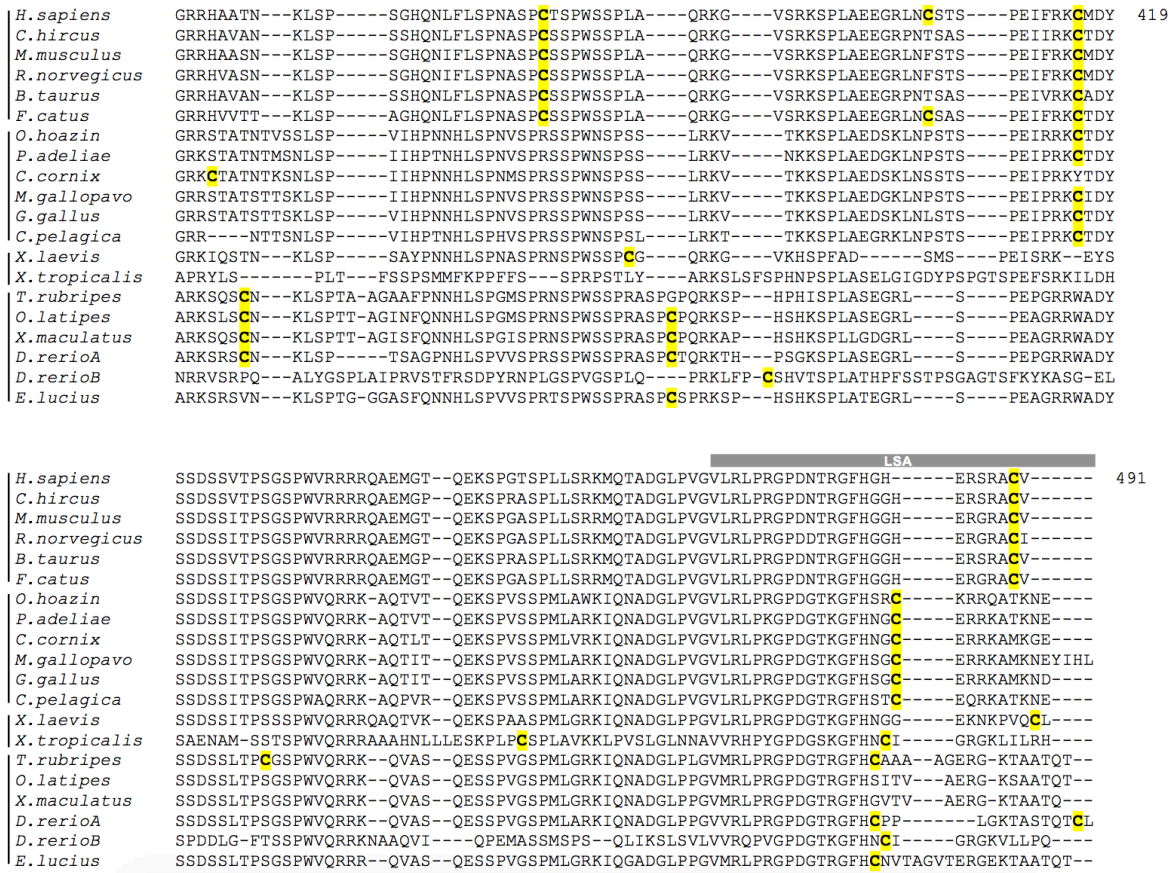

Figure 4. Multiple alignment of vertebrate LARP6 C-terminal sequences. Vertebrate sequences were aligned with PROMALS3D[22], and vertically organized by class as indicated by vertical bars on left; in descending order are mammals, birds, amphibians, and bony fish. The domain topology is marked above the sequences, identifying the LSA motif that was identified via bioinformatics analysis[6]. Cysteine residues are highlighted in bold text and yellow background. The numbering on the right is for the human sequence.

Collagen synthesis is down-regulated when fibroblasts are exposed to hydrogen peroxide, which induces oxidative stress.[23-25] Both the mRNA encoding the alpha subunit of collagen type I (*COL1A1*) as well as the protein were observed to be degraded. It is known that macromolecules are subjected to oxidation by these ROS, which often

can cause conformational changes in the protein. Because LARP6 is a critical regulatory factor for stimulating collagen type I synthesis, we hypothesize that the absence of LARP6 on the *COL1A1* mRNA allows the mRNA to be susceptible to degradation. For this mechanism to function, however, LARP6 would have to be responsive to the redox environment of the cell.

As discussed above, the exact function of the LSA domain has not yet been identified in LARP6. This sequence is also found in cold shock response protein 1 (CSP1) and has been hypothesized to be in involved in ligand selectivity.[6] The predicted disorder of the entire C-terminal domain of LARP6 could allow the structural flexibility that would be required for cysteine residues to interact and form disulfide bonds under oxidative conditions.[16]

Oxidative Regulation via Disulfide Formation

There is considerable precedent for cysteine-containing proteins to respond to oxidative stress via conformational changes. During periods of oxidative stress, reactive oxidative species (ROS) are generated in the cytosol, which can target macromolecules for oxidation.[26, 27] When a protein contains two cysteines, the free sulfhydryls may oxidize and form an intramolecular disulfide linkage. This new disulfide bond alters the structural conformation of the protein and consequently changes the protein's function.[24] This process is reversible in the presence of a reducing agent such as glutathione (GSH) or dithiothreitol (DTT) (**Figure 5**).

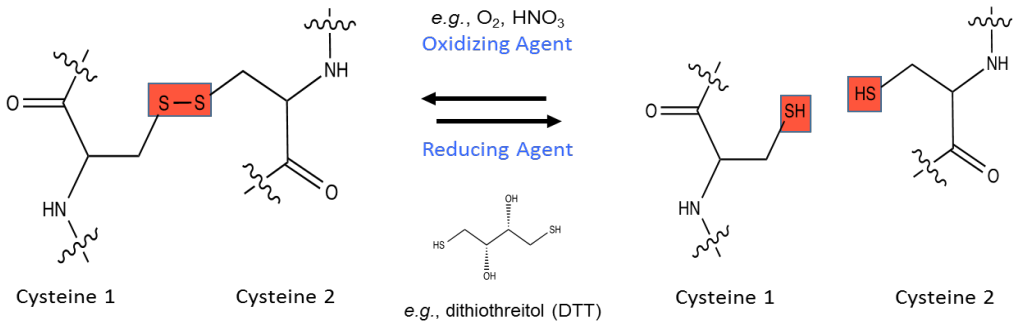

Figure 5. Redox state of cysteine residues. Oxidation of free cysteine (sulfhydryl) functional groups generates a covalent disulfide bond through generation of a thiol anion on one of the free sulfhydryls. Reduction of the disulfide linkage generates two sulfhydryl functional groups.

In normal cytosolic conditions, cysteine residues remain in their reduced sulfhydryl states. The thioredoxin system exploits the chemistry of cysteine residues to protect other cytosolic proteins from aberrant oxidation.[28] This is a classic example of the level of control resulting from the exchange of electronic states of the cysteines present in proteins. However, during periods of oxidative stress, these endogenous reducing systems of the cytosol can be overwhelmed, and cytosolic cysteines can be targeted by oxidants such as reactive oxidative species (ROS), thus generating the aberrant disulfide linkages.[29] Several redox-responsive proteins have evolved to use cysteine chemistry as a sensor of changes in cellular oxidative conditions.

In non-catalytic proteins, the effect of redox-induced disulfide formation could result in either a negative or positive effect. Two examples are the proteins OxyR and Protein Kinase C. In the case of the transcription factor OxyR, disulfide formation found in the dimerization interface temporarily locks the protein in the active conformation for cooperative binding to the RNA polymerase.[30] Physiologically, this conformational

shift promotes the transcription and translation of reducing agents to re-establish homeostasis in the cell, resulting in a positive response by oxidative stress.

In contrast, oxidation of Protein Kinase C disables the signaling activity of the protein.[31] The formation of an intramolecular disulfide bond induces a conformational change that shields the phosphorylation site of the protein. By occluding this activation site, all downstream signaling is blocked. It is hypothesized evolutionarily; the cell developed this form protection, to deter from tumorigenesis caused by over activation.[31]

LARP6 and oxidative stress have opposing effects on collagen type I synthesis. If LARP6 undergoes a conformational switch in response to oxidative stress, it could ultimately be the basis of oxidative regulation of collagen biosynthesis *in vivo*. To determine whether LARP6 directly responds to oxidative stress, we first evaluated whether there are cysteines in the C-terminal portion of LARP6 that could react to oxidizing conditions. If so, the formation of a disulfide bond may alter the structure of the C-terminal domain to affect the phosphorylation state of LARP6, the RNA binding activity, or both.

Research Plan

The goal of this project is to characterize the structural and functional roles of the N-terminal and C-terminal domains in LARP6. Following identification of the La Module structural boundaries in the human (*Hs*), zebrafish (*Dr*), and platyfish (*Xm*)

LARP6 proteins, the RNA binding activity of the isolated domain will be measured in comparison to the full-length protein. These proteins will also be used for future high-resolution structural characterization. Additionally, the contribution of the conserved cysteines in the C-terminal domain to the overall structure and RNA binding activity of LARP6 will be evaluated using the full-length *Hs*LARP6.

Aim 1. Determine the La Module structural boundaries and RNA binding activity.

The putative sequences that code for the La Module in each species will be recombinantly expressed in *E. coli* and purified by nickel-affinity and gel filtration chromatography as previously described [10]. Previous literature determined that the isolated La Module domains were able to fold independently from the rest of the protein.[9] We hypothesized that because the full-length fish LARP6 proteins are more structurally stable than the human (*Hs*) LARP6 protein, the La Module domains would share this characteristic. The La Modules of zebrafish (*Dr*), platyfish (*Xm*) and human (*Hs*) will be molecularly cloned into a pET28a expression plasmid and recombinantly expressed and purified using an N-terminal His_6-tag. To determine where the structural boundaries of each La Module protein begins, each protein will be analyzed by limited proteolysis using the exo-protease aminopeptidase, and expression construct boundaries revised as needed. To establish proper folding of the three La Modules, each module will be assessed for binding activity against the stem-loops found in *COL1a1* and *COL1a2* of each of the three species. The apparent binding affinity, $K_{D,\,app}$, will be measured through electrophoretic mobility shift assay. Additionally, to prepare our proteins for structural

characterization using solution NMR and small-angle X-ray scattering, the N-terminal His_6-tag will be removed using thrombin cleavage.

Aim 2. Identify whether the LARP6 C-terminal domain contains a disulfide bond.

The conserved cysteines in human LARP6 protein reside at positions 258, 378, 410, and 490. Each of these cysteines will be mutated to serine using standard site directed mutagenesis of a previously constructed pET28a-*Hs*LARP6 expression plasmid, which encodes an N-terminal His_6 tag. Serine was selected as the mutation to conserve the overall electrochemical profile, while disabling disulfide bond formation. The mutagenized *Hs*LARP6 coding sequences will be recombinantly expressed in *E. coli*, and the proteins purified by nickel-affinity and gel filtration chromatography as previously described.[10, 16] We hypothesize that disulfide bonds may contribute to LARP6 structure. To determine whether Cys-to-Ser mutations affect protein structure, the apparent molecular weight of the purified protein will be determined by size exclusion chromatography in both the presence and absence of DTT. If the protein forms oligomers in either reducing or oxidizing conditions, the accurate identification of higher-order oligomers will be critical for verifying changes in stability as a potential biological function, as opposed to an experimental artifact.

II. MATERIALS AND METHODS

Site Directed Mutagenesis

Cysteine residues suspected of engaging in disulfide bonding were targeted for site-directed mutagenesis to serine using complimentary primers containing a single point mutation in the codon (**Table 1**). All DNA oligonucleotides were commercially synthesized from Integrated DNA Technology (Coralville, IA). The *Hs*LARP6a-pET28a expression plasmid generously gifted by Dr. Mark Bayfield from York University (Toronto, Ontario).

Introduction of point mutation was performed on pET28a-*Hs*LARP6a expression plasmid construct (Figure 6) using PCR-mediated site-directed mutagenesis. The mutagenized plasmid was selected for by subjecting the amplification product to DpnI restriction enzyme to degrade parental plasmid and transformed into *Escherichia coli* DH5α for amplification and purification. Purified DNA plasmids were then subjected to commercial Sanger sequencing to confirm the incorporation of the intended point mutation.

Table 1. Site-Directed Mutagenesis Primers

Protein	*Primer Name*	*Sequence*	*Purpose*
***Hs*LARP6-C258S**	KAL064	5'-GTGGGGACCCAGGAGAGCGCCATCGTGGAGTTC-3'	Fwd
	KAL065	5'-GAACTCCACGATGGCGCTCTCCTGGGTCCCCAC-3'	Rev
***Hs*LARP6-C378S**	KAL066	5'-GAGAGGAGCAGGGCCTCTGTATAAGCGGCCGC-3'	Fwd
	KAL067	5'-GCGGCCGCTTATACAGAGGCCCTGCTCCTCTC-3'	Rev
***Hs*LARP6-C490S**	KAL068	5'-CCAAATGCCTCCCCGAGCACAAGTCCTTGGAG-3'	Fwd
	KAL069	5'-CTCCAAGGACTTGTGCTCGGGGAGGCATTTGG-3'	Rev

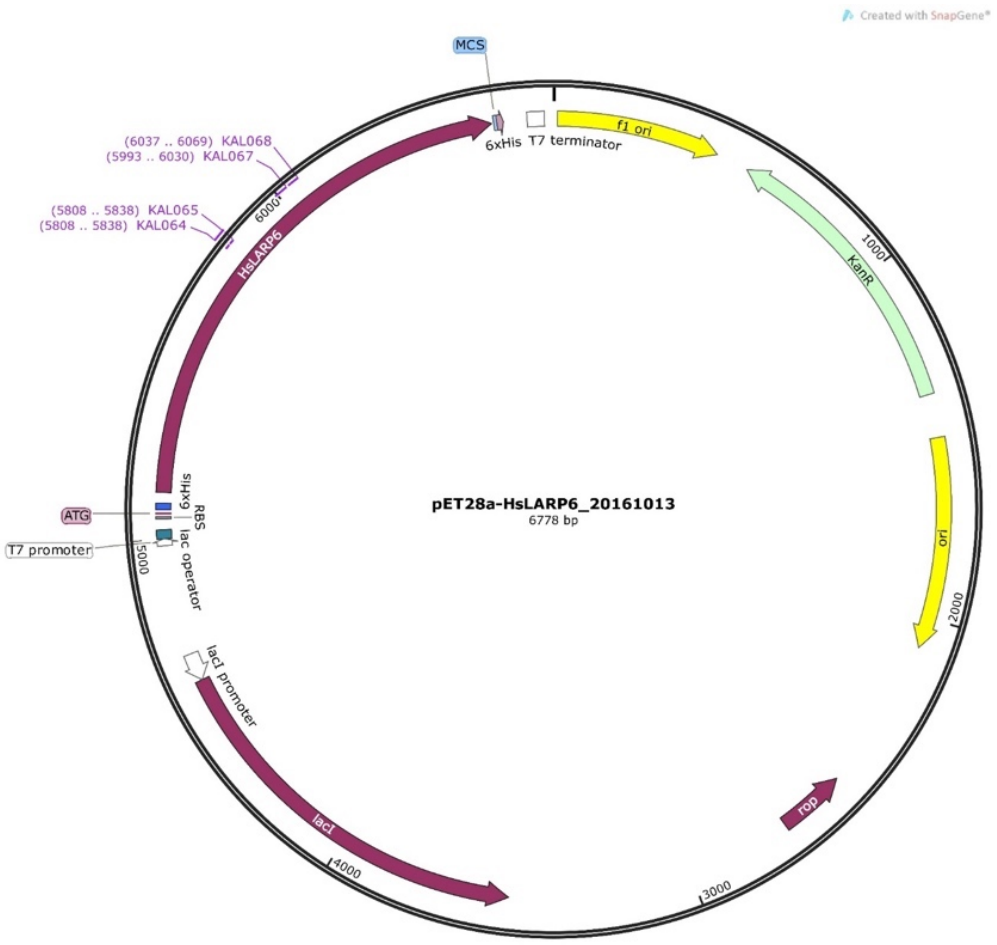

Figure 6. pET28a-HsLARP6 expression plasmid. Previously constructed pET28a-HsLARP6 expression plasmid was generously gifted by Mark Bayfield (University of Toronto). Targeted cysteines were changed to serine using site-directed mutagenesis. Primers KAL064 and KAL065 were used to convert the codon for C258 (TGC) to a serine codon (AGC) at position; primers KAL067 and KAL068 were used to convert the wildtype codon from C490 (TGT) to a serine codon (TCT). Expression of *Hs*LARP6 gene is controlled by the *lac* promoter with kanamycin resistant gene for antibiotic selection.

Bacterial Transformation

Escherichia coli DH5α ultracompetent cells were used for both restriction enzyme-based cloning and amplification of site-directed mutagenesis products. Briefly, the cell/DNA mixture was incubated for 30 min on ice, heat shocked for 90 s at 37 °C, incubated on ice for 2 min and recovered using 700 µL of autoclave-sterile Luria Broth

(LB) at 37 °C for 1 h with shaking, and plated onto a LB-agar plate containing appropriate antibiotics and incubated at 37 °C overnight. A single colony from the plate was used to inoculate an overnight culture of LB with appropriate antibiotic. Plasmid DNA was extracted from 3 mL of that suspension culture using QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's protocol; DNA was eluted in 30 µL of the kit elution buffer.

For protein expression, *E. coli* Rosetta™ (DE3) pLysS competent cells (a kind gift from Dr. Robert McLean, Texas State University Department of Biology). The cells were transformed as following: cell and DNA mixture were incubated on ice for 30 min on ice, heat shocked at 42 °C for 45 s, and incubated on ice for 2 min. Cells were recovered with 900 µL of autoclave-sterile Luria Broth (LB) and allowed to incubate with shaking at 37 °C with shaking. Recovered cells were then plated onto a LB-agar plate containing chloramphenicol and the appropriate plasmid selection antibiotic and incubated overnight at 37 °C.

Table 2. List of plasmid constructs, antibiotic resistance, and expressed protein product

Plasmid Construct	Antibiotic Resistance	Expressed Protein Product
pET28a-*Hs*LARP6C258S	Kanamycin	His_6-*Hs*LARP6-C258S
pET28a-*Hs*LARP6C490S	Kanamycin	His_6-*Hs*LARP6-C490S
pET28a-*Hs*LARP6C258S/C490S	Kanamycin	His_6-*Hs*LARP6-C258S/C490S
pET28a-*Hs*LARP6(70-300)	Kanamycin	His_6-*Hs*LARP6 La Module (70-300)
pET28a-*Dr*LARP6(60-290)	Kanamycin	His_6-*Dr*LARP6 La Module (60-290)

Recombinant Expression and Purification of Full-Length LARP6 proteins

Rosetta™ cells were used for all recombinant protein expression and purification. One to two liters of autoclave-sterile LB with appropriate antibiotics were inoculated with overnight cultures of transformed Rosetta™ cells with a target OD_{600} of 0.5-0.6. Cells were induced with 1 mM isopropyl thiogalactopyranoside (IPTG) and incubated for 8 h at 18 °C. Cells were then collected by centrifugation at 5-8,000 x*g*, 4 °C for 10 min, and the cell pellets stored at -20ºC until lysis. Cell pellets were resuspended using Lysis/Wash #1 (**Table 4**) buffer containing a dissolved protease inhibitor tablet (ThermoFisher Scientific). Cells were lysed by sonication in an ice water bath, using 6

sonication cycles of 20 s each with rest intervals of 30 s at 20% amplitude using the Fisherbrand™ model 505 Sonic Dismembrator (Fischer Scientific; Pittsburgh, PA) The cell debris was pelleted by centrifugation at 18,000 x*g,* 4 °C for 15 min, yielding the clarified lysate as the supernatant.

Nickel Affinity Chromatography

The supernatant from the centrifuged lysate (~ 30 mL) was transferred to a conical vial containing equilibrated Ni^{2+}-NTA agarose beads (twice with 7 CV of ultrapure H_2O, twice with 7 CV of Lysis/Wash#1 Buffer + Protease inhibitor) at 2-2.5 CV and allowed to incubate for 1 h at 4ºC with shaking. The bead/lysate mixture was then transferred to a clean glass flex column and allowed to settle. Once the beads settled uniformly on the bottom of the column, the flow-through was collected. The beads were washed with 20 CV of Lysis/Wash Buffer #1 without protease (**Table 4**) After the first wash, 20 CV Wash Buffer #2 was applied to the column (**Table 4**) and collected. After the second wash, 6 CV of Elution Buffer was applied (**Table 4**) and the eluate was collected in six fractions of 4 – 5 mL each. All fractions were analyzed using SDS-PAGE and Coomassie blue staining as described above.

Size Exclusion Chromatography

A Sephadex S200 column (data are on label on the column) was used for all size exclusion chromatography of the full length HsLARP6 proteins. The column was equilibrated with S200 Equilibration and Storage Buffer which had filtered through a 0.2 μm nitrocellulose membrane to remove particles, degassed, and chilled at 4 °C overnight prior to use. Elution fractions from the nickel-affinity column were pooled based on SDS-

PAGE band intensity and concentrated to 2 – 3.5 mL using a Vivaspin™ centrifugal concentrator (10,000 MWCO, Sartorius; pre-rinsed with polished with water) by centrifugation at 4,000 x*g*, 4 °C for 10 min cycles. Between each cycle, the retentate was gently mixed by aspiration to ensure consistent protein concentration and to monitor for aggregation. The concentrated protein was then filtered with a 0.2 µm syringe filter before loading into the ÄTKA Pure Fast Protein Liquid Chromatography (FPLC) system. The column was set to elute at a total volume of 180 mL at a flow rate of 1 mL per min, with the A_{280} chromatogram being monitored and recorded by the UNICORN software. Once the fractions were collected, the A_{280} chromatogram was used to determine which fractions were needed to be analyzed by SDS-PAGE and Coomassie stain as described above. Fractions were pooled based on band intensity and purity. When the size exclusion chromatography was the last step before storage (full-length protein cysteine mutants and His_6-tagged La Modules), the pooled fractions were then prepared for long-term storage by concentrating 2-3-fold of the total volume, distributed into small aliquots, and snap-frozen using liquid N_2 and then stored at -70 °C.

Recombinant expression and initial purification of the His_6-tagged La Module protein constructs

The expression of the His_6-tagged La Module protein constructs were performed the same as the full length HsLARP6 as described above. The nickel-affinity purification of La Modules was also performed the same as the full length as described above, however for the zebrafish La Module protein construct, the Lysis/Wash #1, Wash #2, and Elution buffer conditions varied (see **Table 4**). The Sephadex S75 size exclusion column was used for all purification of the His_6-tagged La Module protein constructs.

Equilibration and elution of the La Module protein constructs were performed with their respective conditions for the S200/S75 Size Exclusion Equilibration and Storage Buffer as described in **Table 4**. Proper degassing and filtration of these buffers remained the same as the ones for the full length *Hs*LARP6 proteins. After elution of the proteins, 20 µL aliquots of the La Modules were then mixed with 5 µL of 5x SDS Sample Buffer (310.5 mM Tris, 2.5% SDS [w/v], 0.65% glycerol [v/v], 5% [v/v] ß-mercaptoethanol) and run on 10% denaturing SDS-PAGE gel for analysis and stained with coommassie blue stain and destained. Fractions containing higher concentration of the protein of interest were than pooled and subjected to thrombin-mediated removal of the N-terminal His_6-tag.

Removal of the N-terminal His_6 tag from the isolated La Module constructs

To remove the N-terminal His_6-tag on the *Hs*LARP6 and *Dr*LARP6 La Module protein constructs, the protein concentrations of the pooled fractions from size exclusion chromatography were calculated using intrinsic absorbance at 280 nm using the Implen Nanophotometer (Munich, Germany). Thrombin (Sigma-Aldrich) was added to the protein solution at a ratio of 2 U of thrombin per 2.5 moles of protein and allowed to incubate on ice for 2-3 h. Following cleavage, the protein solution was then subjected to a second set of nickel affinity and size exclusion chromatography as described above, to separate the His_6-tag, the cleaved protein, and the non-cleaved protein.

Table 3. Molar extinction coefficients of LARP6 proteins used in this study.

Protein Construct	Expected	Molar Extinction

	Molecular Weight (kDa)	Coefficient (M^{-1} cm^{-1})
His_6-*Hs*LARP6	55.0	46660
His_6-*Hs*LARP6-C258S	55.0	46660
His_6-*Hs*LARP6-C490S	55.0	46660
His_6-*Hs*LARP-C258S/C490S	55 .0	46600
His_6-*Hs*LARP6 La Module (70-300)	30.2	25440
His_6-*Dr*LARP6 La Module (60-290)	29.6	23950
His_6-T7-*Xm*LARP6 La Module (49-290)*	32.0	22460

* = Protein constructs recombinantly expressed and purified by previous graduate student Jose Miguel Castro of the KAL Lab

Denaturing Polyacrylamide Gel Electrophoresis and Gel Staining

Protein samples were prepared using a 5x SDS Sample Buffer (310.5 mM Tris, 2.5% SDS [w/v], 0.65% glycerol [v/v], 5% [v/v] ß-mercaptoethanol), boiled at 90 °C for 5 min and then loaded onto 10-13 % SDS-PAGE gel. Gels were electrophoresed at 200 V for 50-65 min, room temperature (RT) and either stained or transferred to nitrocellulose for western blotting (see below). Two different types of staining methods were used to visualize protein migration in SDS-PAGE gels.

Coomassie blue staining (58.4 µM Coomassie Brilliant Blue R-250, 45.5% [v/v] methanol, 9.1 % Glacial Acetic Acid) was performed for 30 min, followed by a 1 – 1.5 h destain (40% [v/v] methanol, 10% [v/v] acetic acid) period until desired intensity was achieved.

For silver staining, the gel was fixed using 50% EtOH for 30 min, stained with 475 mM silver nitrate, developed using the Developing Solution (with 0.75 % [v/v]

formaldehyde and 0.04 % [v/v] citric acid. The development of signal was stopped by adding an equal volume of Kill Solution (0.20 % [v/v] methanol, 0.04% [v/v] acetic acid). All developed gels were then imaged on the ChemiDoc XRS+ Molecular Imager.

Western Blot

To verify the presence of desired His_6-tagged proteins, a western blot was performed using the Horseradish Peroxidase (HRP)-conjugated α-His probe (Fisher Scientific cat no#). In order to verify the presence of full-length LARP6 proteins, a primary rabbit α-LARP6 antibody (Abcam # ab94640) that targets 15 residue sequences near the C-terminus of the protein was used. This primary antibody was detected with a goat α-rabbit – horseradish peroxidase conjugate.

After gel electrophoresis, proteins were transferred onto nitrocellulose using the Bio-Rad Turbo Transblot® using the Mixed Molecular Weight preset (1 V, 1 A, 7 min). For both the anti-His and anti-LARP6 blots, the nitrocellulose was blocked using 5% BSA suspended in 1X TBS + 0.05% Tween 20.

For the anti-His blots, the blocked membrane was incubated with diluted α-His probe (1:5000 v/v in 1x TBS-T) (ThermoFisher) for 1 hour at RT, washed twice with TBS-Tween 20, washed twice with 1xTBS.

For the anti-LARP6 blots, incubated with diluted 1° antibody: rabbit α-LARP6 (1:1000 v/v in 1x TBS-T) (Abcam cat no# ab94640) for 1 hour at RT, incubated with

diluted 2° antibody: goat α-rabbit (1: 20,000 v/v in 1x TBS-T) (washed twice with TBS-Tween 20, washed twice with 1x TBS.

The membranes were detected using a homemade chemiluminescence detection agent (a.k.a The Juice: 0.1 M Tris [pH 8.8 @ 25 °C, 1.25 mM luminol in DMSO, 2 mM 4-IBPA in DMSO) and 12 μL of stable 30% H_2O_2.[16] Membranes were detected using the chemiluminescence preset in the ChemiDoc XRS+ imager.

Proteolysis using Aminopeptidase, trypsin, and thrombin

Several protocols for limited proteolysis were used in this work. In order to analyze protein preps for stability and also determine boundaries in the La Module constructs, the proteases thrombin (2U: 2.5 moles of protein), aminopeptidase (20 moles protein : 1 mole enzyme), and trypsin (1000:1) were used to assess stability of prepared proteins.

Stability of the protein was assessed by performing digestion reactions in ratios described above. Time points were taken at 0, 1, 2, 5, 10, 20, 30, 45, 60, 120, and 180 min. Each time point was prepared with 5X SDS sample buffer and analyzed by SDS-PAGE and either silver staining or Coomassie blue staining.

In order to test for the possibility of an in-frame, C-terminal His_6-tag read-through product in the *Hs*LARP6 cysteine mutant, a thrombin digest was performed using a molar ratio of 2 U thrombin: 2.5 moles of LARP6 protein. Thrombin was added to all

preparation of *Hs*LARP6 and incubated on ice for 6 hours to allow for sufficient cleavage of the N-terminal His_6-tag. Protein samples were run on triplicate SDS-PAGE gels: one was stained with Coomassie and the other two were detected by western blot for either α-His or α-LARP6 as described above.

A time course series with thrombin was also performed to determine optimal conditions for His_6-tag cleavage during protein purification. His-tagged containing a thrombin cleavage site in LARP6 La Modules were digested using thrombin. The proteins were mixed in a ratio as described above, and time points were taken at 0, 2, 4, and 6 h. Each time point was prepared with SDS-sample buffer and ran through 10-13% SDS-PAGE gel at a constant 200 V, RT for 50-65 min. Each gel was visualized using either Coomassie staining and cleavage confirmation was assessed through α-His probe western blot (as previously described).

Mass Spectrometry

To further confirm the identity of proteins found in the preparative samples, excisions were prepared and sent to collaborators for mass spectrometric analysis (Biomolecular Research Center, Boise State University, Boise, ID). Care was taken to exclude introduction of keratin into our samples by washing all glassware with fresh 20% solution of acetic acid, ultrapure H_2O, and then 95% ethanol, and brand-new razor blades were used for excision of desired bands. Proteins were prepared with SDS sample buffer, run on 10-13% SDS-PAGE gel at 200 V, 50 min, RT, and visualized using staining protocols as previously described with fresh Coomassie blue or silver stain. Once

visualized, high intensity bands were excised using a clean, brand new razor blade, transferred to a sterile micro-centrifuge tube, and submitted for complete trypsinolysis and tandem liquid chromatography-mass spectrometry for protein mapping (Xinzhu Pu, Boise State University).

3' Biotinylation of RNA Ligands

To detect our RNAs in our binding system, ligands were biotinylated using a cytosine-streptavidin ligation at the 3' end using the Pierce 3' End Biotinylation Kit (ThermoFisher). Stock solutions of RNA oligos were brought to 4% (v/v) DMSO, heated to 85 °C for 2 mins and snap cooled on ice (4 °C) for a minimum of 2 minutes. A standard ligation reaction followed manufacturer's protocol. Briefly, each reaction contained 10.3 µM RNA oligo, 1x RNA Ligase Reaction Buffer (from kit), 1 µL of RNase inhibitor, 33.3 µM biotinylated cytidine bisphosphate, 15% PEG, and 1.33 U/µL T4 RNA Ligase were incubated at 30 °C overnight in a heating block. Reaction was recovered from heating block, at which point 70 µL of nuclease free water (from kit) was added, then 100 µL of a chloroform: isoamyl solution (24:1) which the solution was vortexed for 10 min. The mixture was then centrifuged at 16,000 x*g* at 25 °C for 3 mins, after which the separated aqueous phase was transferred to a sterile, newly labeled microcentrifuge tube. The solution was then precipitated in 240 mM sterile NaCl, 1 µL of 20 mg/mL glycogen, and 60% v/v EtOH for 3 hours on ice (4 °C).

RNA oligos were then recovered from the precipitation solution by pelleting via centrifuged for 15 min at ≥13,000 *xg*, 4 °C. Supernatant was carefully removed as to not

disturb the pellet. Pellet was washed with 300 µL of cold 70% ethanol, centrifuged at ≥ 13,000 x*g*. Supernatant of the wash was removed, and the pellet was allowed to air dry before resuspending in 20 µL of sterile 0.5X TBE.

Electrophoretic Mobility Shift Assay

In order to determine the binding activity of purified LARP6, an *in vitro* binding system was created in the following conditions: 10 mM Tris-HCl at pH 7.4 at 4 °C, 20 mM KCl, 1 mM Mg_2Cl, and 100 mM NaCl. Binding reactions were performed with the previously identified stem loop structure found in the 5' untranslated region of collagen type I mRNA.

The biotinylated RNA ligands were diluted to a working dilution of 1 pM in Tris binding buffer (15% glycerol, 10 mM Tris, 20 mM KCl, 1 mM $MgCl_2$, 100 mM NaCl). The proteins were serially diluted in Tris Binding Buffer. The working stocks of protein and RNA were mixed 1:1 allowed to equilibrate for 1 h before separation on a 6.5% native PAGE gel (6.5% (29:1 bis-acrylamide: acrylamide), 15% glycerol, 1X TBE (89.2 mM Tris, 89.0 Boric Acid, 4.0 EDTA [pH 8.0]) at 200 V for 15 min in 1X TBE running buffer. The separated RNAs were transferred to a Hybond N^+ nitrocellulose membrane using the Transblot® Turbo™ module for 30 min at 25V, 10A in 1x TBE (no glycerol). The nitrocellulose membrane with transferred RNA-protein binding complex was UV-crosslinked at 120 mJ for 45 s. The cross-linked membrane was then left to dry for up to 1 week before proceeding to detection.

Detection of Biotinylated RNA on Hybond+ membrane

All reagents were part of the Thermofisher Chemiluminescent Nucleic Acid Detection Module Kit. Prior to using, the Nucleic Acid Blocking Buffer and 4x Wash Buffer were warmed for 30 min at 37 °C, and the Substrate Equilibration Buffer was warmed at room temperature for 30 mins. The membrane was blocked with 20 mL of the Nucleic Acid Binding Buffer with oscillation for 15 min. The membrane was then incubated with 20 mL of Nucleic Acid Blocking Buffer + 66.7 µL of the streptavidin-Horseradish Peroxidase(HRP) conjugate at room temperature for 15 min with oscillation. The membrane was then briefly washed with 20 mL of a 1x Wash Buffer (1:3 4x Wash Buffer from the kit: ultrapure H_2O and discarded before continuing with four subsequent washes with the same buffer each incubated for 5 min with oscillation at room temperature and discarded after each wash. Then, the membrane was allowed to incubate with 30 mL of the Substrate Equilibration Buffer with oscillation for 5 min. The Hybond+ membrane was blotted dry on the side as to not let the paper towel to touch the face of the membrane. The membrane was lay on clean cling wrap and 5 mL of the Substrate Working Solution (1:1 Luminol/Enhancer Solution: Stable Peroxide Solution from the kit), allowed to incubate for 2 min in a darkened box, and imaged immediately using the Chemiluminescence preset of the ChemiDocXRS+. Exposure times are denoted in the legends of the results.

Quantification of EMSA and $K_{D,app}$ fitting

After detection of the RNA on the Hybond+ membrane, the bound and unbound regions of the gel shift were quantified using the Volume Tools in the Image Lab Software (Bio-rad). Once quantified, calculation of the fraction bound ($\frac{[RNA]_{Bound}}{[RNA]_{Total}}$) was

performed using Microsoft Excel. The data was then plotted in a graph of concentration of protein vs ligand in Sigma Plot (UK) or Kaleidagraph (Synergy Software). The $K_{D,\,app}$ was then fitted using either the Simplified Binding Isotherm (**Equation 1**) or Quadratic Form of Binding (**Equation 2**). [32, 33]

Equation 1. Simplified Binding Isotherm:

$$\frac{[\mathrm{PL}]}{[\mathrm{L}]_{\mathrm{T}}} = \mathrm{S}\left(\frac{[\mathrm{P}]_{\mathrm{T}}}{[\mathrm{P}]_{\mathrm{T}} + \mathrm{K}_{\mathrm{D,app}}}\right) + 0$$

Equation 2. Quadratic Binding Isotherm:

$$\frac{[PL]}{[L]_T} = \frac{\left(K_{D,app} + [P]_T + [L\}_T\right) - \sqrt{(K_{D,app} + [P]_T + [L\}_T)^{\,2} - 4[P]_T[L\}_T}}{2[L\}_T}$$

Table 4. Protein Purification Buffers

wildtype *Hs*LARP6 ***Hs*LARP6-C258S** ***Hs*LARP6-C490S** ***Hs*LARP6-C258S/C490S** ***Hs*LARP6 La Module (70-300)**	***Dr*LARP6 La Module (60-290)**

Lysis/ Wash Buffer #1	50 mM sodium phosphate pH 7.4 @ 4 °C 300 mM NaCl 10 mM imidazole	50 mM sodium phosphate pH 8.0 200 mM NaCl 500 mM glucose 300 mM NaI 10 mM imidazole
Wash Buffer #2	50 mM sodium phosphate pH @ 4 °C 300 mM NaCl 30 mM imidazole	50 mM sodium phosphate pH 8.0 200 mM NaCl 500 mM glucose 300 mM NaI 30 mM imidazole
Elution Buffer	50 mM sodium phosphate pH 7.4 @ 4 °C 300 mM NaCl 300 mM imidazole	50 mM sodium phosphate pH 8.0 200 mM NaCl 500 mM glucose 300 mM NaI 300 mM imidazole
Size Exclusion (S200/S75) Equilibration and Storage Buffer	50 mM Tris+ HCl pH 7.4 @ 4 °C 100 mM NaCl 5% glycerol	50 mM sodium phosphate pH 8.0 200 mM NaCl 500 mM glucose 300 mM NaI

III. IDENTIFICATION OF THE LA MODULE DOMAIN BOUNDARIES FOR STRUCTURAL ANALYSIS

Recombinant expression and purification of His_6-tagged Human La Module

The pET28-His_6-LaModule expression constructs were previously cloned in the lab by K.A. Lewis, J.M. Castro, and H. Kulkoyluoglu (unpublished). The pET28-His_6-*Hs*LARP6-La Module expression plasmid was transformed into Rosetta™ DE3 *E. coli* cells. Cells were then selected for using appropriate antibiotics (**Table 2**) and grown into large culture. Protein expression was induced using 1 mM final concentration IPTG, and 1 mL aliquots were taken at regular intervals of 0, 2, 4, 6, and 8 hours and pelleted. Cell pellets were then lysed in 1X SDS sample buffer and analyzed using an anti-His_6 probe western blot. A band can be seen in the anti-His_6 blot with a molecular weight of 25 kDa (**Figure 7**) with increasing intensity over time. While that is 5 kDa less than the expected molecular weight of the La Module construct, this result suggests an increase in expression of a His_6-tagged protein over time.

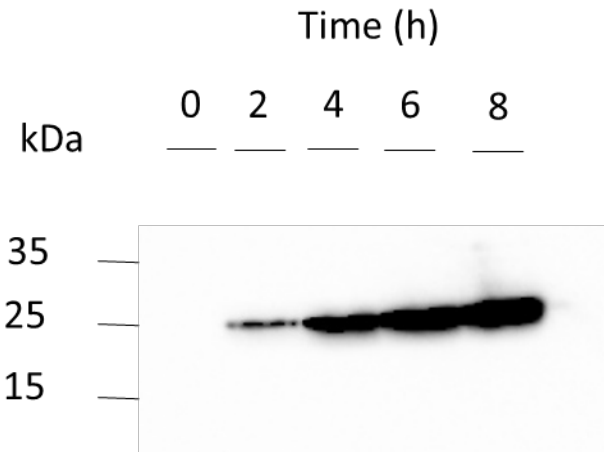

Figure 7. Expression of *Hs*LARP6 La Module (70-300) construct in Rosetta™ (De3) pLysS competent cells. *H. sapiens* LARP6 La Module construct was mutagenized and cloned from stock pET28a-HsLARP6 plasmid construct, recombinantly expressed as a N-terminally His_6-tagged protein construct in Rosetta™ *E. coli* cells. Following induction with 1 mM IPTG, cells were incubated at 18 °C with expression being monitored at regular intervals of 2, 4, 6, and 8 h, after which cells were pelted, lysed with SDS-PAGE sample buffer, and analyzed on 10% SDS-PAGE gel followed by α-His6 western blot. Bands were seen to increase in intensity as time progresses, indicating that the His_6-HsLARP6 La Module (~25 kDa) is being expressed.

Cultured cells were pelleted and lysed with sonication for purification of the His_6-HsLARP6 La Module (70-300). Cleared lysate was subjected to initial purification using nickel-affinity chromatography; fractions were analyzed on denaturing SDS-PAGE and stained with Coomassie blue stain (**Figure 8**). A faint band can be seen at the expected molecular weight of 25 kDa in the last two wash fractions. A high intensity band can be seen in Elution fraction 1 and 2, with decreasing intensity as elution fractions were collected. Elution fractions 1-3 were selected for injection onto the S75 size exclusion column.

The S75 size exclusion column was monitored using UV/Vis spectroscopy absorbance at 280 nm. The UV chromatogram shows the presence of two peaks at 44 mL

and 54 mL, and the corresponding fractions were analyzed using denaturing SDS-PAGE (**Figure 9**). The gel was visualized using Coomassie blue stain, where the presence of a strong band at the expected molecular weight of 30 kDa was seen in all elution fractions (**Figure 10**). No other co-eluting species was observed in either of the elution fractions. Because the peak can be seen at the void volume (46.25 mL), it may be the result of aggregated protein. In contrast, the second peak corresponds more closely to the apparent molecular weight of the protein (expected, 30 kDa; apparent, 57 kDa), fractions 16-20 were selected, concentrated with a 10,000 molecular weight cut off centrifugal filter and stored in -70 °C. A third peak can be seen at ~130 mL corresponding to a molecular weight less than 6.5 kDa, which may be the result of imidazole from the nickel affinity purification prior to injection.

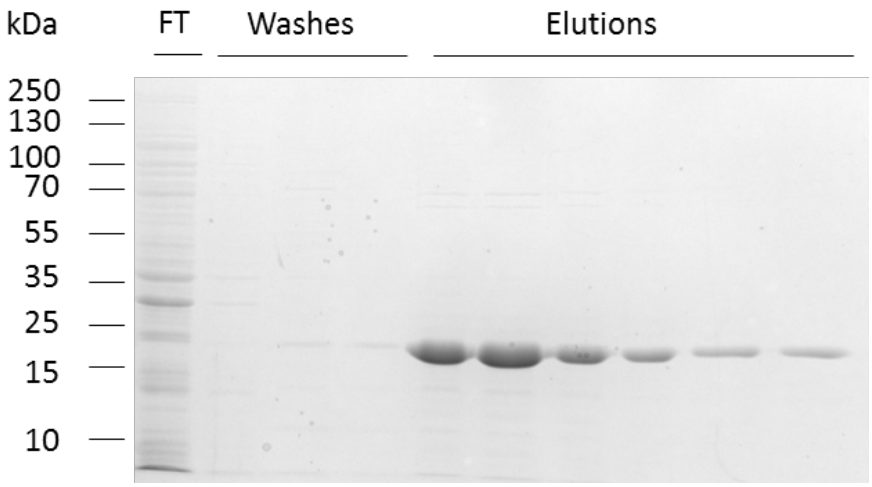

Figure 8. His_6-*Hs*LARP6 La Module (70-300) affinity chromatography fractions. Cells containing the His_6-tagged human LARP6 La Module construct were lysed and purified using nickel-affinity chromatography. Aliquots were separated by gel electrophoresis on denaturing 10% SDS-PAGE gel and stained with Coomassie blue stain. Elution fractions containing the protein of interest. Elution fractions 1-3 were collected and prepared for further purification.

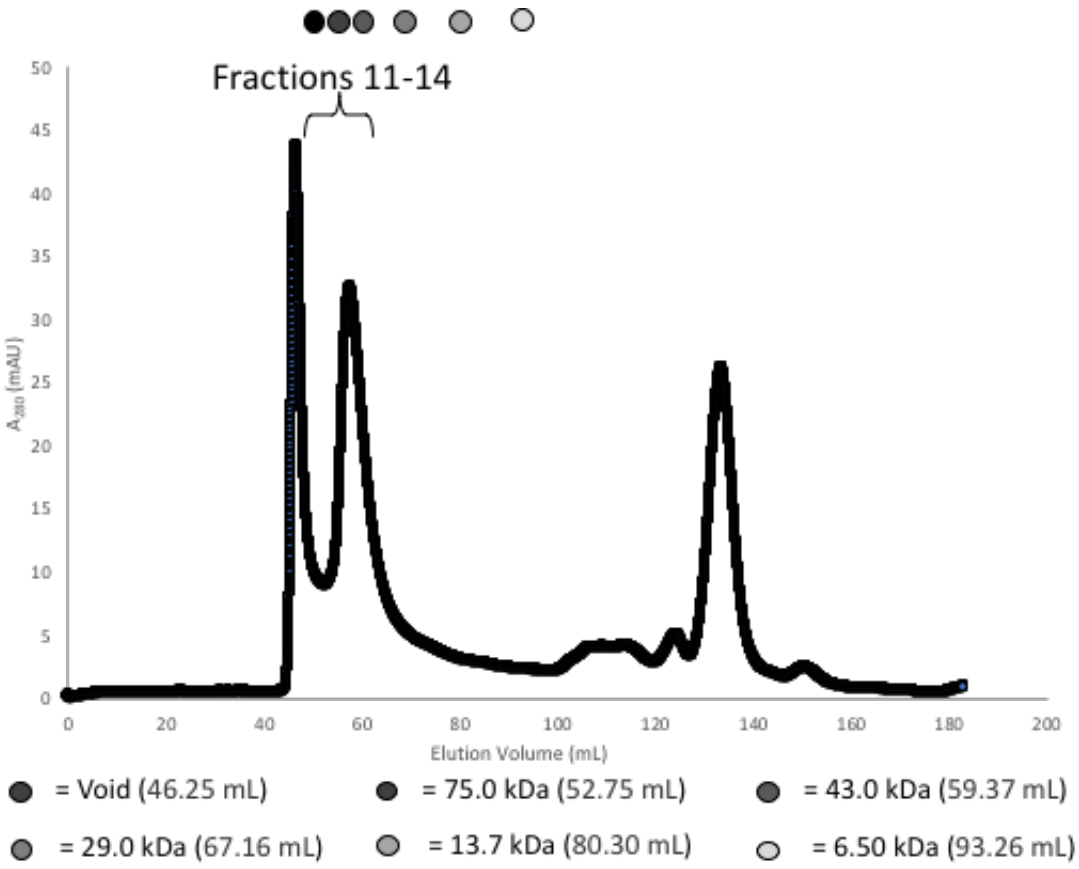

Figure 9. S75 size exclusion chromatogram of His_6-tagged *Hs*LARP6 La Module (70-300). Pooled fractions from the affinity chromatography were concentrated against a 10,000 MW cut off filter and injected into the AKTA FPLC for size exclusion through Sephadex 75 gel filtration column. Peak 1 (fraction 11-14, 45 – 52 mL) and Peak 2 (fraction 15-22, 53-67 mL) were collected for analysis through denaturing gel.

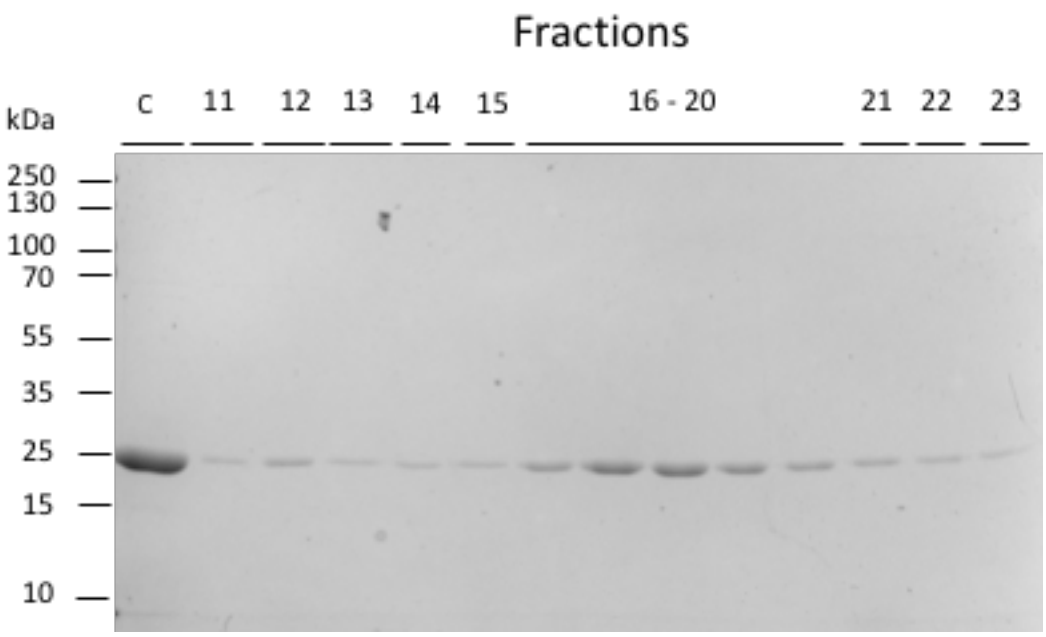

Figure 10. SDS-PAGE analysis of His_6-tagged *Hs*LARP6 La Module (70-300) size exclusion fractions. Fractions 11-23 (45 – 67 mL) were selected from the size exclusion for analysis through denaturing gel electrophoresis and stained with Coomassie blue stain. Elution fractions 16-20 (55 – 63 mL) were pooled, concentrated as previously described, and stored as 50 µL aliquots at -70 °C for further experiments.

Recombinant expression and purification of His_6-tagged zebrafish LARP6 La Module

Cultured cells containing the pET28a-His_6-*Dr*LARP6 La Module (60-290) were lysed and subjected to nickel affinity chromatography. Fractions were collected and aliquoted for analysis using denaturing SDS-PAGE (**Figure 11**). A high intensity band at the expected molecular weight of 25 kDa were seen in elution fractions 1 and 2, with decreasing intensity over the elution fractions, indicating the presence of our desired His_6-tagged protein. Faint bands were detected in the fractions corresponding with the last two wash fractions of the purification, indicating that small amounts of desired protein were washed from the column. This phenomenon may be the result of high concentrations of protein saturating the nickel-NTA beads, thus allowing unassociated protein to travel through the column. In order to deter concentration-dependent protein

aggregation, elution fractions 1 and 2 were selected for injection into the S75 size exclusion column.

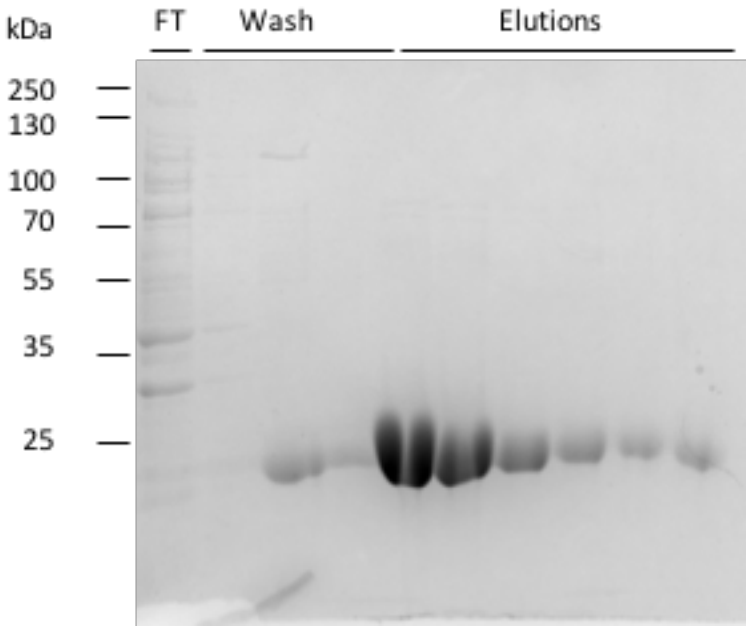

Figure 11. Denaturing gel analysis of His_6-*Dr*LARP6a (60-290) affinity chromatography fraction. Cells containing the His_6-tagged zebrafish LARP6 La Module construct were lysed and purified using nickel-affinity chromatography. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE gel and stained with Coomassie blue. Elution fractions 1-6 contained a band that migrated around 25 kDa, which is only 5 kDa below the expected for this protein construct.

The S75 size exclusion UV chromatogram (**Figure 12**) shows the presence of a single high intensity peak at elution volume 54 mL with high baseline resolution. A lower intensity peak can be seen at the void volume (46.25 mL) which may be the result of a small amount of aggregated protein. Based on the UV chromatogram, the fractions 11-21 were selected for analysis using denaturing SDS-PAGE gel (**Figure 13**). For fractions 16-

21, a single high intensity band at a molecular weight of 25 kDa was observe. Fainter bands were seen in fractions 11-15 with increasing intensity. Fractions 16-20 were selected for concentration against a 10,000 molecular weight cut off filter and stored in -70 °C.

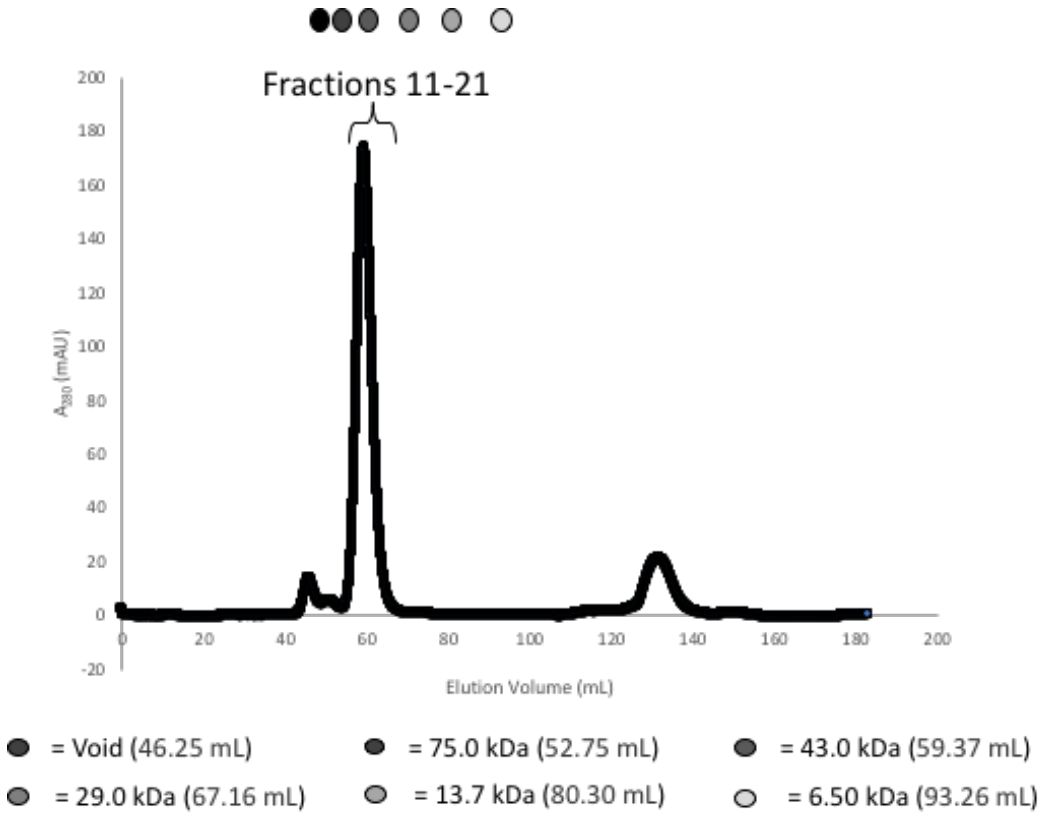

Figure 12. S75 size exclusion chromatogram of His_6-*Dr*LARP6a La Module (60-290). Pooled and concentrated fractions from the affinity chromatography were injected onto the Sephadex 75 gel filtration column. Elution volumes were monitored via absorbance at 280 nm. Fractions 11 – 21 (45-65 mL) were selected for analysis through denaturing SDS-PAGE.

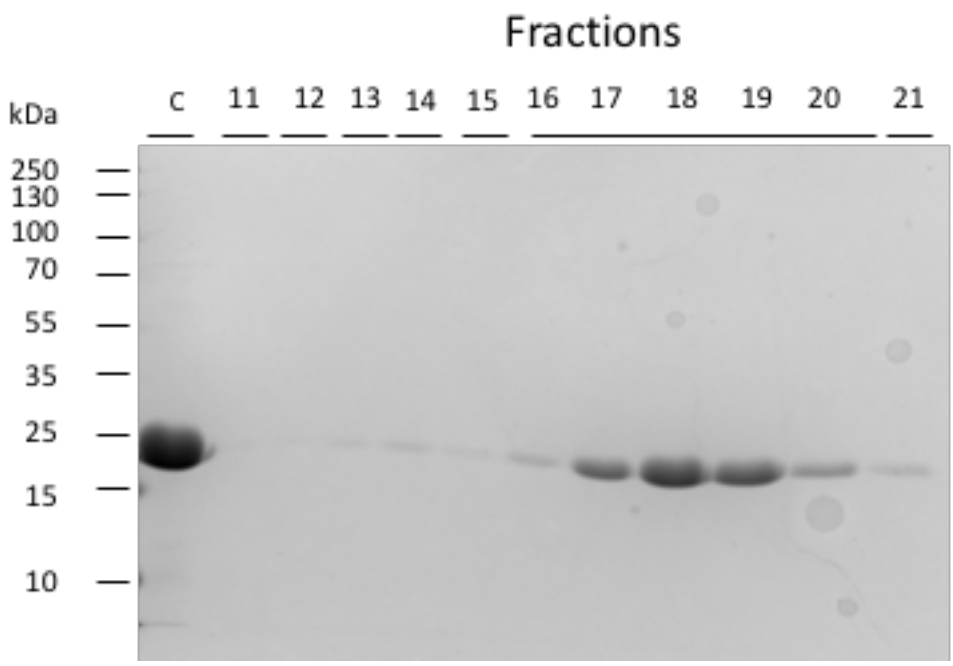

Figure 13. Denaturing gel analysis of S75 size exclusion fractions of His_6-*Dr*LARP6a La Module. Fractions 11-21 (45- 65 mL) were selected for denaturing gel analysis. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE and stained using Coomassie blue. Elution fractions 16-20 containing high concentration of the protein of interest (~25 kDa) were pooled, concentrated against a 10,000 molecular weight cutoff filter, and store in 50 µL at – 70 °C for future experiments.

Global stability of La Module constructs

In order to assess the stability of the human, zebrafish, and platyfish La Module constructs, limited proteolysis reactions were performed and analyzed on denaturing SDS-PAGE and visualized using silver staining as previously described above.

When the human La Module was subjected to trypsinolysis (**Figure 14**, *top right*), the appearance of a lower molecular weight band was seen to increase in intensity as time increased. The presence of two bands corresponding with a molecular weight between 15-25 kDa was observed in all time points with the exception of the 120th minute. The

presence of a faint band at a molecular weight corresponding to 15 kDa was observed from the 2nd-60th minute, and the presence of a 10 kDa molecular weight band between the 10th-60th minutes. These lower molecular weight bands may be the result of regions in the protein that contained accessible trypsin cut sites that were being degraded as time progressed.

When the platyfish La Module was subjected to the same proteolytic reaction using trypsin (**Figure 14**, *top left*), the two bands were present with high intensity at molecular weights corresponding to 35 kDa and 55 kDa. The absence of any lower molecular weight, time-dependent bands suggests this protein construct is resistant to proteolytic cleavage by the trypsin.

As the trypsinolysis digestion of the zebrafish La Module progressed (**Figure 14**, *bottom*), the high intensity band at a molecular weight of ~35 kDa was observed. While multiple bands at molecular weights between 15-18 kDa were observed, none show a trend of steady increase or decrease over time. Unexpectedly, a high intensity band was observed in the 60-minute time point at a molecular weight of about 100 kDa. A similar-size band is seen in all time points in the zebrafish; however, for this specific time point the intensity is significantly greater than that of the other time points.

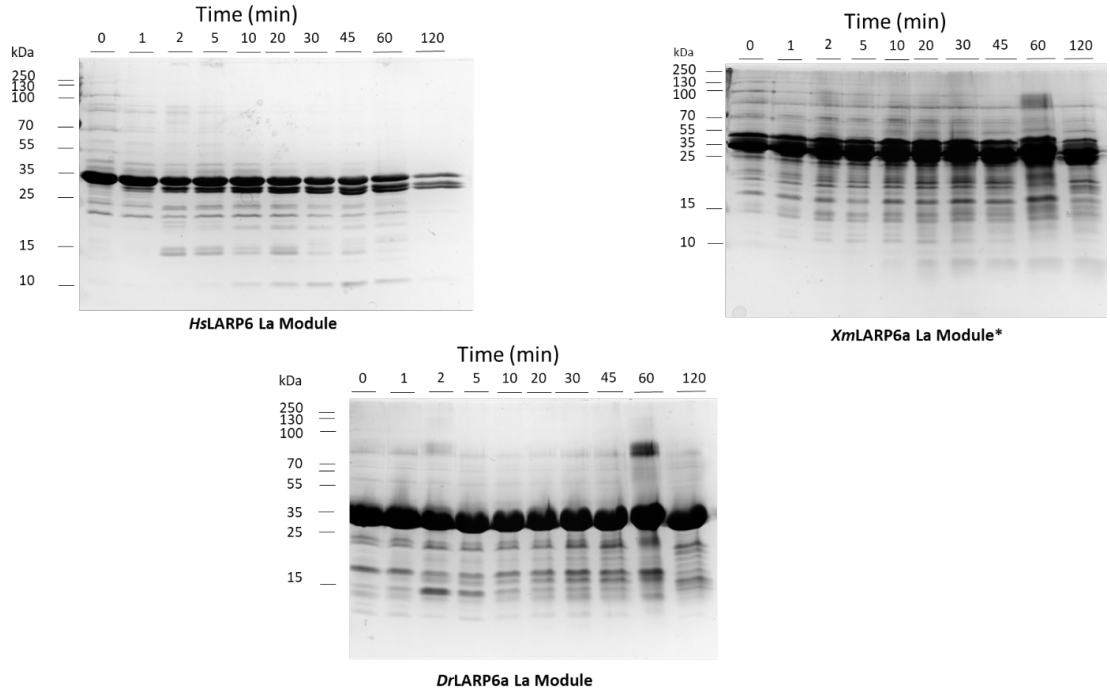

Figure 14. Limited proteolysis of LARP6 La Modules using trypsin. Limited trypsinolysis of the human (*Hs*), zebrafish (*Dr*), and platyfish (*Xm*) LARP6 La Modules. All proteins were incubated with 1:1000 molar ratio of trypsin and incubated on ice (4 °C). Aliquots of 0, 1, 2, 5, 10, 20, 30, 45, 60, and 120 mins were taken and subjected to gel electrophoresis using denaturing SDS-PAGE and visualized using silver staining.

Suitability of La Modules for structural analysis

To determine whether the purified La Module protein constructs were suitable for structural analysis using the solution NMR and/or small-angle X-ray scattering, further limited proteolysis was done using the exo-protease aminopeptidase, which begins digestion of polypeptides at the N-terminus of the protein. Because the protease is not sequence-specific in its substrate, this allowed for an assessment of the N-terminal boundaries of the La Modules. The human La Module construct (**Figure 15**, *top left*) could be seen to contain the high intensity band at 25 kDa corresponding to our protein of

interest. The presence of 3 faint bands can be seen beginning at the 30th minute time point with a molecular weight of ≤15 kDa. For the platyfish La Module (**Figure 15**, *top right*) the 25 kDa band corresponding to our protein of interest could be seen in all time points with an additional co-eluting band at a molecular weight band of about 30 kDa. The intensity of a 15 kDa band can be seen to increase as time progresses with a peak intensity to appear at the 120th minute, indicating increase of concentration of this peptide fragment. For the zebrafish La Module construct subjected to the same proteolysis using aminopeptidase (**Figure 15**, *bottom*), the 25 kDa band can be seen in all time points of the digestion reaction. A ~12 kDa band can been seen to increase over time with a peak intensity at the 120th minute time point. As in the previous trypsinolysis, a 100 kDa band can be seen at the 30th, 60th, and 120th minute time point.

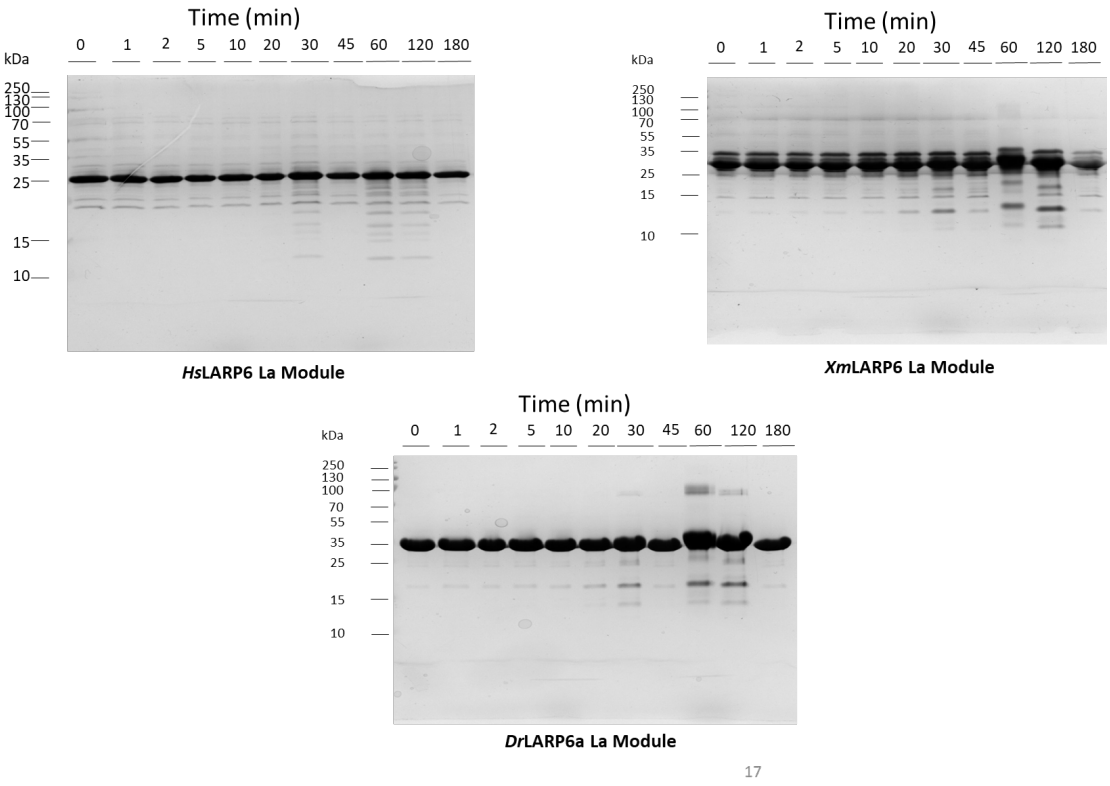

Figure 15. Aminopeptidase digest of LARP6 La Modules. Limited proteolysis using aminopeptidase of the human (*Hs*), zebrafish (*Dr*), and platyfish (*Xm*) LARP6 La Modules. All proteins were incubated with a molar ratio of 20:1 protein: protease and incubated on ice (4 °C). Aliquots of 0, 1, 2, 5, 10, 20, 30, 45, 60, and 120 mins were taken and subjected to gel electrophoresis using denaturing SDS-PAGE, and visualized using silver staining.

In summary, when the human La Module is subjected to limited proteolysis using trypsin, the presence of a band can be seen in the denaturing gel analysis. When the protein was digested using the aminopeptidase, no such band was produced. Because trypsin digests at assessible lysines and arginines, while aminopeptidase digest a protein non-specifically beginning at the N-terminus, it can be concluded that the human La Module may have C-terminal residues that are not part of the stable folded structure of the RRM. In the sequence of the human La Module, there are lysines at position 270, 282, 286, 293, and 296, 297, and 298. Since the difference between these undigested and

trypsinized bands is ~ 5 kDa, the beginning of the region that is not well-protected as part of the RRM fold may begin at position 270, 282, or 286.

Optimization of Thrombin Cleavage of His_6 tags

The His_6-LARP6 La Module constructs contain a thrombin cleavage site downstream of His_6-tag, and this site was exploited to remove the His_6 tag for later use in structural characterization by solution NMR and small angle X-ray scattering. To identify suitable cleavage conditions, the purified human (*Hs*) and zebrafish (*Dr*) La modules were subjected to a molecular ratio of the 2 U thrombin per 2.5 moles of protein and incubated on ice with regular time points taken at 0, 1, 2, 3, 4, 5, and 6 hours. The time points were analyzed by anti- His_6 western blot (**Figure 16,** panels *A* and *B*), where signal can be seen to decrease after the 2nd hour for the *Hs*LARP6 La Module (Left). Residual signal can be seen, indicating the incomplete cleavage of the His_6 for the protein construct. This suggests that a longer incubation time may be necessary for complete cleavage of the His_6-tag. For the *Dr*LARP6A La Module (**Figure 16**, panels *B* and *D*), the chemiluminescence signal was no longer present after the 1st hour of the cleavage, indicating that complete removal of the His_6-tag by the thrombin protease was achieved. Confirmation that the thrombin-mediated cleavage did not compromise the overall structural stability of the protein constructs can be seen by the presence of a consistent band throughout the time points in the duplicate gel electrophoresis that was visualized using Coomassie blue stain (**Figure 16,** panels *C* and *D*). Interestingly, in the thrombin-mediated cleavage of the HsLAR6 La Module, a shift can be seen in which the presence of the ~34 kDa band can be seen to decrease in intensity as time progresses, and the

increased production of a lower molecular weight band (~30 kDa) can be seen from the first hour until the end of the experiment.

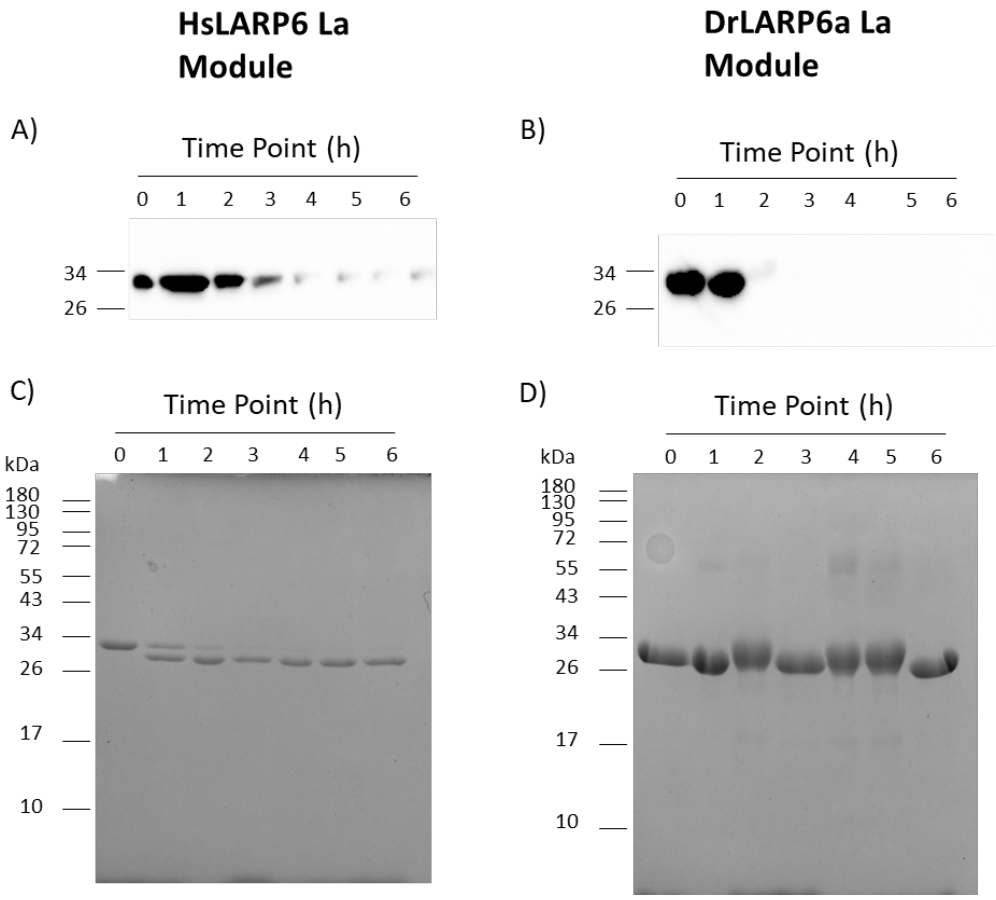

Figure 16. Thrombin digest of the La Module constructs. In order to remove the His_6-tag of the protein fusion, the thrombin protease was used to cleave at a site downstream from the His_6-tag in the human (*Hs*) and zebrafish (*Dr*) LARP6 La Module constructs. All proteins were incubated with a protein: thrombin ratio (2.5 moles: 2 U) and incubated on ice (4 °C). Aliquots were taken at regular time intervals of 0, 1, 2, 3, 4, 5, and 6 hours and subjected to separation on denaturing SDS-PAGE gel electrophoresis and transferred onto nitrocellulose and detected using the α-His_6 probe western blot.

Purification of Human La Module without tag

Pelleted cells containing the His_6-*Hs*LARP6 La Module were lysed and subjected to purification using nickel affinity chromatography as previously described. Fractions

collected were analyzed on denaturing SDS-PAGE and visualized using Coomassie blue stain (**Figure 17**). A high intensity band can be seen with a molecular weight of about 30 kDa in the elution fractions 1 and 2, with decreasing intensity in the succeeding elutions. The presence of other co-eluting bands cannot be seen in this gel. Elution fractions 1-3 were selected for concentration against a 10,000 Da molecular weight cut off and injected into the S75 size exclusion column for further purification.

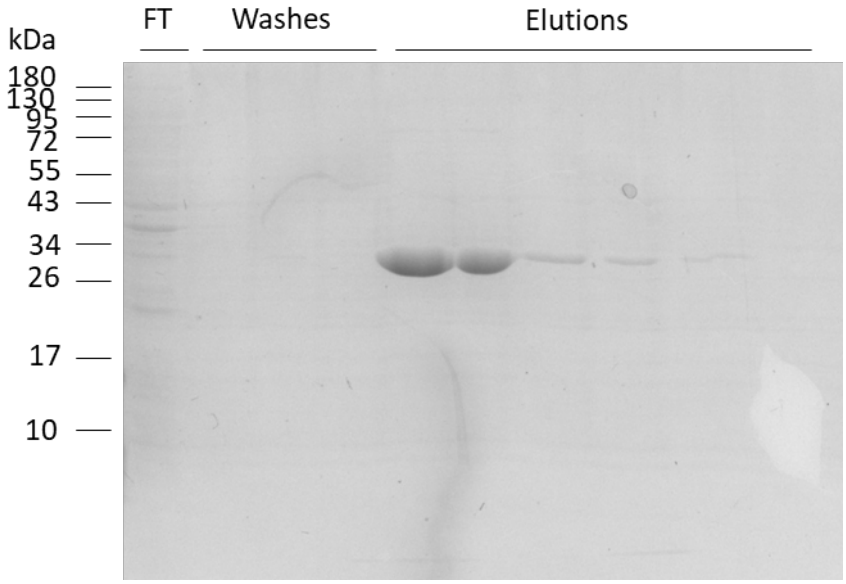

Figure 17. Denaturing SDS gel analysis of affinity chromatography purification fractions of HsLARP6 La Module (70-300). Rosetta™ DE3 *E. coli* cells containing the His_6-*Hs*LARP6 La Module (70-300) constructs were lysed and purified using nickel affinity chromatography. The flowthrough (FT), washes, and elutions were collected, aliquoted, separated on 10% SDS-PAGE gel, and stained with Coomassie. Elution fractions 1-5 contained the protein of interest (~30 kDa), however fractions 1-3 were pooled, concentrated against a 10,000 molecular weight cut off until reached a total volume of ~2 mLs, and loaded onto the AKTA FPLC for gel filtration.

Two peaks at elution volumes were observed at 50 and 65 mL in the size exclusion UV chromatogram. Fractions corresponding to these peaks (11-23) were selected for analysis on SDS-PAGE gel (**Figure 18**), where a single band corresponding to a molecular weight of 32 kDa was observed in all fractions with peak intensity across

seen in fractions 16-20. These five fractions were then pooled for His_6-tag removal using thrombin.

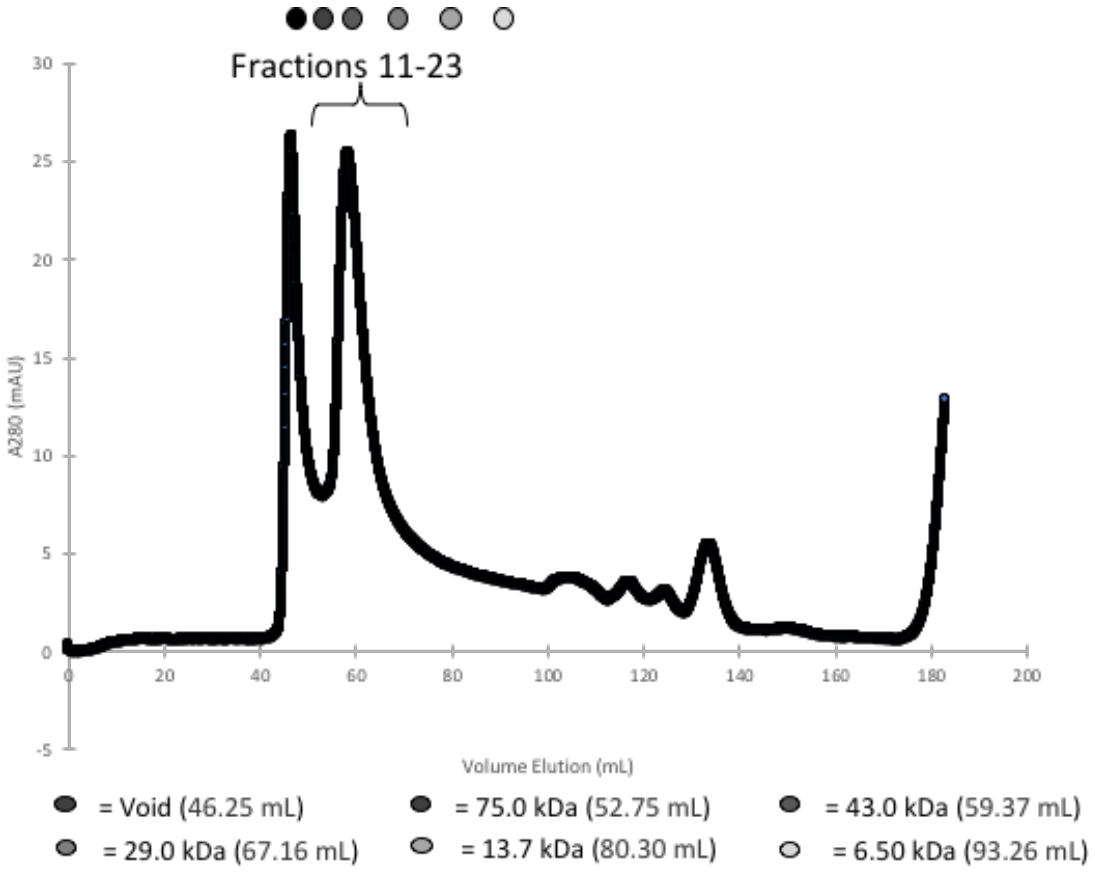

Figure 18. S75 size exclusion chromatogram of His_6-*Hs*LARP6 La Module (70-300). Pooled fractions from the affinity chromatography were concentrated against a 10,000 molecular weight cut off filter until reached a final volume of ~2 mL, and injected into the ÄKTA FPLC for gel filtration through the Sephadex S75 size exclusion column. Monitoring of the elutions was done using a UV absorbance at 280 nm, in which fractions 11-23 (45-70 mL) were collected, aliquoted, and analyzed on denaturing SDS-PAGE gel.

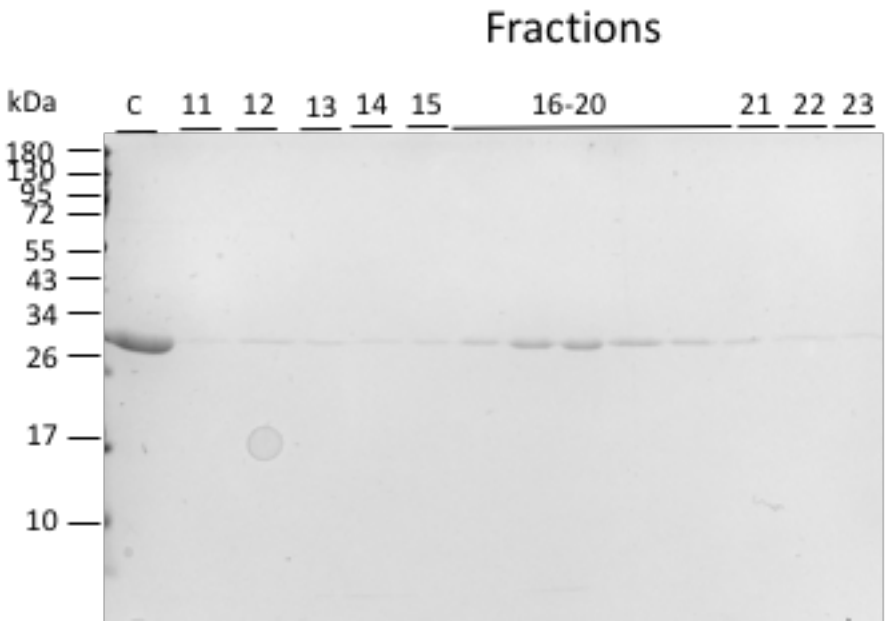

Figure 19. SDS-PAGE analysis of His_6-*Hs*LARP6 La module (70-300) post size exclusion. Based on the UV chromatogram fractions 11-23 were selected for denaturing analysis. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE gel and stained with Coomassie blue stain. Fractions 16-20 were pooled for further preparation.

A ratio of 2.5 moles: 2U of protein:thrombin was incubated on ice (4 °C) for 4 hours. Nickel affinity chromatography was used to separate non-cleaved protein from cleaved protein. Fractions of the nickel affinity were subjected to analysis using SDS-PAGE and visualized using Coomassie blue stain (**Figure 19**). Because the La Module constructs should not contain the His_6-tag, it was expected that the 25 kDa band would be present in either the flowthrough (FT) or wash fractions. Interestingly enough, the expected molecular weight band did not appear in either the flowthrough or wash fractions, but instead in the elution fractions (**Figure 19**). In order to preserve the protein, the elution fraction was pooled and concentrated against 10,000 Da molecular weight cut off and injected onto the S75 size exclusion column.

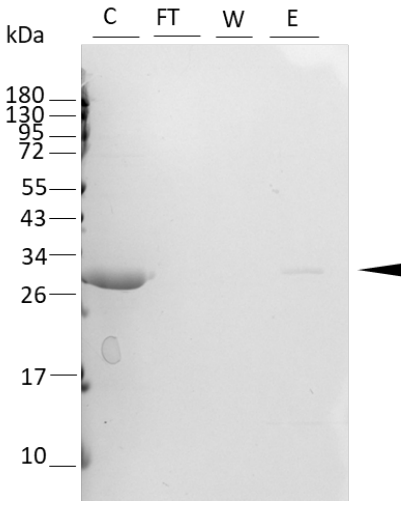

Figure 20. Affinity chromatography of thrombin-cleaved *Hs*LARP6 La Module (70-300). Pooled size exclusion purified protein constructs were incubated at a molar ratio of protein: thrombin (2.5 moles: 2 U) and allowed to incubate on ice (4 °C) for 4 h. Proteolysis reaction incubated with nickel-NTA beads and further nickel affinity purification was performed for isolation of the cleaved protein product. The flowthrough (FT), wash (W), and elution (E) were aliquoted and separated with an aliquot of pre-thrombin cleaved control (C). A faint band (indicated by the black arrow) can be seen to migrate at about ~30 kDa

Two low intensity peaks could be seen in the S75 UV chromatogram of the protein beginning at 45 and 55 mL (**Figure 21**). Fractions 11-23 were selected for analysis using SDS-PAGE gel and stained using Coomassie blue stain (**Figure 22**), where the band at 25 kDa can be faintly seen fractions 21-23. These fractions were pooled, concentrated against a 10,000 Da molecular weight cut off and stored in -70 °C.

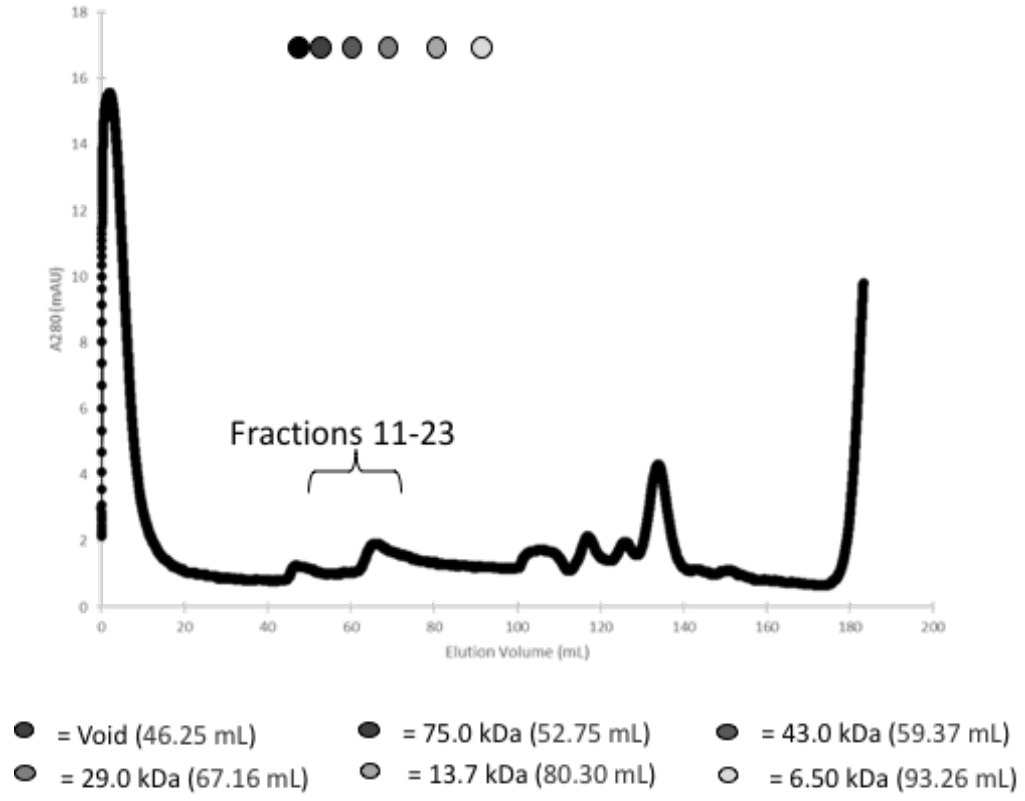

Figure 21. S75 size exclusion chromatogram of *Hs*LARP6 La Module (70-300) post thrombin-mediated removal of His_6-tag. Pooled fractions from the nickel-affinity purification of the La Module without the His_6-tag were concentrated to 2 mL final volume using a 10,000 molecular weight cut off filter and subsequently filtered and degassed through a 0.20 µm filter and injected into the pre-equilibrated (see Table 4 for buffer conditions) S75 size exclusion column. The size exclusion run was monitored by measuring the absorbance at 280 nm (A_{280}) through UV spectrometry.

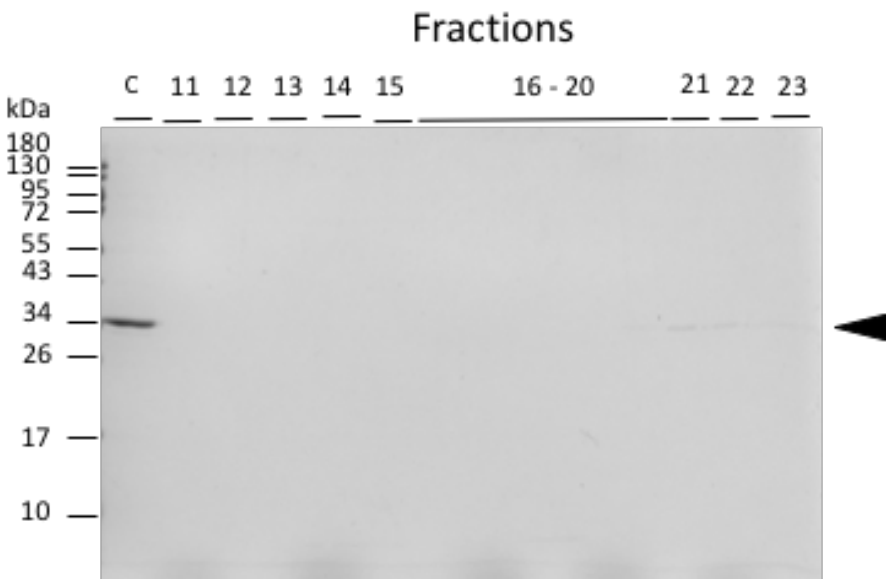

Figure 22. SDS-PAGE analysis of thrombin-cleaved *Hs*LARP6 La Module (0-300) post size exclusion purification. Fractions 11- 23 (45-69 mL) were selected for denaturing analysis. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE gel and stained with Coomassie blue stain. Faint bands (shown by black arrow) can be observed beginning at fraction 16-20 (49 –70 mL), which were then pooled, concentrated against a 10,000 molecular weight cut off down to a final volume of ~2 mL, and stored at 50 µL aliquots at -70 °C for further experiments.

Purification of zebrafish La Module without tag

For preparation of the zebrafish (*Dr*) LARP6A La Module (60-290) without the His_6-tag, an initial nickel affinity chromatography purification was performed with the clarified cell lysate and analyzed on the denaturing SDS-PAGE gel (**Figure 23**). A very high intensity band at 34 kDa can be seen in elution 1 with decreasing intensity as elution fractions were collected, indicating the presence of our protein of interest in high concentrations. The same 34 kDa band could be seen in the precursor wash fractions, which was consistent with earlier affinity chromatography purification of the His_6-*Dr*LARP6 La Module protein construct (**Figure 11**). In order to reduce the risk of concentration dependent protein aggregation, fractions 1 and 2 only were pooled and directly injected into the S75 size exclusion column.

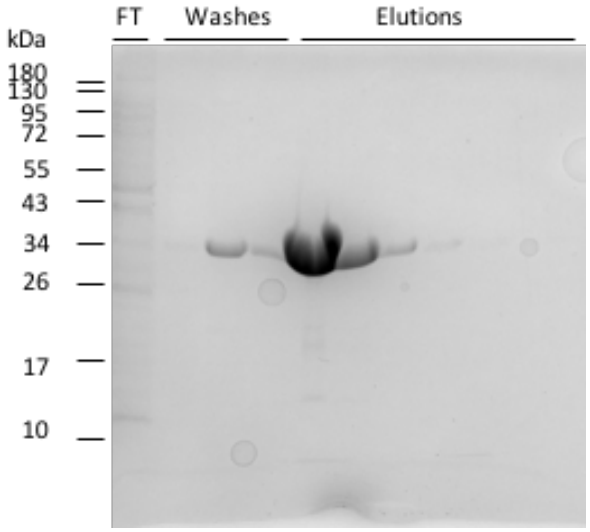

Figure 23. Affinity Chromatography of His_6-*Dr*LARP6 La Module (60-290). Cells containing the recombinantly expressed His6-tagged Zebrafish LARP6 La Module construct were lysed and purified using nickel-affinity chromatography. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE gel and stained with Coomassie blue stain. Elution fractions containing the protein of interest (30 kDa). Elution fractions 1 and 2 were collected, concentrated against a 10,000 molecular weight

cut off to ~2 mL, filtered through a 0.2 µm filter, and injected into the Ätka FPLC for gel filtration.

The S75 UV chromatogram (**Figure 24**) shows a single high intensity peak, high baseline resolution peak beginning at 55 mL of elution volume. The fractions 13-25 were selected for analysis on SDS-PAGE, where a single band could be seen with increasing intensity from fractions 13-21, with the highest peak intensity seen at fractions 18 and 19. The fractions 16-20 were pooled for further purification.

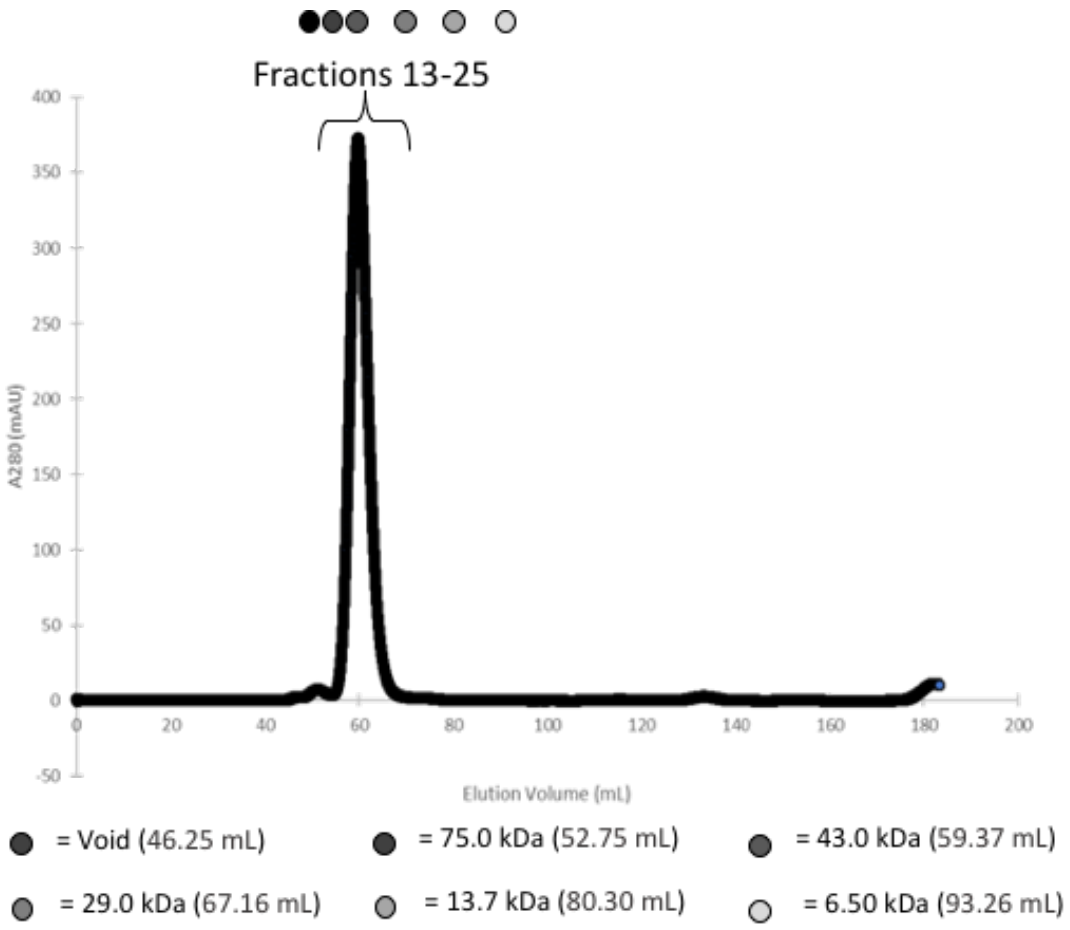

Figure 24. S75 size exclusion chromatogram of His_6-*Dr*LARP6a La Module (60-290). Pooled and prepared fractions of the affinity chromatography. Pooled, prepared, and injected fractions from the affinity purification were run on the Sephadex S75. A steep peak beginning from an elution volume of 57 mL was seen and the fractions 17-20 (55-

65 mL) of the column were collected and aliquoted for analysis through denaturing SDS-PAGE gel.

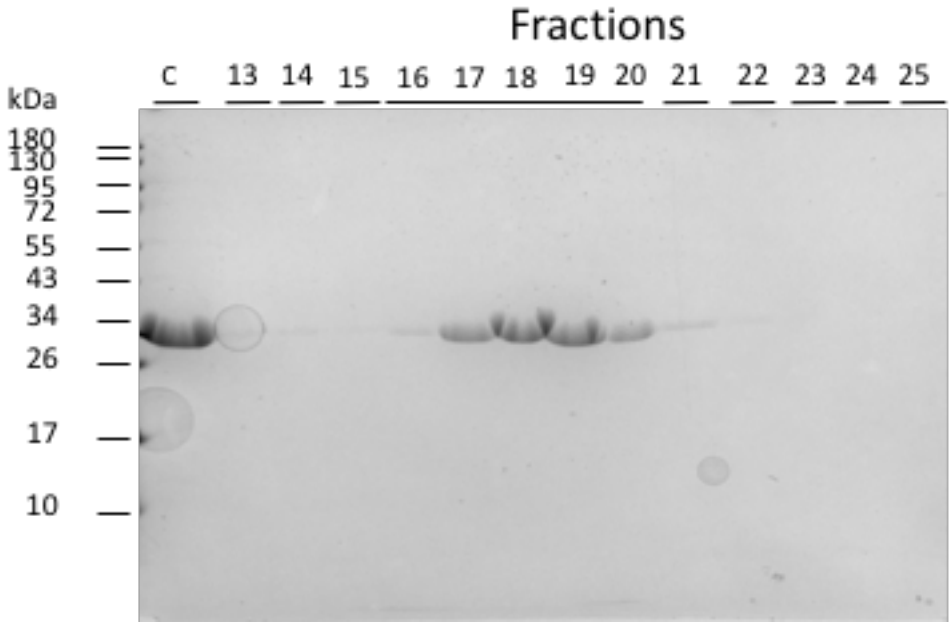

Figure 25. SDS-PAGE analysis of His_6-*Dr*LARP6a La Module (60-290) post size exclusion purification. Fraction 13-25 (49-69 mL) were selected for denaturing analysis. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE and stained with Coomassie blue stain. Fractions 16-20 were pooled for thrombin-cleavage of the His_6-tag.

To cleave the His_6-tag from the protein construct, thrombin was allowed to incubate with the pooled fractions from the previous size exclusion chromatography. A second nickel-affinity chromatography was done to separate the cleaved and non-cleaved products (**Figure 26**), where a single 26 kDa band was seen in elutions 1 and 2. It was expected that the single band would be seen in the flowthrough fraction. However, because the flowthrough did not contain imidazole, it is possible that non-specific interaction occurred between the Nickel-NTA beads and the protein.[34] Because the band seen in elutions 1 and 2 corresponded to the expected molecular weight of the cleaved protein construct (without the His_6-tag, 25 kDa), these fractions were collected and prepared for re-injection into the S75 size exclusion column.

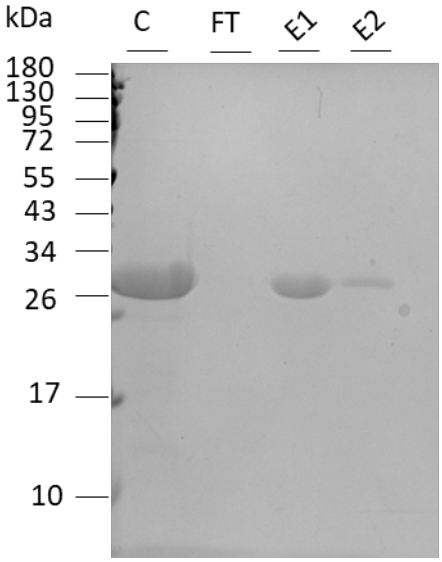

Figure 26. Affinity chromatography of thrombin-cleaved *Dr*LARP6a La Module (60-290). Pooled size exclusion fractions from the previous gel analysis (fractions 16-20) were pooled and allowed to incubate with thrombin at a molar ratio of protein: thrombin (2.5 moles:2 U) on ice for 4 °C for 1 hour. Proteolysis reaction incubated with the Nickel-NTA bead and further affinity purification was performed for isolation of cleaved protein product. The elution fractions were pooled and concentrated against a 10,000 molecular weight cut off filter down to a total of 2 mL, filtered against a 0.2 µm filter, and loaded into the ÄKTA FPLC for gel filtration through the Sephadex S75 size exclusion column.

The S75 size exclusion UV Chromatogram of the post-thrombin cleaved DrLARP6 La Module protein (**Figure 27**) shows a high intensity peak beginning at 59 mL of elution volume, corresponding to an approximate molecular weight of 49.5 kDa, which is an approximate 20 kDa increase from the expected molecular weight (~30 kDa). Fractions 13-25 were selected for analysis on denaturing SDS-page gel (**Figure 28**), where a single band at 34 kDa can be observed between fractions 13-20, with increasing intensity, which is close to the expected molecular weight of the protein (30 kDa). Because fractions 16-20 corresponded with the fractions covering the peak observed in

the UV chromatogram, these fractions were pooled, concentrated, and stored in -70 °C for further experiments.

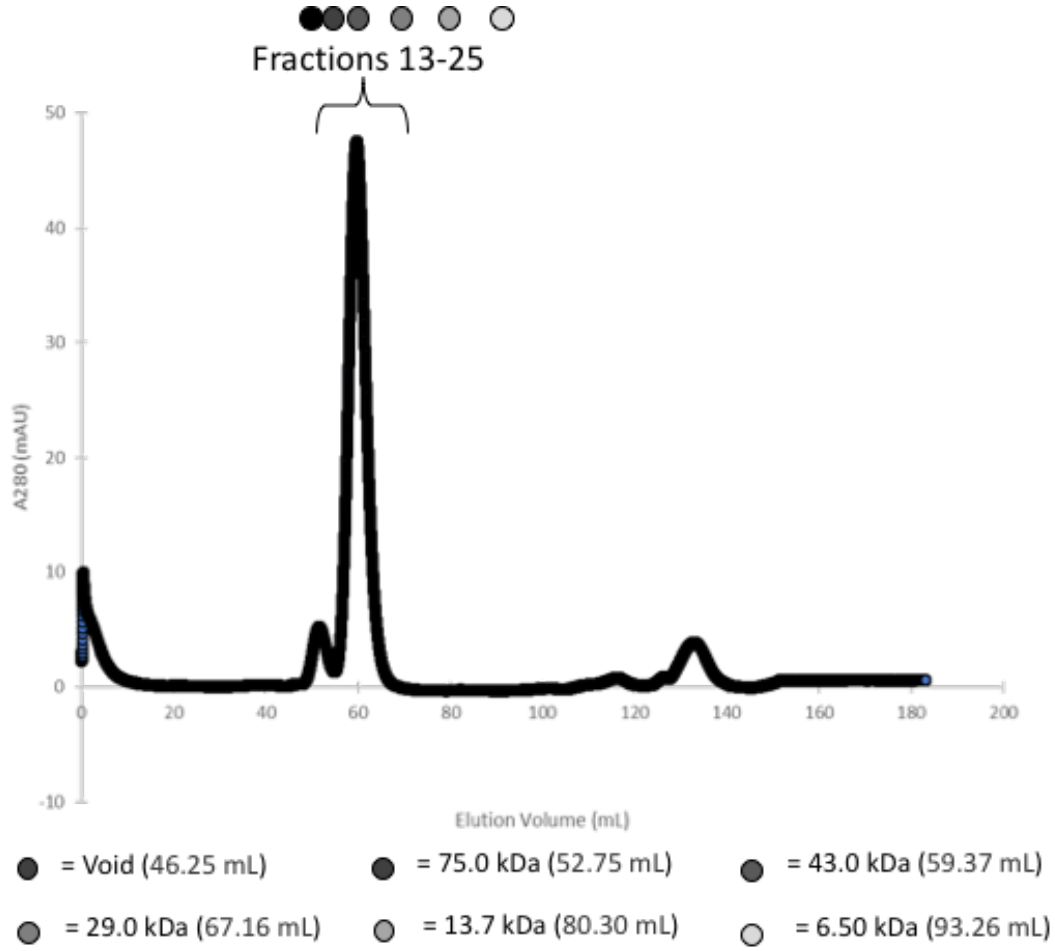

Figure 27. Gel filtration UV chromatogram of thrombin-cleaved *Dr*LARP6a La Module (60-290). The pooled nickel-affinity fractions were concentrated against a 10,000 molecular weight cutoff filter until reached a final volume of 2 mL, filtered through a 0.20 µm, and injected into the ÄKTA pure FPLC for gel filtration using the Sephadex S75 size exclusion chromatogram. A large peak can be seen starting at elution volume 58 mL (fraction 16). A secondary low intensity band can be seen in at the void volume (46.25 mL).

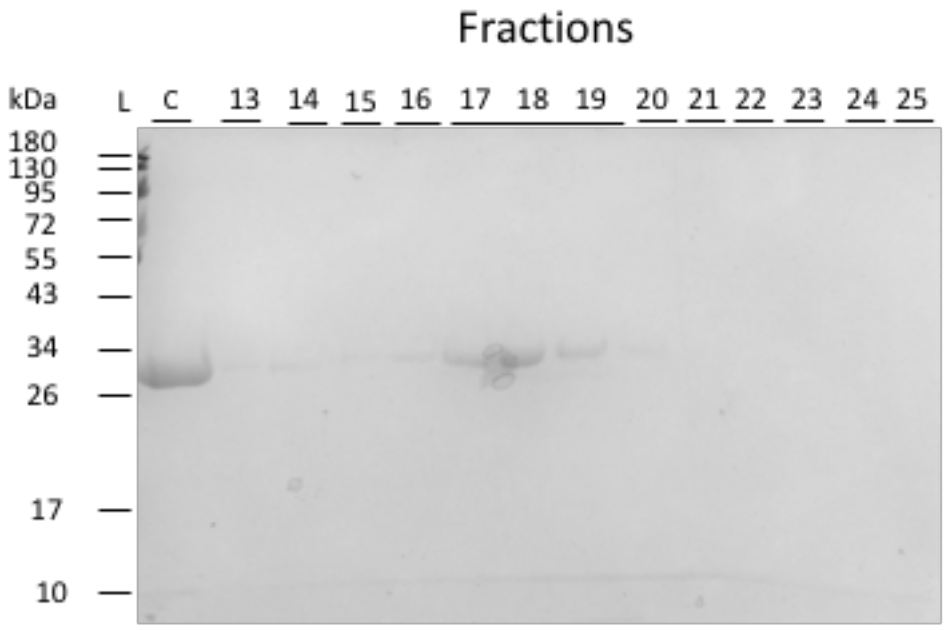

Figure 28. Denaturing gel analysis of thrombin-cleaved *Dr*LARP6a La Module (60-290) post size exclusion purification. The post-thrombin cleaved affinity purified zebrafish LARP6 La Module protein injected onto the Sephadex S75. The column was monitored using an absorbance at 280 nm, in which fractions 13-25 (49-69 mL) were selected for analysis through the denaturing SDS-PAGE gel with an (C) aliquot of sample that was loaded to serve as a control. The gel was visualized with Coomasie blue stained. Bands at about 30 kDa can be seen in the gel in correspondence to our protein of interest (30 kDa) in fractions 17-19. These fractions were pooled, concentrated down to 2 mL against a 10,000 molecular weight cut off, and stored at 50 µL aliquots at -70 °C for future experiments.

Since the second nickel-affinity purification of the untagged protein resulted in our protein of interest associating with the column, a comparison of the purified untagged and tagged *Dr*LARP6 La Module was performed (**Figure 29**). The anti-His_6 probe relies on the electrostatic interaction between a nickel cation and the lone pair on the histidine residues, which allows for more specific association. Therefore, the lack of signal seen in the thrombin-treated *Dr*LARP6 La Module protein confirms the successful removal of the His_6-tag.

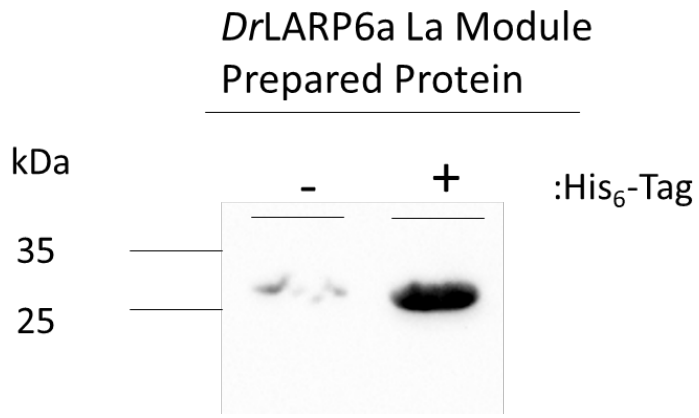

Figure 29. Confirmation of His_6-tag removal from *Dr*LARP6a La Module construct. The previously purified zebrafish La module with the His_6-tag (+) as the preparation with the added thrombin digest (-) in the purification were run on denaturing SDS-PAGE gel, transferred onto nitrocellulose, and detected using the α-His_6 western blotting.

Biochemical Characterization of La Module constructs

In preparation for the biochemical characterization of the La Module constructs, the human and zebrafish *COL1a1* and *COL1a2* RNA stem-loop sequences were commercially synthesized (IDT) and then enzymatically biotinylated on the 3' end. The efficiency of the biotinylation reaction was determined by spot testing the biotinylated

oligos and serial dilutions of the 25% diluted commercially biotinylated IRE control for comparison using 2 µL spots (**Figure 30**). The signal of the serial dilutions of the 25% diluted IRE control appears to be weak, direct comparison of the efficiency of biotinylation could not be obtained. However, based on the intensity of signal, it appears that the zebrafish (*Dr*) *COL1a1* and *COL1a2* RNA ligands were more efficiently labeled than the human (*Hs*) *COL1a1* and *COL1a2* RNA ligands.

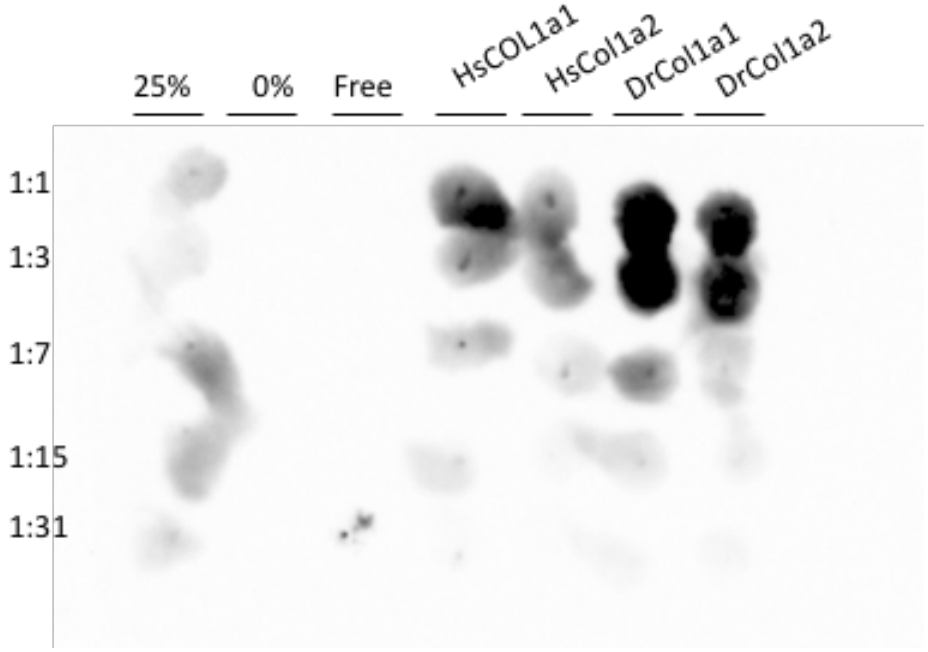

Figure 30. Biotinylation efficiency of *COL1a1* and *COL1a2* RNA ligands from human and zebrafish. All RNA oligo ligands were obtained from Integrated DNA Technologies (Skokie, IL) and biotinylated using the Thermofisher Scientific 3' end RNA biotinylation kit as previously described above. 400 pM of RNA oligo ligands were biotinylated. Efficiency was determined by performing a 1:2000 dilution of the biotinylated ligands and compared to a 0% and 25% solution (dilution of 100% stock) of the IRE commercially biotinylated RNA control. Two µL of the 100% commericially biotinylated IRE control is the 5th spot of the "Free" lane. The diluted ligands and the commercially bionylated contorl were then serially diluted by a factor of 1:1 until a final dilution of 1:31 was achieved. The serially diluted samples were than blotted at an aliquot of 2 µL onto a pre-equilibrated Hybond+ membrane and detected using the

ThermoScientific® Chemiluminescence Nucleic Acid detection module as previously described.

Because the molecular weights of the La Module protein constructs are significantly smaller than that of the full-length proteins (a difference of ~2 fold difference), it was necessary to optimize native gel condition to clearly resolve La Module-bound RNA from unbound RNA. For these protein constructs, comparison of 6.0% and 6.5% native polyacrylamide gels were prepared with an abbreviated six-point binding reaction performed using the *Hs*LARP6 La Module protein construct against the *HsCOL1a1* RNA, where the relative distance of separation of the bound and unbound signal was determined. The gels were transferred onto Hybond+ membrane, crosslinked, and detected using the RNA Light Shift Chemiluminescence Kit (**Figure 31**). The separation of the bound and free RNA appears to be greater in the 6.5% native gel than in the 6.0 % native gel, and so it was determined that the 6.5% native gels would be used for the La Module EMSAs.

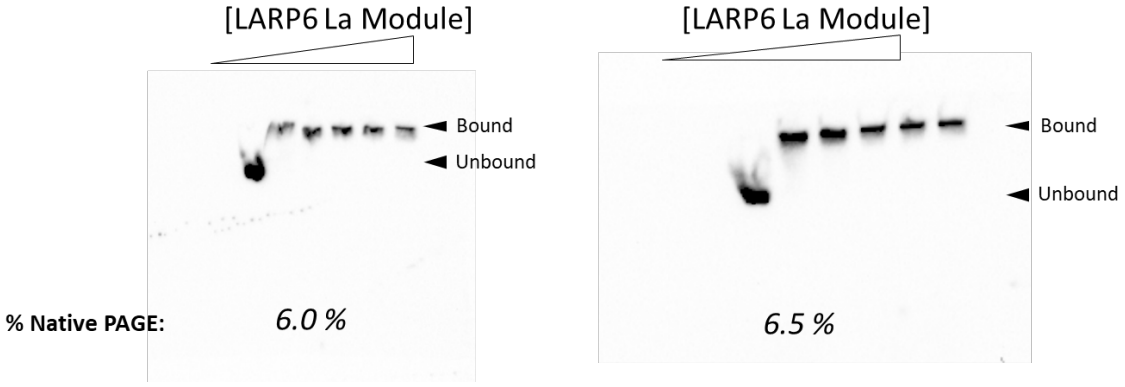

Figure 31. Native PAGE optimization for EMSAs of the LA Module protein constructs. 6.0% and 6.5% native polyacrylamide gels were used for direct comparison. The binding reaction was a previously purified His_6-HsLARP6 La Module (70-300) and the biotinylated *Hs*COL1a1 ligand. The reactions were allowed to incubate in 1x Binding Buffer on ice for 1 hour before 20 µL of each reaction was loaded onto each gel, run at 90 V, 30 mins, in cold 1x TBE. The gels were processed as described in Methods and

detected with the ThermoScientific® Chemiluminescence Nucleic Acid detection module as described.

In order to determine a starting comparative $K_{D,\ app}$ value, reactions using the previously characterized *HsCOL1a1* RNA ligand was used. The separation of the binding reaction involving the human La Module showed the expected separation of bound and unbound RNA (**Figure 32**). However, for the binding reactions involving the platyfish (*Xm*) and zebrafish (*Dr*) La Modules, a band is visible above the "bound" region. This band is characteristic of a well shift, of which the most likely causes could be either protein aggregation or oligomerization.[32] These well shifts, however, were still included in the gel densitometry quantification, as these RNA ligands appear to still be complexed, regardless of the phase the protein constructs appear to have adopted.

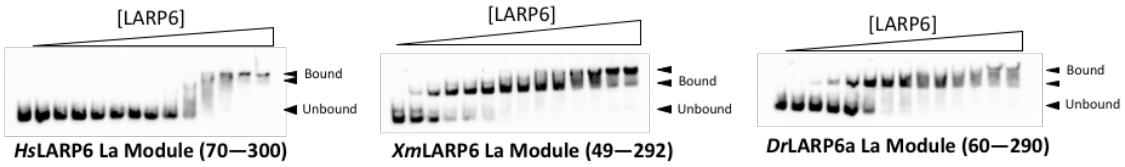

Figure 32. Representative EMSAs of La Modules binding to RNA. Serial dilutions of the human (Hs), platyfish (Xm), and zebrafish (Dr) La Modules were allowed to incubate with 1.25 nmoles of biotinylated HsCOL1a1 RNA oligo, and then bound and free RNA separated on a 6.5% Native PAGE. In the gels shown here, the reactions were separated at either 90 V for 30 mins (HsLARP6 La Module and DrLARP6 La Module) or 200 V, 15 min (XmLARP6 La Module). The RNA was then detected using the ThermoScientific® Chemiluminescence Nucleic Acid LightShift kit.

The fraction of total RNA bound in reaction was quantified using gel densitometry and then plotted as a function of the LARP6 La Module concentration (**Figure 33**). In order to fit the $K_{D,\,app}$ the Simplified Isotherm Equation for binding was first used:

$$\frac{[PL]}{[L]_T} = S\left(\frac{[P]_T}{[P]_T + K_{D,app}}\right) + 0$$

where [PL] is the concentration of the protein ligand complex, $[L]_T$ is the total concentration of ligand, and $[P]_T$ is the total protein in the system. The fitted $K_{D,app}$ of the human LARP6 La Module was 170 nM (± 31.0) which is a ~3-fold difference than previously reported value of 48 nM.[9] Interestingly enough, the measured $K_{D,\,app}$ for both fish La Module protein constructs were in the low nM range (*Dr*LARP6a La Module, $K_{D,\,app}$ = 430 pM [± 267]; *Xm*LARP6 La Module, $K_{D,\,app}$ = 256 pM [± 72]) (**Figure 33**). The baselines of the binding reactions of the fish proteins were not fully achieved in the plotting of the binding data. When the concentration of the biotinylated RNA stocks was quantified, it was determined that the concentration of the RNA ligands was higher than what was expected. To be able to properly negate the influence the ligand concentration has to the bound versus free state of the protein, the concentration of the ligand must be at least 10-fold that of the $K_{D,app}$ of the protein.[32]

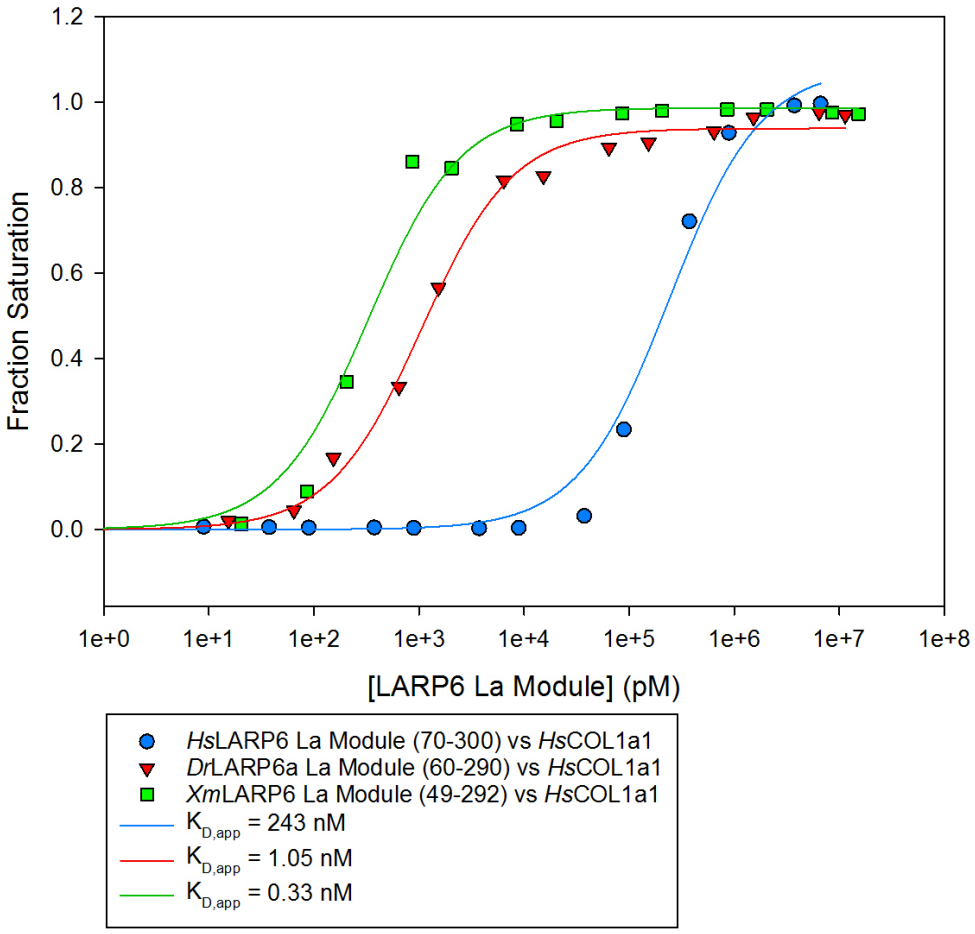

Figure 33. Representative graph for La Module protein constructs vs *Hs*COL1a1 RNA oligo ligands EMSAs. The quantification from each gel was plotted in Sigma Plot (London, UK) as the fractional saturation ($[RNA]_{bound}/[RNA]_{free}$) vs. the concentration of LARP6 La Module (in nM). The plot was fitted using the simplified binding isotherm equation. The fitted $K_{D,\,app}$ are as follows: for the human (*Hs*LARP6 La Module, $K_{D,\,app}$ = 243 nM), zebrafish (*Dr*LARP6 La Module, $K_{D,\,app}$ = 1.05 nM), and platyfish (*Xm*LARP6 La Module, $K_{D,\,app}$ = 0.33 nM). Consistent with what has been observed with the full length, the two fish LARP6 La module protein constructs seem to be binding to the *HsCOL1a1* ligand with higher apparent affinity.

$K_{D,app}$ *fitting with known concentration of the Ligand*

Because the apparent K_D of the fish appears to be at or near the concentration of the ligand, it is not appropriate to use the simplified version of the binding isotherm. Therefore, we attempted to fit the data using the quadratic form of the binding isotherm [33]:

$$\frac{[PL]}{[L]_T} = \frac{\left(K_{D,app} + [P]_T + [L\}_T\right) - \sqrt{(K_{D,app} + [P]_T + [L\}_T)^2 - 4[P]_T[L\}_T}}{2[L\}_T}$$

where $\frac{[PL]}{[L]_T}$ is the fraction saturation, $[P]_T$ is the total concentration of the protein, and $[L]_T$ is the total concentration of ligand in the system. This was used in attempt in to remedy the issue with using the Simplified Binding Isotherm equation. The measured mean $K_{D,\ app}$ using this equation for both fish La Module protein constructs still appear to reside in the subnanomolar range (*Dr*LARP6a La Module, $K_{D,\ app}$ = 0.180 ± 0.16 nM; *Xm*LARP6 La Module, $K_{D,\ app}$ = 0.128 ± 0.15) (**Table 5**). Because the fitting for the $K_{D,\ app}$ for the two fish is significantly lower than the concentration of ligand used in this study, we were unable to properly fit the data using either equation. However, the concentration of the ligand was well below the measured $K_{D,\ app}$ for human La Module. It appears that fitting of the data for the human La Module using either form of the equation gives us a consistent value, therefore we are able to confidently associate a $K_{D,app}$ of 170 ± 31 nM.

Table 5. Fitted $K_{D,app}$ of La Module protein constructs vs *Hs*COL1a1 using Quadratic Form

Protein Name	*Mean* $K_{D,\,app}$ *(*nM*)*	*Error (S.E.M.)*	*N=*
*Hs*LARP6 La Module (70-300)	170	31	3
*Dr*LARP6a La Module (60-290)	0.180	0.160	3
*Xm*LARP6 La Module (49-292)	0.128	0.150	3

These data are consistent with the observation by Jose M. Castro with the full-length LARP6 protein.[16] However, a molecular explanation for this 1000-fold difference in apparent binding remains to be identified. To explore whether this is the result of a difference in the mechanism of binding between the orthologs, further structural characterization and perhaps dynamic modeling of the protein-RNA interaction will be necessary.

IV. TESTING THE ROLE OF CYSTEINES IN THE STRUCTURE AND FUNCTION OF LARP6

Recombinant expression of serine mutants of human LARP6

In order to determine whether cysteines in the C-terminal domain are responsible for structural stability of LARP6, the conserved cysteines in the wildtype human LARP6 gene were targeted for mutagenesis in a pET28a vector (**Figure 34**). The human LARP6 proteins will hereafter be referred to as "*Hs*LARP6", to distinguish from the fish homologs that are described in Chapter III. Three cysteines were initially chosen for mutagenesis in the human ortholog: C258 within the RNA Recognition Motif subdomain that was fully conserved in all vertebrates, C378 upstream of the LSA, and C490 near the extreme *C*-terminus of the protein. To conserve the overall electrochemical property of these positions in the protein, but still disable the formation of covalent binding, these cysteine residues were targeted using site-directed mutagenesis to convert them into serines. However, only the single point mutation in the in the coding sequence of the full length *Hs*LARP6 protein for the C258S and C490S were successfully achieved. In order to progress the project, these mutations were the only ones that were purified and characterized.

Sequence-verified pET28a-*Hs*LARP6-C258S, pET28a-*Hs*LARP6-C490S, and pET28a-HsLARP6-C258S/C490S (double mutant) expression plasmids were then transformed into Rosetta™ DE3 *E. coli* cells and expression was induced with IPTG. Expression of the wildtype HsLARP6 (**Figure 34A**), *Hs*LARP6-C258S protein (**Figure 34B**), and HsLARP6-258S/C490S double mutant (**Figure 34D**) all showed increased

over time after induction, with a small faint band being visible after 2 hours of induction, indicating expression of the *Hs*LARP6-C258S mutant. The *Hs*LARP6-C490S (**Figure 34C**) mutants showed the same trend, however protein production was detectable until after the fourth hour.

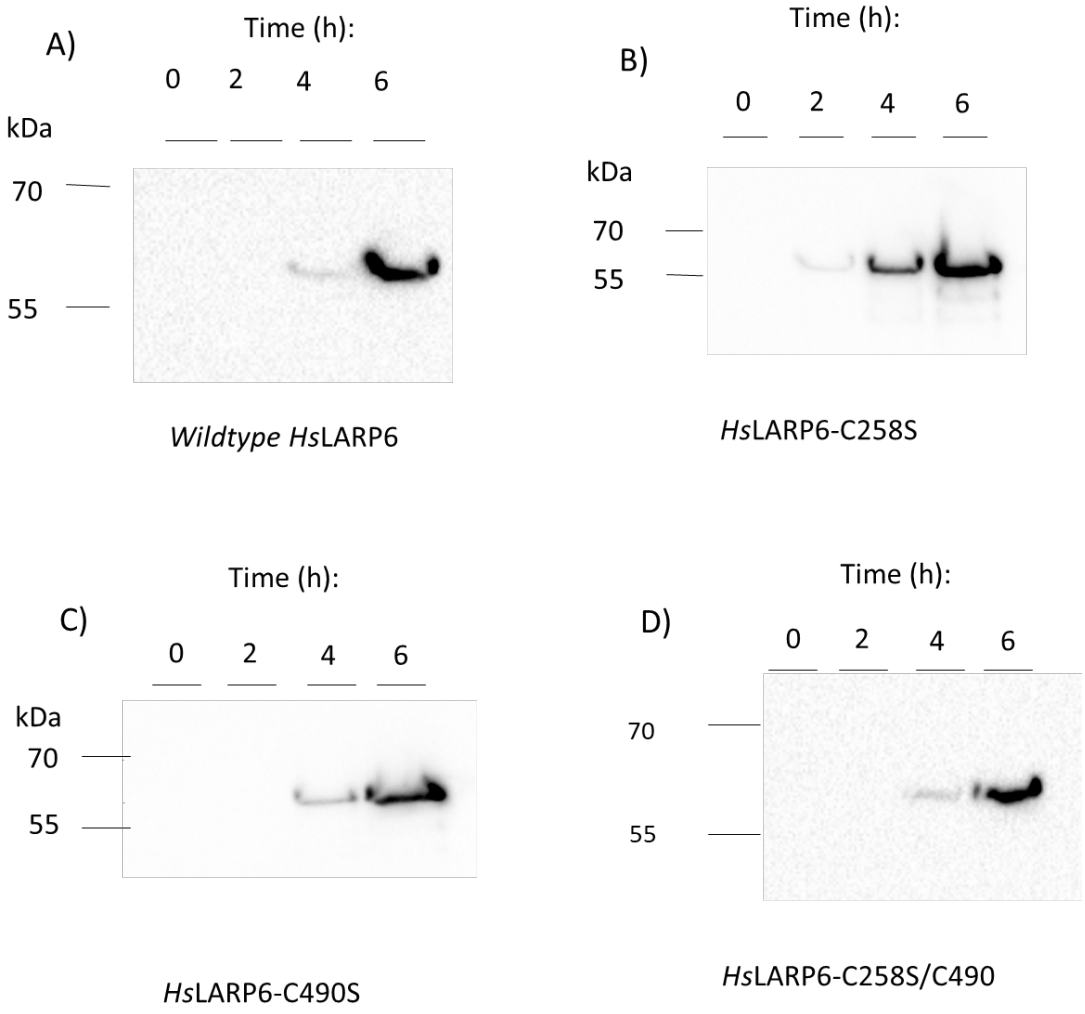

Figure 34. Expression trial of *Hs*LARP6 mutants. Plasmids containing the cysteine-to-serine mutants in the *Hs*LARP6 sequence were transformed into Rosetta (DE3) cells and grown in LB containing kanamycin and chloramphenicol. Culture was grown in large scale in 1 L LB Broth containing appropriate antibiotics until the culture for wildtype *Hs*LARP6, *Hs*LARP6-C258S, *Hs*LARP6-C490S, and *Hs*LARP6-C258S/C490S (double mutant) reached an OD600 of 0.5-0.7 before induction with IPTG, and 1 mL aliquots removed to assay expression by western blot at 0, 2, 4, and 6 h of induction. Production of the *Hs*LARP6 mutant proteins was verified through interaction of an anti-His_6 probe with the in-frame N-terminal His_6-tag.

Reducing agent affects apparent R_h of wildtype HsLARP6

Preliminary data in our laboratory indicated that the hydrodynamic radius of *Hs*LARP6 is sensitive to reducing agent (**Figure 35**). During protein purification, a shift in elution volume can be observed in the preparative S200 size exclusion column when DTT is added to the elution buffer. In both elutions, there are two peaks observed: an earlier peak ("Peak 1") at 69 mL, representing an approximate globular molecular weight of ~256 kDa, and a later peak ("Peak 2") at 87 mL (~60 kDa). When the elution buffer is doped with the reducing agent DTT, the population of Peak 2 (~60 kDa) decreases, and the population in Peak 1 increases. Because size exclusion separates by the overall hydrodynamic radius, it can be assumed that this shift in the high intensity peak corresponds to a shift in a conformation that allows for a much larger hydrodynamic radius when DTT is present to a conformation in which the LARP6 protein adopts a smaller hydrodynamic radius.

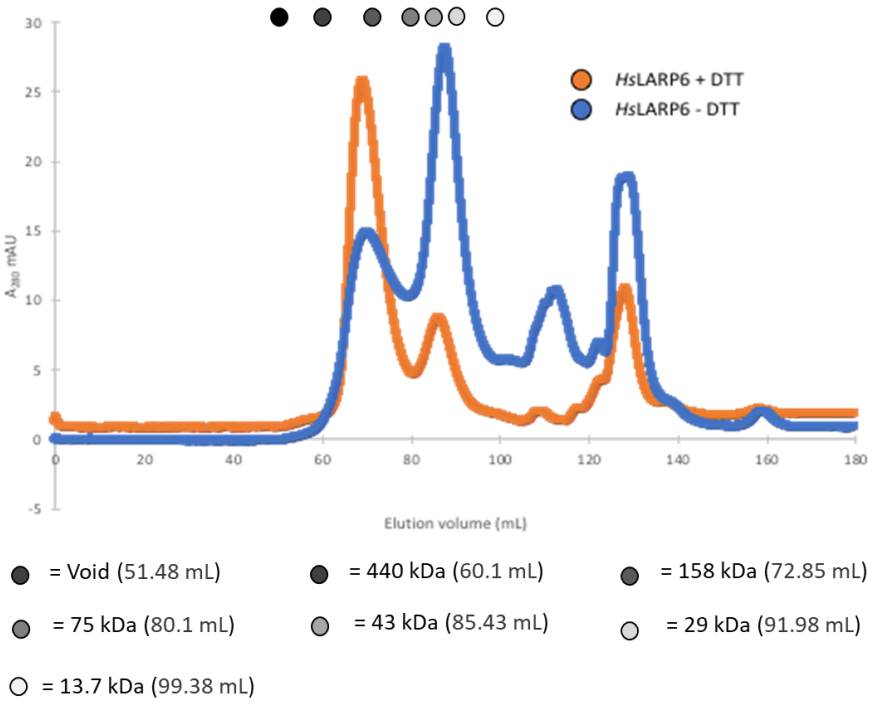

Figure 35. Reducing agent alters apparent hydrodynamic radius of *Hs*LARP6. Affinity-purified protein sample was prepared and purified using the Sephadex S200 preparative size exclusion column and monitored elution using A_{280} (UNICORN software, GE Life Sciences). Two peaks are observed: Peak 1 (~60 mL; 256 kDa) to Peak 2 (~85 mL; 60 kDa). A Clear shift of the areas under the curve was observed, indicating a distinct difference of the hydrodynamic radius when DTT is present in the storage buffer. The expected molecular weight of *Hs*LARP6 proteins is 55 kDa.

Purification of HsLARP6-C258S mutant

Cells containing the *Hs*LARP6-C258S protein construct were lysed and purified using affinity chromatography (**Figure 36**). When analyzing the elution fractions through denaturing SDS-PAGE gel electrophoresis, all of them contained a high intensity band that migrated at a distance in the gel corresponding to a ~60 kDa globular protein. However, a second band was also observed in some fractions that corresponded to a protein of ~55 kDa. The presence of the secondary co-eluting species was seen diminish after the second elution fraction. Unfortunately, because the fractions containing the two-migrating species also contained high concentrations of our protein of interest, these fractions were chosen for further purification by size exclusion chromatography.

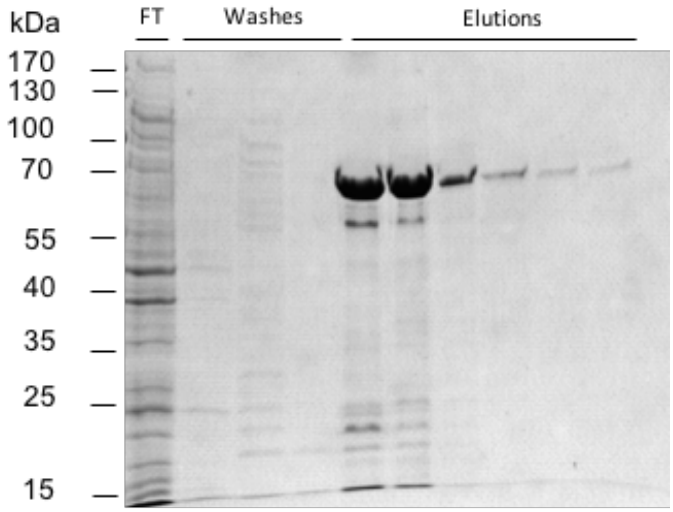

Figure 36. Affinity Chromatography of His_6-*Hs*LARP6-C258S. Cells containing the His_6-HsLARP6-C258S mutant were recombinantly expressed in Rosetta™ DE3 E. coli, lysed and purified using nickel affinity chromatography. The flowthrough (FT), washes, and elutions were collected and aliquots were analyzed on denaturing 10% SDS-PAGE gel electrophoresis and visualized using Coomassie blue stain. Elutions 1-4 were collected for further purification.

The protein was prepared for size exclusion purification in a storage buffer that did not contain the reducing agent DTT, in order to mimic oxidizing condition. The UV chromatograph (**Figure 37**) showed a distribution like that of the wildtype LARP6 in the absence of DTT, indicating a smaller hydrodynamic radius. The elution fractions 22- 34 were selected based on the UV chromatogram for analysis using SDS-PAGE. The presence of the co-eluting species was observed for fractions 23-28 (**Figure 38**). Without an obvious way to explain this co-eluting species, these fractions were collected and stored for further analysis.

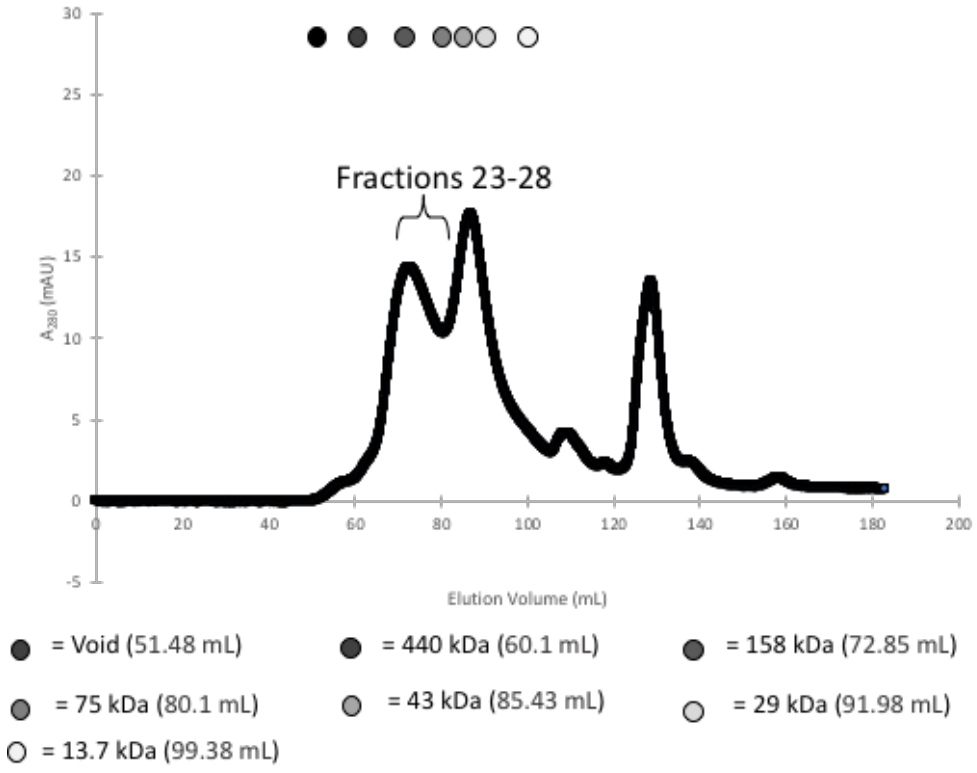

Figure 37. Sephadex S200 size exclusion chromatogram of His_6-tagged *Hs*LARP6 C258S in the absence of DTT. Flowthrough and elutions of the Sephadex S200 column were collected and concentrated against a 10,000 MW cutoff filter before loading into the FPLC. This preparative column was eluted with S200 storage buffer (**Table 4**) without DTT. Peak fractions were collected for analysis by denaturing gel electrophoresis.

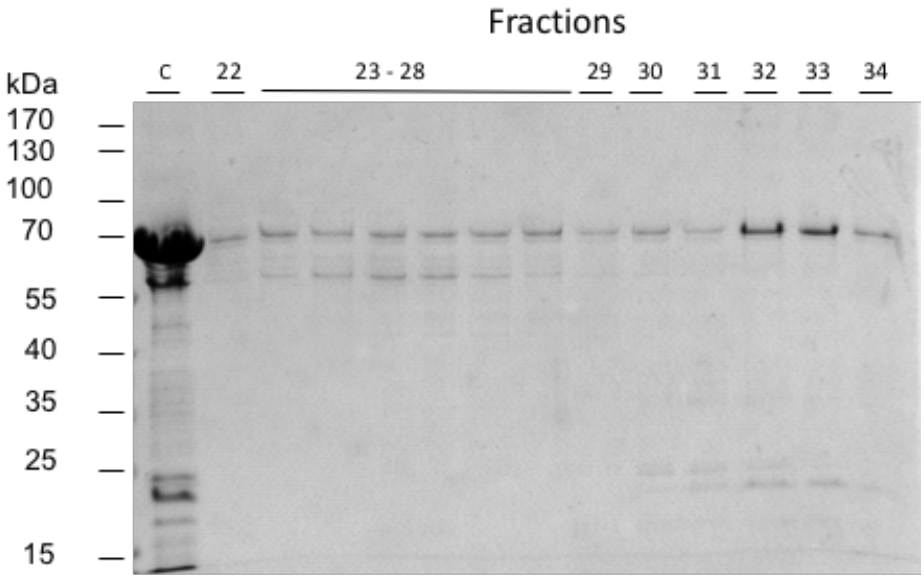

Figure 38. SDS-PAGE analysis of *Hs*LARP6-C258S S200 fractions in the absence of DTT. Fractions 22-34 (66-92 mL) were separated by gel electrophoresis on denaturing 10% SDS-PAGE gel and stained with Coomassie blue stain. Fractions 23-28 (69-80 mL) were pooled, concentrated using a 10,000 kDa molecular weight cutoff centrifugal concentrator, and stored in 50 µL aliquots at -70 °C or further experiments.

In order to determine effects of DTT on the *Hs*LARP6-C258S mutant, a separate purification was performed, with DTT added to only the storage buffer. The *Hs*LARP6-C258S was expressed in Rosetta® DE3 *E. coli* cells, pelleted, and lysed as previously described. Isolation of the His_6-tagged *Hs*LARP6 C258S from the crude lysate was done through nickel-affinity chromatography and the aliquots of each fraction were separated on denaturing SDS-PAGE (**Figure 39**). The appearance of high intensity band migrating at ~65 kDa can be seen beginning in the first elution fraction, with decreasing intensity throughout the progression of the elution fractions. The presence of the secondary co-eluting species at 55 kDa can be seen in the elution fractions. This is consistent with the nickel-affinity chromatography analysis from the previous *Hs*LARP6-C258S purification.

Lower molecular weight bands can be seen in the first 2 elution fractions which is characteristic of degradation products that co-eluted from the column.

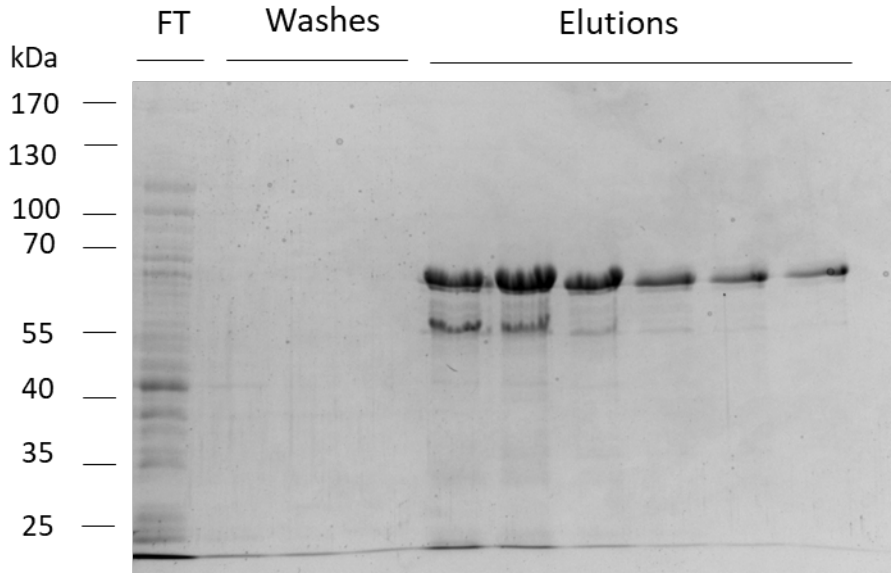

Figure 39. Affinity chromatography of His_6-*Hs*LARP6C258S. Cells containing the His6-HsLARP6C258S mutant were recombinantly expressed in Rosetta™ DE3 E. coli, lysed and purified using nickel-affinity chromatography. The flowthrough (FT), washes, and elutions were collected and aliquots were analyzed on denaturing 10% SDS-PAGE gel electrophoresis. Elutions 1-4 were collected for further purification.

The pooled nickel elutions were then purified by size exclusion chromatography. The UV chromatogram (**Figure 40**) of the S200 elution showed a population distribution similar to that of the wildtype + DTT (**Figure 35**). The large peak at 64 mL indicated a larger hydrodynamic radius. Fractions 11-23 were selected for further analysis using SDS-PAGE. The presence of the co-eluting species was seen in fractions 22 and 23, indicating that the addition of DTT in the storage buffer did not eliminate this species (**Figure 41**). Fractions 19-23 were pooled, concentrated, and stored at -70ºC.

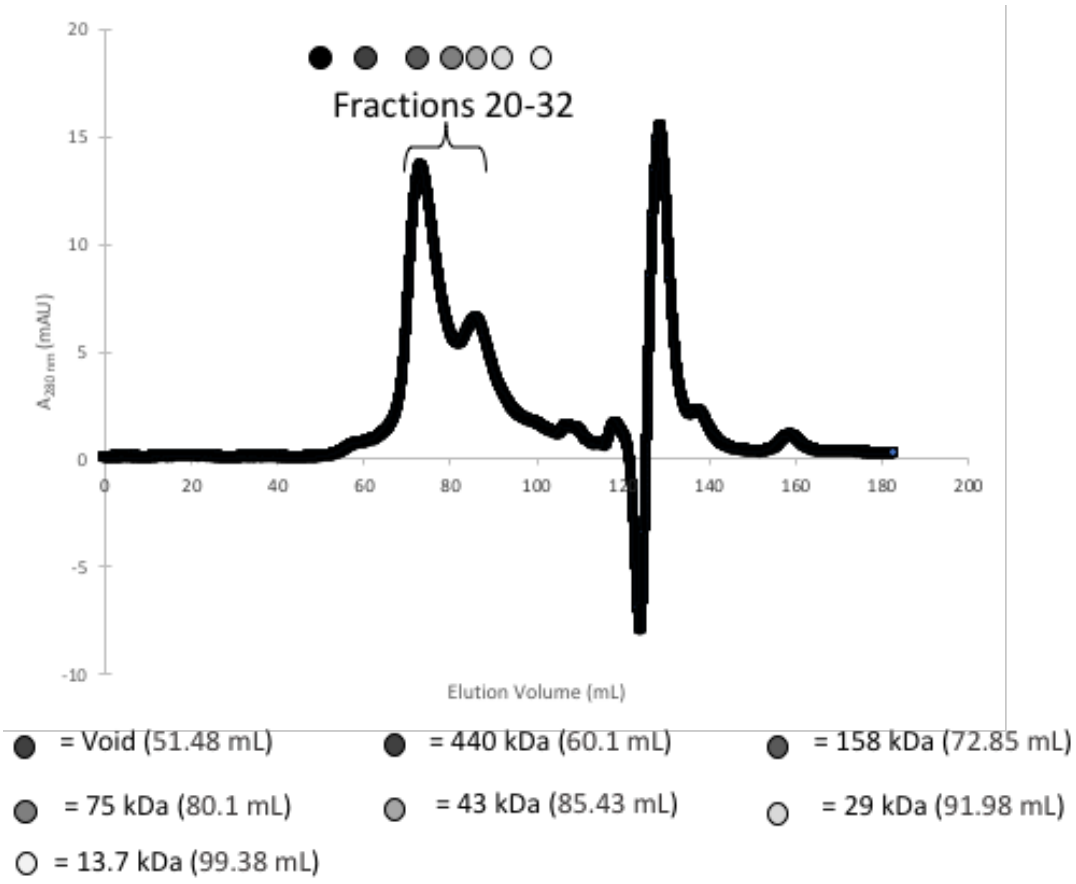

Figure 40. S200 size exclusion chromatogram of His_6-tagged *Hs*LARP6-C258S+DTT. The pooled fractions of the affinity chromatography were concentrated against a 10,000 MW cut off filter and loaded onto the S200 column. This preparative column was run in S200 Storage Buffer (Table 4) with 1 mM DTT. Peak 1 (fractions 21-29, 64-82 mL) and Peak 2 (fractions 30- 34, 83-93 mL) were observed and collected for analysis through denaturing gel.

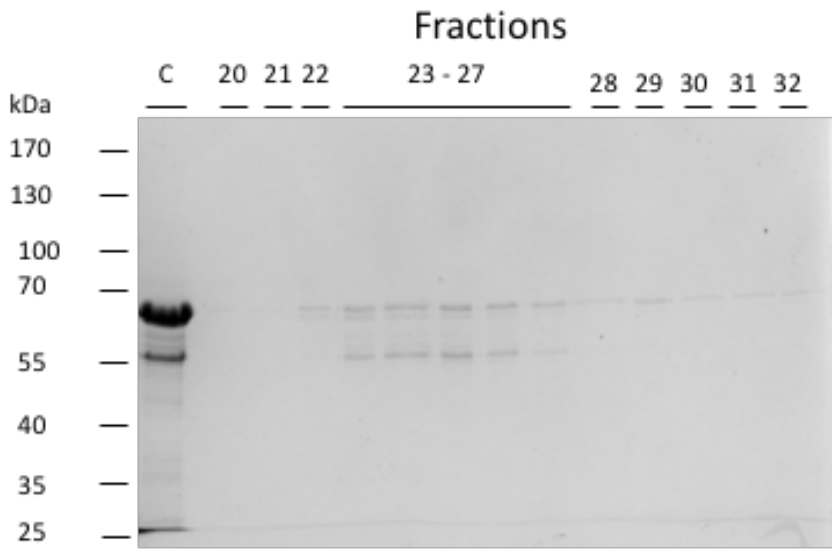

Figure 41. SDS-PAGE analysis of size exclusion chromatography of *Hs*LARP6-C258S +DTT. Fractions 20 - 32 (69 – 93 mL) were selected for denaturing analysis. Aliquots were separated by gel electrophoresis on denaturing SDS-PAGE gel and stained with Coomassie blue stain. Fractions 23-27 were pooled, concentrated against a 10,000 molecular weight cut off, and stored at 50 µL aliquots at -70 °C for further experiments.

Purification of* Hs*LARP6-C490S mutant

The *Hs*LARP6-C490S mutant was subjected to the same purification process as the previous mutant, beginning with using purification buffer without DTT. Initial nickel-affinity chromatography (**Figure 42)** showed the presence of the smaller ~55 kDa species in elution fractions 1-3. The first four fractions were selected for injection into the S200 size exclusion column for further purification.

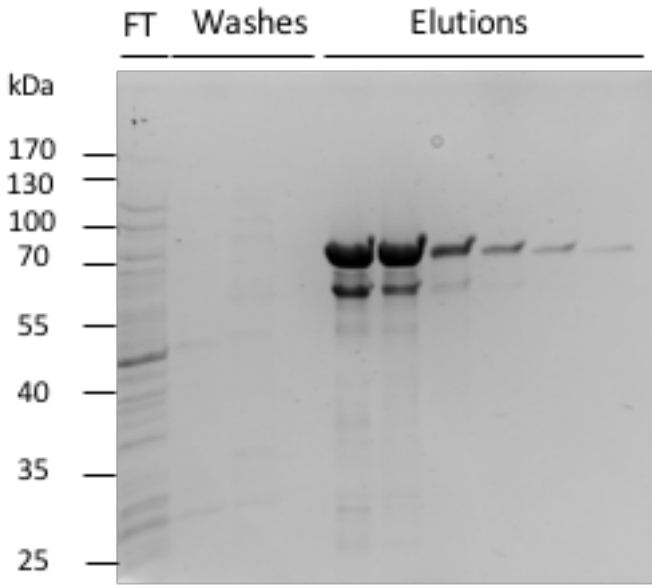

Figure 42. SDS-PAGE analysis of affinity chromatography purification fractions of *Hs*LARP6-C490S. Cells containing the His_6-tagged HsLARP6-C490S mutant were recombinantly expressed in Rosetta™ DE3 E. coli and purified using nickel-affinity chromatography. The Flowthrough (FT), washes, and elutions were collected and aliquots were analyzed on denaturing 10% SDS-PAGE gel electrophoresis. Elutions 1-4 were collected for further purification

The S200 size exclusion UV chromatogram (**Figure 43**) shows an elution volume for *Hs*LARP6-C490S in the *absence* of DTT that is similar to that of the wildtype *Hs*LARP6 in the *presence* of DTT. This result suggests that the cysteine at position 490 may play a role in the formation of the structure that produces the larger apparent hydrodynamic radius of *Hs*LARP6+DTT.

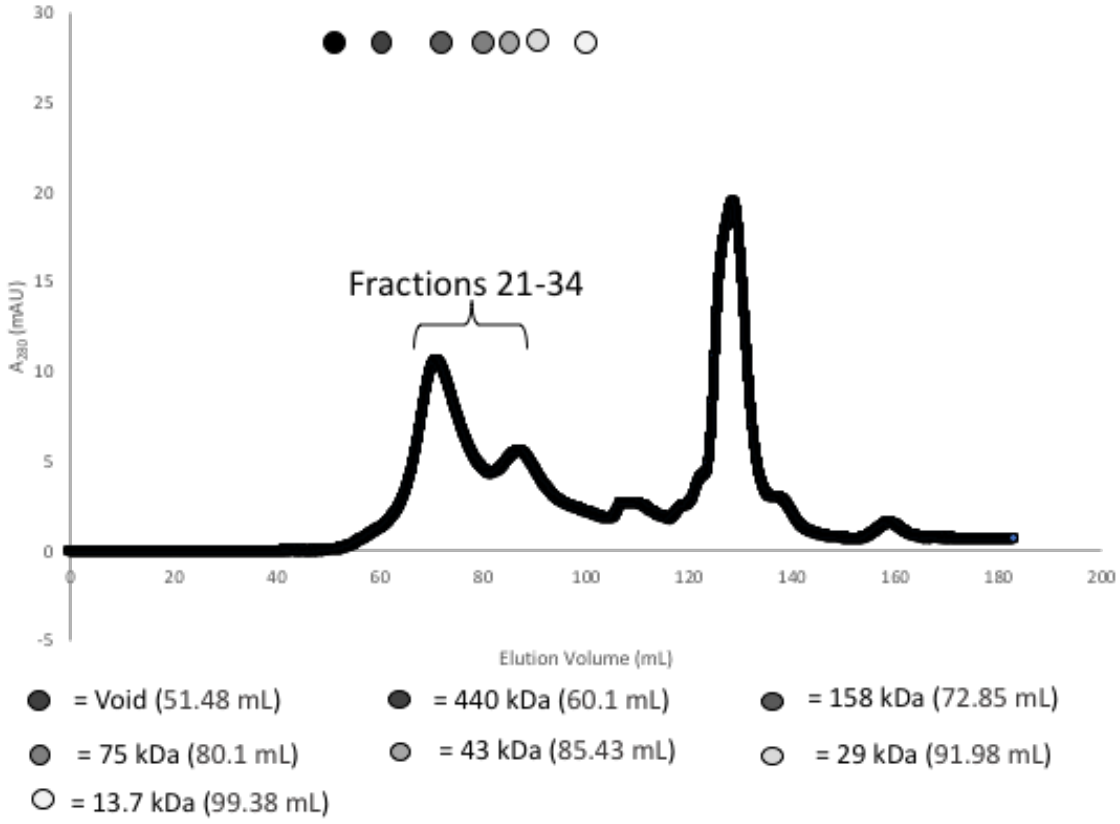

Figure 43. S200 size exclusion chromatogram of His_6-tagged *Hs*LARP6-C490S-DTT. Pooled fractions of the affinity chromatography column were collected and concentrated against a 10,000 molecular weight cut off filter, and injected the Atka FPLC, and run through the Sephadex 200 gel filtration column. Fractions 21-34 (65 – 93 mL) were collected and analyzed via electrophoresis on denaturing SDS-PAGE.

Fractions 21-33 were selected for further analysis using SDS-PAGE (**Figure 44**). The smaller co-eluting species was again present in the fractions 22-27. These fractions were than pooled, concentrated against a 10,000 molecular weight cut off filter down from an initial 12 mL to final volume of 4 mL, aliquoted in 50 µL aliquots, and snap frozen using liquid nitrogen.

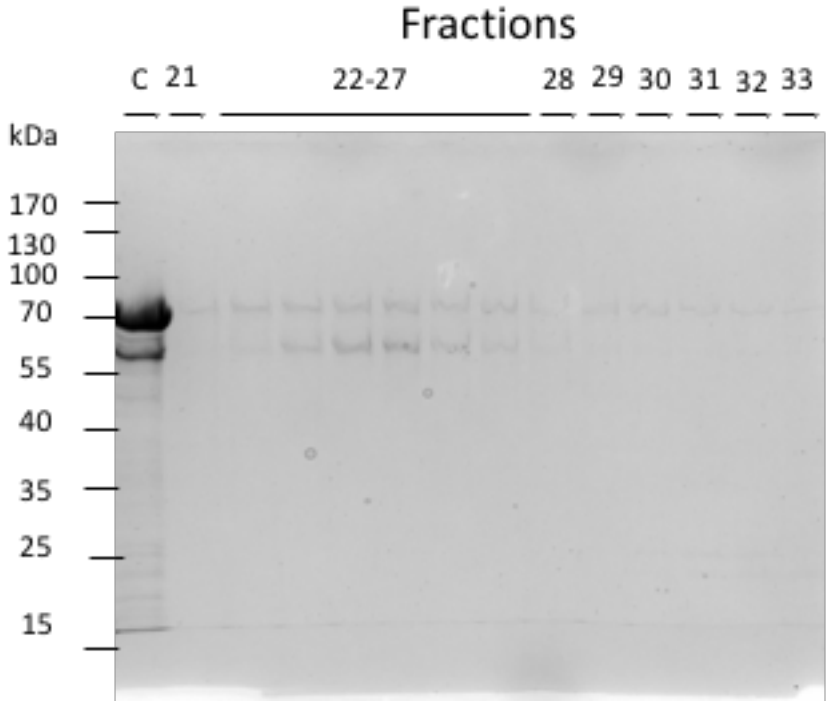

Figure 44. S200 gel filtration analysis of *Hs*LARP6-C490S-DTT. Fractions 21 – 33 (65 – 92 mL) and a sample of what was loaded were analyzed via electrophoresis through denaturing gel and stained with Coomassie blue. Fractions 22-27 contained the protein of interest that migrated at 55 kDa, however the presence of the co-eluting were present in elution fractions 21-33. The fractions 22-27 were pooled and concentrated against 10,000 molecular weight cut off and stored at 50 µL aliquots at -70 °C for further experiments.

To determine how the addition of the reducing agent DTT would affect the apparent hydrodynamic radius of the *Hs*LARP6-C490S mutant, a purification was performed with the presence of 1 mM DTT in the S200/Storage buffer. The initial nickel-affinity purification (**Figure 45**) The S200 chromatogram (**Figure 46**) shows a distribution similar to those observed for *Hs*LARP6(+DTT) and *Hs*LARP6-C490S(-DTT), again indicating that this mutant is eluting at a hydrodynamic radius that would correspond to a protein (or complex) that is five times the expected molecular weight of the monomer. This data demonstrates that the presence or absence of DTT in the storage buffer does not affect the elution volume of the *Hs*ARP6-C490S.

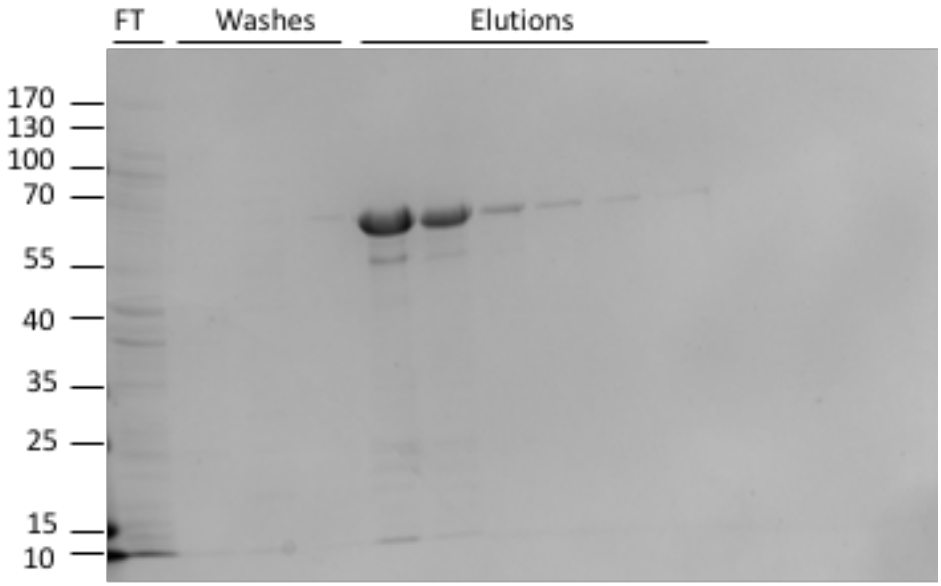

Figure 45. Denaturing gel analysis of Nickel Affinity purification of *Hs*LARP6-C490S. The flowthrough (FT), washes, and elution fractions were collected from the nickel-NTA beads, aliquoted at 20 µL + 5 µL 5x SDS Sample Buffer and ran on denaturing SDS PAGE gel. A high intensity migrating at ~65 kDa can be found in all elution fractions. The appearance of the 55 kDa band can be seen in elutions 1 and 2. Elutions 1-3 were pooled, concentrated against a 10,000 molecular weight cut off until reached a total volume of ~2 mL, and injected onto the Sephadex S200 size exclusion column.

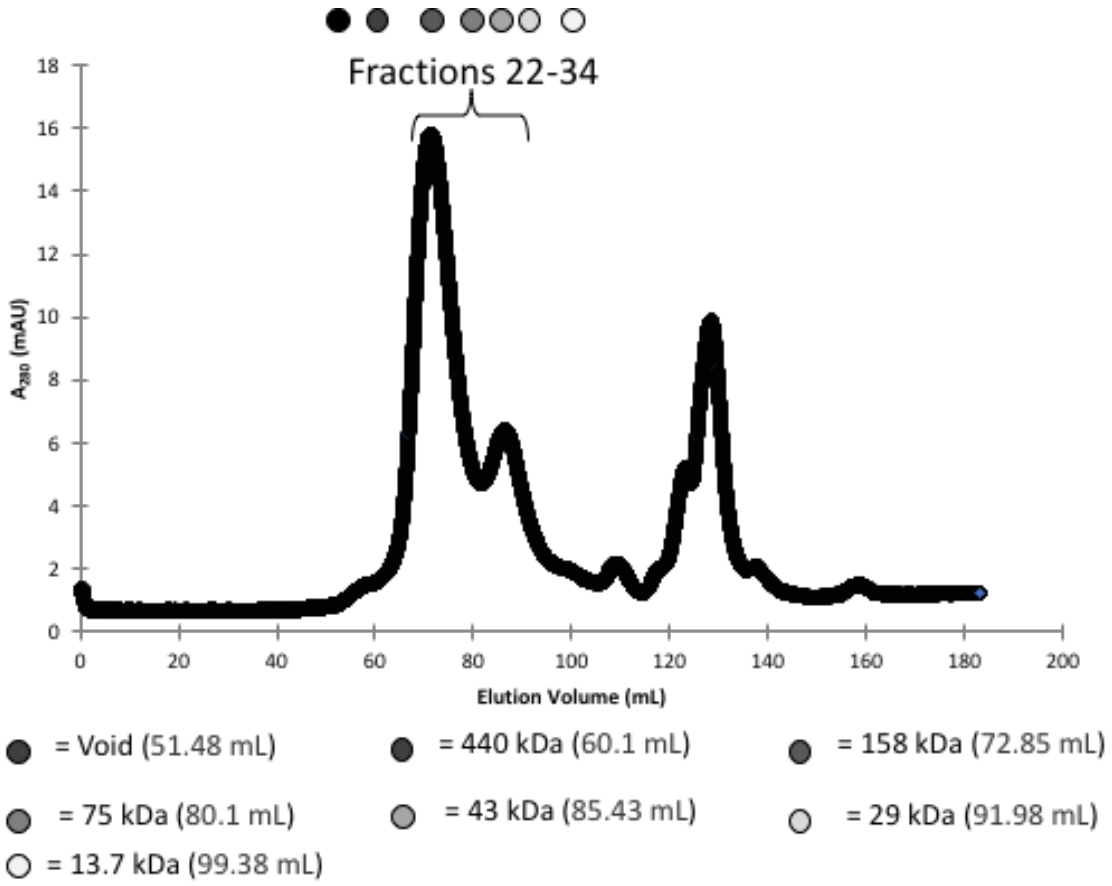

Figure 46. S200 gel filtration chromatogram of the His_6-tagged *Hs*LARP6-C490S in the presence of DTT. Pooled fractions from the affinity chromatography were concentrated against a 10,000 molecular weight cut off filter and injected into the Atka FPLC for gel filtration through the Sephadex 200. The intensity of Peak 1 appears to be significantly larger than that of Peak 2, indicating that the population of protein is adopting a form concurrent with a protein of a higher R_h value. Fractions 22 – 34 (66- 93 mL) were selected for the analysis on denaturing gel electrophoresis.

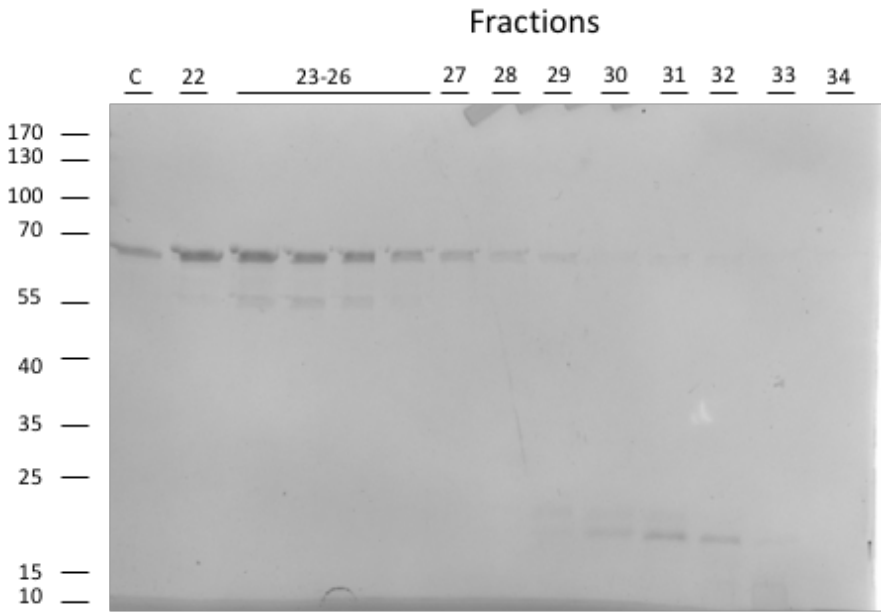

Figure 47. Denaturing gel analysis of *Hs*LARP6-C490+DTT post gel filtration. Fractions 22 – 34 (xx-92 mL) and a loading control (2 mL of sample injected into the FPLC) were collected from the S200 gel filtration and analyzed via electrophoresis through denaturing gel and stained with Coomassie blue. Fractions 23-26 contained the protein of interest that migrated at 55 kDa, however the presence of the co-eluting were pooled and concentrated against 10,000 molecular weight cut off and stored at 50 µL aliquots at -70 °C for further experiments.

Purification of the HsLARP6-C258S/C490S double mutant

Since the trend of peak shifting upon addition of DTT was seen in the *Hs*LARP6-C258S and wildtype HsLARP6 protein constructs but not in the *Hs*LARP6-C490S, the C490S mutation was combined with the C258S to form a double mutant, *Hs*LARP6-C258S/C490S. In collaboration with undergraduate LaPatience Lane, this construct was expressed, purified, and subjected to the same analysis as above to determine whether the addition of DTT affects the population of the Peak 1 and Peak 2 in the elution profile of the preparative column. Initial purification of the HsLARP6-C258S/C490S double

mutant began with the nickel-affinity chromatography (**Figure 48**). The two co-eluting species previously found in the single mutants can be seen with here in elution fractions 1 and 2.

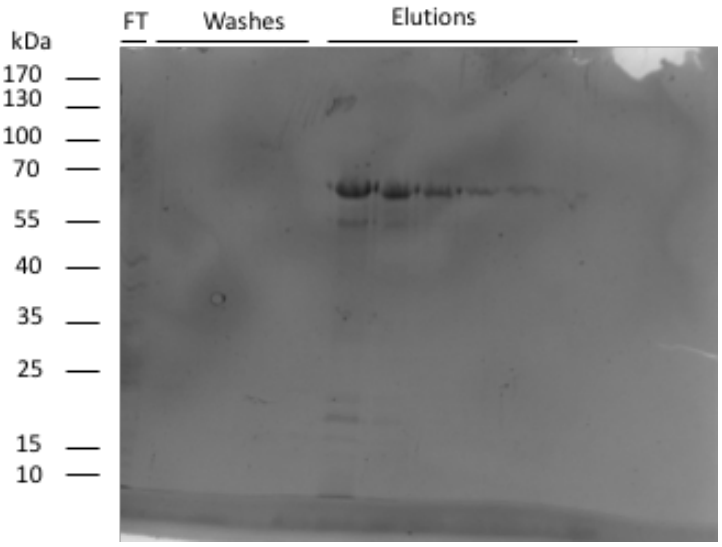

Figure 48. Denaturing gel analysis of post-nickel-affinity purification of *Hs*LARP6-C258S/C490S. Cells containing the His_6-tagged HsLARP6-C258S/C490S protein construct were lysed in *Hs*LARP6 Lysis/Wash Buffer (**Table 4**) and then incubated with equilibrated nickel-NTA beads. The lysate + nickel-NTA bead mixture was decanted and allowed to settle in a glass gravity flow column. The flowthrough (FT), washes, and elution fractions were collected, and aliquots were taken at 20 µL + 5 µL of 5x SDS sample buffer and separated on denaturing SDS-PAGE. The presence of a ~65 kDa can be seen in elution fractions 1-5, with decreasing intensity. The presence of a ~55 kDa band was seen in fractions 1 and 2.

Further purification using the Sephadex S200 size exclusion column indicated the presence of the two elution peaks. The trend in the elution UV chromatogram is consistent with that has been shown in purification of the single mutants and the wildtype HsLARP6, where the population of the peak favors the earlier elution peak (Peak 1) (**Figure 49**). Denaturing gel analysis of fractions 22-35 reveal the presence of a 70 kDa band that is consistently present in all fractions (**Figure 50**). The presence of the 55 kDa

band (expected) is present in only fractions 24-27. Because fractions 22-25 correlated with Peak 1, these fractions were selected for concentration and storage.

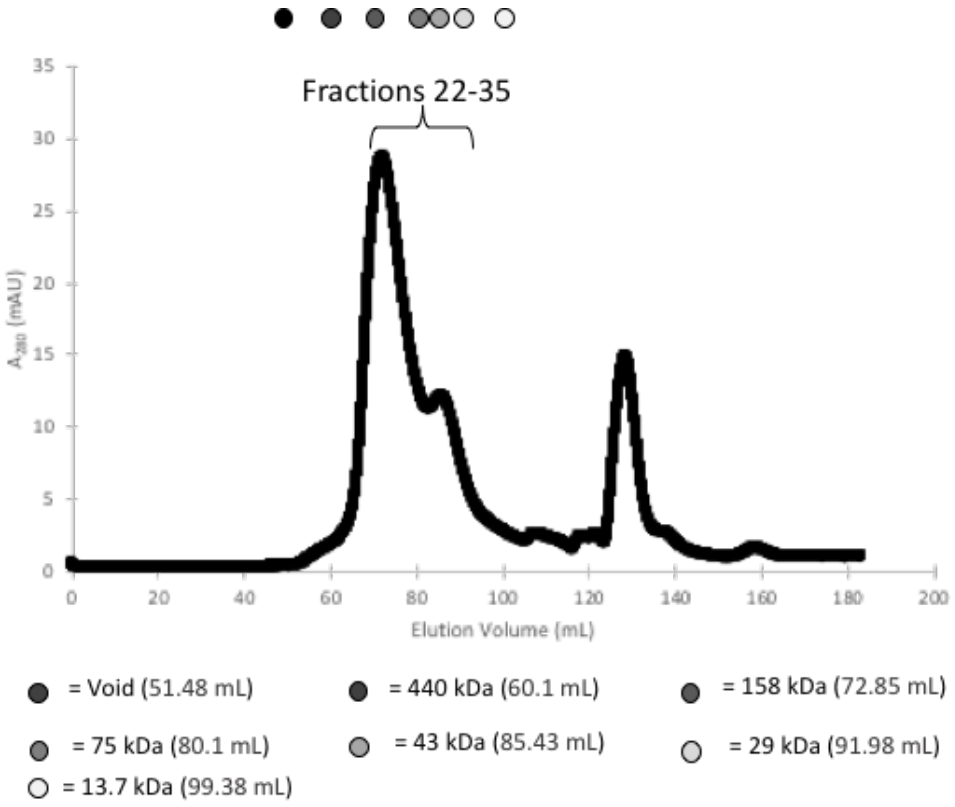

Figure 49. Gel filtration UV chromatogram of *Hs*LARP6-C258S/C490S+DTT. The pooled fractions of the nickel-affinity chromatography were concentrated against a 10,000 Da molecular weight cut off filter, filtered and degassed using a 0.20 µm filter, and injected onto the Sephadex S200 size exclusion column. The column was monitored using UV spectroscopy. The both Peak1 and Peak 2 can be seen in the UV chromatogram. Fractions 22-35 (67-94)

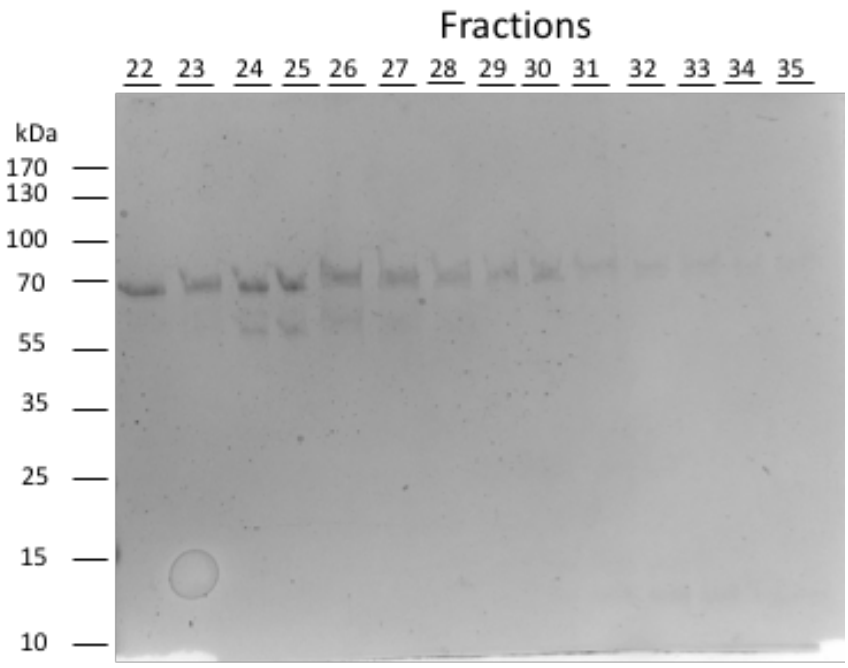

Figure 50. Denaturing gel analysis of SEC purified *Hs*LARP6-C258S/C490S+DTT. Based on the UV chromatogram, fractions 22-35 were selected for denaturing gel analysis. The presence of a 70 kDa band can be seen in all fractions that were analyzed. The 55 kDa can be seen in fractions 24-27. Because fractions 22-25 correspond with the first peak, they were pooled, concentrated against a 10,000 molecular weight cut off filter from 6 mL to 1.5 mL, aliquoted at 50 µL aliquots and stored in -70 °C.

To quantitatively analyze the effect of DTT on the distribution of LARP6 between Peak 1 (69 mL; ~256 kDa) and Peak 2 (87 mL; ~60 kDa), the area under the curve for each peak was plotted as a fraction of the total area for each preparation of *Hs*LARP6 (**Figure 51**). In the wildtype *Hs*LARP6 protein and *Hs*LARP6-C258S mutant, the population of the protein appears to inhabit Peak 2, which is associated with the smaller approximate molecular weight (black). However, for the *Hs*LARP6-C490S and *Hs*LARP6-C258S/C490S double mutant, the area under the curve of Peak 2 is seen to be unaffected. These data indicate that the mutation of C490S has made these two mutants resistant to the apparent size changes induced by DTT. One potential explanation is that

C490 may be engaged in a disulfide bond with another cysteine residue somewhere else in the protein. Further analysis of how the individual C490S mutation affects the overall stability of the protein may give insight into the role of this cysteine residue in the overall function of the protein.

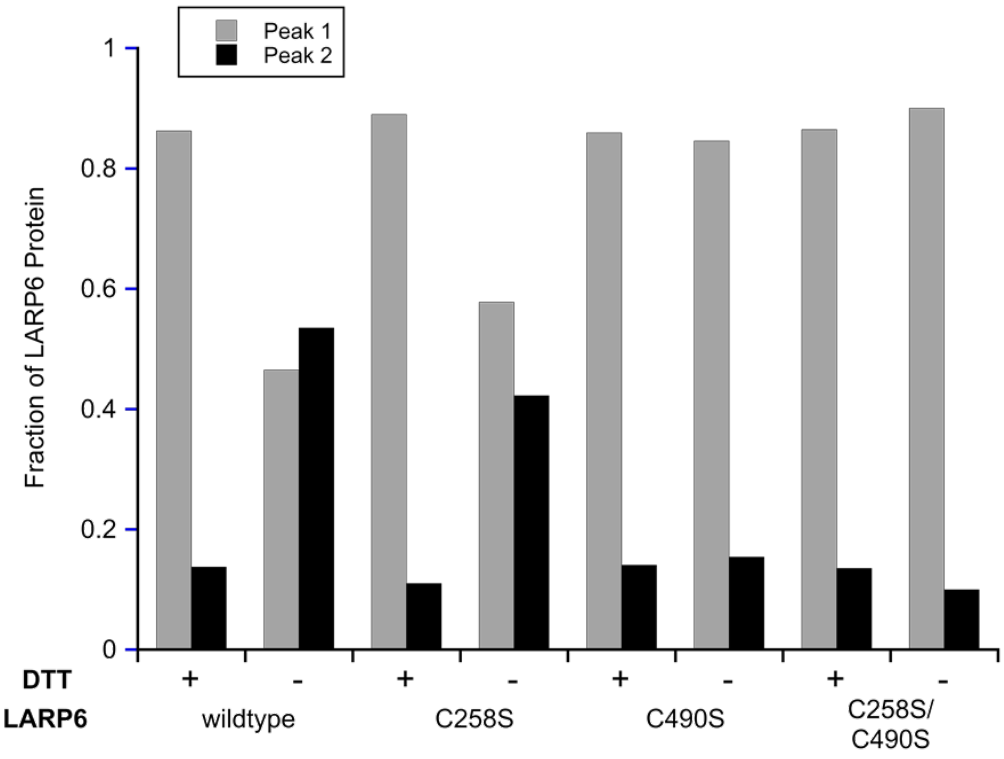

Figure 51. Effect of DTT on the elution peak ratio of *Hs*LARP6 cysteine mutants. The area under the curve was quantified for both peaks observed in the elution profile of the preparative column and plotted as a fraction of observed area over the total area between both peaks. The *H*sLARP6-C258S can be seen to have an increase in the area under second peak when DTT is removed from the storage buffer, which follows the same phenomenon seen with the wildtype *H*sLARP6. However, no effect of DTT on the second peak was observed for the *Hs*LARP6-C490S nor the double mutant containing the same mutation (*Hs*LARP6-C258S/C490S).

Identification of the two species in* Hs*LARP6 protein preps

Because all the preparations of *Hs*LARP6 contained two apparent protein species when analyzed by SDS-PAGE, it was important to identify each species. To identify the N- and C-termini of these species, western blots were performed using two detection reagents: an anti-His_6 probe to detect the N-terminal His_6 tag, and a primary anti-LARP6 antibody to detect the C-terminus of LARP6.

The anti-His western blot detected both species in almost all of the *Hs*LARP6 proteins (**Figure 52**). The exceptions were three samples from the *Hs*LARP6-C258S mutant preps: Peak 1 (–DTT) and Peaks 1 and 2 (+ DTT). The *Hs*LARP6C258S Peak 2 preparation showed only 1 band at 65 kDa. The goat anti-LARP6 antibody recognizes a 15-amino acid epitope at the extreme C-terminus of the *Hs*LARP6 protein. This probe was used to determine whether the C-terminus was intact in either species of *Hs*LARP6 in each preparation. In all preparations of *Hs*LARP6, both the ~65 kDa and the ~55 kDa species react with the anti-C-terminal antibody. These data demonstrate that not only do both bands contain LARP6, but they both contain the C-terminus of LARP6. An additional result is that three samples (Peak 2 *Hs*LARP6-C258S(-DTT) and Peaks 1 and 2 of *Hs*LARP6-C258S(+DTT) show both the 65 kDa band and 55 kDa band, but these bands were not detectable in the anti-His_6 probe western blot.

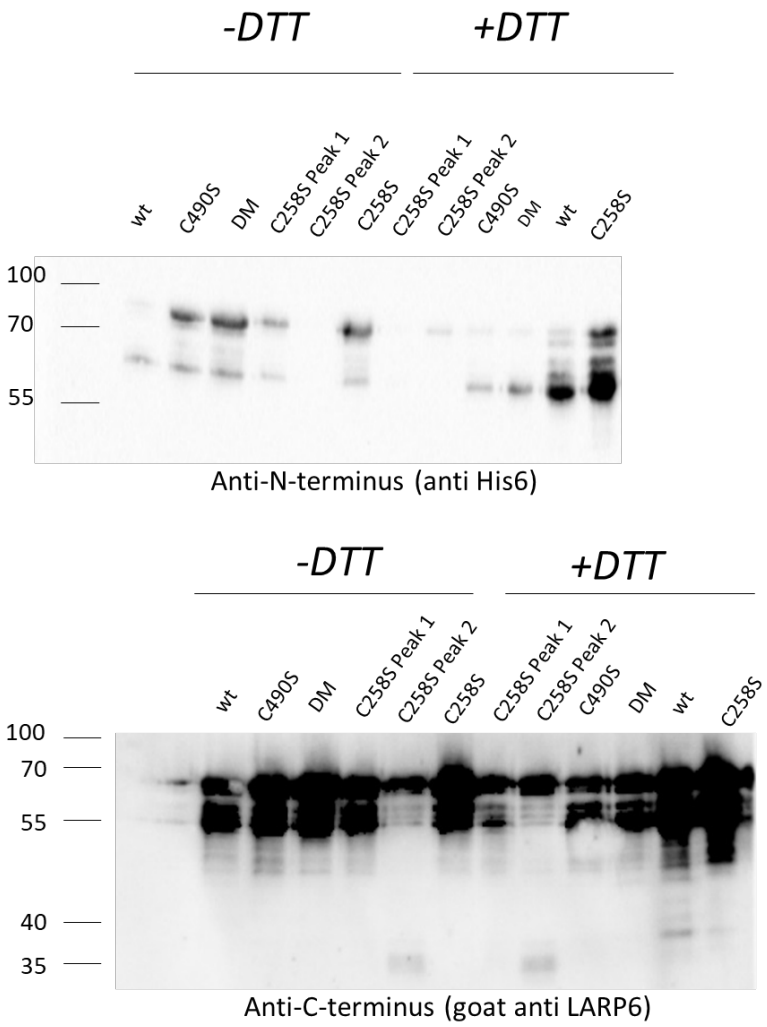

Figure 52. Stability of *Hs*LARP6 constructs after long term freezer storage. Proteins were thawed from the -70 °C storage and subjected to denaturing SDS-PAGE in duplicate. After transfer to nitrocellulose, one membrane was probed with anti-His_6 probe (*top*) and the other was probed with anti-LARP6 secondary antibody (*bottom*). Both probes were detected using chemiluminescence as previously described

We first hypothesized that the larger molecular weight species may be the result of ribosomal read-through of the mRNA. To test this hypothesis, we took advantage of two features of the cloned plasmid vector. First, following the stop codon but upstream of the polymerase termination signal, there is an in-frame C-terminal His_6-tag coding sequence. Second, the N-terminal His_6-tag is separated from the LARP6 sequence by a thrombin protease cleavage sequence. If translational read-through is responsible for creating the 65 kDa species, incubation with thrombin should remove the N-terminal His_6-tag on both the 55 kDa and the 65 kDa bands. However, the C-terminal His6 tag would remain on only the 65 kDa band, which would show signal when detected with the anti-His_6 probe. After a 6-hour incubation with thrombin, none of the protein preparations showed anti-His reactivity (**Figure 53**). To control for any possibility of the proteins being degraded by the thrombin or other contaminating proteases, a duplicate gel was run, transferred, and subjected to western blot with the C-terminal anti-LARP6. This experiment confirmed that LARP6 was still present in all thrombin-treated samples. Therefore, both the 55 kDa and the 65 kDa species have only the N-terminal His_6-tag, and the larger species is not the result of translational read through. Further characterization was needed to identify the different between these bands.

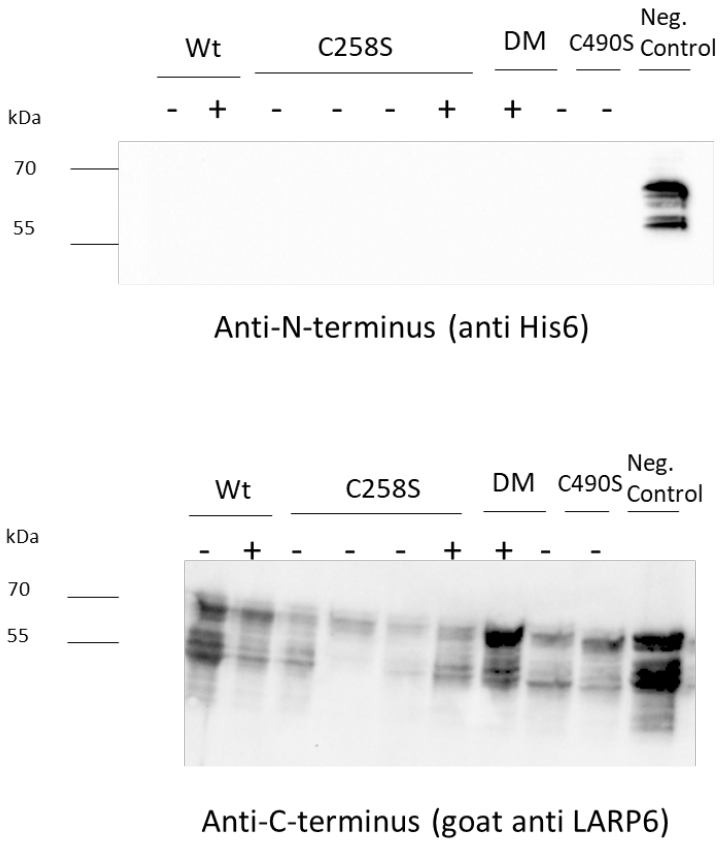

Figure 53. Thrombin Digest of LARP6 proteins. The N-terminal His_6-tag was removed from the LARP6 protein constructs by digesting for 6 hours on ice with thrombin protease. Digestion reactions were separated on 10% SDS-PAGE gel at 200 V for 55 min. Gels were transferred onto nitrocellulose and blotted for either the His tag (*top*) or LARP6 C-terminus (*bottom*).

ß-mercaptoethanol (BME) is a common denaturing reagent used in SDS-PAGE sample buffer, though the BME is often regarded as a mild denaturing agent and certain folding structures in the protein could resist complete reduction of disulfide bonds by BME. We next hypothesized that replacement of BME with the strong denaturing agent TCEP (tris (2-caboxyethyl) phosphine) would resolve the two species into one single band. To test this hypothesis, wildtype *Hs*LARP6 samples were prepared for SDS-PAGE using either the standard BME-containing 5x SDS-PAGE sample buffer or a 5x SDS-PAGE sample buffer prepared with 50 mM TCEP in place of the BME. Both samples were heated at 90 °C and run in triplicate for detection with either Coomassie blue stain, anti-His western blot, or anti-LARP6 western blot.

The Coomassie blue stain (**Figure 54,** *left*) shows that the expected collapsing of the two co-eluting bands into one 55 kDa band could not be achieve when ß-ME was replaced with TCPE in either wildtype HsLARP6 preparation (presence or absence of DTT). Unexpectedly, the TCEP-containing samples show a new band (**Figure 54**, *arrow*). This band reacts with both the anti-His_6 probe and anti-LARP6 C-terminal antibody. The two primary bands from previous gels (55 kDa and 65 kDa) remain unaffected. For both detections of the *Hs*LARP6 protein (**Figure 54**, *middle* and r*Right*), the same two migrating species were identified. It is apparent that the two-original species in the purification were unaffected by the addition of TCEP and may be overall resistant to reduction.

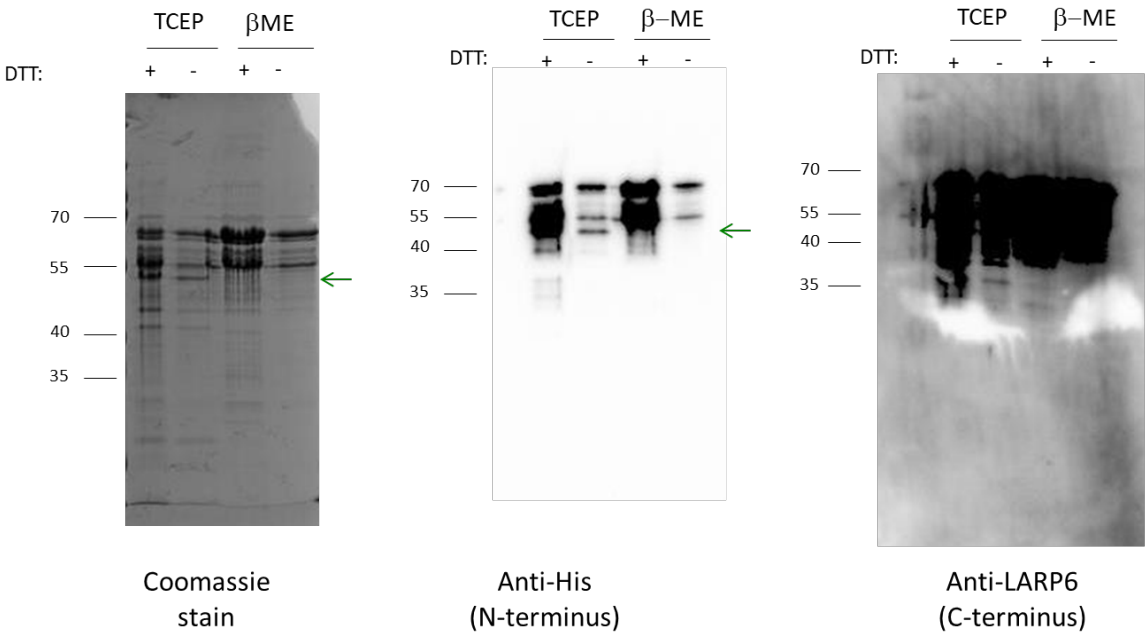

Figure 54. Reduction of wildtype *Hs*LARP6 protein preps using either TCEP or ß-mercaptoethanol. To test if wildtype LARP6 protein contains reductant-resistant disulfide bonds, the ß-mercaptoethanol (ß-ME) in the standard SDS sample buffer was replaced with TCEP (tris(2-carboxyethyl) phosphine) at a final concentration of 50 mM Samples from two preparations (with or without DTT) were incubated at 90 °C for 10 min in either the TCEP or ß-ME sample buffer, and then analyzed by 10% SDS-PAGE in triplicate. (*Left*) Coomassie blue staining of the protein was used to identify total protein content. Western blots were performed on the other two gels. (*Middle*) α-His probe was used to target the N-terminal His6-tag and (*Right*) rabbit α-LARP6 targeted the C-terminal 15 residues of LARP6. Contrary to our prediction, the multiple bands in each protein preparation did not collapse into a single species. In fact, the TCEP buffer appeared to induce the formation of a third migration species in both the +DTT and -DTT purification products (green arrow)

To fully confirm that these two species are indeed solely LARP6 protein that contains both the N- and C-terminus, we analyzed the bands by mass spectrometry. The wildtype *Hs*LARP6 proteins stored in the presence and absence of DTT were separated using denaturing SDS-PAGE gel (**Figure 55**). The two co-eluting species (~65kDa and ~55 kDa band) were excised in preparation for in-gel trypsineolysis and LC-MS by our collaborator Xinzhu Pu (Boise State University). The recovered peptides mapped to the human LARP6 sequence shows the overall identified sequence coverage for each of the 65 kDa band in the presence and absence of DTT was 77% (**Figure 55**). The identified sequence coverage for each of the 55 kDa bands in the presence and absence of DTT was ~61% and 74%, respectively. No peptides from other proteins were detected, proving that both bands contain full-length human LARP6. The reason for the differential migration in SDS-PAGE remains to be determined.

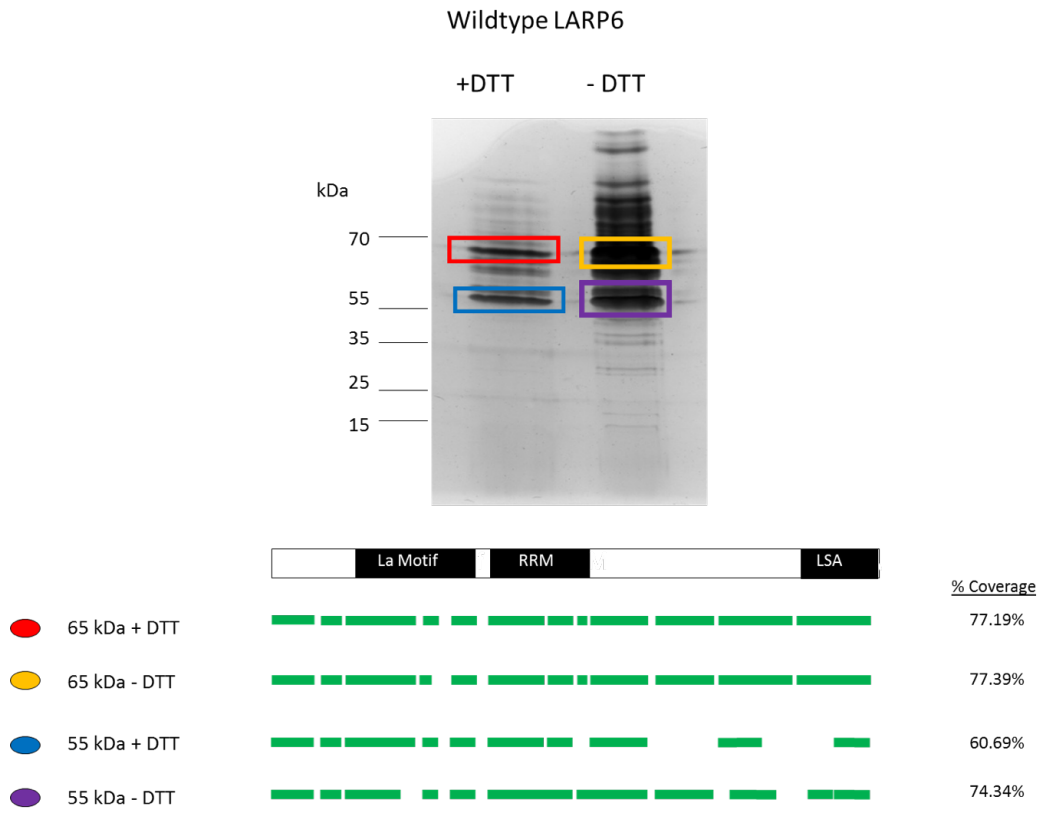

Figure 55. Protein mapping of *Hs*LARP6 species. Protein samples from +DTT and -DTT preparations were separated on SDS-PAGE and visualized by silver stain. The two-major species were excised and subjected to complete trypsin digest. The fragments were then analyzed by tandem liquid chromatography-mass spectrometry, and then detected peptides were mapped back onto the LARP6 sequence. All four species contain full-length LARP6, suggesting that the difference in migration is due to differences in conformation and/or covalent modification.

Conclusion

For this project, the mutagenesis of the C258 and C490 residues into serines was successful. While the cysteine mutants were successfully expressed, the resulting preparations do not appear pure by SDS-PAGE. A 55 kDa band in present is all protein preps and appears to be mostly associated with the Peak 1 found in the size exclusion elution UV chromatogram. Both the 65 kDa and the 55 kDa species in the denaturing SDS-PAGE appear to be the full-length human LARP6 protein. There does not appear to be any presence of a C-terminal read-through product or unique susceptibility or resistance to reducing agents. Currently, further characterization of these two bands is necessary to identify the cause of these two species of LARP6 that migrate differently in SDS-PAGE, and how the mutation of cysteine to serine may affected the production of these two species.

V. CONCLUSIONS AND DISCUSSION

Characterization of Vertebrate LARP6 La Modules

In this work, the human and zebrafish LARP6 La Modules were successfully recombinantly expression and purified. Interestingly, two peaks were observed in the preparative size exclusion elution chromatogram of the *Hs*LARP6 La Module. In contrast, a single, high resolution peak was seen in the *Dr*LARP6 protein, which was similar to that observed for the purification of the platyfish (*Xm*) La Module (previously achieved in the lab by J. M. Castro).[16] While each elution peak for the two fish proteins corresponded almost precisely to the expected molecular weight, the first peak observed in the human La Module preparation corresponded to a molecular weight that was 5-fold greater than the expected molecular weight. Because there was only a single band in the SEC gel analysis, this result may indicate that the human La Module may experience oligomerization. Alternatively, the presence of the early elution peak may also be an artifact of protein aggregation, as it appears to elute near the void volume (46.25 mL). To conclusively determine oligomerization states, the actual molecular mass of the human LARP6 La Module would need to be analyzed by analytical ultracentrifugation.

To better prepare the La Module constructs for structural characterization using solution NMR and small-angle X-ray scattering, it was necessary to establish a protocol for the removal of the His_6-tag to reduce the amount of unstructured polypeptide. We successfully established a His_6-tag cleavage purification of the human and zebrafish La Modules with His_6-tag removal as evidence by diminished signal of the anti-His_6

detection of the La Module constructs after the purification including the thrombin-cleavage.

Initial RNA binding assays with the human, zebrafish, and platyfish La Modules against the *HsCOL1a1* RNA ligand were performed. The binding affinity of the two fish La Modules initially appeared to be much tighter than that of the human La Module. This trend is consistent with the binding data for the full length human and fish LARP6 proteins, which were also assayed using the *HsCOL1a1* ligand.[16] Together with the full-length data, these data indicate that the fish LARP6 proteins are able to bind tighter to the human Col1a1 mRNA ligand than the human LARP6 protein.

In fact, the fitted $K_{D,\,app}$ for the fish La Modules was approaching the assumed RNA concentration in the assay of ~1.25 nM. In order to appropriately apply the simplified form of the binding isotherm equation, it is crucial to be able to make the assumption that the known concentration of the RNA ligand is a minimum of 10-fold below the $K_{D,app}$ value, so that we can assume that $[RNA]_{free} \sim [RNA]_{total}$. However, when the fitting of the approximate K_D value approaches the known concentration of the RNA, that assumption is likely not valid. To confirm this, the concentration of the biotinylated RNA ligand stocks was directly quantified using UV spectroscopy. These data showed that the concentration of the biotinylated RNA ligands was too high to make that assumption at a concentration of 61.8 μM. Therefore, in order to proceed this issue must be addressed. There are two ways to do this: either change the experimental conditions to use a working concentration of the RNA that meets the criteria stated above or employ

the quadratic form of the binding isotherm (**Equation 2**) for fitting of the $K_{D,\,app}$ to the experimental data. It was decided that the latter option would be attempted first as reducing the concentration RNA in the system could result in a higher signal to noise ratio.

When the experimental data was fitted using the quadratic form of the binding, it was conclusive that the two fish La Modules were indeed binding with tighter apparent affinity to the Hs*COL1a1* mRNA ligand. The $K_{D,app}$ reported for this study for the human La Module was 170 ± 31 nM, which is ~3 times weaker than the $K_{D,app}$ of 48 nM reported by Martino *et al*.[9] It was hypothesized that this may be the result of the presence of contaminating *E. coli* RNA in our protein preparation from the purification. In order to test this, a UV absorbance spectrum of each protein preparation was taken from 200-400 nm to quantify the $A\frac{260}{280}$ ratio. The human La Module had an A_{260}/A_{280} of 0.94 which did not indicate nucleic acid contamination. However, upon analyzing the spectra (**Figure 56**), a large absorbance between 210-245 nm was observed in the two fish La Module preparations, strongly suggesting the presence of a contaminant (**Figure 56,** *top right* and *bottom*). The identity of this contaminant remains to be characterized, but this large absorbance also likely skewed the A_{260}/A_{280} and the A_{280} measurements of these protein samples. In order to reduce the presence of this contaminant in the future and to ensure an accurate measurement of protein concentration and therefore apparent binding affinity, a different storage buffer may be necessary for future protein preparations.

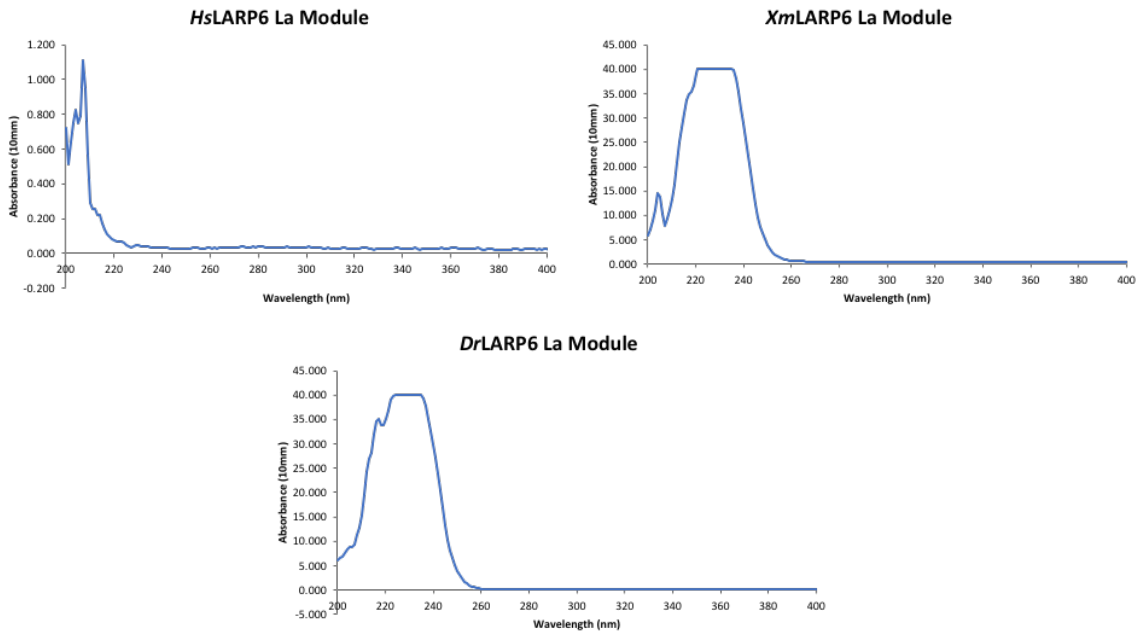

Figure 56. UV absorbance spectra of La Module proteins. The absorbance spectra of the protein preparations were blanked with 1 µL of the appropriate storage buffer in a final concentration of 7.5 M urea. Each protein preparation was denatured in 7.5 M urea for 30 min before the absorbance spectrum was measured using the Implen Nanophotometer. The human (*Hs*) La Module (*Top Left*) can shows a high absorbance at ~200 nm, corresponding to the absorbance of the carbonyls in urea. The zebrafish (*Dr*; *bottom*) and platyfish (*Xm*; *top right*) La Modules show high absorbance between 210-245 nm. The identity of the highly-absorbing contaminant remains unknown.

Cysteine 490 in human LARP6 may be involved in disulfide bonding

Initial studies indicated that the apparent hydrodynamic radius of LARP6 may be affected by reducing agent. To test this, two highly conserved cysteines were mutated to serine. When comparing the elution profiles of the wildtype protein against all mutants, it was observed that the apparent hydrodynamic radius of any mutants containing the C490S mutation was unaffected by DTT. Since DTT acts as a reducing agent for the cysteine residues, it can be inferred that where a cysteine residue was replaced by a serine residue, the ability to form disulfide bonds with one or more other cysteine residues was

disabled. Without a disulfide bond compacting the structure, a less constrained conformation of the protein with a larger hydrodynamic radius would be observed.

The cysteine mutant purifications were complicated by the consistent presence of two peaks that were observed in the elution profiles of the preparative size exclusion column. The first peak corresponded to a 5-fold larger apparent molecular weight than expected, while second peak only exhibited a ~10 kDa difference from the expected molecular weight. It is possible that the 5-fold difference observed in the first peak was the result of oligomerization between the full-length protein molecules. A similar trend was also seen in the elution profiles of the wildtype human La Module protein, which may suggest that the human La Module is a key domain involved in oligomerization. If that is the case, the C490S mutant may have induced and favored oligomerization of the protein, as that mutant predominantly eluted in the first peak. Further investigation would be necessary to determine how the C490 residue and the La Module play a role in oligomerization. Subjecting the protein to analytical ultracentrifugation can give us a better understanding of what oligomeric states, if any, the protein adopts. If native oligomerization can be verified, further investigation of the role of the La Module and C490 residue in the oligomerization of the human LARP6 will be required to fully understand the structure of this protein.

The presence of two co-eluting species in all stages of purification of the full-length cysteine mutants, including post size exclusion purification, was troublesome. However, all our studies strongly indicate that these two co-eluting species are indeed

both full-length, and only full-length, *Hs*LARP6. It is conclusive that the larger band is not a read-through product, and additionally appears to not be affected in the presence of the strong reducing agent DTT. The structural identity of these two species remains unknown, although it is hypothesized that these two species are the result of proteins that are highly resistant to reduction. In order to fully test this, a screening assay must be performed on the proteins using various different reducing agents. It is also possible to determine how various reducing agents effect the R_h of the *Hs*LARP6 through a doping experiment using analytical HPLC. Attempts to isolate the two co-eluting species through purification using a low flow-rate (low pressure) for better separation may also help better characterize the two species without influence from the each other.

In summary, these studies have laid a foundation for further explorations for the roles of the N-terminal (NTD) and C-terminal (CTD) domains in LARP6 structure and function. Structural and biochemical characterization of the isolated NTD, CTD, and LSA domains will establish the contribution of these components to the structure and function of the full-length LARP6 proteins of vertebrates.

REFERENCES

1. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.

2. Hocine, S., R.H. Singer, and D. Grunwald, *RNA processing and export.* Cold Spring Harb Perspect Biol, 2010. **2**(12): p. a000752.

3. Licatalosi, D.D. and R.B. Darnell, *RNA processing and its regulation: global insights into biological networks.* Nat Rev Genet, 2010. **11**(1): p. 75-87.

4. Alves, L.R. and S. Goldenberg, *RNA-binding proteins related to stress response and differentiation in protozoa.* World J Biol Chem, 2016. **7**(1): p. 78-87.

5. Maraia, R.J., et al., *The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives.* Wiley Interdiscip Rev RNA, 2017.

6. Bousquet-Antonelli, C. and J.M. Deragon, *A comprehensive analysis of the La-motif protein superfamily.* Rna, 2009. **15**(5): p. 750-64.

7. Bayfield, M.A., R. Yang, and R.J. Maraia, *Conserved and divergent features of the structure and function of La and La-related proteins (LARPs).* Biochim Biophys Acta, 2010. **1799**(5-6): p. 365-78.

8. Maraia, R.J., et al., *The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives.* Wiley Interdiscip Rev RNA, 2017. **8**(6).

9. Martino, L., et al., *Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module.* Nucleic Acids Res, 2015. **43**(1): p. 645-60.

10. Hussain, R.H., M. Zawawi, and M.A. Bayfield, *Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7.* Nucleic Acids Res, 2013. **41**(18): p. 8715-25.

11. Cai, L., et al., *Binding of LARP6 to the conserved 5′ stem-loop regulates translation of mRNAs encoding type I collagen.* J Mol Biol, 2010. **395**(2): p. 309-26.

12. Zhang, Y. and B. Stefanovic, *LARP6 Meets Collagen mRNA: Specific Regulation of Type I Collagen Expression.* Int J Mol Sci, 2016. **17**(3): p. 419.

13. Cai, L., et al., *Nonmuscle myosin-dependent synthesis of type I collagen.* J Mol Biol, 2010. **401**(4): p. 564-78.

14. Gerstberger, S., M. Hafner, and T. Tuschl, *A census of human RNA-binding proteins.* Nature Reviews Genetics, 2014. **15**: p. 829.

15. Parsons, C.J., et al., *Mutation of the 5'-untranslated region stem-loop structure inhibits alpha1(I) collagen expression in vivo.* J Biol Chem, 2011. **286**(10): p. 8609-19.

16. Castro, J.M., et al., *Recombinant expression and purification of the RNA-binding LARP6 proteins from fish genetic model organisms.* Protein Expr Purif, 2017. **134**: p. 147-153.

17. Kozlov, A.G., et al., *Glutamate promotes SSB protein-protein Interactions via intrinsically disordered regions.* J Mol Biol, 2017. **429**(18): p. 2790-2801.

18. Reed, B.J., M.N. Locke, and R.G. Gardner, *A Conserved Deubiquitinating Enzyme Uses Intrinsically Disordered Regions to Scaffold Multiple Protein Interaction Sites.* J Biol Chem, 2015. **290**(33): p. 20601-12.

19. Ishino, S., et al., *Multiple interactions of the intrinsically disordered region between the helicase and nuclease domains of the archaeal Hef protein.* J Biol Chem, 2014. **289**(31): p. 21627-39.

20. Zhang, Y. and B. Stefanovic, *Akt mediated phosphorylation of LARP6; critical step in biosynthesis of type I collagen.* Sci Rep, 2016. **6**: p. 22597.

21. Zhang, Y. and B. Stefanovic, *mTORC1 phosphorylates LARP6 to stimulate type I collagen expression.* Sci Rep, 2017. **7**: p. 41173.

22. Pei, J., B.H. Kim, and N.V. Grishin, *PROMALS3D: a tool for multiple protein sequence and structure alignments.* Nucleic Acids Res, 2008. **36**(7): p. 2295-300.

23. Liu, C., et al., *Collagen metabolic disorder induced by oxidative stress in human uterosacral ligamentderived fibroblasts: A possible pathophysiological mechanism in pelvic organ prolapse.* Mol Med Rep, 2016. **13**(4): p. 2999-3008.

24. Siwik, D.A., P.J. Pagano, and W.S. Colucci, *Oxidative stress regulates collagen synthesis and matrix metalloproteinase activity in cardiac fibroblasts.* Am J Physiol Cell Physiol, 2001. **280**(1): p. C53-60.

25. Watari, T., et al., *Evaluation of the effect of oxidative stress on articular cartilage in spontaneously osteoarthritic STR/OrtCrlj mice by measuring the biomarkers for oxidative stress and type II collagen degradation/synthesis.* Exp Ther Med, 2011. **2**(2): p. 245-250.

26. Cremers, C.M. and U. Jakob, *Oxidant sensing by reversible disulfide bond formation.* J Biol Chem, 2013. **288**(37): p. 26489-96.

27. Schippers, J.H., et al., *ROS homeostasis during development: an evolutionary conserved strategy.* Cell Mol Life Sci, 2012. **69**(19): p. 3245-57.

28. Cumming, R.C., et al., *Protein disulfide bond formation in the cytoplasm during oxidative stress.* J Biol Chem, 2004. **279**(21): p. 21749-58.

29. Sevier, C.S. and C.A. Kaiser, *Formation and transfer of disulphide bonds in living cells.* Nat Rev Mol Cell Biol, 2002. **3**(11): p. 836-47.

30. Zheng, M., F. Aslund, and G. Storz, *Activation of the OxyR transcription factor by reversible disulfide bond formation.* Science, 1998. **279**(5357): p. 1718-21.

31. Gopalakrishna, R. and S. Jaken, *Protein kinase C signaling and oxidative stress.* Free Radic Biol Med, 2000. **28**(9): p. 1349-61.

32. Altschuler, S.E., K.A. Lewis, and D.S. Wuttke, *Practical strategies for the evaluation of high-affinity protein/nucleic acid interactions.* J Nucleic Acids Investig, 2013. **4**(1): p. 19-28.

33. Ream, J.A., L.K. Lewis, and K.A. Lewis, *Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions.* Anal Biochem, 2016. **511**: p. 36-41.

34. Bornhorst, J.A. and J.J. Falke, *[16] Purification of proteins using polyhistidine affinity tags*, in *Methods in Enzymology*. 2000, Academic Press. p. 245-254.